

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

LA REPRÉSENTATION SOCIALE DE LA VIE PRIVÉE DANS UN CONTEXTE
D'USAGE DES TECHNOLOGIES DE L'INFORMATION ET DES
COMMUNICATIONS

THÈSE PRÉSENTÉE
COMME EXIGENCE PARTIELLE DU
DOCTORAT CONTINUUM D'ÉTUDES EN PSYCHOLOGIE
(PROFIL RECHERCHE)

PAR
EUDES LOUBET

OCTOBRE 2023

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire, de cette thèse ou de cet essai a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire, de sa thèse ou de son essai.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire, cette thèse ou cet essai. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire, de cette thèse et de son essai requiert son autorisation.

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

DOCTORAT CONTINUUM D'ÉTUDES EN PSYCHOLOGIE
(PROFIL RECHERCHE) (Ph. D.)

Direction de recherche :

Michael Cantinotti, Ph. D. directeur de recherche
Université du Québec à Trois-Rivières

Elias Rizakallah, Ph. D. codirecteur de recherche
Université du Québec à Montréal

Jury d'évaluation :

Michael Cantinotti, Ph. D. directeur de recherche
Université du Québec à Trois-Rivières

Elias Rizakallah, Ph. D. codirecteur de recherche
Université du Québec à Montréal

Marie-Pierre Gagnon-Girouard, Ph. D. présidente du jury
Université du Québec à Trois-Rivières

Marcos Balbinotti, Ph. D. évaluateur interne
Université du Québec à Trois-Rivières

Evan Light, Ph. D. évaluateur externe
Université York

Thèse soutenue le 29/08/2023

Sommaire

De nombreux chercheurs dans le domaine de la vie privée en contexte technologique se penchent depuis plusieurs années sur un comportement apparemment paradoxal de la population. D'un côté, les individus se disent inquiets pour leur vie privée, mais de l'autre côté, ils ne manifestent pas des comportements de protection qui semblent en adéquation avec ces préoccupations. Malgré la pluralité des théories qui tentent d'expliquer le phénomène, aucune n'arrive à éclaircir de manière satisfaisante les causes du déséquilibre entre les préoccupations importantes des individus et les faibles comportements de protection qu'ils mettent en place. Pour ces raisons, les chercheurs qualifient le phénomène de « paradoxe de la vie privée ». Or, il est possible de s'interroger sur le prisme de lecture des chercheurs qui conceptualisent ce phénomène comme « paradoxal » ou comme une incohérence de la part des utilisateurs des technologies. En effet, les expériences concluant à un phénomène paradoxal sont essentiellement élaborées en faisant abstraction de la perspective subjective des utilisateurs. Se pourrait-il que l'absence de prise en compte de cette subjectivité des utilisateurs dans le cadre d'analyse des chercheurs sous-tendent en partie cette apparence de paradoxe de la vie privée? Cette thèse propose d'intégrer une théorie permettant d'avoir accès à la subjectivité des utilisateurs pour étudier la conceptualisation que ceux-ci entretiennent à propos de la vie privée en contexte technologique. La théorie des représentations sociales permet justement d'interroger la subjectivité des individus tout en considérant le rapport qu'ils entretiennent avec leur environnement technologique et social. L'objectif de cette thèse vise à recueillir la représentation sociale de la vie privée en contexte technologique chez des étudiants de

l'enseignement supérieur français afin de savoir si les éléments constitutifs de la représentation ainsi que leurs organisations appuient le concept de paradoxe de la vie privée. Au-delà de cette mise en relation des résultats avec le paradoxe de la vie privée, l'objectif poursuivi est plus large consiste également à déterminer les éléments incontournables de la représentation (éléments centraux) et ceux de moindre importance, mais qui permettent de contextualiser le système d'interprétation de la réalité que constitue la représentation pour les étudiants français. La particularité de cette thèse est qu'elle repose sur une lecture du phénomène à la jonction entre la psychologie individuelle et sociale. Ce regard a été opérationnalisé par l'approche structurale des représentations sociales théorisée initialement par le chercheur français Jean-Claude Abric. Notre recherche a procédé en deux études, la première concernant 369 participants et 163 pour la seconde. Conformément à l'approche retenue, une première collecte visait la libre génération par les participants de mots ou expressions après leur avoir présenté l'expression inductrice « la vie privée en contexte d'usage des technologies de l'information et des communications ». Par la suite, les participants classaient les termes qu'ils avaient générés en les ordonnant selon l'importance qu'ils leur attribuaient pour représenter la vie privée en contexte technologique. Une fois cette première collecte de données terminée, 11 items (p. ex., confidentialité, protection, intimité) ont pu être repérés comme susceptibles d'être incontournables pour les individus lorsqu'ils se représentent la vie privée en contexte technologique. La seconde étude, plus confirmatoire, s'est centrée sur ces 11 mêmes items afin d'examiner s'ils étaient incontournables pour se représenter la vie privée en contexte technologique. Pour ce faire, l'expérience testait si les 11

éléments restaient incontournables en tout temps et en toute situation pour se représenter la vie privée. Lorsque cette condition était remplie, alors l'élément était considéré comme incontournable. L'étude visait aussi à déterminer les différentes relations que certains mots préalablement sélectionnés (p. ex., pour leur importance ou leur thème) entretenaient entre eux, dans le but de connaître l'organisation de la représentation. Les résultats font apparaître que les mots les plus cités se recoupent sous des thèmes importants dans la représentation comme : la protection de la vie privée ainsi que les dangers qui pèsent sur elle, l'intimité et la technologie. De manière plus spécifique, les mots incontournables de la représentation étaient majoritairement reliés au thème de la protection de la vie privée. L'étude a également révélé que les mots reliés au thème du danger n'étaient pas incontournables pour représenter la vie privée en contexte technologique. De plus, l'organisation de la représentation s'est avérée relativement incertaine, car peu de mots étudiés conservaient leurs relations durant les différentes analyses. Cependant, un des rares liens persistants lors des analyses concernait deux mots reliés aux thèmes de la protection de la vie privée et de surcroît, considérés comme incontournables : confidentialité et protection. Ce dernier résultat souligne l'importance des mots liés aux thèmes de la protection dans la représentation de la vie privée en contexte technologique. En somme, ces résultats suggèrent que les individus se représentent la vie privée comme intrinsèquement protégée, mais ils suggèrent aussi que la notion de danger concernant la vie privée n'est pas nécessaire pour entretenir une représentation de la vie privée en contexte technologique. Ces deux constats suggèrent que le concept de paradoxe de la vie privée pourrait constituer une forme d'artefact qui découle de certaines méthodes de

recherche utilisées. De plus, l'organisation incertaine de la représentation suggère que les individus rencontrent des difficultés à se représenter clairement la vie privée en contexte technologique.

Mots-clés : Vie privée, paradoxe de la vie privée, artefact, représentation sociale, approche structurale, analyse prototypique.

Table des matières

Sommaire	iii
Liste des tableaux	x
Liste des figures	xiii
Remerciements	xvi
Introduction	1
Contexte théorique	8
Réflexion sur les définitions de la vie privée et de privacy	9
Historique de la vie privée en France et de la privacy aux États-Unis	11
Débuts de la vie privée dans les sciences sociales	14
Difficultés à définir la vie privée en contexte technologique	17
Dimensions psychologique et sociale de la vie privée en contexte d'utilisation des TIC	22
Représentations sociales : genèse et conceptualisation.....	30
École d'Aix-en-Provence et approche structuraliste.....	39
Problématique	50
Objectif	55
Questions de recherche et hypothèses.....	55
Méthode.....	57
Participants.....	58
Première collecte de données.....	58
Seconde collecte de données.....	64
Instruments.....	67

Déroulement.....	81
Résultats	85
Présentation des analyses.....	86
Analyse prototypique.....	86
Question du seuil d'appartenance au noyau central et confirmation de l'appartenance au noyau central dans une recherche sur les représentations sociales.....	99
Analyse de similitude.....	109
Analyses descriptives.....	114
Analyse descriptive du questionnaire Privacy Orientation Scale	115
Analyse descriptive du questionnaire Security Behavior Intentions Scale.....	119
Analyse descriptive du Test d'indépendance au contexte	124
Analyse descriptive du Questionnaire de caractérisation	130
Validation des questionnaires et comparaison de moyennes.....	134
Analyse	149
Étape préliminaire à la découverte du noyau central : l'établissement des hypothèses de structure dégagées par l'analyse prototypique et thèmes émergents	149
Validation du noyau central avec le Test d'indépendance au contexte	160
Relation entre les items de la représentation avec l'analyse de similitude.....	177
Discussion.....	190
Conclusion	224
Références.....	235
Appendice A. Message de recrutement.....	254

Appendice B. Tableau 1. Résumé des variables sociodémographiques de la première collecte de données	256
Appendice C. Tableau 2. Résumé des variables sociodémographiques de la seconde collecte de données.....	261
Appendice D. Précaution à l'utilisation de l'analyse en composante principale	266
Appendice E. Questionnaires des collectes de données.....	268
Appendice F. Tableau 11. Résumé des variables sociodémographiques du questionnaire Privacy Orientation Scale	302
Appendice G. Tableau 13. Résumé des variables sociodémographiques du questionnaire Security Behavior Intentions Scale.....	307
Appendice H. Tableau 14. Résumé des variables sociodémographiques du Test d'indépendance au contexte	312
Appendice I. Tableau 16. Résumé des variables sociodémographiques du Questionnaire de caractérisation	317

Liste des tableaux

Tableau

1	Résumé des variables sociodémographiques de la première collecte de données	257
2	Résumé des variables sociodémographiques de la seconde collecte de données	262
3	Extrait de la base de données fictives représentant les 10 premiers participants ...	90
4	Base de données fictives représentant les associations précédentes ordonnées	91
5	Fréquence des mots	92
6	Distribution des rangs d'importance du mot « ordinateur ».....	94
7	Distribution des rangs d'importance du mot « tablette »	94
8	Importance des mots.....	96
9	Exemple illustratif	103
10	Test t sur échantillon unique des différentes sous-échelles de la Privacy Orientation Scale de l'étude avec les moyennes dans l'article de validation originale	116
11	Résumé des variables sociodémographiques du questionnaire Privacy Orientation Scale	303
12	Test t sur échantillon unique des différentes sous-échelles de la Security Behavior Intentions Scale de l'étude avec les moyennes dans l'article de validation originale	122
13	Résumé des variables sociodémographiques du questionnaire Security Behavior Intentions Scale.....	308
14	Résumé des variables sociodémographiques du Test d'indépendance au contexte.....	313
15	Caractéristiques des scores de chaque item au Test d'indépendance au contexte.....	127

Tableau

16	Résumé des variables sociodémographiques du Questionnaire de caractérisation	318
17	Données descriptives concernant le Questionnaire de caractérisation.....	133
18	Saturation des items de la Privacy Orientation Scale selon l'analyse.....	136
19	Alpha de Cronbach des sous-échelles de la Privacy Orientation Scale	138
20	Comparaison selon le genre pour les différents scores des sous-échelles de la Privacy Orientation Scale avec un test t pour échantillons indépendants.....	139
21	Comparaison selon la déclaration ou non de l'orientation politique et les différents scores des sous-échelles de la Privacy Orientation Scale avec un test t pour échantillons indépendants	140
22	Corrélation entre la tendance politique et les différents scores des sous-échelles de la Privacy Orientation Scale	141
23	Corrélation entre le nombre de réponse « Je ne souhaite pas répondre » et les différents scores des sous-échelles de la Privacy Orientation Scale	141
24	Saturation des items du questionnaire Security Behavior Intentions Scale.....	143
25	Alpha de Cronbach des sous-échelles du questionnaire Security Behavior Intentions Scale.....	145
26	Comparaison selon le genre pour les différents scores des sous-échelles de la Security Behavior Intentions Scale avec un test t pour échantillons indépendants	146
27	Comparaison selon la déclaration ou non de l'orientation politique et les différents scores des sous-échelles de la Security Behavior Intentions Scale avec un test t pour échantillons indépendants.....	147
28	Corrélation entre la tendance politique et les différents scores des sous-échelles de la Security Behavior Intentions Scale	148
29	Corrélation entre le nombre de réponses « Je ne souhaite pas répondre » et les différents scores des sous-échelles de la Security Behavior Intentions Scale	148
30	Zone n°1 candidat au noyau central (fréquence élevée et rang faible)	152

Tableau

31	Zone n°2 première périphérie (fréquence élevée et rang élevé).....	153
32	Zone n°3 zone mixte (fréquence faible et rang faible)	154
33	Zone n°4 seconde périphérie (fréquence faible et rang élevé)	155
34	Résumé des scores de chaque item du Test d'indépendance au contexte	162
35	Résumé des scores de chaque item du Test d'indépendance au contexte	177

Liste des figures

Figure

1	Nombre d'articles avec la recherche « privacy paradox » dans les champs titre d'article, résumé et mots clés de la banque de donnée Scopus de 1977 à 2022.....	23
2	Schéma récapitulatif adapté de Wagner et al. (1999, p. 98) [traduction libre].....	32
3	Principe du panoptique	38
4	Schéma comparatif des représentations sociales du terme « Magrébin » adapté de Abric (2003, p. 79).....	48
5	Représentation formelle des hypothèses adaptée de Abric (2003, p. 64).....	72
6	Schéma récapitulatif	80
7	Récapitulatif de la démarche	82
8	Diagramme à barres des fréquences terminologiques (mots centraux hachurés)...	93
9	Diagramme à barres de la distribution des rangs du mot « ordinateur ».....	95
10	Diagramme à barres de la distribution des rangs du mot « tablette ».....	95
11	Diagramme à barres de l'importance.....	97
12	Histogrammes des réponses pour les 4 items relatifs à la représentation du groupe d'amis adapté de Baggio et Renaud (2015, p. 43).....	108
13	Distributions des 4 items précédents (en noir) et distributions théoriques de la loi exponentielle pour chacun des 4 items précédents (en rouge) adapté de Baggio et Renaud (2015, p. 45).....	110
14	Typologie des graphes	111
15	Algorithme de l'arbre des maximums	114
16	Diagramme à points de l'item « anonymat ».....	129

Figure

17	Diagramme à points de l'item « confidentialité »	129
18	Diagramme à points de l'item « protection »	130
19	Représentation des hypothèses structurales de la représentation	151
20	Graphique de la distribution de l'item « confidentialité »	164
21	Diagramme à points de la distribution de l'item « confidentialité »	164
22	Graphique de la distribution de l'item « donnée personnelle »	165
23	Diagramme à points de la distribution de l'item « donnée personnelle »	165
24	Graphique de la distribution de l'item « intimité »	166
25	Diagramme à points de l'item « intimité »	166
26	Graphique de la distribution de l'item « droit »	168
27	Diagramme à points de l'item « intimité »	168
28	Graphique de la distribution de l'item « protection »	169
29	Diagramme à points de la distribution de l'item « protection »	169
30	Graphique de la distribution de l'item « respect »	170
31	Diagramme à points de la distribution de l'item « respect »	170
32	Graphique de la distribution de l'item « sécurité »	171
33	Diagramme à points de la distribution de l'item « sécurité »	171
34	Graphique de la distribution de l'item « anonymat »	172
35	Diagramme à points de la distribution de l'item « anonymat »	173
36	Graphique de la distribution de l'item « choix »	173

Figure

37	Diagramme à points de la distribution de l'item « choix ».....	174
38	Graphique de la distribution de l'item « liberté ».....	174
39	Diagramme à points de la distribution de l'item « liberté »	175
40	Graphique de la distribution de l'item « secret ».....	175
41	Diagramme à points de la distribution de l'item « secret »	176
42	Matrice de similitude de l'analyse de similitude comprenant les 21 items et leurs Tau C de Kendall	179
43	Arbre maximum avec les valeurs algébriques	180
44	Arbre maximum avec les valeurs absolues.....	181
45	Graphe seuil à 0,27	185
46	Graphe seuil à 0,18	186
47	Graphe seuil à 0,11	187
48	Graphe seuil à 0,08	189

Remerciements

Merci aux professeurs Michael Cantinotti et à Élias Rizkallah sans qui cette aventure n'aurait été possible. Merci à vous deux pour vos multiples relectures, votre pédagogie et votre patience. Merci de ne pas trop m'avoir bridé sur ma curiosité intellectuelle et de toujours avoir répondu à mes questions. Un remerciement tout particulier pour Michael qui m'a donné la chance de réaliser cette thèse, à renforcer mon appétit pour les statistiques et les méthodologies de recherche. Encore merci pour ta grande patience, ta pédagogie et ta bienveillance.

Merci aux professeurs Marie-Pierre Gagnon-Girouard et Marcos Balbinotti de m'avoir fait l'honneur de participer à mon comité de thèse ainsi que leur aide bienveillante.

Enfin, merci au professeur Evan Light pour m'avoir fait l'honneur d'accepter l'évaluation de la présente thèse.

Je tiens à remercier mes parents sans qui je n'aurais pu mener ce long projet que fut ma thèse et plus généralement ma réorientation en psychologie. Je ne vous remercierai jamais assez d'avoir cru en moi tout le long de cette aventure. Même si celle-ci m'a éloigné géographiquement de vous, je pense chaque jour à vous et à quel point vous avez été et êtes toujours des parents inspirants.

Encore merci à ma mère de m'avoir transmis son sens de l'écoute et sa posture de non-jugement. Merci de m'avoir transmis ta curiosité pour les arts et plus généralement pour le monde sensible. Merci à mon père de m'avoir transmis sa curiosité scientifique, le goût de la résolution de problèmes, la persévérance. Encore merci papa pour m'avoir transmis ta combativité pour les causes sociales. Merci à vous deux de m'avoir transmis votre humanité.

Merci à Charlotte pour sa patience, son soutien, sa compréhension. Je suis heureux de partager ma vie avec toi et que tu assistes à la fin de ce chantier que fut ma thèse. Je te promets que cette thèse sera la seule et unique.

Merci à Samuel sans qui je n'aurai jamais entrepris et réussi cette aventure québécoise. Merci pour ta présence, ton soutien et ton amitié sans cesse renouvelée depuis nos onze ans.

Merci à Pierre et Amélie pour leur soutien moral indéfectible dans les moments difficiles et leurs conseils avisés durant ces mêmes périodes. Merci encore pour votre précieuse amitié malgré les kilomètres qui nous séparent.

Merci à Carl et Cédric pour leurs amitiés et surtout leurs saillies aussi drôles que percutantes.

Merci à Amandine pour son amitié, ses conseils, sa bienveillance et sa franchise.

Merci à mes « compagnons de galères » et amis Marie-Anne et Dimitri. Merci d'avoir pris le temps de me relire.

Introduction

Cette thèse aborde le thème de la vie privée en contexte d'utilisation des technologies de l'information et des communications (TIC), ce que nous appellerons « contexte technologique ». Depuis l'affaire Snowden (Lyon, 2014), les enjeux relatifs au respect de la vie privée lors de l'utilisation des technologies se sont accrus et diversifiés (Stuart & Levine, 2017). Parmi eux, on retrouve des interrogations sur la surveillance du citoyen grâce à la reconnaissance faciale par l'intelligence artificielle (Naker & Greenbaum, 2017), des enjeux criminels avec les vols d'identité et les fraudes bancaires (Bergström, 2015) et même des enjeux éthiques avec l'exploitation non consentie des données personnelles (Isaak & Hanna, 2018).

Outre le fait que ces phénomènes soient d'actualité, les enjeux concernant la vie privée en contexte technologique sont présents dans le quotidien de chacun. Chaque utilisation d'un appareil technologique (téléphone intelligent, ordinateur, tablette, caméra de surveillance intelligente, etc.) est une occasion d'y être confronté, car leur utilisation est souvent conditionnelle à l'accord de l'utilisateur de donner accès à ses données personnelles transitant par ces appareils technologiques.

En France, les statistiques du Centre de recherche pour l'étude et l'observation des conditions de vie sur les TIC pour 2018 renseignent sur l'ampleur de la population pouvant être touchée : 75 % des Français ont un téléphone intelligent, 78 % disposent d'au moins

un ordinateur à leur domicile, 41 % possèdent une tablette et enfin 86 % d'entre eux disposent d'une connexion Internet (Credoc, 2018). De plus, certains appareils regroupent plusieurs fonctionnalités, rendant leur usage encore plus omniprésent dans la vie quotidienne. Par exemple, le téléphone peut être utilisé en tant que réveil, appareil photo, montre, agenda, GPS, lecteur MP3 et même parfois comme quotidien de presse (Micheau, 2017). Par ailleurs, 23 % des Français n'éteignent sous aucun prétexte leur téléphone, soulignant ainsi l'aspect incontournable de cet outil technologique (Micheau, 2017).

Bien que le taux d'équipement et les formes d'usage des technologies soient des données intéressantes, elles ne révèlent pas l'intensité avec laquelle est utilisé l'objet pour communiquer. Pour en avoir une approximation, il est possible d'observer ce qui passe sur l'Internet mondial en une minute. Selon l'entreprise Domo (2020), il s'échangeait durant une minute dans le monde en 2020 environ 41,7 millions de messages sur WhatsApp, environ 208 000 personnes étaient en réunion sur Zoom et environ 1,3 million de personnes passaient un appel vidéo ou vocal. Globalement, de 2011 à 2016, le trafic de l'Internet mobile a été multiplié par 18.

Dès lors, on peut comprendre les inquiétudes que relevait la Commission nationale d'informatique et liberté (CNIL) chez les Français¹. Cette dernière relevait en 2018 qu'une

¹ La Commission Nationale de l'Informatique et des Libertés (CNIL) est un organisme public français créé en 1978 qui a une mission d'alerte, de conseil et d'information et qui veille à la protection des données personnelles contenues dans les fichiers informatiques ou papiers, publics ou privés. (<https://www.cnil.fr/fr/cnil-direct/question/la-cnil-cest-quoi>)

majorité de Français étaient inquiets concernant la vie privée en contexte technologique (CNIL, 2018a). Par exemple, ils étaient près de 66 % à se déclarer plus sensibles qu'auparavant à la protection des données personnelles. Ils étaient aussi 59 % à trouver la protection des données insuffisantes. Pour autant, est-ce que la population adopte des comportements de protection de la vie privée lorsqu'elle utilise la technologie?

Si on s'en tient à l'affaire « Ashley Madison », nous pourrions répondre par la négative. Pour rappel, Ashley Madison n'est pas un site de rencontre classique. Il se centre sur les relations extraconjugales. En 2015, lorsque le site fut victime d'une cyberattaque, les usagers ont pu saisir l'ampleur des problématiques de vie privée et les conséquences du manque de sécurisation de celle-ci. À cette occasion, les données personnelles de près de 33 millions de comptes utilisateurs ont été piratées, puis exploitées et parfois exposées publiquement (Mao & Saintourens, 2016). Les utilisateurs du site ont pris un risque en mettant des informations privées et authentiques sur ce site. Cependant, ils n'imaginaient sans doute pas les risques encourus en fournissant leur nom, leur adresse ou d'autres informations qui peuvent paraître anodines en apparence. Quoiqu'intéressant, l'exemple d'Ashley Madison concerne principalement les personnes pratiquant l'adultère. En revanche, les cambriolages assistés par les médias sociaux peuvent affecter l'ensemble de leurs utilisateurs. En effet, les voleurs peuvent utiliser les médias sociaux pour faciliter le cambriolage d'un domicile (Donoghue, 2017), notamment en surveillant les statuts Facebook, les *stories* Instagram, etc. Ainsi, ils peuvent constater que la personne est en vacances et le domicile sans occupant.

Au regard de ces exemples, il semble y avoir un comportement apparemment paradoxal de la population. D'un côté, les individus se disent inquiets pour leur vie privée et de l'autre côté, il semblerait qu'ils ne protègent pas suffisamment leur vie privée. Ce phénomène qui peut paraître surprenant est bien connu des chercheurs depuis les années 2000. Il porte même un nom : le *privacy paradox* (Brown, 2001), terminologie traduite en français par « paradoxe de la vie privée ».

La psychologie sociale permet d'étudier ce genre de phénomène. Parmi les méthodes utilisées dans le domaine, celle des représentations sociales nous est apparue particulièrement indiquée. En effet, l'approche structurale des représentations sociales définit celle-ci comme étant « un système d'interprétation de la réalité qui régit les relations des individus à leur environnement physique et social » (Abric, 1994, p. 13).

Durant notre recherche, nous avons poursuivi plusieurs objectifs. L'un d'eux visait à approfondir les connaissances sur la vie privée en contexte technologique, et ce, autant d'un point de vue social que psychologique. L'approche structurale des représentations sociales nous permet cela en adoptant un regard à la croisée de la psychologie et de la sociologie (Abric, 2003). De plus et au meilleur de nos connaissances, aucune recherche n'a été entreprise pour identifier cette représentation sociale. Enfin, dans son dernier objectif, cette étude visait à contraster les modèles établis dans la littérature et de celui issu des représentations sociales afin de souligner les faiblesses des premiers et envisager des pistes d'explication novatrice. À cette fin, les différents modèles explicatifs de la vie

privée ainsi que leurs enjeux ont été explorés, et l'approche structurale des représentations sociales comme modèle alternatif a été détaillée afin d'expliquer sous un nouveau jour les problématiques de la vie privée.

Pour cela, il est capital de recueillir la représentation de la vie privée en contexte d'usage des TIC. En effet, comme précédemment indiqué, la représentation guide l'action de l'individu dans son environnement. Dans l'approche structurale, une partie spécifique de la représentation joue un rôle important dans ces aspects opératoires (Abric, 2003). Cette partie est dénommée le « noyau central » et contient les éléments les plus stables et les plus essentiels de la représentation (Abric, 2003). Par conséquent, si aucun des éléments du noyau central ne concerne les thèmes de la protection, de la sécurité ou du danger, alors nous pourrions comprendre différemment l'absence de comportement de protection de la vie privée. En effet, pourquoi les individus protégeraient-ils leur vie privée si dans leurs représentations, elle n'a pas besoin de l'être? Si tel est le cas, nous apporterons un éclairage nouveau sur ces comportements supposément paradoxaux.

Afin de répondre au mieux à ces objectifs, nous explorons dans une première partie de la thèse la genèse de la vie privée, les différentes modélisations qui ont pu en découler et les enjeux associés à ces modélisations. Ensuite, nous nous attarderons sur les motivations qui nous ont dirigés vers l'approche des représentations sociales, leur histoire et leurs spécificités dans l'approche structurale. Dans un deuxième temps, nous

présenterons notre étude divisée en deux expériences et nous discuterons des résultats et des possibles recherches qu'il serait intéressant de mener à la lumière des résultats.

Contexte théorique

Préalablement à notre problématique de recherche, le contexte dans laquelle s'inscrit cette dernière sera présenté. Afin d'appréhender notre objet de recherche qu'est la vie privée en contexte technologique, une première partie permettra de mettre en relation ses définitions, son apparition dans l'histoire et les lois, sa genèse dans les sciences sociales, ses difficultés de conceptualisation dans le contexte technologique et ses spécificités dans ce dernier. Une seconde partie consistera en la présentation de notre cadre d'analyse de la vie privée en contexte technologique à travers la présentation des représentations sociales. Cette partie fera l'objet d'un développement autour de la genèse des représentations sociales et des spécificités du courant de l'approche structuraliste choisie pour l'analyse.

Réflexion sur les définitions de la vie privée et de privacy

L'étymologie et les définitions du dictionnaire permettent de recueillir des conceptualisations communes concernant la vie privée. Pour le Centre national des ressources textuelles et lexicales (CNRTL, 2017), l'étymologie du mot « privé » est à chercher dans les termes latins *privatus* et *privare*. Le sens de *privatus* est celui de « propre, particulier, individuel » alors que *privare* signifie « écarter de, ôter de, dépouiller » (CNRTL, 2017, en ligne). Le dictionnaire Larousse (2017, en ligne) propose une définition du mot « privé » : « Qui ne concerne pas le public, qui se fait sans témoins, en dehors d'un cadre officiel ». Le Robert (2017), quant à lui, propose des définitions semblables. Une première définition est proposée comme étant « Où le public n'a pas

accès, n'est pas admis » (Le Robert, 2017, en ligne) et une seconde définition est donnée comme « Individuel, particulier » (Le Robert, 2017, en ligne). Nous souhaitons aussi définir l'équivalent anglais du terme « privé », à savoir *privacy*. En effet, outre l'omniprésence de l'anglais dans les publications scientifiques, il s'est avéré que certaines publications en langue française utilisaient autant le terme anglais que français (Casilli, 2013; Pucheral et al., 2016; Rallet et al., 2015; Rey, 2012). Comme le terme « privé », le terme « *privacy* » présente lui aussi des racines communes avec le terme latin « *privare* » (Online Etymology Dictionary, 2017). Concernant la définition de *privacy*, le dictionnaire en ligne de Oxford (2017) propose une première définition : « A state in which one is not observed or disturbed by other people » (Un état dans lequel on n'est pas observé ou dérangé par d'autres personnes [traduction libre]). Une seconde définition est proposée, où le dictionnaire (Oxford, 2017) renvoie à « The state of being free from public attention » (Un état à l'abri du regard des autres [traduction libre]).

À la lecture des définitions, aucune différence n'est vraiment apparente et bien que certains auteurs francophones utilisent *privacy* et vie privée sans distinction (Casilli, 2013; Pucheral et al., 2016; Rallet et al., 2015; Rey, 2012), d'autres auteurs précisent les différences entre ces deux termes (Halpérin, 2005, 2015; Perbal, 2018; Pierre, 2014). Par exemple, Pierre (2014, p. 153) déclare que « la notion de *privacy* dépasse d'ailleurs la simple sphère de l'intime : il ne faut pas se contenter de traduire le terme par “vie privée”, et réduire à cette sphère ce qui devrait être protégé ». Comme l'indique Halpérin (2005), il serait naïf de résumer l'implantation du terme « *privacy* » dans la littérature scientifique

en langue française comme l'indice de l'américanisation¹ de la société française. Il nous semble donc nécessaire d'explorer davantage l'histoire des concepts de vie privée et de *privacy* pour comprendre les points de vue divergents sur les deux termes.

Historique de la vie privée en France et de la *privacy* aux États-Unis

Whitman (2004) repère l'origine de la *privacy* et de la vie privée dans la notion de « dignité » présente sous l'Ancien Régime (référence au Royaume de France précédant la Première République française de 1792). Historiquement, le concept de vie privée est antérieur à celui de *privacy*, car il apparaît avec l'adoption de la Constitution française de 1791 alors que le concept de *privacy* apparaît en 1890 (Halpérin, 2015). La Constitution française de 1791 vise à encadrer la liberté de la presse et notamment les risques de calomnies ou d'injures contre les citoyens (Halpérin, 2015; Perbal, 2018). Plus précisément, l'article 17 du chapitre 5 de la Constitution française de 1791 spécifie que : « Les calomnies et injures contre quelques personnes que ce soit, relatives aux actions de leur vie privée, seront punies sur leur poursuite » (Perbal, 2018, p. 11). Durant les changements politiques des décennies suivantes (10 régimes politiques différents de 1789 à 1868), la vie privée fait de rares apparitions, toujours enchâssées dans des lois concernant la presse. Par exemple, lorsque la loi du 17 mai 1819 sur la liberté de la presse a été débattue, le député Royer Collard avait défendu la vie privée (Halpérin, 2015). Par

¹ L'américanisation en France peut être définie (Kuisel, 1990) comme l'importation en France d'un modèle économique, social et culturel américain. Durant l'après-guerre, l'américanisation a pu se traduire en France par l'importation de modèles de production et de consommation standardisés, un style de vie différent comme des : « nouveaux modèles de dépense, des niveaux de salaire plus élevés, une plus grande mobilité sociale, une vie centrée sur la consommation et une philosophie matérialiste » (Kuisel, 1990, p. 1). Sur le plan culturel, celle-ci s'est cristallisée avec l'importation de la culture de masse comme avec les films d'Hollywood (Kuisel, 1990). Renvoyons à Bagola (2004) pour l'américanisation des termes francophones sur Internet.

la suite, la vie privée a aussi été défendue lors de la promulgation de la loi sur la liberté de la presse du 11 mai 1868. Finalement, c'est au XX^e siècle que la vie privée devient un objet de droit à part entière en sortant des lois sur la liberté de la presse. En effet, la loi du 17 juillet 1970 indique que (Halpérin, 2015, p. 63) « chacun a droit au respect de sa vie privée ».

Quant à la naissance de la notion de *privacy* aux États-Unis, elle dérive en partie du concept français de la vie privée. En effet, l'article à l'origine du droit à la *privacy* (Warren et Brandeis, 1890) cite la loi française du 11 mai 1868 comme exemple de loi sur la vie privée. Tout comme les législateurs français, Warren et Brandeis (1890) s'inquiètent pour la dignité de leurs concitoyens, notamment à travers la publication de la vie intime des personnes par la presse. Cette inquiétude est renforcée par l'émergence de nouvelles technologies comme l'appareil photo portable. Plus précisément, Warren et Brandeis (1890) s'inquiètent des conséquences délétères des photos publiées dans la presse et prises à l'insu de la personne. Toutefois, en 1931, les révélations publiques d'informations privées issues d'écoutes téléphoniques gouvernementales provoqueront beaucoup plus de questionnement chez les citoyens concernant leur droit à ne pas être surveillés par leur gouvernement qu'au sujet de leur dignité. Ainsi, la condamnation de ces révélations est parfois considéré comme un point de bascule du droit à la *privacy* de la protection de la dignité vers la protection de la liberté (Halpérin, 2015).

L'année 1931 signe donc l'émergence d'une différence entre les concepts français et étasunien, où la défense de la dignité n'apparaît plus comme la valeur sous-jacente

dominante dans le concept de *privacy* (Halpérin, 2015). Rappelons que le système légal français, et plus généralement celui européen, vise à défendre la dignité du citoyen alors que la justice américaine vise plutôt à protéger la liberté du citoyen face aux intrusions de l'État (Halpérin, 2015; Levin & Nicholson, 2005; Whitman, 2004). Dans son article, Whitman (2004) explique que les différences culturelles peuvent expliquer ces divergences. D'une part, il souligne l'importance de l'honneur dans les pays européens, où il est essentiel d'éviter les humiliations publiques (Whitman, 2004). Sa conclusion est que la loi de la protection de la vie privée permet de condamner l'humiliation pour défendre la dignité de la personne. D'autre part, cet auteur rapporte l'importance de conserver l'autonomie individuelle pour le citoyen américain (Whitman, 2004). Selon lui, la loi assurerait cette mission à travers la défense de la liberté.

On pourrait conclure des travaux de Whitman (2004) que la législation des deux pays présente des objectifs différents. Certains auteurs (Pierre, 2014; Smith et al., 2011) partagent d'ailleurs cette perspective. Pourtant, l'article de Whitman se conclut tout en nuances. En effet, cet auteur explique que les valeurs de dignité et de liberté coexistent dans les pays, peu importe le continent. Pour Whitman, la différence n'est pas tant dans l'absence de l'une ou l'autre des valeurs (dignité c. liberté), mais dans l'intensité avec laquelle le législateur défend ces deux valeurs lorsqu'il édicte les lois de protection de la vie privée.

En conclusion, l'évolution historique des deux concepts dans les traditions juridiques française et américaine (vie privée et *privacy*) a permis de tracer certaines de leurs

divergences et de souligner leurs recoupements. Après ce premier éclairage historique et légal, il nous semble pertinent de continuer ce travail sur l'origine de la conceptualisation actuelle de la vie privée. À ce titre, les sciences sociales ont joué un rôle décisif.

Débuts de la vie privée dans les sciences sociales

Selon plusieurs auteurs (Acquisti et al., 2015; Barth & de Jong, 2017; Dienlin & Trepte, 2015; Jeong & Kim, 2017; Laufer & Wolfe, 1977; Li, 2012; Paine et al., 2007; Pierre, 2014), l'un des auteurs ayant eu un impact majeur sur la recherche concernant la vie privée est le psychologue social américain Altman. Pour Altman (1975), la vie privée se définit par le contrôle sélectif de l'accès au soi d'une personne et comporte plusieurs caractéristiques.

Premièrement, la vie privée est un processus interpersonnel qui se déroule avec au minimum deux acteurs (Altman, 1975). À ce propos, Altman (1975) propose deux types d'acteurs : une personne ou un groupe.

Deuxièmement, la vie privée est pour cet auteur un processus bidirectionnel impliquant des interactions entrantes et sortantes. Les interactions sociales entrantes sont des stimulations venant de l'environnement (groupe ou personne seule) vers la personne, par exemple lorsqu'elle reçoit un appel téléphonique (Altman, 1975). Les interactions sociales sortantes sont des stimulations émanant de la personne en direction des autres (groupe ou personne seule), par exemple dans le cas d'une demande d'aide pour la résolution d'un problème personnel (Altman, 1975).

Troisièmement, pour Altman (1975), la vie privée constitue un processus dialectique où des forces d'ouverture et de fermeture se rencontrent de manière simultanée et dynamique pour arriver à un équilibre dans le meilleur des cas, ou à un déséquilibre dans les autres cas. Cet équilibre est défini par deux variables. La première est le niveau souhaité d'interaction, lequel est déterminé préalablement et mentalement par les individus. Par exemple, avant même de communiquer avec un inconnu, une personne pourrait ne pas avoir envie de dévoiler son salaire à son interlocuteur. Ce niveau souhaité d'interaction va être confronté à la seconde variable qui est le niveau réel d'interaction. Par exemple, une personne en colocation pourrait interagir minimalement avec ses colocataires avant de s'enfermer dans sa chambre. Ainsi, elle peut maintenir une ambiance conviviale dans la colocation tout en répondant à son besoin d'intimité. En d'autres termes, cette étape vise à trouver un équilibre entre aspiration et réalité, comme dans notre exemple. Cependant, la confrontation entre les aspirations et la réalité ne débouche pas toujours sur un équilibre, comme par exemple, des parents qui s'introduisent dans la chambre d'un adolescent alors qu'il cherche à s'isoler. Dans ce cas, il s'agira d'un déséquilibre entre aspiration et réalité.

Dernièrement, la vie privée est un processus dynamique qui s'actualise en continu. Pour Altman (1975), cela est possible grâce à deux variables d'ajustement du niveau d'interaction réelle. La première variable est l'intensité des échanges des individus en interaction (entrantes ou sortantes). Cette intensité des échanges peut être située à trois niveaux : adéquat, excessif ou insuffisant. Pour reprendre le premier exemple, demander le salaire à un inconnu pourrait être jugé excessif. La seconde variable est la perméabilité

des frontières personnelles de chaque individu en interaction. On peut retrouver des frontières personnelles plus ou moins ouvertes/fermées ou accessibles/inaccessibles. Ces états représentent la volonté d'accepter ou non les interactions entrantes de chaque individu en interaction. Pour continuer avec notre exemple, il se peut que certaines personnes plus enclines à se dévoiler acceptent malgré tout de préciser le montant de leur salaire à un inconnu.

Un second modèle de référence est proposé par Westin (1967), professeur de droit américain. Westin a défini la vie privée comme l'état volontaire de retrait d'un individu par rapport à l'univers public. Contrairement à Altman (1975), son modèle ne spécifie pas un processus, mais quatre états possibles de la vie privée. Ces quatre états sont déterminés selon le besoin de la personne. Le premier proposé est l'état de solitude où l'individu est à l'abri des observations des autres (Westin, 1967). Le second est celui d'intimité défini par un isolement en petits groupes afin que les membres puissent établir une relation étroite (Westin, 1967). Un troisième état est l'anonymat. Westin le définit comme l'absence d'identification et de surveillance en public. Enfin, le dernier état est la réserve. Elle représente le désir de limiter les divulgations aux autres tout en exigeant que les autres reconnaissent et respectent ce désir dans les lieux publics ou pour des actes publics (Westin, 1967).

Enfin, les autres champs des sciences sociales proposent des définitions similaires à celle de Westin (1967) et d'Altman (1975) pour définir la vie privée. Elles se basent sur le concept de séclusion : « isolement protecteur d'un organisme par rapport au milieu »

(CNRTL, 2018, en ligne). Cette idée est reprise par le philosophe Schoeman (1984) et le juriste Weinstein (1971) qui la définissent respectivement comme un état limitant l'accès au *soi* et comme un repli sur soi.

Même si les auteurs s'intéressant à la vie privée s'accordent sur l'idée de capacité de contrôle d'accès au soi pour tenter de définir la vie privée (Culnan, 1993; Margulis, 1977, 2003; Paine et al., 2007; Solove, 2006; Wagner & Eckhoff, 2018; Zureik et al., 2010), certains problèmes demeurent. En effet, il reste difficile de départager les éléments relevant de la vie privée et ceux n'en relevant pas. Solove (2006) est d'ailleurs l'un des premiers auteurs à soulever ce problème. Selon lui, la difficulté pour les chercheurs à définir avec précision la vie privée vient justement du fait que la définition doit couvrir des situations et des éléments très divers, allant par exemple de l'intimité d'un couple aux renseignements confidentiels (Solove, 2006). Les chercheurs s'intéressant à la vie privée en contexte technologique relèvent eux aussi cette difficulté à proposer une définition qui couvre l'entièreté des situations et des éléments (Moore, 2021; Mulligan et al., 2016; Smith et al., 2011; Wagner & Eckhoff, 2018; Xu et al., 2011), mais ils remarquent aussi des enjeux qui sont propres aux contextes technologiques (Smith et al., 2011; Wagner & Eckhoff, 2018; Xu et al., 2011).

Difficultés à définir la vie privée en contexte technologique

Un des premiers enjeux que Smith et al. (2011) remarquent est la polysémie que peut prendre le terme « vie privée » (« *privacy* ») dans les articles concernant la vie privée en contexte technologique. En effet, dans les recherches sur la vie privée en contexte

technologique, ce concept et celui de confidentialité des renseignements (*privacy* c. *information privacy*) s'utilisent de manière indifférenciée (Smith et al., 2011). Smith et al. soulignent que les contradictions engendrées par cette utilisation indifférenciée ont été peu discutées. Lors de leur revue de la littérature, ils constatent même une quasi-transposition de la définition de la vie privée d'Altman à la confidentialité des renseignements : « The state of limited access was translated to the state of limited access to the information » (L'état d'accès limité a été traduit en état d'accès limité à l'information [traduction libre]) (Smith et al., 2011, p. 995). Cela pose un problème pour plusieurs raisons.

D'une part, l'utilisation systématique et sans distinction de ces deux concepts (*privacy* et *information privacy*) crée de la confusion (Smith et al., 2011) et d'autre part, l'impact de cette utilisation n'est pas anodin. En effet, les TIC apportent des spécificités qui leur sont propres. Par exemple, on pourrait s'interroger si l'introduction d'un assistant domotique (p. ex., Google home ou Alexa) change la gestion de la vie privée d'un domicile familial, notamment à travers le prisme de la protection des informations. En effet, l'assistant domotique est obligé de collecter et de traiter l'information pour effectuer la tâche demandée. Le traitement ne s'effectue pas dans la machine, mais sur le serveur de l'entreprise, d'où la nécessité d'une connexion Internet pour faire fonctionner ces appareils. Or, les spécificités liées à l'utilisation des TIC n'ont pas été considérées dans les travaux originaux d'Altman (1975) et de Westin (1967). Il nous semble donc problématique de ne pas tenir compte de la dimension technologique et de ses spécificités.

L'anecdote de Solove (2007) sur *the dog poop girl* éclaire de façon sommaire les enjeux dus aux spécificités des TIC. L'anecdote retrace l'histoire d'une personne coréenne harcelée de manière durable sur Internet, car n'ayant pas voulu ramasser les déjections de son chien dans la rame de métro de Séoul (Solove, 2007). Cette anecdote met en exergue le rôle de l'archivage systématique de l'information qui modifie la façon d'envisager le concept de la vie privée en contexte technologique.

En effet, l'archivage systématique entrave les mécanismes naturels de l'oubli et rend les données difficilement effaçables (Acquisti et al., 2015). Des dizaines d'années après l'incident, il est encore facilement possible de retracer de nombreux éléments sur l'affaire comme : un film basé sur l'évènement (Hutuleac, 2021), sa mention dans des thèses (Jeong, 2020; Kim, 2015) ou des articles (Billingham et Parr, 2020; Hendricks, Lumadue et Waller, 2012) et de trouver encore des photos de l'incident¹. Même s'il est difficile de savoir si les photos retrouvées sont authentiques, car certaines ont visiblement été retouchées, le visage d'une femme reste identifiable sur certaines d'entre elles. Par conséquent, la femme présente sur les photos ne bénéficie pas de l'oubli alors que cette histoire et sa protagoniste auraient pu bénéficier de l'oubli dans un contexte sans technologie (Solove, 2007).

Si la situation de *the dog poop girl* permet de prendre la mesure de la persistance de l'information, elle permet aussi de prendre conscience de la grande accessibilité de

¹ Pour les images : <https://web.archive.org/web/20070125082945/http://english.chosun.com/w21data/html/news/200507/200507080017.html>

l'information grâce aux TIC de sorte qu'avant l'arrivée des TIC, les informations privées ou publiques concernaient un nombre restreint de personnes (parents, amis, collègues, voisins, connaissances) alors qu'aujourd'hui, elles en concernent des centaines, voire des milliers, dès le moment où d'autres personnes relaient l'information initialement publiée.

En outre, les TIC n'ont pas uniquement modifié les caractéristiques de l'information comme leur potentiel d'effacement ou le nombre de personnes touchées par un message. Dans un contexte technologique, la dyade émetteur/récepteur (Jakobson, 1963) est remplacée par une triade émetteur/récepteur/plateforme. Ainsi, la plateforme modifie la dynamique de communication, car elle recueille des informations pour faire fonctionner le service dont l'émetteur et le récepteur se servent, mais elle peut aussi partager les informations recueillies pour les revendre comme dans l'affaire Cambridge Analytica (Isaak & Hanna, 2018). Pour rappel, cette société avait effectué un projet de recherche visant à corrélérer le modèle Océan (taxonomie des traits de personnalité en cinq facteurs, connus sous le nom des « *Big Five* » en anglais) et l'activité du profil Facebook. L'objectif était d'influencer un grand nombre de personnes à travers des messages et vidéos sur Facebook ciblés lors des élections américaines et du référendum sur le Brexit (Isaak & Hanna, 2018).

Dès lors, lorsqu'une personne communique à travers Facebook par exemple, a-t-elle toujours la possibilité de limiter l'accès à sa personne ou à ses données et par conséquent, est-elle toujours maîtresse de sa vie privée? On pourrait facilement croire qu'il s'agit là d'un problème purement technique, pouvant être aisément résolu en utilisant un « suicide

informationnel » appelé aussi « une machine à suicide ». Un tel programme permet de supprimer les informations d'un profil de façon permanente (Ayalon & Toch, 2016). On pourrait aussi instaurer une date de péremption dans les données (Mayer-Schönberger, 2011) ou encore les chiffrer de façon automatique passé un certain délai. Ces solutions techniques sont cependant imparfaites, car une fois qu'une autre personne accède à l'information, celle-ci peut les partager de nouveau. De plus, une solution technique ne considère pas la dynamique psychologique et sociale de la problématique, notamment la psychologie des utilisateurs.

D'une part, les utilisateurs ont une tendance à ne pas supprimer ou à ne pas modifier le contenu passé. À titre d'exemple, dans l'étude de Ayalon et Toch (2016), seulement 18 % de l'échantillon voulait cacher une publication antérieure jugée inadéquate. D'autre part, le fait qu'Internet, à travers des plateformes telles que Facebook ou Instagram, devienne pour l'utilisateur un lieu d'archivage pose problème (Ayalon & Toch, 2016). En effet, ces plateformes n'ont pas comme unique fonction l'archivage, elles permettent une définition de soi à travers des valeurs personnelles (Ayalon & Toch, 2016). Ainsi, la décision de supprimer du contenu archivé, comme des photos ou des vidéos, devient coûteuse à l'utilisateur autant sur le plan social que psychologique (Ayalon & Toch, 2016; Xuan et al., 2013). Par conséquent, il nous paraît opportun de situer la problématique dans une dimension autant humaine que technique, même lorsqu'elle s'inscrit dans un contexte technologique.

Dimensions psychologique et sociale de la vie privée en contexte d'utilisation des TIC

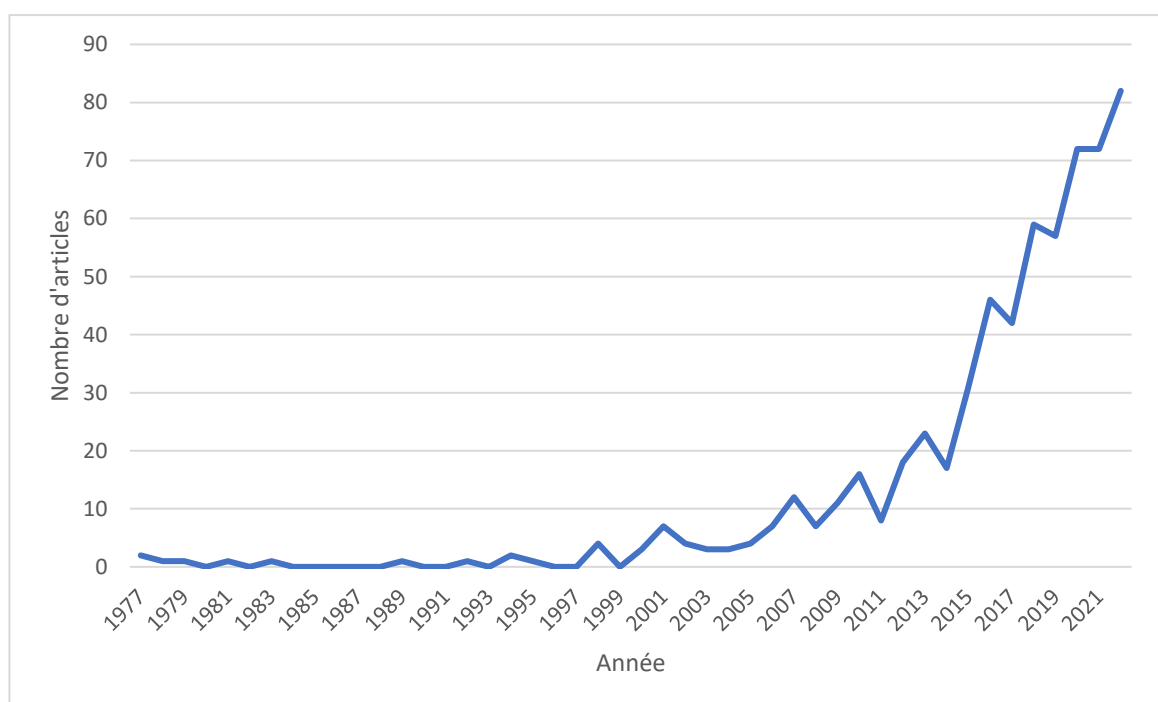
À leurs débuts, les recherches sur la vie privée en contexte technologique étaient restreintes, car peu de technologies numériques étaient accessibles (Westin, 1967). Par exemple, il fallut attendre les années 1980 pour voir l'apparition des premiers ordinateurs personnels chez les particuliers (Xanthos, 2020). Avec le temps, de nouvelles technologies sont apparues, et de nouvelles problématiques se sont manifestées concernant la préservation de la vie privée. Sur le plan technique, des problématiques propres à la technologie ont émergé comme avec l'Internet des objets (Vasilomanolakis, Daubert, Luthra, Gazis, Wiesmaier et Kikiras, 2015; Weber, 2010), le nuage informatique (AlSudiari et Vasista, 2012; Takabi, Joshi et Ahn, 2010), les villes intelligentes (Eckhoff et Wagner, 2017) ou plus récemment la norme de réseau de téléphonie mobile 5G (Khan, Kumar, Jayakody et Liyanage, 2019).

Sur le plan humain, ce sont principalement les aspects psychologiques et sociaux des préoccupations de protection de la vie privée lors de l'utilisation des TIC qui ont été étudiés (Barth & de Jong, 2017; Baruh et al., 2017; Bergström, 2015; Jeong et Kim, 2017; Lutz & Strathoff, 2014; Thomson et al., 2015). Les préoccupations examinées étaient multiples, allant des risques et des avantages de l'utilisation de Facebook pour la vie privée (Debatin et al., 2009) aux différents risques encourus par les utilisateurs selon le type de données personnelles rendues publiques (Gross & Acquisti, 2005). Cependant, peu importe le type de préoccupations étudiées, un phénomène était de plus en plus retrouvé dans les études : le paradoxe de la vie privée (Barth & de Jong, 2017; Baruh &

Cemalcılar, 2014; Baruh et al., 2017; Bergström, 2015; Debatin et al., 2009; Gross & Acquisti, 2005; Jeong & Kim, 2017; Lutz & Strathoff, 2014). L'intérêt pour ce phénomène se fit de plus en plus croissant comme l'illustre la Figure 1 qui représente l'augmentation du nombre d'articles sur le sujet entre 1977 et 2022 (Scopus, 2021).

Figure 1

Nombre d'articles avec la recherche « privacy paradox » dans les champs titre d'article, résumé et mots clés de la banque de donnée Scopus de 1977 à 2022



Comme mentionné en introduction, le paradoxe de la vie privée est l'observation simultanée chez l'utilisateur d'une absence ou d'un faible niveau de protection vis-à-vis de la vie privée avec des inquiétudes apparentes concernant celle-ci (Brown, 2001). Plusieurs théories sont avancées pour expliquer ce phénomène.

L'une d'elles, le *privacy calculus*, avance que l'utilisateur effectue un calcul basé sur la comparaison des risques encourus et des bénéfices possibles de la révélation d'informations personnelles (Acquisti et al., 2015; Barth & de Jong, 2017; Gerber et al., 2018; Kokolakis, 2017; Lee et al., 2013). Il est important de noter que le *privacy calculus* repose sur un amalgame de théories, à savoir la théorie de la maximisation de l'utilité, la théorie des attentes et enfin, la théorie de la valeur espérée (Li, 2012). Selon Li (2012), il est d'ailleurs difficile de les discerner, car toutes les trois se basent sur un calcul coût/bénéfice pour expliquer le comportement de l'utilisateur. L'auteur prend l'exemple de la théorie de la maximisation de l'utilité où l'utilisateur trouve un équilibre entre sa satisfaction à utiliser les TIC et les coûts perçus pour sa vie privée. L'auteur la compare avec la théorie des attentes qui suppose que l'individu évalue sa mise en action selon l'effort à engager pour obtenir un certain bénéfice. Enfin, l'auteur termine ce regroupement avec la théorie de la valeur espérée dans laquelle les individus entretiennent des croyances sur les comportements à mettre en place (p. ex., protéger leur vie privée). Ces croyances vont favoriser les comportements attendus comme bénéfiques et défavoriser les comportements attendus comme délétères (Li, 2012). Ainsi les comportements envisagés comme ayant le plus de bénéfices sont mis en place alors que les autres sont abandonnés. En résumé, la théorie du *privacy calculus* et celles qui s'y rattachent envisagent l'individu comme un être rationnel qui calcule préalablement les coûts et les bénéfices, avant de décider de fournir ou non l'accès à des éléments concernant sa vie privée, au profit de l'utilisation des TIC.

Une seconde théorie questionne la rationalité de l'individu en envisageant la possibilité qu'elle soit tronquée, faisant ainsi apparaître l'individu comme potentiellement irrationnel. La théorie de la rationalité limitée et des biais décisionnels postule que les décisions des utilisateurs se basent sur des informations incomplètes ou limitées (Kokolakis, 2017). En effet, les études relèvent souvent que l'utilisateur ne lit pas les politiques de confidentialité (Steinfeld, 2016), même s'il s'agit du stockage de données sensibles comme son patrimoine génétique (Christofides & O'Doherty, 2016). Ainsi, ce manque d'information vient fausser son estimation de l'impact de son comportement sur sa vie privée (Gerber et al., 2018). Par exemple, un utilisateur peut soutenir que sa vie privée est importante tout en fournissant des données sensibles sans considérer qu'elles seront utilisées à des fins qu'il n'approuverait pas forcément s'il avait pris connaissance des implications de sa décision. D'une part, lorsqu'il juge que ces données sont uniquement traitées par l'entreprise avec qui il fait commerce, l'utilisateur sous-estimerait le danger encouru pour ses données. D'autre part, cela l'amènerait à surestimer la confiance qu'il accorde à cette entreprise. Selon cette approche, c'est ainsi que la surestimation et la sous-estimation attribuables à des informations incomplètes s'articuleraient pour générer le paradoxe de la vie privée.

Enfin, une autre théorie proposée dans la revue de littérature de Gerber et al. (2018) envisage le paradoxe de la vie privée comme un artéfact méthodologique (Dienlin et Trepte, 2015). Un artéfact est un phénomène artificiel attribuable soit à la méthode d'étude d'un phénomène (facteurs techniques), soit à des facteurs humains. Par exemple, lorsque le patient bouge (facteur humain) lors d'un scanner, il peut créer un artéfact sur l'image

en faussant ses contours, car le scanner ne dispose pas d'un système de correction de l'image (facteur technique).

Dienlin et Trepte (2015) avancent que le paradoxe de la vie privée serait généré par un certain nombre de défauts de conceptualisation des expérimentations. Par exemple, ils questionnent l'impact d'interroger une seule et unique dimension des comportements relatifs à la vie privée lors des recherches, comme la dimension informationnelle (p. ex., combien d'informations sociodémographiques sont publiques sur votre profil Facebook?). En effet, Dienlin et Trepte (2015) rappellent que les comportements relatifs à la vie privée sont au contraire de nature multidimensionnelle et comprennent des dimensions sociales (p. ex., à quel point avez-vous filtré les personnes ayant accès à votre profil Facebook?) ou psychologiques (p. ex., à quel point vous avez personnalisé votre profil Facebook?). Ainsi, cibler une seule dimension revient à éluder une partie constitutive du phénomène, ce qui pourrait expliquer la manifestation apparemment paradoxale d'un comportement lorsqu'il est décontextualisé.

Un autre défaut relevé concerne la mesure des variables utilisées dans les études, notamment l'échelle de mesure dichotomique. En effet, l'évaluation binaire des comportements de protection implique une variance réduite qui diminue la puissance statistique. Par conséquent, Dienlin et Trepte (2015) soutiennent l'hypothèse que l'absence de lien entre comportement de protection et les préoccupations potentielles sur le sujet de la protection de la vie privée pourrait découler de l'opérationnalisation artificielle des variables.

Bien qu'il existe d'autres théories qui tentent d'expliquer le paradoxe de la vie privée, les revues de littérature sur la question relèvent une absence de consensus scientifique sur ce concept (Barth & de Jong, 2017; Gerber et al., 2018; Kokolakis, 2017). Les trois théories précédentes proposent cependant des pistes de réflexion sur les raisons de cette absence de consensus.

Dans la théorie du *privacy calculus* et de la rationalité limitée, les chercheurs peuvent conceptualiser l'utilisateur comme un être rationnel pour la première et quasi rationnel pour la seconde. Un premier problème est que ces théories ne permettent pas aux chercheurs d'évaluer directement ce que pense l'individu ou ce qu'il fait. En effet, ce sont les chercheurs qui définissent autant la vie privée que les concepts qui s'y rattachent ainsi que leurs opérationnalisations. Par le même fait, les réponses fournies par les participants ne peuvent que s'insérer dans la conceptualisation des chercheurs.

Un second problème est que les attitudes et les préoccupations concernant la vie privée sont hétérogènes, car elles varient en fonction du type de données concernées, celui-ci étant lui-même hétérogène (Kokolakis, 2017). On peut effectivement comprendre qu'il y ait une différence de préoccupations pour un utilisateur entre une photo de lui nu et un message envoyé à sa grand-mère. Or, comme le souligne Kokolakis (2017), il n'y a, à ce jour, aucune théorie qui permette de prendre en compte l'impact de la diversité des données sur les attitudes ou les préoccupations. Si l'impact de la diversité des données n'est pas pris en compte par la théorie, alors on peut s'attendre à ce qu'il en soit de même lorsque les chercheurs opérationnalisent leurs concepts lors de leurs expériences. En effet,

comment opérationnaliser la diversité des données si celle-ci n'a été pas conceptualisé dans la théorie? En résumé, on peut s'interroger si ce manque théorique n'implique pas des lacunes lors de l'opérationnalisation, voire des absences qui expliqueraient le paradoxe de la vie privée. Dans l'affirmative, une autre question se pose. On peut se demander dans le cas où certains comportements ne sont pas détectés, si cela découle de modalités de réponse qui offrent trop peu de variabilité comme l'énoncent Dienlin et Trepte (2015) dans la théorie de l'artefact ou même si la modalité de réponse que l'utilisateur énoncerait spontanément est disponible dans le questionnaire.

De plus, un troisième problème est soulevé par Gerber et al. (2018) qui est lui aussi lié aux définitions de la vie privée et des concepts rattachés. Ces auteurs expliquent en effet que les chercheurs ne distinguent pas automatiquement les construits qui pourraient être différenciés et vont utiliser conjointement plusieurs construits sans forcément clarifier en quoi ils sont liés. Cela rajoute des difficultés pour arriver à un consensus sur les causes théoriques présumées du paradoxe de la vie privée. Afin de répondre aux problèmes conceptuels et méthodologiques soulevés précédemment, deux étapes semblent nécessaires.

Premièrement, il est pertinent d'accorder moins de place aux questionnaires autorapportés concernant la vie privée et d'instaurer des méthodologies permettant d'observer directement les comportements réels des utilisateurs concernant leur vie privée. Ainsi, l'on évite les préconceptions du chercheur sur la vie privée. Cette recommandation est d'ailleurs proposée par Kokolakis (2017) qui conseille la collecte de

comportements directement observables plutôt que des mesures autorapportées. On peut imaginer le recueil des données publiques de pages Facebook (avec l'accord des participants) sans que le chercheur ait décidé de certains types de données ou encore demander aux participants de rapporter directement le paramétrage de son compte. Une alternative qui nous semble aussi intéressante serait de donner la possibilité au répondant de signifier sa volonté de ne pas répondre. Ce choix pourrait être considéré comme un comportement observable de protection de la vie privée. Cependant, et à notre connaissance, aucune étude ne propose cette possibilité.

Deuxièmement, il serait utile d'adopter un paradigme qui permet un recueil plus approfondi de la vie privée pour que l'opérationnalisation des construits soit plus fidèle à la réalité de l'utilisateur. Pour ce faire, les études qualitatives ou mixtes, notamment avec des questions ouvertes, sont une avenue pertinente pour adopter un point de vue centré sur l'utilisateur. Il s'avère qu'il y a relativement peu d'études de ce type sur le paradoxe de la vie privée. À titre d'exemple, lors de sa revue de littérature, Kokolakis (2017) a trouvé uniquement six études avec un paradigme qualitatif sur 59 articles.

Les représentations sociales proposent une méthodologie à la croisée du quantitatif et du qualitatif qui permet l'exploration des interactions entre l'objet social (la vie privée), l'individu et son environnement. Comme le dit Markovà (1996), accéder aux représentations sociales revient à accéder à la connaissance du quotidien et à l'agencement de la réalité des individus ou des groupes. Ainsi, on est en mesure de comprendre comment les utilisateurs rendent leur réalité stable, prévisible et quel type de logique a

été mis en place pour y arriver (Markovà, 1996). Comme nous allons le voir dans la partie suivante, les différentes étapes de la conceptualisation des représentations sociales permettent de mieux comprendre sa pertinence pour répondre à nos objectifs.

Représentations sociales : genèse et conceptualisation

Le concept de représentations sociales trouve ses racines dans la sociologie française, plus exactement dans l'article de Durkheim (1898) intitulé « Représentations individuelles et représentations collectives ». Dans ce document, Durkheim présente le concept de représentation collective et les oppose aux représentations individuelles. Celles-ci seraient propres à chaque individu alors que les représentations collectives seraient propres à la société. Aussi, les représentations individuelles seraient caractérisées par leur instabilité. En effet, pour Durkheim, le contenu de ces dernières serait variable et éphémère, à la différence des représentations collectives qui seraient stables. Selon Durkheim, les fonctions des représentations collectives consistent à préserver le lien social et amener les personnes à réfléchir et agir de façon homogène. Durkheim souligne aussi la composante transgénérationnelle des représentations collectives, puisqu'elles seraient partagées par toute la société. Ainsi, il défend l'idée qu'il existe des représentations collectives qui ne sont pas une simple extension des représentations individuelles. Ces représentations collectives ayant leurs propres modalités de fonctionnement, Durkheim va proposer d'étudier les représentations collectives et par extension, les mécanismes de la pensée collective de façon indépendante.

Cette notion de représentation collective fut relativement oubliée jusqu'en 1961, où le psychologue social et historien des sciences Serge Moscovici entreprend de créer une psychologie sociale des représentations afin d'étudier la psychanalyse, son image et son public. Il étoffe la notion de représentation collective à la lumière de la psychologie du développement de Piaget (1932) et de la psychanalyse de Freud (1908).

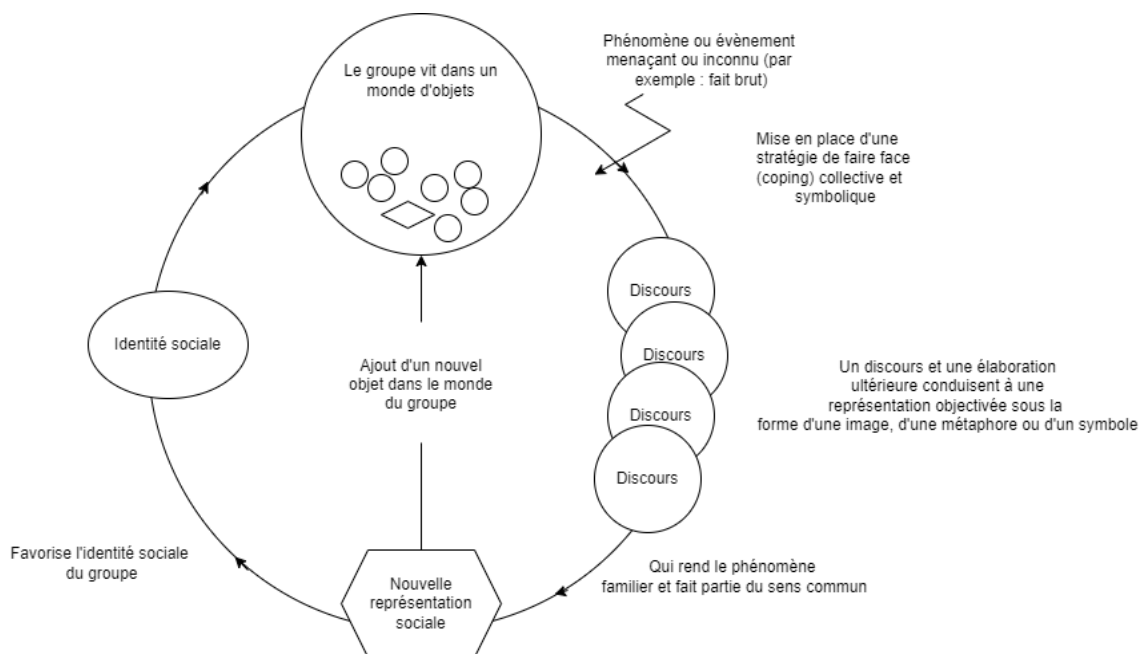
Pour Moscovici (1961, p. 300), les représentations sociales peuvent être définies comme faisant « partie d'une famille de concepts – idéologie, vision du monde, idée-force, mythe, utopie – qui se réfèrent tous à une élaboration théorique censée refléter les rapports sociaux tout en contribuant à les édifier ». Il spécifie qu'il « considère comme vision du monde, idéologie, tout ce qui n'est pas science » (p. 300).

Lors de l'élaboration de la théorie des représentations sociales, Moscovici (1961) se concentre sur la genèse d'une représentation sociale. Selon lui, la naissance d'une représentation sociale se confond avec l'apparition d'une situation inhabituelle.

Wagner et al. (1999) partagent cette définition de la représentation, car pour eux aussi, celle-ci émerge lorsqu'un groupe social rencontre un phénomène, un événement ou une idée nouvelle qui ne lui est pas familière, en d'autres termes un nouvel objet social en devenir. Cependant, Wagner et ses collaborateurs tentent de rendre compte de l'entière de la genèse d'une représentation sociale en proposant le schéma suivant (voir Figure 2).

Figure 2

Schéma récapitulatif adapté de Wagner et al. (1999, p. 98) [traduction libre]



Pour Wagner et ses collaborateurs (1999), la naissance d'une représentation passe par une phase de communication symbolisée dans le schéma par l'enchaînement des sphères « discours ». En effet, au début de la naissance d'une représentation, l'information est éparse. Certains groupes auront plus d'informations que d'autres. Par exemple, pour la représentation sociale de la vie privée en contexte technologique, un groupe de citoyens lambda aura sans doute moins d'informations sur la vie privée en contexte technologique qu'un groupe d'informaticiens ou que les employés d'une entreprise de cybersécurité. Peut-être par contre qu'un groupe de juristes spécialisés aura d'autres informations que les deux groupes précédents n'auront pas en leur possession. Le fait que l'information soit disséminée dans plusieurs groupes est appelé « dispersion de l'information » par Moscovici (1961).

La nouvelle situation va aussi impliquer que les membres du groupe aient une connaissance minimale de ladite situation. Par exemple, si on s'intéresse au règlement général sur la protection des données de l'Union européenne (CNIL, 2018b), un groupe d'amis pourrait être au courant qu'une loi sur la vie privée et les technologies a été votée récemment, alors que le service informatique d'une entreprise connaîtrait déjà le règlement en question, mais sans toutefois en connaître nécessairement l'application concrète. Quoi qu'il en soit, lors de la genèse d'une représentation sociale, les connaissances sont incomplètes et dispersées dans chacun des groupes sociaux qui s'intéressent à ce nouveau phénomène. C'est d'ailleurs parce que l'information est incomplète et récente qu'elle obligera les différents acteurs du groupe à en discuter. Cette nouvelle situation va donc susciter une certaine préoccupation, une certaine attention, qui viendront rompre le quotidien des individus. Cette nouveauté qu'occasionne le nouvel objet, quel qu'il soit (phénomène, idée, événement), provoquera une forte activité cognitive de la part des individus. Celle-ci visera à comprendre, maîtriser ou se défendre du nouvel élément à travers des communications interpersonnelles ou médiatiques entre membres d'un ou de groupes. Moscovici (1961) définira cette réaction comme la « pression à l'inférence ». Face à son ignorance et à la nouveauté de l'objet social, les membres du groupe se sentiront « pressés » et ressentiront le besoin d'élaborer un discours, un système de croyances concernant le nouvel objet.

Enfin, l'apparition de cette représentation sociale partagée par les membres du groupe est facilitée par le phénomène de « focalisation » (Moscovici, 1961). Celui-ci consiste en

la tendance du groupe à se centrer sur un aspect particulier de l'objet ou de la situation, selon ses attentes ou ses orientations.

En résumé, pour Moscovici (1961), la genèse d'une représentation sociale repose sur trois phénomènes : la dispersion de l'information, la pression à l'inférence et la focalisation. Ces trois phénomènes se développent grâce à deux processus majeurs qui sont l'objectivation et l'ancrage. C'est d'ailleurs grâce à ces deux processus que chaque représentation sociale est propre à un groupe social donné (Abric, 2003). En effet, c'est à travers eux que le groupe va exprimer ses valeurs et son propre référentiel de connaissances. L'ancrage et l'objectivation s'articulent l'un et l'autre lors de la période de communications à propos de l'objet (Moscovici, 1961). Pour mieux saisir la genèse d'une représentation, il est important d'apporter des précisions sur ces deux mécanismes.

Lors de la naissance d'une représentation sociale, l'ancrage permet à un groupe de rattacher un nouvel objet à des éléments plus familiers pour les membres du groupe (Jodelet, 1991). Pour l'illustrer, penchons-nous sur l'apparition de l'ordinateur personnel. Les scientifiques de l'époque ont peut-être rapproché celui-ci de la machine à différence de Babbage¹ ou l'Eniac² (Haigh & Priestley, 2015) alors que les acheteurs de l'ordinateur l'ont sans doute rapproché d'une calculatrice n'ayant pas d'autre référentiel. En ce sens, le processus d'ancrage permet au groupe d'évaluer le nouvel événement à la lumière de

¹ La machine à différence de Babbage est une calculatrice mécanique conçue pour calculer des tables de fonctions polynomiales construite en 1821 par Charles Babbage.

² Eniac est l'ancêtre de l'ordinateur créé en 1945. Eniac est l'acronyme pour : « *Electronic Numerical Integrator and Computer* ».

son propre référentiel et de l'ensemble de ses valeurs (Jodelet, 1991). Ces valeurs sont hiérarchisées dans un système préexistant de la pensée cohérent et propre à chaque groupe. Ainsi, ce dernier permet d'attribuer des significations à l'évènement afin d'en réduire la nouveauté (Jodelet, 1991). Cette familiarisation facilitera l'intégration de ce nouveau fait dans ce système (Jodelet, 1991). Cette étape ne se fait pas sans heurt, car elle peut se confronter au système de pensée préexistant (Jodelet, 1991).

Mais l'ancrage ne se limite pas uniquement à l'attribution de sens face à une situation nouvelle ou à l'intégration d'un nouvel objet dans un système de pensée existant. En effet, selon Jodelet (1984), l'ancrage exprime aussi l'identité du groupe à travers son système de valeurs qui produit des significations uniques à celui-ci comme pourrait le faire une signature ou un étendard. Par exemple, il y a plus de chance qu'une personne technophobe voit l'apparition de l'ordinateur personnel d'un mauvais œil et développe un discours en conséquence comparativement à une personne technophile.

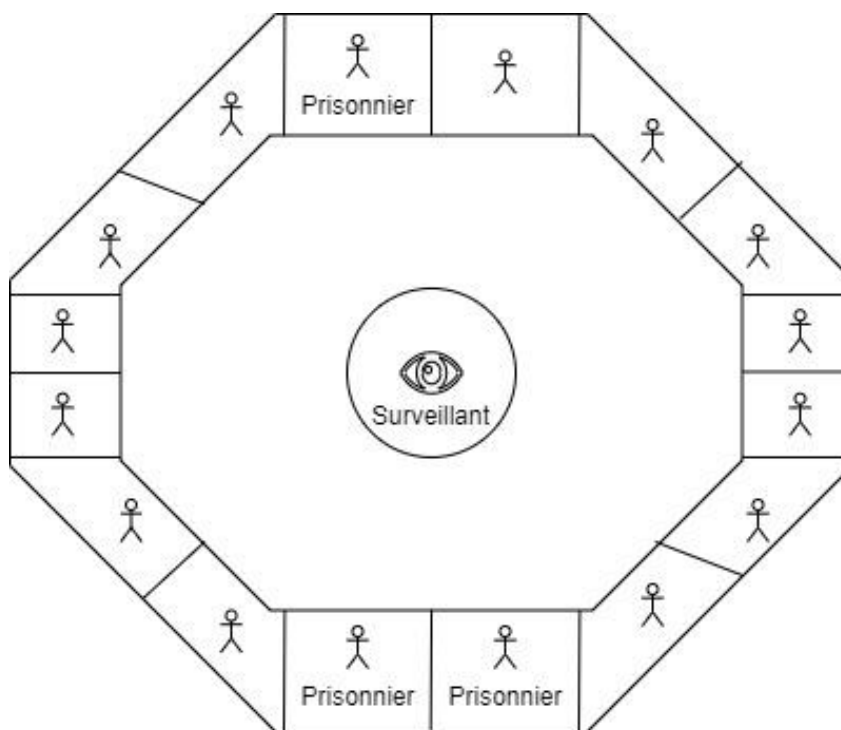
Enfin, l'ancrage permet de sélectionner des éléments pertinents pour l'objectivation. À ce propos, Moscovici (1961) parle de référentiel de valeurs et de connaissances permettant un choix préalable d'éléments à conserver pour l'objectivation. Celle-ci se déploie à travers trois phases que Jodelet détaille (1989, p. 73) : « construction sélective – schématisation structurante – naturalisation ».

Durant la phase de construction sélective, les savoirs scientifiques sont alors décontextualisés et sélectionnés pour créer de l'information facilement assimilable par

l'individu hors du champ scientifique. Cette sélection se fait à la lumière de l'ancrage précédemment opéré. Par exemple, un vulgarisateur scientifique peut donner une idée de l'augmentation du trafic de données mobiles en affirmant que celle-ci correspond à la capacité de plus d'un milliard de DVD, au lieu de dire qu'il est passé de 400 pétaoctets (10^{15}) à 7200 pétaoctets de 2011 à 2016 (Cisco, 2017). Ainsi, en décontextualisant et sélectionnant les informations qui caractérisent la capacité de stockage d'un DVD, le vulgarisateur peut ainsi représenter par analogie l'augmentation du trafic de données. Une personne centrée sur le concept d'écologie aurait pu choisir un équivalent en nombre de feuilles de papier comprenant chacun un certain nombre de mots. On remarque donc que l'ancrage a préalablement orienté la sélection des éléments, s'ajustant bien avec le système de pensée préexistant du groupe, système ici lié à l'informatique. Cette construction sélective est aussi tributaire d'éléments culturels et normatifs (Jodelet, 1991). Les éléments culturels correspondent à l'accès à l'information selon le groupe d'appartenance de l'individu (Jodelet, 1991). Dans l'exemple précédemment développé, on peut supposer qu'informaticien et écologiste n'aient pas accès au même type d'information du fait qu'ils appartiennent à deux groupes sociaux différents. Par exemple, il est peu probable qu'un informaticien sache qu'un téraoctet est équivalent à 1300 armoires de stockage papier, soit l'équivalent de 6,5 millions de pages de document PDF (Dropbox, 2021). Quant à l'aspect normatif, il est décrit par Jodelet (1991) comme tout élément n'entrant pas en contradiction avec le présent système de valeur. On peut donc considérer que la magie ne soit pas considérée comme un élément explicatif du fonctionnement de l'ordinateur au XXI^e siècle, car elle rentrerait en contradiction avec la science moderne alors qu'elle aurait pu l'être au XVI^e siècle.

Une fois cette phase terminée, la phase de schématisation structurante consiste en un réagencement des informations qui les rendent concrètes dans un noyau figuratif. Ce noyau permet à l'individu de réduire la complexité des informations accumulées à travers un modèle basique, réel et imagé tout en conservant une cohérence au regard des normes sociales et de la culture (Abric, 1994). Souvent, la métaphore et la métonymie sont citées comme exemple d'illustration imagée du savoir scientifique lors de l'objectivation (Moliner, 2015). C'est donc à travers ce noyau que l'individu peut interpréter les nouveaux événements auxquels il est confronté, les communiquer et agir sur ces derniers.

Par exemple, une étudiante en science politique qui devrait faire un exposé sur les possibilités de surveillance offertes par les technologies numériques pourrait examiner le fonctionnement de Facebook et le rapprocher de celui d'un panoptique. Dans ce dernier, chaque prisonnier est surveillé par un surveillant, mais peut aussi l'être par ses codétenus, comme le montre la Figure 3. De façon comparable sur Facebook, chaque personne expose sa vie aux yeux de tous et peut accéder à la vie d'autrui, la juger par des *likes* ou même la commenter. L'entreprise Facebook fait office de surveillant général en recueillant l'ensemble des données personnelles que les gens consignent sur ce site. À travers la métaphore du panoptique, l'étudiante pourrait ainsi souligner le risque de surveillance pour l'individu qui l'utilise, au lieu de faire référence à des détails techniques comme l'architecture informatique de Facebook, à ses algorithmes ou aux moyens techniques de collecte des données personnelles.

Figure 3*Principe du panoptique*

Enfin, selon Jodelet (1991), la phase de naturalisation repose sur la transformation des informations vagues en informations concrètes (facilement signifiante comme un symbole). C'est à ce moment que le noyau figuratif revêt son statut d'évidence pour l'individu : « il [le noyau figuratif] est pour lui la réalité même, il constitue le fondement stable autour duquel va se construire l'ensemble de la représentation » (Abric, 1994, p. 21). À la fin de cette étape, les savoirs scientifiques devenus évidents rejoignent le sens commun. Le caractère évident qu'obtient la représentation sociale à la fin du processus fait apparaître le raisonnement individuel comme une « réalité extérieure et objective » (Moliner, 2015, p. 83).

Dans une ultime phase, l'ancrage se déploie de nouveau à la suite du processus d'objectivation. Rappelons que l'individu s'est déjà constitué une grille d'interprétation permettant la création de sens pour le nouvel objet qu'il rencontre dans les précédentes étapes. Ce système n'appréhende pas seulement les objets rencontrés par les individus. En donnant du sens au nouveau, l'ancrage permet d'expliquer et de gérer le monde dans lequel évolue le groupe en classant autant les nouveaux événements que les autres groupes. Il permet aussi à l'individu de communiquer à l'intérieur de son propre groupe.

Si les travaux de Moscovici (1961) sont les bases théoriques des représentations sociales connues, ils marquent aussi le fondement de l'approche sociogénétique. Mais les travaux de Moscovici servent aussi de fondations aux nouveaux courants comme les courants sociodynamique (Doise, 1986), dialogique (Marková, 2016) et structuraliste (Abric, 2003). C'est d'ailleurs ce dernier courant des représentations sociales que nous avons choisi pour la présente étude. La partie suivante vise à le détailler.

École d'Aix-en-Provence et approche structuraliste

En 1976, Jean-Claude Abric soutient sa thèse de doctorat intitulé « Jeux, conflits et représentations sociales ». Il y défend le concept du noyau central qui prend racine dans les travaux de Moscovici sur les représentations sociales, notamment le processus d'objectivation et le noyau figuratif. Abric (1994) reprend l'ensemble des fonctions d'une représentation permis par les processus d'ancrage et d'objectivation. Il les réagence en quatre fonctions cognitives : celle de savoir, d'orientation des conduites, identitaire et

justificatrice. Pour ne pas être trop répétitif, voici succinctement les définitions des fonctions données par Abric (1994) :

- la fonction de « savoir » qui permet de saisir la réalité et de l'interpréter;
- la fonction d'« identité » qui permet la caractérisation de l'identité et la défense des spécificités des groupes;
- la fonction d'« orientation » qui permet de gérer et d'agir sur le monde à travers des comportements et des pratiques concrètes;
- la fonction « justificatrice » qui permet de légitimer les comportements et les prises de position.

Abric (1994) étoffe le concept de noyau figuratif avec les théories de Asch (1946), Rosch (1973) et Heider (1925) qui l'incitent à penser le noyau comme une structure d'éléments hiérarchisés. En effet, Abric (1994) pense que la représentation est constituée de différents éléments qui n'auront ni le même rôle ni la même importance. Selon ces deux caractéristiques (rôle et importance), les éléments se placent dans différentes parties de la représentation qu'il convient de décrire.

Selon Abric (1994), les éléments les plus importants, aussi appelés « centraux », vont se regrouper dans une partie unifiée, stable, cohérente, nommée le noyau central ou système central (Abric, 1994). Sa constitution est déterminée par le contexte social dans lequel le groupe évolue (faits historiques, sociologiques et idéologiques vécus par le groupe), par les normes et les valeurs du groupe. Dans la représentation sociale, le noyau

central permet de maintenir la cohérence et la pérennité de celle-ci. Selon Abric (2003), ces deux capacités peuvent être expliquées grâce à deux de ses fonctions :

- La fonction génératrice de sens grâce à laquelle le noyau détermine le sens des éléments hors noyau. Le noyau peut ainsi attribuer ou même changer le sens des éléments hors noyaux. Il assure ainsi la cohérence de la représentation.
- La fonction organisatrice grâce à laquelle le noyau explicite la nature des liens des différents éléments. Il garantit ainsi la stabilité et l'unification de la représentation.

À titre d'exemple, imaginons la situation où l'on s'intéresse à la représentation sociale des TIC. Admettons que le chercheur a récupéré l'ensemble des éléments de la représentation, notamment son noyau central. Comme nous le disions précédemment, la fonction génératrice de sens va créer le sens de la représentation et celui des éléments moins importants de celle-ci. Imaginons pour les besoins de l'explication que l'élément *cookies* soit recueilli comme élément non central des TIC. À travers la fonction génératrice de sens, le noyau va expliciter le sens du mot « *cookies* » afin que les individus comprennent qu'il s'agit d'un témoin de connexion et non d'un dessert. Quant à la fonction organisatrice, elle ordonne les différents éléments afin que des éléments comme « *cookies* » (témoin de connexion) soient considérés comme moins importants dans la représentation. On peut imaginer que le mot « informatique » soit par exemple un mot décrit comme plus important pour décrire les TIC que « *cookie* ».

Concernant les autres éléments moins importants de la représentation, ils se rassemblent dans le système périphérique, d'où leurs appellations d'éléments

périphériques. Ce système s'occupe de créer la jonction entre la réalité et le noyau central (Abric, 1994). Le système périphérique remplit trois fonctions, que Abric (1994) définit comme suit :

- la fonction de « concrétisation » qui permet à la représentation d'être concrète, compréhensible et facilement transmissible entre les individus;
- la fonction de « régulation » qui permet, quant à elle, une certaine adaptation de la représentation au changement de contexte. Grâce à cette fonction, la représentation peut s'enrichir de nouvelles informations ou de changements dans la réalité (p. ex., ajout d'Internet dans les TIC);
- la fonction de « défense » qui agit comme « un parechoc » (Abric, 1994, p. 26). Le système périphérique agit comme une zone tampon où l'importance des éléments peut changer, en interprétant les nouvelles informations et intégrant de façon conditionnelle des éléments contradictoires (p. ex., des limites technologiques qui ont pu paraître contradictoires à un moment et qui ne le sont plus grâce au progrès).

Pour illustrer la fonction de concrétisation, la recherche de Béch  (2017) qui s'int resse aux TIC   l' cole l'exemplifie de fa on int ressante. Dans cette recherche, la fonction de concr tisation s'illustre avec d'un c t  des  l ments centraux de la repr sentation comme *communication, recherche, technologie informatique* et d'un autre c t , des  l ments p riph riques plus concrets comme *Facebook, clavardage, jeu, musique, film, blogue, site pornographique*. Concernant la fonction de r gulation et de d fense, il est plus dur de trouver des illustrations dans la recherche. Cependant, on peut imaginer un exemple tel que la cr ation de Facebook. Dans les ann es 2000, la

représentation sociale d'Internet ne comportait pas encore Facebook. On peut imaginer que cela a pu changer et que Facebook s'est vu intégré au fil du temps grâce au système périphérique. De nos jours, il est possible que Facebook fasse partie de la représentation sociale d'Internet au vu de l'importance toujours plus grande qu'il a pris. Ainsi, un élément nouveau de la réalité d'Internet, à savoir Facebook (créé en 2004), a pu progressivement intégrer la représentation. Cela illustre autant la fonction de régulation avec l'intégration d'un nouvel élément Facebook que la fonction de défense où l'importance de l'élément a pu être calibrée en fonction de la réalité d'Internet. C'est pourquoi, à la suite de l'exposition des différentes parties de la représentation et leurs fonctions, il est important que les chercheurs suivent une méthode afin d'y accéder.

Abric (1994) propose trois étapes à ce sujet. Une première étape consiste en un recueil du contenu de la représentation. Une seconde étape est l'étude de l'organisation des contenus dans les différentes zones de la représentation à travers l'importance des mots et leurs hiérarchies ainsi que leurs relations. Enfin, une dernière étape est celle de l'évaluation des éléments centraux.

La première étape initiale du repérage du contenu de la représentation sociale peut s'effectuer de plusieurs manières (Abric, 1994, 2003; De Rosa, 2003; Lo Monaco, Delouée et al., 2016; Moliner & Lo Monaco, 2017). Parmi elles, on retrouve (Moliner & Lo Monaco, 2017) l'entretien (Jodelet, 1989), les questionnaires (Moscovici, 1961), l'observation (Jodelet, 1989), les images (Moliner, 1996, 2016) ou dessins et supports graphiques (Abric, 1994), l'analyse documentaire (Abric, 1994; Moliner & Lo Monaco,

2017) et enfin, les méthodes associatives qui sont très largement représentées (Abric, 1994, 2003; De Rosa, 2003; Flament & Rouquette, 2003; Lo Monaco, Delouée et al., 2016; Moliner & Lo Monaco, 2017; Vergès, 1992).

Parmi les méthodes associatives, on retrouve dans une faible mesure la carte associative (De Rosa, 2003) et dans une large mesure l'association libre (Abric, 1994, 2003; Flament & Rouquette, 2003; Lo Monaco, Delouée et al., 2016; Moliner & Lo Monaco, 2017; Vergès, 1992). La carte associative consiste à demander un nombre déterminé d'associations libres à la suite d'un mot inducteur utilisé. À la suite de cette première série d'associations, on demande au participant de faire une seule et unique association libre avec chacun des mots précédemment donnés par le participant. Concernant l'association libre, celle-ci consiste à demander au participant un certain nombre de mots en lien avec le mot inducteur. Elle peut être enrichie d'un second temps qui est appelé évocation hiérarchisée, où le chercheur demande à l'individu de classer par ordre d'importance les mots qu'il a produits dans l'association libre. Cette possibilité fut proposée en 1994 lorsque Abric a questionné le choix du rang d'apparition des mots comme un indicateur de leur importance. Il explicite que le rang d'apparition n'est pas nécessairement garant de l'importance du mot. Afin de répondre à cette critique, il proposa de hiérarchiser les mots dans une seconde tâche. À la suite de l'association libre ou de l'évocation hiérarchisée, le chercheur obtient des hypothèses sur la structure de la représentation qu'il vérifie dans la dernière étape de l'étude des représentations.

La seconde étape de l'étude des représentations sociales est l'analyse de l'organisation des contenus dans les différentes zones de la représentation. La hiérarchie, l'importance des mots recueillis ainsi que les relations qu'ils entretiennent peuvent s'évaluer d'une façon quantitative ou qualitative (Moliner, 1994).

Dans le cadre d'une évaluation quantitative, deux indicateurs s'offrent au chercheur : la saillance et la connexité. Concrètement, la saillance s'évalue par la fréquence d'un élément. Plus celle-ci est grande, plus l'élément est saillant et par conséquent, plus il est susceptible d'être central. Concernant la connexité, c'est la fréquence des liens d'un élément avec les autres qui permet l'évaluation (Moliner, 1994), que ce soit par cooccurrence d'éléments associés par la fréquence des liens interéléments fournis par les participants directement lors d'une carte associative ou indirectement lors d'un *Questionnaire de caractérisation*. Plus cette fréquence est grande, plus l'élément est connexe et par conséquent, plus il est susceptible d'être central. Si le chercheur veut évaluer la connexité, il peut aussi se servir d'un *Questionnaire de caractérisation* (Abric, 2003). À partir des mots recueillis dans la première étape, on demande au participant de constituer des ensembles de mots selon qu'il les considère comme caractéristiques de la représentation ou non. On utilise ensuite l'analyse de similitude pour traiter les résultats et mettre en exergue les liens les plus importants (Flament, 1981; Vergès & Bouriche, 2001).

Dans le cadre d'une évaluation qualitative, deux autres indicateurs sont possibles : la valeur symbolique et l'associativité. La valeur symbolique est la capacité d'un élément

central à faire référence implicitement ou explicitement à sa représentation sociale, par exemple en l'absence de l'élément central en question, l'individu ne reconnaît plus l'objet de la représentation sociale. Dans son article, Moliner (2016) prend la représentation sociale des études où l'élément « permet d'obtenir un diplôme » fait directement référence aux études alors que les éléments d'« activité transitoire et valorisante » pourraient faire référence aux études, comme au travail ou au bénévolat. Pour évaluer la valeur symbolique, Moliner (1988) propose d'utiliser le *Questionnaire de mise en cause* pour évaluer cet aspect. Ce questionnaire consiste à remettre en cause chaque élément susceptible d'être central à travers des tournures négatives. Du fait de sa propension à inclure une double négation qui nuit à la tâche, une technique plus directe et sans double négation a été proposée par Lo Monaco et al. (2008) : le *Test d'indépendance au contexte*. Enfin, l'associativité s'évalue en demandant au participant d'identifier le type de lien qui existe entre les éléments (du mot inducteur aux mots induits). À ce jour, une seule technique permet de faire cela, il s'agit du *Questionnaire des schèmes cognitifs de base* de Rateau (1995).

La dernière étape consiste à tester les éléments recueillis ultérieurement et susceptibles d'être centraux. Une fois celle-ci effectuée, le chercheur obtient les éléments centraux constituant la représentation. Pour ce faire, le chercheur a trois possibilités pour tester la centralité des éléments : le *Test de mise en cause*, l'*Introduction du scénario ambigu* ou le *Test d'indépendance au contexte*. Par exemple, le *Test d'indépendance au contexte* permet de savoir si l'élément est « trans-situationnel » et donc central (Lo Monaco et al., 2008). L'aspect trans-situationnel peut se comprendre comme la capacité

de l'élément central à rester toujours valable, peu importe le contexte ou la situation (Lo Monaco et al., 2008).

Après avoir examiné l'ensemble des étapes de l'étude d'une représentation sociale, il nous semble important de connaître l'impact du contexte sur la production des éléments recueillis durant la recherche. En effet, le contexte peut modifier les éléments recueillis lors des différentes étapes de la recherche (Abric, 2003; Piermattéo et al., 2018; Wolter, 2011). Bien que l'effet de contexte ne soit pas défini par les auteurs (Abric, 2003; Piermattéo et al., 2018; Wolter, 2011), on pourrait l'énoncer comme suit : il s'agit de mesurer l'effet des variations du contexte de production des éléments recueillis pour l'étude des représentations (p. ex., éléments produits lors de l'association libre) et d'étudier l'impact de ces effets sur les résultats de l'étude. Au meilleur de nos connaissances, nous avons trouvé deux types de méthode pour étudier cet effet. La première méthode demande de faire l'exercice d'association libre du point de vue d'un autre groupe social; on parle dans ce cas de technique de substitution (Abric, 2003). Par exemple, dans son étude, Abric (2003) demande à l'individu non pas d'associer librement des mots au terme « Magrébin », mais plutôt d'associer librement des mots en se mettant à la place d'un Français. La consigne pourrait être : « À votre avis, si on demandait aux Français en général (au « grand public ») d'associer cinq mots ou expressions avec le mot « Magrébin », quels seraient-ils? ». La seconde méthode modifie le statut de l'expérimentateur, on parle alors de décontextualisation normative (Abric, 2003). Concrètement, si l'on reste sur l'exemple de l'étude de la représentation des Magrébins, le chercheur propose deux types de contexte (Abric, 2003), soit celui dit normatif, où le

chercheur se présente comme appartenant au même groupe que les personnes constituant l'échantillon (Abric, 2003), soit celui dit décontextualisé, où le chercheur se présente comme proche du groupe, mais dont les valeurs sont perçues comme différentes (Abric, 2003). Les résultats obtenus varient d'une condition à l'autre comme en témoigne la Figure 4.

Figure 4

*Schéma comparatif des représentations sociales du terme « Magrébin »
adapté de Abric (2003, p. 79)*

	Contexte normatif		Décontextualisation normative	
	Premier rangs	Derniers rangs	Premier rangs	Derniers rangs
Très fréquent	Géographique Culture Racisme Chaleureux	Religion	Géographique Délinquance Racisme Immigration	Victimes de racisme Pour l'intégration
Peu fréquent	Immigration Pour l'intégration Différence Communauté	Égaux Délinquance Cités On ne veut pas les intégrer Victimes de racisme		Ne veulent pas s'intégrer Nourriture Bas niveaux économiques

Enfin, l'étude de Wolter (2011) et celle de Piermattéo et al. (2014) observent toutes deux que les situations dans lesquelles sont produits les éléments ont un impact sur la représentation. Par exemple, Piermattéo et al. (2014) ont étudié la représentation sociale des Gitans. Selon les conditions de cueillette, les résultats étaient différents : le mot

« voleur » était rapporté 11 fois en contexte privé avec un expérimentateur non-Gitan contre trois fois en contexte public avec un expérimentateur non-Gitan.

Au meilleur de nos connaissances, aucune étude ne propose d'énoncer un contexte donné pour l'objet de représentation (ici : le contexte technologique pour la vie privée). Or, comme nous l'avons vu dans la partie précédente, les chercheurs s'accordent pour dire que la notion de vie privée dans un contexte technologique s'articule différemment que celle dans un contexte hors technologie (Acquisti et al., 2015; Gross & Acquisti, 2005; Kokolakis, 2017; Li, 2012; Smith et al., 2011). Au vu de l'impact que peut avoir le contexte, il nous semble capital pour notre étude de préciser celui-ci dans la consigne pour éviter toute ambiguïté aux participants.

Pour conclure, le fait que les méthodologies des représentations sociales permettent de préciser le contexte (technologique) de l'objet de représentation (la vie privée) évite de tomber dans l'écueil d'étudier uniquement la vie privée de façon générale (Smith et al., 2011) et permet l'accès à ses spécificités. Il est évident qu'il ne s'agit pas du seul avantage à utiliser cette méthodologie. Comme nous l'avons vu à travers les différentes étapes de l'étude d'une représentation, nous avons accès directement à l'agencement de la réalité des individus ainsi qu'à leurs connaissances. Ainsi, le principal avantage que nous voyons à utiliser la théorie des représentations dans son approche structurale est d'avoir accès directement aux quatre fonctions que la représentation possède. En particulier, nous avons accès à la fonction de savoir qui permet d'interpréter le monde et la fonction d'orientation qui permet d'agir sur le monde. Par conséquent, si les éléments constituant la

représentation n'incluent pas d'éléments permettant d'appréhender les dangers pour la vie privée en contexte technologique ou la nécessité de la protéger, alors il est possible que cette absence explique la relative faiblesse ou le manque de comportement protecteur de la vie privée. Par la même, cela offre une piste intéressante pour expliquer le paradoxe de la vie privée. En effet, pourquoi l'individu protégerait-il quelque chose pour lequel il ne percevrait aucun danger?

Problématique

Dès le début des recherches sur la vie privée (Altman, 1975; Westin, 1967), des interrogations ont émergé sur la possibilité d'une définition consensuelle du concept (Burgoon, Parrott, Le Poire, Kelley, Walther et Perry, 1989; Parent, 1983; Schoeman, 1984). Peu importe le domaine, il existe une pluralité de points de vue concernant les problématiques de la vie privée. Il est ainsi possible de répertorier minimalement 14 théories qui visent à modéliser la vie privée en contexte technologique (Li, 2012). Ce foisonnement théorique et l'absence de définition consensuelle constituent un enjeu scientifique qui pourrait se résumer ainsi : quand nous abordons le concept de vie privée, référons-nous à une seule et même réalité? (Kokolakis, 2017; Paine et al., 2007; Smith et al., 2011). De plus, parmi ces théories, celles étudiant l'aspect social de la question restent rares.

Un autre enjeu que nous avons abordé concerne un phénomène fréquemment mentionné dans la littérature sur la vie privée et qui demeure difficilement explicable : le

phénomène du « paradoxe » de la vie privée¹. L'intérêt envers ce paradoxe s'observe à travers la dizaine de théories qui tente de l'expliquer (Barth et de Jong, 2017; Gerber, Gerber et Volkamer, 2018). Encore une fois et malgré leur nombre, ces recherches abordent peu l'aspect social du paradoxe de la vie privée.

Or, l'analyse sociale de la vie privée et de ses problématiques nous paraît indispensable, ou tout au moins complémentaire à l'analyse individuelle, ne serait-ce qu'en raison du fait que le pendant de la vie privée, soit la « vie publique », s'avère intrinsèquement sociale. Il faut donc tenir compte de la nécessité d'avoir une double lecture, à la fois individuelle et sociale, pour bien comprendre le phénomène. De plus, notre problématique prenant place dans le contexte TIC, il nous semble difficile de rendre compte de la question en faisant l'économie de l'aspect communicationnel. En effet, cet aspect nous semble lui aussi intrinsèquement social : sans interlocuteur, il n'y a pas de communication possible.

Même si l'absence d'un prisme de lecture sociale est un enjeu important des recherches portant sur la vie privée et son paradoxe, il ne faut pas pour autant exclure le prisme de lecture plus individuel. Toutefois, malgré l'avantage d'adopter cette double lecture, on relève peu d'études abordant la problématique dans cette perspective. En effet, Bélanger et Crossler (2011) qui ont recensé 142 études en relèvent 88 qui présentaient un niveau d'analyse individuelle, aucune n'avait un niveau d'analyse groupal, 32 qui avaient

¹ Pour rappel, le paradoxe de la vie privée est un phénomène présumé par de nombreux chercheurs et qui consiste en une faible protection de la vie privée alors que les préoccupations la concernant seraient importantes chez les individus (Brown, 2001).

un niveau organisationnel, neuf présentaient un niveau sociétal et le reste présentait des niveaux d'analyses mixtes (11 études).

Enfin, dans les revues de littérature (Barth et de Jong, 2017; Baruh, Secinti et Cemalcilar, 2017; Bélanger et Crossler, 2011; Gerber, Gerber et Volkamer, 2018; Kokolakis, 2017; Smith, Dinev et Xu, 2011) ainsi que les articles qu'elles répertoriaient, une perspective d'expert émanant du chercheur domine dans l'étude des standards de comportements de protection et des préoccupations concernant la vie privée. Peu d'articles prenaient en compte la perspective de l'utilisateur pour déterminer si l'attitude apparaissent paradoxale dans la perspective de l'utilisateur. En effet, il peut sembler évident pour un chercheur qu'une certaine pratique numérique puisse s'avérer risquée envers la vie privée. Pour autant, est-ce que l'utilisateur a conscience du risque? Dans la négative, il semble subjectif de qualifier le comportement de « paradoxal ». Or, en inventoriant les recherches qualitatives comportant des questions ouvertes qui permettraient d'accéder directement à la vision de l'utilisateur, il apparaît qu'elles demeurent marginales par rapport à l'ensemble des études (Baruh, Secinti et Cemalcilar, 2017).

Pour examiner ces enjeux, prendre en compte les conceptions qu'entretiennent les utilisateurs, avec une approche centrée sur l'utilisateur, pourrait s'avérer pertinente pour expliquer une des problématiques récurrentes (Barth et de Jong, 2017) dans le champ de la vie privée et des technologies, celle du « paradoxe » de la vie privée qui occupe d'ailleurs une place croissante dans la littérature scientifique (voir Figure 1).

Il nous paraît donc important d'aborder la question de la vie privée en contexte technologique à travers une vision psychosociale laissant place à la fois à l'analyse individuelle et à l'analyse sociétale, tout en accordant un maximum d'espace à la parole de l'utilisateur. Dans cette optique, les méthodologies des représentations sociales sont pertinentes, car elles se centrent sur l'acteur social, sa conception d'un objet et son contexte. Par exemple, la possibilité qu'offre l'approche des représentations sociales de recueillir du matériel grâce à des questions ouvertes (associations libres) permet davantage de prendre en compte la singularité de l'expérience de l'utilisateur.

La présente thèse s'inscrit dans l'approche structurale des représentations. Nous avons donc choisi la définition de Abric (1994) comme point de départ. Pour cet auteur, une représentation est « un système d'interprétation de la réalité qui régit les relations d'individus à leurs environnements physique et social » (Abric, 1994, p. 13) et qui va déterminer leurs comportements ou leurs pratiques. La représentation est « un guide pour l'action, elle oriente les actions et les relations sociales. Elle est un système de prédécodage de la réalité, car elle détermine un ensemble d'anticipations et d'attentes » (Abric, 1994, p. 13). Le choix des représentations sociales comme cadre théorique dans cette recherche se justifie ainsi par sa position à la jonction entre la psychologie individuelle et la psychologie sociale, et ce faisant, s'adapte à la nature complexe de notre objet de recherche.

Par conséquent, les méthodes et techniques de recherche de cette approche permettront de dépasser la focalisation essentiellement individuelle des études

précédemment évoquées. Ainsi, à travers celles-ci, nous souhaitons aborder la question dans une perspective plus proche de l'individu et éviter au maximum les préconceptions du chercheur, minimisant la possibilité de capter un artéfact comme le soulignent Dienlin et Trepte (2015). De plus, la méthodologie de l'approche structurale permet de rendre compte de la réalité à laquelle est confronté l'individu grâce à la représentation sociale. En effet, son étude approfondie permet de faire émerger ses différentes parties pour mieux comprendre le système de pensée des individus (fonction de savoir), ainsi que leurs actes (fonction d'orientation). Enfin, grâce à l'analyse du noyau central et de ses éléments, nous aurons accès au sens des éléments de la représentation sociale (fonction génératrice de sens), mais aussi à la façon dont les éléments sont organisés, liés et l'importance qu'ils peuvent avoir dans la représentation (fonction organisatrice).

De cette façon, nous souhaitons savoir si le paradoxe de la vie privée relevé par les chercheurs est conscientisé chez le participant en étant déjà présent au sein même de la représentation sociale de la vie privée, plus précisément par l'observation de l'absence ou de la présence d'éléments centraux ayant comme thématique la protection de la vie privée en contexte technologique. Si la présence de tels éléments se confirme dans la représentation, alors le constat des chercheurs concernant le paradoxe serait confirmé et pourrait être approfondi selon la prépondérance du thème de la protection ou du danger au sein du noyau. À l'inverse, si ces éléments sont absents du noyau central, cela pourrait expliquer que les protections des utilisateurs soient minimales ou absentes. Si de tels résultats émergeaient, nous pourrions éclairer le paradoxe de la vie privée sous un autre

angle : si les utilisateurs ne conçoivent pas d'enjeux de sécurité dans la vie privée en contexte technologique, il ne serait pas paradoxal qu'ils ne la protègent pas.

Objectif

Un des objectifs de cette thèse est d'approfondir les connaissances sur la vie privée en contexte technologique en proposant une grille d'analyse psychosociale du phénomène centrée sur les conceptions de l'individu à travers l'approche structurale (Abric, 2003).

La première collecte de la thèse s'inscrit dans une perspective exploratoire. La seconde collecte de la thèse comporte une partie confirmatoire, notamment avec le *Test d'indépendance au contexte*, et exploratoire, avec l'analyse de similitude. Enfin, un troisième objectif porte sur une présentation didactique du courant structuraliste des représentations sociales. En effet, durant le traitement des résultats de la première collecte de la thèse, des questionnements conceptuels sont apparus sur le traitement des données habituellement préconisé, et font l'objet d'une réflexion et ainsi que des propositions visant à solutionner les problèmes soulevés.

Questions de recherche et hypothèses

Dans la partie exploratoire de la thèse, plusieurs questions de recherche portent sur la vie privée :

- Quelles sont les représentations sociales de la vie privée dans un contexte d'usage des technologies de l'information et des communications d'étudiants universitaires français?

- Quels sont les différents items qui constituent le noyau central de la représentation?
- Comment s'organisent les items du noyau central entre eux et avec d'autres items de la représentation?

D'autres questions sont plus centrées sur les méthodes utilisées :

- Quelle statistique représente le meilleur critère de décision pour opérationnaliser l'analyse prototypique?
- Quels éléments théoriques et statistiques peuvent être considérés pour éviter de faire reposer l'analyse sur un seuil arbitraire pour définir l'appartenance des items au noyau central?

Concernant la partie plus confirmatoire de la thèse, les hypothèses de recherche sont les suivantes :

- Le noyau central de la représentation sociale comportera des items liés au thème de la protection.
- À l'intérieur du noyau central, les items ayant un lien avec la protection seront majoritaires comparativement aux autres thèmes des différents items.
- À titre exploratoire, certaines variables sociodémographiques ont été testées avec les différents résultats utilisés pour évaluer l'orientation concernant la vie privée ainsi que les comportements sécuritaires en lien avec la technologie.

Méthode

Afin de répondre au mieux aux questions et hypothèses de recherches, nous avons élaboré deux collectes de données. Dans les sections suivantes, nous présenterons les caractéristiques de nos échantillons ainsi que les méthodes qui nous ont permis d'effectuer nos collectes de données. Dans les sections suivantes, nous présenterons les caractéristiques de nos échantillons ainsi que les méthodes qui nous ont permis d'effectuer nos collectes de données.

Participants

La présente section vise à détailler la procédure de recrutement des échantillons ainsi que leurs caractéristiques.

Première collecte de données

L'échantillon du premier temps est un échantillon de convenance constitué d'étudiants de l'enseignement supérieur français. Cinq compensations de 20€ en cartes Amazon étaient prévues pour favoriser le recrutement. Le recrutement s'est déroulé sur le réseau social Facebook, par une publicité ciblée et dans plusieurs groupes Facebook d'étudiants de l'enseignement supérieur de la France. Nous avons élaboré la publicité (voir Appendice A) pour qu'elle rejoigne un maximum d'étudiants et qu'elle apparaisse dans leur fil d'actualités. Le questionnaire a été mis en ligne via la plateforme Limesurvey cloud et hébergé sur celle-ci (Schmitz et LimeSurvey Project Team, 2012).

Parallèlement aux campagnes publicitaires, 52 groupes Facebook ont été choisis pour recruter des participants. La sélection des groupes s'est faite en entonnoir. Premièrement, nous avons commencé par cibler les groupes généraux et avec un grand nombre de personnes inscrites. Une première cible a donc été des groupes d'étudiants de grandes et moyennes villes françaises comme Paris, Marseille, Lyon, Rennes ou Montpellier. Deuxièmement, et une fois que nous avons atteint une dizaine de groupes, nous avons commencé à prospecter, en plus des précédents groupes, des groupes légèrement plus restreints en choisissant les groupes des universités. Lors du choix de ces groupes, il a fallu demeurer vigilant à la taille de la ville de l'université ciblée pour éviter les erreurs d'échantillonnage. En effet, si la ville dans laquelle était implantée l'université était de petite taille ou de taille moyenne, l'université était souvent la seule implantée et par conséquent, avait des domaines d'études assez diversifiés. En revanche, si la ville était de taille conséquente, il fallait rechercher le nombre exact d'universités qu'elle avait, car dans ce cas, les disciplines sont partagées selon leurs disciplines, par exemple des arts, lettres et langues ou des sciences dures. Il était donc important de cibler les groupes des autres universités de la ville afin d'échantillonner une variété de disciplines d'études et ne pas se retrouver avec des étudiants issus exclusivement de domaines littéraires ou d'administration, par exemple. Cependant, il n'a pas toujours été possible d'intégrer les groupes des universités. Pour ces raisons, nous avons fait le choix, dans un troisième temps, de cibler des groupes de domaines plus spécifiques afin d'équilibrer les étudiants des diverses disciplines.

Une fois l'intégration dans les groupes effectués, nous avons commencé notre recrutement de la façon suivante : un message standard était posté dans chaque groupe (voir Appendice A) et toutes les publications avaient lieu entre 11 h et midi (heure canadienne) pour correspondre à 17 h et 18 h heure française. La collecte a eu lieu de janvier à mars 2020.

Le nombre de participants a pu être défini en fonction du but de la première collecte. Ce but consistait à obtenir un maximum de variétés dans les réponses des participants afin d'établir les items qui seraient candidats au noyau central. Une fois les réponses recueillies, une analyse prototypique était effectuée (Vergès, 1992). L'analyse prototypique est une opération de réordonnement des items recueillis, selon leurs fréquences et leurs importances déclarées par les participants. L'analyse prototypique ne repose pas sur un test statistique et n'a donc pas d'hypothèses statistiques à tester. Pour cette raison, il n'est pas possible de calculer la puissance statistique a priori afin de déterminer une taille d'échantillon à collecter. Pour pallier cela, nous nous sommes donc référés au nombre de participants utilisé dans d'autres études antérieures. Celui-ci est d'environ 150 personnes (Baggio et Renaud, 2015; Lo Monaco, Lheureux et Hamili-Falkowicz, 2008; Vergès, 1992). Comme l'objectif consiste à recueillir un maximum de variétés, un échantillon d'environ 300 participants a été visé, par principe de précaution.

En fin de compte, 378 questionnaires partiellement remplis ont été récoltés et 256 participants ont terminé le questionnaire. Les 378 participants avaient au moins

effectué la tâche d'association libre et la tâche de hiérarchisation des associations. Rappelons aux lecteurs que les items candidats au noyau central sont déterminés uniquement à l'issue de la tâche de hiérarchisation des associations.

Nous présentons ci-dessous l'échantillon qui nous a permis d'établir la liste des items candidats au noyau central, soit les 378 participants qui ont terminé au moins les tâches d'association et de hiérarchisation.

Nous avons dû exclure neuf participants qui ne respectaient pas les conditions d'inclusion. Soit les participants avaient moins de 18 ans ($n = 7$), soit ils n'étaient pas étudiants dans l'enseignement supérieur ($n = 2$). Une fois la base de données nettoyée, nous avons donc 369 participants ayant au moins effectué l'expérimentation jusqu'à la tâche de hiérarchisation permettant d'établir les hypothèses du noyau central de la représentation.

Il s'agissait principalement de jeunes adultes ($M = 22,46$ ans, $ÉT = 5,61$), majoritairement de sexe féminin ($n = 280$) et célibataire ($n = 266$). La grande majorité des participants avait uniquement la nationalité française ($n = 320$). Les étudiants ayant un diplôme de licence (équivalent au baccalauréat québécois) représentaient la majorité des participants ($n = 143$), suivis de près par 123 participants ayant obtenu leur baccalauréat (équivalent au DEC.), 60 étudiants avaient déjà un master complété (équivalent d'une maîtrise québécoise) et trois avaient un doctorat complété. Concernant

les domaines d'études, les étudiants étaient principalement en psychologie ($n = 70$) et en droit ($n = 38$).

À propos des questions politiques, les participants se disaient principalement peu intéressés ($n = 142$) ou assez intéressés ($n = 136$) par celles-ci. Lorsqu'on demandait aux participants de se positionner sur une échelle visuelle analogique où 0 représente un engagement politique à gauche et 10 à droite, on obtenait une tendance plutôt à gauche. En effet, le mode était de 3 ($n = 70$) et la moyenne était de 3,59 ($ÉT = 2,54$). Ce résultat était cohérent avec l'affiliation aux partis politiques qui étaient majoritairement à gauche : Europe Écologie Les Verts (EELV) avec 88 participants et le parti socialiste avec 43 participants. Cependant, il est important de nuancer ces chiffres, car il s'agit de la question où les gens ont le plus explicitement refusé de répondre, avec 98 personnes qui ont voulu garder cette information privée.

Les catégories socioprofessionnelles des parents étaient recueillies. Celle paternelle était principalement « Employés » ($n = 70$), suivie de très près par « Cadres » ($n = 69$). Concernant la catégorie socioprofessionnelle maternelle, le mode était « Employés », avec 115 participants. Enfin, le revenu des participants était majoritairement compris entre 0 et 5000€ avec 176 participants.

Enfin, attendu que les quatre dernières questions permettant de décrire l'échantillon étaient placées à la fin du questionnaire, un grand nombre de données manquantes a été

relevé parmi ces dernières ($n = 118$, soit 31,98 %). La majorité de l'échantillon ne faisait partie d'aucune association défendant la vie privée ($n = 247$, soit 66,94 %). Concernant les questions portant sur CNIL et malgré le nombre conséquent de données manquantes ($n = 118$, soit 31,98 %), 98 participants connaissaient la CNIL (26,56 %) et 108 appliquaient quelques-uns de ses conseils (31,98 %). Pour terminer, les participants savaient ce qu'est un VPN (*Virtual private network*, ou réseau virtuel privé [RVP], $n = 132$, soit 35,77 %)¹, mais plusieurs personnes n'en utilisaient jamais ($n = 117$, soit 31,71 %).

Finalement, toutes les questions sociodémographiques étaient obligatoires et avaient la modalité de réponse suivante : « Je ne souhaite pas répondre ». Cocher cette option de réponse signifiait un refus de répondre à la question concernée et a donc été interprété comme un marqueur d'une information relevant de la vie privée. Il nous paraît donc important de rapporter aux lecteurs les quatre variables sociodémographiques dont le taux de réponse « Je ne souhaite pas répondre » était supérieur à 5 % : l'orientation politique (8,13 %), le parti politique (26,56 %), la catégorie socioprofessionnelle du père (7,32 %) et les ressources financières annuelles, c'est-à-dire le salaire, l'aide parentale, les bourses, etc. (8,67 %) et l'application des conseils de la CNIL en matière de protection de la vie

¹ Un VPN ou [RVP] est une connexion préalable chiffrée qui fait la jonction entre l'appareil utilisé et Internet, ce qui accroît la sécurité et la confidentialité de la navigation sur Internet. Le RVP est particulièrement utile lorsque vous vous connectez à un réseau public (gare, aéroport, bibliothèque, etc.), car il garantit le chiffrement des données comme vos identifiants et mots de passe ou vos données bancaires s'ils sont utilisés afin qu'ils ne puissent pas être récupérés et utilisés.

privée (7,86 %). L'ensemble des données est accessible dans un tableau récapitulatif dans l'Appendice B (voir Tableau 1).

Seconde collecte de données

Concernant la seconde collecte de données, le recrutement était toujours volontaire et les compensations étaient elles aussi identiques à la première collecte. Les deux seuls paramètres ayant changé durant cette récolte de données furent le nombre de groupes prospectés (54, au lieu de 52 dans la première), la date et la durée de recrutement, lequel s'est déroulé du 25 août 2020 au 9 septembre 2020.

Le nombre de participants a pu être défini en fonction de l'objectif de la seconde collecte. En raison du type de tests et des méthodologies utilisées lors de l'analyse, à savoir le *Test de Kolmogorov-Smirnov* (Kolmogorov, 1933) et le *Test de Cramer von Mises* (Cramér, 1928) pour le *Test d'indépendance au contexte*, ainsi que l'analyse de similitude pour le *Questionnaire de caractérisation*, aucun calcul de puissance statistique n'a été mené. En effet, il ne semble pas y avoir de logiciel permettant d'effectuer un calcul de puissance a priori pour le *Test de Kolmogorov-Smirnov* et le *Test de Cramer von Mises*. À la lecture d'articles testant les éléments centraux, nous avons décidé de viser un échantillon minimal de 150 personnes (Methivier, 2010; Pawlowski et Jung, 2015; Sanchez *et al.*, 2016). Quant à l'analyse de similitude, celle-ci est basée sur la théorie des graphes (visée descriptive) et non sur un modèle statistique (Vergès et Bouriche, 2001). Il n'est donc pas possible d'effectuer un calcul de puissance préalable. À la lecture d'articles

effectuant une analyse de similitudes (Aim, Decarsin, Bovina et Dany, 2018; Lheureux et Guimelli, 2009; Salesses, 2005b), nous avons décidé de viser un échantillon minimal de 150 personnes.

Les critères d'inclusion furent les mêmes que pour la première collecte : être majeur et étudiant dans l'enseignement supérieur français. Aucun participant n'a été exclu. À la fin de la collecte, l'échantillon se composait de 227 participants avec des questionnaires partiellement complétés. Les participants ayant effectué le protocole jusqu'aux *Questionnaires de caractérisation* qui permettait de quantifier les relations entre les items étaient au nombre de 172. Les participants ayant rempli le questionnaire suivant, qui était le *Test d'indépendance au contexte* qui permettait de confirmer la centralité des items candidats au noyau central, étaient au nombre de 163.

L'échantillon de 163 participants avait principalement les mêmes caractéristiques que lors de la première collecte, à savoir : jeunes ($M = 22,24$ ans, $ÉT = 4,48$), de sexe féminin ($n = 112$), célibataire ($n = 111$) et de nationalité française ($n = 123$). Comme précédemment, ce sont les étudiants détenteurs d'un diplôme de licence (équivalent au baccalauréat québécois) qui représentaient les catégories les plus fréquentes ($n = 62$), suivis de près par 54 bacheliers (équivalent au DEC); 29 étudiants avaient déjà un master complété (équivalent d'une maîtrise) et un participant avait un doctorat complété. Comme dans le précédent échantillon, ce sont les étudiants en psychologie ($n = 25$) et en droit

($n = 19$) qui étaient majoritaires. Pareillement au Tableau 1, les 40 autres domaines d'études ayant moins de cinq représentants n'ont pas été affichés.

Concernant les questions politiques, les participants rapportaient principalement être assez intéressés ($n = 56$) ou peu intéressés ($n = 54$). Lorsqu'on demandait aux participants de se positionner sur une échelle visuelle analogique où 0 représente un engagement politique à gauche et 10 à droite, on obtenait une tendance plutôt à gauche. Le refus de répondre était la modalité de réponse la plus fréquente ($n = 24$) pour cette question. La seconde modalité la plus fréquente était 4 ($n = 20$) et indique un échantillon plutôt à gauche sur le spectre politique. La moyenne sur cette échelle était de 4,02 ($ET = 2,53$). Ce résultat était cohérent avec le parti politique des participants, qui était majoritairement à gauche : Europe Écologie Les Verts (EELV) avec 29 participants et le parti socialiste avec 7 participants. Ces résultats sont à nuancer, car 44,17 % de l'échantillon a choisi de ne pas divulguer le parti qui le représentait le mieux. Cela revient à 72 personnes qui ont voulu garder cette information privée. Enfin, les catégories socioprofessionnelles des parents étaient recueillies. Autant pour celle du père, que de la mère, la modalité principale était « Employés ». On retrouvait 33 personnes dont le père était employé et 42 dont la mère était employée. La modalité du revenu des participants le plus fréquemment trouvée était de 0 et 5000€.

Pour conclure, la quasi-totalité de l'échantillon ne faisait partie d'aucune association défendant la vie privée ($n = 147$, soit 90,18 %). Concernant les questions portant sur

CNIL, beaucoup de participants connaissaient cet organisme public ($n = 60$, soit 36,81 %) et 59 participants (36,20 %) appliquaient quelques-uns de ces conseils. Pour terminer, les participants rapportaient connaître ce qu'est un VPN ($n = 91$, soit 55,83 %), mais de nombreuses personnes rapportaient n'en faire jamais usage ($n = 74$, soit 45,40 %).

Comme précédemment, nous souhaitons mettre en évidence les variables ayant obtenu plus de 5 % de modalités « Je ne souhaite pas répondre » : le statut marital (6,13 %), l'orientation politique (14,72 %), le parti politique (44,17 %), la catégorie socioprofessionnelle du père (13,50 %), celle de la mère (11,66 %), les ressources financières annuelles (19,02 %) et l'application des conseils de la CNIL (12,88 %). L'ensemble des données est accessible en Appendice C dans un tableau récapitulatif (voir Tableau 2).

Instruments

Certains outils étaient communs aux deux phases de l'expérience : le *Questionnaire sociodémographique* et les tâches d'association libre et de hiérarchisation des associations. Afin d'avoir plus de précision sur le rapport qu'entretenait le participant avec l'objet de la représentation (la vie privée en contexte technologique), nous avons administré les deux questionnaires suivants : la *Privacy Orientation Scale* (Baruh et Cemalcılar, 2014) et la *Security Behavior Intentions Scales (Sebis)* (Egelman et Peer, 2015b). Ils ont été utilisés durant les deux collectes.

Le *Questionnaire sociodémographique* visait à relever les caractéristiques suivantes : l'âge, le sexe, le statut marital, la nationalité, le niveau d'études, le domaine d'études, l'intérêt pour la politique, l'orientation politique, le parti politique représentant le mieux l'individu, la catégorie socioprofessionnelle du père et de la mère, et enfin, les ressources financières annuelles. Le but de ces nombreuses questions sociodémographiques était d'identifier les groupes sociaux précis (Mariotti, 2003; Moliner, 1993; Moliner et Guimelli, 1994) sur plusieurs des niveaux d'analyses de Doise (1982) : intraindividuel, interindividuel, positionnel et idéologique. En quelques mots, Mariotti (2003, p. 4) décrit le niveau intra individuel comme les « variables de personnalité [l'] histoire individuelle », les variables interindividuelles comme celle renvoyant à des « situations de co-action ou d'interaction », les variables positionnelles comme celle renvoyant à des « positions sociales », par exemple, les positions hiérarchiques et enfin, les variables idéologiques qui renvoient à des variables comme des « normes, des croyances, des valeurs ».

Le *Questionnaire d'orientation envers la vie privée* ou *Privacy Orientation Scale* en langue originale (Baruh et Cemalcılar, 2014) évaluait le positionnement de l'individu face à la vie privée. La cohérence interne (alpha de Cronbach) de chaque dimension est bonne, oscillante entre 0,83 et 0,88. Avec l'analyse en composante principale (ACP), les auteurs ont identifié quatre facteurs qui expliquent 69,9 % de la variance. Les quatre facteurs étaient (1) la vie privée en tant que droit; (2) la préoccupation pour la vie privée d'autrui; (3) la préoccupation concernant la protection de la vie privée; et (4) la vie privée comme conditionnel aux autres. L'analyse confirmatoire avait une bonne adéquation (Marsh, Hau

et Wen, 2004) avec les données (CFI = 0,97, RMSEA = 0,05), aucun item n'était complexe et chaque item cotait donc sur un seul facteur. Il est important de signaler que ce questionnaire a nécessité une traduction, car il n'était pas disponible en langue française. Nous renvoyons le lectorat à l'Appendice D pour les enjeux rencontrés liés à cette traduction.

Le questionnaire *Security Behavior Intentions Scales* (Egelman et Peer, 2015b) évalue les comportements sécuritaires vis-à-vis des TIC. La cohérence interne de chaque dimension est moins bonne que celle du précédent questionnaire, mais reste acceptable. Celle-ci oscille entre 0,64 et 0,76. Avec l'ACP avec une rotation Varimax, les auteurs avaient révélé quatre facteurs qui expliquent 55,6 % de la variance. Ces quatre facteurs étaient la sécurité de l'outil, la génération de mots de passe, la conscience proactive (pour la sécurité) et les mises à jour (pour évaluer si l'individu mettait à jour ses logiciels). L'analyse confirmatoire présentait une bonne adéquation avec les données (Marsh, Hau et Wen, 2004) (CFI = 0,92, RMSEA = 0,058) et aucun item n'était complexe, chacun cotant sur un seul des facteurs. Ce questionnaire a nécessité lui aussi une traduction.

Des questions que nous avons conçues ont été ajoutées afin de mieux décrire l'échantillon. Par exemple, nous avons interrogé les participants sur leur éventuelle appartenance à une association défendant les droits à la vie privée. Les participants ont aussi été questionnés sur leurs connaissances de la CNIL ainsi que sur les pratiques qu'elle conseillait. La CNIL prodigue 10 conseils simples pour protéger la vie privée (CNIL,

2016) : (1) réfléchir avant de publier; (2) faire preuve de courtoisie et de respect envers les autres; (3) ne pas tout divulguer; (4) sécuriser ses comptes; (5) créer plusieurs adresses courriel; (6) être vigilant(e) avec les photos et les vidéos; (7) utiliser un pseudonyme; (8) respecter les règles de création d'un mot de passe; (9) nettoyer son historique; et enfin (10) vérifier les traces laissées sur les réseaux. Ces conseils sont tous accompagnés d'un guide pratique pour les opérationnaliser. De surcroît, nous avons aussi évalué si les participants connaissaient et utilisaient un réseau virtuel privé (*Virtual Private Network* ou VPN) comme protection supplémentaire. En conclusion de cette série de questions, les préoccupations du participant furent interrogées. Afin d'éviter que le participant se sente jugé concernant ses éventuelles préoccupations, celles-ci furent questionnées de façon indirecte. Pour cela, nous avons demandé la fréquence à laquelle l'entourage faisait des commentaires sur les préoccupations de la vie privée reliées au participant.

Concernant les outils spécifiques des représentations, la tâche d'association et la tâche de hiérarchisation ont été mises en place afin d'obtenir les items constituant la représentation ainsi que l'importance de chacun d'entre eux au sein de celle-ci. La tâche d'association consistait donc à proposer aux participants un mot inducteur ou une expression inductrice afin qu'ils génèrent un nombre déterminé de mots ou expressions, à la façon d'une association libre. Pour notre recherche, la consigne précisait le contexte (technologique) et le mot inducteur était la vie privée¹. Pour cette tâche, les auteurs

¹ Voir Appendice E pour la consigne exacte.

(Flament et Rouquette, 2003; Piermattéo, Lo Monaco et Girandola, 2016; Piermattéo, Tavani et Lo Monaco, 2018) conseillent entre 3 et 5 items. Pour ces raisons, nous avons donc demandé aux participants de générer quatre items pour l'association. Par la suite, la tâche de hiérarchisation consistait à demander aux participants de hiérarchiser leurs items par ordre croissant d'importance¹. Ainsi, le participant attribuait un rang aux items qu'il avait précédemment formulés, allant de la première place (la plus importante) à la quatrième place (la moins importante).

Une fois les données récupérées, nous avons effectué une analyse prototypique afin d'avoir une première idée de la répartition des items au sein de la représentation. À ce stade, la répartition des items dans la représentation reste une supposition. En effet, l'analyse prototypique ne constitue pas un test statistique. Il s'agit en fait d'une opération d'ordonnement des données visant à obtenir la structure supposée de la représentation sociale entre les différentes zones de la représentation (Abric, 2003).

Selon Abric (2003), l'ordonnement des mots s'effectue en croisant la fréquence moyenne des items avec la moyenne totale de l'importance moyenne des items, toutes deux calculées grâce aux données de la phase de hiérarchisation. Précisons que la moyenne totale de l'importance des items est elle-même calculée grâce à l'importance moyenne de chaque item. Ainsi, le chercheur obtient l'emplacement hypothétique de

¹ Voir Appendice E pour la consigne exacte.

chaque élément, soit dans le noyau central, soit dans le système périphérique. Attendu que cette structure est dite hypothétique, sa validation fait l'objet d'un second temps.

Abric (2003) a proposé un rectangle (voir Figure 5) qui formalise les emplacements que chaque élément peut occuper. Dans le schéma suivant, la zone grisée représente le noyau central et les zones en blanc représentent le système périphérique.

Figure 5

Représentation formelle des hypothèses adaptée de Abric (2003, p. 64)

		Importance des mots cités par les individus	
		Inférieure à la moyenne ¹	Supérieure à la moyenne
Fréquence des mots cités par les individus	Supérieure à la moyenne	CASE 1 – ZONE DU NOYAU CENTRAL	CASE 2 – 1^{re} PÉRIPHÉRIE
	Inférieure à la moyenne	CASE 3 – ÉLÉMENTS CONTRASTÉS	CASE 4 – 2^e PÉRIPHÉRIE

En résumé, la case n°1 regroupe les items qui ont été cités par de très nombreux participants, puisque la fréquence de citations de ces mots dépasse la citation moyenne.

Les mots de la case n°1 sont aussi considérés comme importants par les participants, car

¹ Il est important de noter qu'un mot important a une moyenne faible concernant son rang d'importance alors qu'un mot sans importance aura une moyenne forte concernant son rang d'importance. Cela vient du fait que les mots importants sont très souvent placés en première position, ce qui fait que leur moyenne tend vers 1. Inversement les mots les moins importants sont souvent placés en dernière position. Dans notre cas, la dernière position étant 4, les mots peu importants présentent une moyenne qui tend vers 4.

leur importance est inférieure à la moyenne du fait qu'ils soient souvent placés dans les premières positions¹. Pour ces raisons, les items de case 1 sont considérés comme des candidats au noyau central. Les cases 2, 3, 4 appartiennent quant à elles au système périphérique avec quelques nuances pour les distinguer. La case 2 par exemple, regroupe les items que les participants n'ont pas considérés comme importants, mais ayant été cités de nombreuses fois. En revanche, dans la case n°3, nous retrouvons des éléments considérés comme saillants par peu de personnes, mais cités peu de fois. Enfin, dans la case n°4, nous retrouvons des items souvent placés en dernière place, car peu importants et cités par peu de monde et donc peu fréquents. Une fois les mots recueillis puis classés, nous avons pu créer les deux questionnaires de la seconde collecte.

Le Test d'indépendance au contexte (Lo Monaco, Lheureux et Hamili-Falkowicz, 2008) est l'un de ces deux questionnaires. Ce questionnaire vise à confirmer ou infirmer l'appartenance du candidat au noyau central à celui-ci. Cependant, il a été nécessaire d'effectuer un tri dans les items, car tous n'étaient pas exploitables tel quel. Nous avons donc élaboré une série de critères afin de sélectionner les candidats au noyau central jusqu'à en arriver à une décision.

- La redondance des termes : est-ce que le terme se répète avec un ou des autres termes? Si les candidats au noyau central étaient jugés redondants, ils pouvaient être fusionnés sous le terme qui les représentait le mieux. Par exemple, les items « donnée » et « donnée personnelle » ont été jugés redondants. Ils ont donc été regroupés sous le terme « donnée personnelle », celui-ci étant spécifique de la vie

privée en contexte technologique. Ce critère seul ne réglant pas tous les cas, d'autres lui furent adjoints.

- L'observation de la fréquence : dans le cas où il y avait redondance et où nous n'avions pu trancher comme précédemment, le mot ayant la plus forte fréquence était conservé. Par exemple, le terme « confidentiel » (13 citations) a été moins cité que « confidentialité » (60 citations). Il se pouvait que les termes précédemment fusionnés soient quand même redondants comme données personnelles et informations. Le même procédé a été effectué dans le cas où les termes avaient été fusionnés. Par exemple, « donnée » (27 citations) et « donnée personnelle » (18 citations) avaient été précédemment fusionnées. Le terme « donnée personnelle » ayant été conservé, il devenait redondant avec le terme « information ». Attendu que le terme fusionné « donnée personnelle » avait plus de citations (27+18) que le terme « information », ce dernier a été supprimé.
- La compatibilité avec la formulation du questionnaire : certains items ne pouvaient s'agencer correctement avec la formulation proposée du *Test d'indépendance au contexte*¹. Ce fut le cas des items « personnel » et « privé ».
- La polysémie, ou à l'opposé, spécificité du terme par rapport au champ de la vie privée; ce point était divisé en deux critères :

¹ La tournure originale de Lo Monaco, Lheureux et Hamili-Falkowicz (2008) était : « À votre avis, la VIE PRIVÉE dans un contexte d'usage des technologies de l'information et des communications, est-ce toujours, dans tous les cas, "personnel"? », avec 4 modalités de réponse : Certainement non, Plutôt non, Plutôt oui, Certainement oui.

- Considérer si l’usage de la synapsie¹ aurait généré une fréquence suffisante pour inclusion dans les candidats au noyau central. Par exemple, « droit » avait deux synapsies différentes, soit « droit à l’oubli » et « droit à l’image » qui auraient pu devenir des éléments centraux si elles étaient en plus grand nombre que « droit » tout court.
- Considérer si le non-usage de la synapsie favorise ou défavorise le fait de retenir certains termes. Est-ce que « droit » reste toujours un élément central sans les synapsies citées précédemment?
- La présence d’artéfact² dû à la tâche : Nous pensons que le mot « privé » est l’exemple type d’un artéfact de la tâche d’association où le participant répète une partie du mot inducteur « vie privée ».

Sur l’ensemble des items, 11 ont répondu à ces critères et ont pu être utilisés pour le *Test d’indépendance au contexte*. Le but du *Test d’indépendance au contexte* consiste à tester les items candidats à la centralité. Il permet d’éprouver la propriété « trans-situationnelle » des éléments centraux. En d’autres termes, on observe si l’élément central reste toujours valide selon les circonstances ou le contexte (Lo Monaco, Lheureux et Hamili-Falkowicz, 2008) alors que les éléments périphériques ne le sont pas. C’est ainsi

¹ Une seule « Unité sémantique composée de plusieurs morphèmes lexicaux (p. ex., *pomme de terre, pince à linge, chemin de fer*) » (Larousse, 2021).

² Phénomène d’origine humain ou artificiel (Robert, 2021). Dans notre cas il s’agirait d’un phénomène dû à la tournure de notre question et de son dernier mot lors de l’association libre : « Dans un contexte d’usage de technologies de l’information et des communications, quels sont les 4 termes et/ou expressions qu’évoque pour vous "VIE PRIVÉE"? ». Nous pensons que la tournure et la consigne ont provoqué cette écholalie.

que le *Test d'indépendance au contexte* permet de mettre à l'épreuve le caractère systématique et non circonstancié des éléments considérés comme centraux. Dans son format original, la tournure des questions aurait été la suivante : « À votre avis, la VIE PRIVÉE dans un contexte d'usage des technologies de l'information et des communications, est-ce toujours, dans tous les cas, l'intimité? ». Les modalités de réponse qui accompagnent la question auraient été au nombre de quatre : « Certainement non », « Plutôt non », « Plutôt oui », « Certainement oui » (Lo Monaco, Lheureux et Hamili-Falkowicz, 2008).

Nous avons opté pour une manière différente du questionnaire original qui offrait plus de possibilités lors des analyses statistiques. Dans le format retenu, les questions se présentaient ainsi : « Dans un contexte d'usage des technologies de l'information et des communications, la “Vie privée” concerne-t-elle le terme “Intimité”? ». Cette formulation permettait au participant de répondre sur une échelle visuelle analogique possédant les bornes inférieures et supérieures suivantes : « Jamais, dans aucun cas » jusqu'à « Toujours, dans tous les cas ». Les scores associés à ces bornes allaient de 0 à 10.

Il est important de préciser que l'échelle visuelle analogique ne présentait à l'écran aucun score lorsque le participant déplaçait le curseur sur l'échelle visuelle, et seules les bornes textuelles aux deux extrémités permettaient au participant de se repérer¹. Cette précaution fut prise, car la visibilité des scores peut introduire un effet plancher pour les

¹ Voir Appendice E pour le visuel du questionnaire.

valeurs faibles et nulles (Paul-Dauphin, Guillemin, Virion et Briançon, 1999). Nous avons aussi randomisé la séquence d'apparition des mots pour chaque participant pour éviter tout effet d'ordre. Les résultats du *Test d'indépendance au contexte* furent analysés à l'aide d'un test d'ajustement de distribution (p. ex., Kolmogorov-Smirnov) afin de connaître le statut des items candidats à la centralité. Ainsi, on peut tester si la distribution de nos items centraux est comparable à la distribution théorique d'un mot considéré comme central. Ainsi, deux choix sont possibles : soit le chercheur choisit une distribution théorique équivalente à 100 % de l'échantillon qui aurait déclaré l'item testé comme central (Abrieu, 2003; Lo Monaco, Lheureux et Hamili-Falkowicz, 2008; Moliner, 2016), soit le chercheur choisit une distribution théorique exponentielle; ce qui reviendrait à effectuer une comparaison avec une distribution où l'échantillon aurait très largement déclaré l'item testé comme central (Baggio et Renaud, 2015).

Quant à lui, le *Questionnaire de caractérisation* est le second questionnaire élaboré grâce aux résultats de l'analyse prototypique. Il permet d'établir la structure et l'organisation de la représentation (Vergès et Bouriche, 2001). Afin, d'évaluer la caractérisation d'un item, on soumet les éléments à la classification des participants et plus exactement à un questionnaire de type *Q-sort* (Stephenson, 1953) avec trois modalités possibles. Chaque élément doit être classé comme « le plus caractéristique », « le moins caractéristique », ou encore « plus ou moins caractéristique ». Selon le classement, un score de 1 à 3 est attribué pour calculer un score de similitude interitems. À la suite de

cela, il est possible de résumer la structure de la représentation sous la forme d'un graphe¹ avec des sommets et des arêtes. Pour construire ce graphe, il est obligatoire que le nombre d'items soit un multiple de trois. En effet, il faut un nombre égal de mots sous chacune des trois modalités afin de ne pas créer de déséquilibre qui impacterait les résultats². Attendu les trois modalités de réponse, 21 items furent choisis et répartis en trois classes de sept items.

Pour obtenir ces 21 items, plusieurs critères ont été mis en place. Les 11 candidats au noyau central qui avaient été sélectionnés pour le questionnaire précédent ont été conservés. En effet, comme les éléments qui composent le noyau central sont censés être les plus caractéristiques de la représentation (Abric, 2003), leur sélection était donc nécessaire.

Cependant, 10 autres items étaient nécessaires afin d'arriver aux nombres de 21 items. Comme le *Questionnaire de caractérisation* ne teste pas la centralité, la sélection a pu s'opérer dans le système périphérique. Nous avons donc repris les critères utilisés précédemment en rajoutant trois autres préalablement. Le premier a été d'avoir une fréquence au-dessus de sept, car nous voulions que des mots appartenant aux systèmes

¹ Dans notre cas, un sommet représentera un item et une arête valuée représentera la relation entre deux items.

² Voir Appendice E pour le visuel du questionnaire.

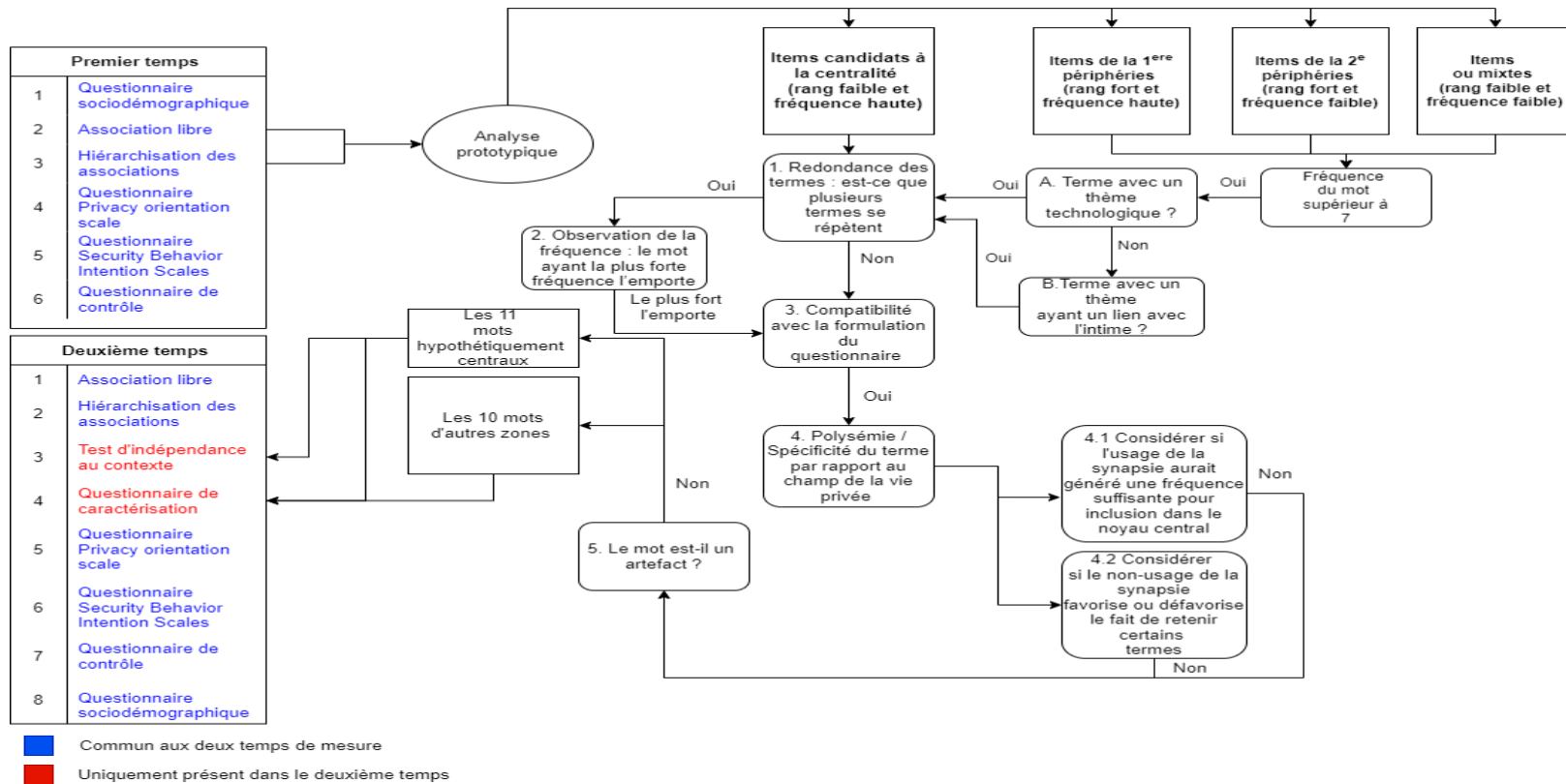
périphériques soient relativement cités. En choisissant la valeur sept comme seuil, un peu moins que 15 % des mots les plus cités dans le système périphérique étaient retenus.

Deux autres critères ont aussi permis de sélectionner les items : soit l'item représentait un thème en lien avec la technologie, soit il comportait un thème en lien avec l'intimité.

Afin de faciliter la compréhension de l'expérimentation dans son ensemble, nous avons résumé les deux collectes sous forme de schéma (voir Figure 6).

Figure 6

Schéma récapitulatif¹



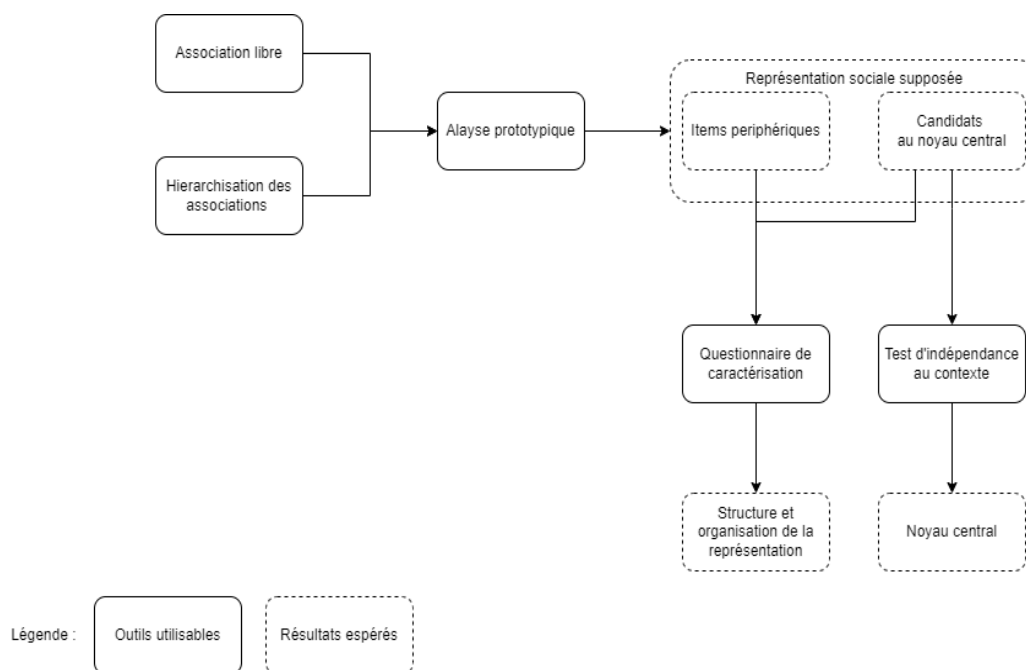
¹ Analyse prototypique (Vergès, 1992), *Test d'indépendance au contexte* (Lo Monaco, Lheureux et Hamili-Falkowicz, 2008), *Questionnaire de caractérisation* (Vergès et Bouriche, 2001), *Privacy Orientation Scale* (Baruh et Cemalcılar, 2014), Sebis (Egelman et Peer, 2015b)

Déroulement

Comme l'a laissé entrevoir le schéma précédent, l'approche structurale se déroule en deux temps, notamment quand aucune recherche n'a déjà été entreprise sur la représentation sociale à l'étude (Abric, 2003; Lo Monaco, Delouvé et Rateau, 2016; Lo Monaco, Piermattéo, Rateau et Tavani, 2016). Afin que le lecteur se fasse une idée plus précise de la forme que peut prendre nos questionnaires, nous le renvoyons aux appendices correspondant à chacun des deux questionnaires (voir Appendices F et G).

La première collecte vise à recueillir le contenu d'une représentation sociale. Durant la première collecte, le chercheur va repérer les éléments constituant une représentation. Cependant, il n'ordonne pas et ne structure pas les éléments. Grâce à l'analyse prototypique qui croise la fréquence des mots recueillis et leur rang d'importance, le chercheur établit des hypothèses de centralité de la représentation et un aperçu des éléments constituant la représentation. Les items candidats à la centralité sont testés ultérieurement pour établir leur centralité. Lors de notre deuxième collecte, le *Test d'indépendance au contexte* teste la centralité des items candidats. Les items récupérés préalablement peuvent aussi être soumis à un *Questionnaire de caractérisation* afin d'établir la structure et l'organisation de la représentation. Nous avons résumé la démarche dans la Figure 7 ci-après.

Figure 7

Récapitulatif de la démarche

Comme décrit précédemment, les participants furent recrutés sur le réseau social Facebook. Le recrutement s’effectuait de deux façons : soit par publicité sur Facebook, soit en affichant le lien sur des murs de groupes. Un message invitant à la participation était systématiquement diffusé pour accompagner la publication.

Si le participant avait eu accès au questionnaire grâce à la publicité sur Facebook, il devait répondre à deux questions préliminaires au consentement. Les questions étaient posées afin de déterminer si l’individu était bien majeur et étudiant. Une fois le consentement lu et accepté, le participant pouvait commencer à répondre aux questions.

Dans le premier questionnaire, l'ordre des questions était le suivant : le *Questionnaire sociodémographique*, le *Questionnaire d'association libre et d'association hiérarchique*, le *Questionnaire d'orientation à la vie privée*, le questionnaire *Security Behavior Intentions Scales* (Egelman et Peer, 2015b), et enfin, les *Questions de contrôle et de rapport à l'objet* ont été placées en fin de questionnaire.

Pour répondre aux objectifs de la seconde collecte de données, l'ordre des questions a dû être différent du premier questionnaire. En effet, un premier objectif était de tester les items candidats à la centralité grâce au *Test d'indépendance au contexte*. Quant au second objectif, il s'agissait d'établir la structure de la représentation grâce au *Questionnaire de caractérisation*. Le *Test d'indépendance au contexte* et le *Questionnaire de caractérisation* ont été placés au début du questionnaire pour éviter au maximum l'attrition. L'ordre des questionnaires était donc le suivant : le *Questionnaire d'association libre et d'association hiérarchique*, le *Test d'indépendance au contexte*, le *Questionnaire de caractérisation*, le questionnaire *Privacy Orientation Scale*, le questionnaire *Security Behavior Intentions Scales*, les *Questions de contrôle et de rapport à l'objet* et, enfin, le *Questionnaire sociodémographique*.

Dans la présente thèse, l'analyse prototypique a été effectuée avec le logiciel en ligne SÉMATO (Plante, 2020) et la validation des questionnaires a été effectuée avec SPSS 28. L'analyse sociodémographique, le *Test d'indépendance au contexte* et le *Questionnaire*

de caractérisation ont été réalisés avec le progiciel R (Ihaka et Gentleman, 1996). La version 4.0.3 de R a été utilisée.

Résultats

Comme annoncé lors de la problématique, cette recherche comporte une partie théorique portant sur un questionnement méthodologique, une première collecte à visée exploratoire et une seconde à visée confirmatoire. La première section de cette partie aura visé à répondre aux questions méthodologiques soulevées durant notre problématique. Une seconde partie portera sur l'analyse descriptive de nos données tandis que la troisième portera sur l'ensemble des résultats exploratoires et confirmatoires concernant la représentation de sociale de la vie privée.

Présentation des analyses

Cette première section vise à répondre à l'ensemble des questions méthodologiques du courant structuraliste des représentations sociales soulevées dans la problématique. Cette section se veut aussi une partie didactique qui vise à réunir les informations nécessaires pour comprendre les méthodes utilisées dans la présente thèse. En effet, ces informations sont souvent développées de manières succinctes dans les monographies (Abric, 2003). Afin de répondre à cette visée pédagogique, la présente partie sera donc jalonnée d'exemples illustratifs.

Analyse prototypique

L'analyse prototypique constitue une opération d'ordonnement des données visant à obtenir les hypothèses de structure de la représentation sociale, notamment les items

composant le noyau central (Abric, 2003). Avant de pouvoir l'effectuer, il faut préalablement récolter les données selon un processus en deux temps. Une fois le matériel colligé, l'analyse prototypique peut être effectuée.

Cependant, une partie du processus n'est pas clairement explicité, ce qui pose certains problèmes. Il est nécessaire de revenir brièvement sur l'article original de Vergès (1992), afin de localiser le manque d'explications sur le processus. Dans son article, Vergès décrit en partie la procédure afin de mener à bien l'analyse. Attendu la nouveauté de la méthode présentée, aucun canon n'est explicitement donné. D'ailleurs, la dimension novatrice de la proposition se ressent avec les différents essais effectués pour l'analyse. Par exemple, Vergès analyse uniquement des mots dont la fréquence est supérieure à dix où il effectue une analyse thématique des termes recueillis. Aussi, il choisit de traiter l'ensemble des données avec la moyenne comme seuil, autant pour la fréquence que pour le rang. Cependant, il omet de détailler s'il établit le rang moyen sur l'ensemble des mots recueillis, ou s'il effectue un premier rang moyen par type de termes puis calcule la moyenne avec ces derniers. Uniquement Flament et Rouquette (2003) détaillent la méthode au complet ; ils choisissent de procéder en deux temps pour les rangs avec un rang moyen pour chacun des termes, puis ils effectuent la moyenne des rangs moyens. Enfin, concernant la fréquence, ils choisissent la moyenne. Quand bien même ces auteurs précisent la méthode dans son entièreté, ils ne justifient pas l'emploi de la moyenne. Globalement, on se rend compte que la moyenne est devenue une norme d'usage à travers les écrits suivants où l'emploi de cette dernière n'est ni justifié, ni questionné (Abric,

2003; Flament et Rouquette, 2003; Lo Monaco, Delouée et Rateau, 2016). Cela nous semble problématique, car la moyenne nécessite une distribution normale pour être utilisée comme indice de tendance centrale pertinent (Leys, Ley, Klein, Bernard et Licata, 2013). En sélectionnant la moyenne autant pour les fréquences que pour les rangs d'importance, cela implique que la distribution générale des mots suit une loi normale. Or, le modèle du noyau central des représentations sociales suppose des distributions différentes de la loi normale; par exemple, le noyau central serait constitué de peu d'éléments ayant été fréquemment cités et considérés comme importants par les participants (Abric, 2003), ce qui sous-entend une distribution asymétrique, et ce, autant pour les fréquences que pour l'importance de ces mots. Bien que la moyenne soit utilisée comme seuil de détermination et non pas comme indicateur de tendance centrale, il nous semble réducteur de suggérer d'utiliser uniquement cette dernière comparativement à d'autres statistiques comme la médiane ou les scores centiles.

Comme illustration du propos, prenons la situation où la personne responsable de la recherche a fait passer une association libre avec le terme inducteur « nouvelle technologie ». La consigne indiquait de générer quatre mots en relation avec le terme inducteur. Conformément à la tâche de hiérarchisation qui suit, il était nécessaire d'ordonner les termes générés du plus important au moins important. Concrètement, le terme le plus important est placé en première place et le moins important en quatrième place. Le responsable de la recherche a réussi à recruter 100 participants. On peut imaginer que le chercheur se retrouve avec une base de données représentée dans le Tableau 3.

Une fois ordonnée, cette base de données pourrait ressemblerait comme montrée dans le Tableau 4.

Une fois la réorganisation faite sur l'ensemble des 100 associations données par les participants, nous avons imaginé le Tableau 5¹ et la Figure 8 qui représentent la distribution générale des mots selon la fréquence. On remarque que seuls les mots faisant possiblement partie du noyau et de la première périphérie présentent une fréquence supérieure à la moyenne, ce qui empêche d'obtenir une distribution d'allure normale, comme le montre le graphique de la Figure 8.

¹ Pour des raisons pédagogiques, nous avons imaginé que les 100 participants ont généré uniquement dix mots différents.

Tableau 3*Extrait de la base de données fictives représentant les 10 premiers participants*

Participant	1 ^{er} mot	2 ^e mot	3 ^e mot	4 ^e mot
Participant 1	Ordinateur	Téléphone intelligent	Réseau social	Internet
Participant 2	Téléphone intelligent	Ordinateur	Réseau social	Internet
Participant 3	Téléphone intelligent	Ordinateur	Internet	Réseau social
Participant 4	Ordinateur	Téléphone intelligent	ChatGPT	Réseau social
Participant 5	Téléphone intelligent	Ordinateur	Imprimante 3D	Tablette
Participant 6	Ordinateur	Téléphone intelligent	Tablette	Tablette
Participant 7	Ordinateur	Téléphone intelligent	Intelligence artificielle	Internet
Participant 8	Ordinateur	Réseau social	Téléphone intelligent	Internet
Participant 9	Ordinateur	Téléphone intelligent	Ordinateur quantique	Internet
Participant 10	Ordinateur	Téléphone intelligent	Wifi	Internet

Tableau 4*Base de données fictives représentant les associations précédentes ordonnées*

1 ^{er} mot	2 ^e mot	3 ^e mot	4 ^e mot
Ordinateur	Téléphone intelligent	Réseau social	Internet
Ordinateur	Téléphone intelligent	Réseau social	Internet
Ordinateur	Téléphone intelligent	ChatGPT	Internet
Ordinateur	Téléphone intelligent	Internet	Internet
Ordinateur	Téléphone intelligent	Intelligence artificielle	Internet
Ordinateur	Téléphone intelligent	Imprimante 3D	Internet
Ordinateur	Ordinateur	Ordinateur quantique	Réseau social
Téléphone intelligent	Ordinateur	Téléphone intelligent	Réseau social
Téléphone intelligent	Ordinateur	Tablette	Tablette
Téléphone intelligent	Réseau social	Wifi	Tablette

Tableau 5*Fréquence des mots*

Mot	Position supposée	Fréquence
Ordinateur	Noyau central	100*
Téléphone intelligent	Noyau central	90
Internet	Périphérique n°1	47
Réseau social	Périphérique n°1	63
Tablette	Périphérique n°1	60
Ordinateur quantique	Contrastée	10
Téléphone	Contrastée	12
Chat GPT	Périphérique n°2	7
Wifi	Périphérique n°2	4
Imprimante 3D	Périphérique n°2	7
Moyenne		40
Total		400**

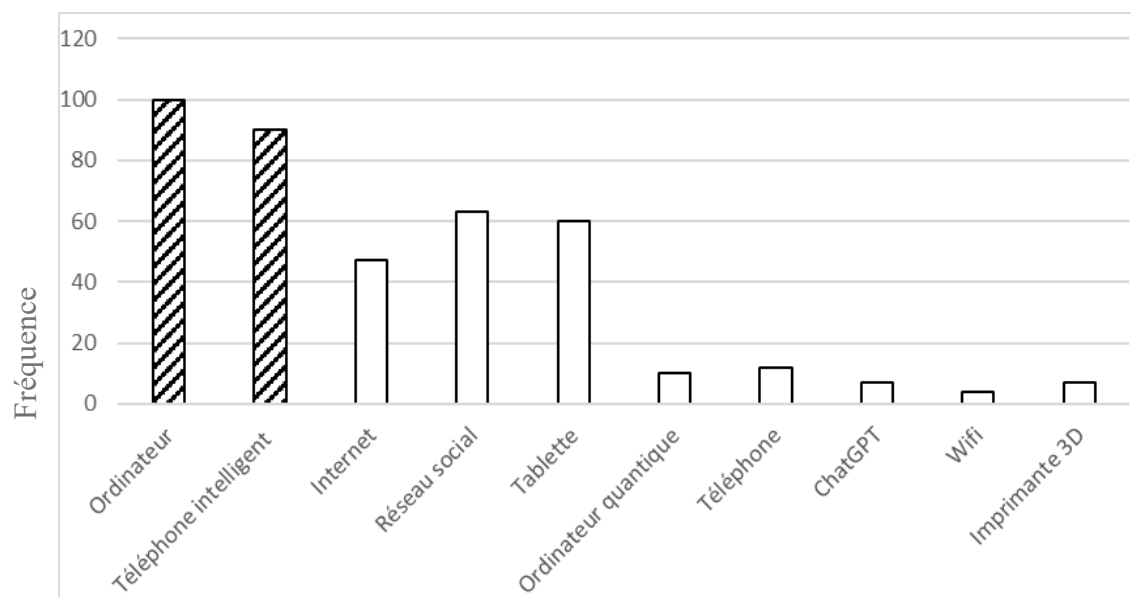
Note. * ceci signifie que 100 % des participants ont généré ce mot en association avec le terme « nouvelle technologie ».

** Comme il y avait 100 participants qui ont généré 4 mots chacun, un total de 400 mots a été généré.

Les données en gras dans le tableau représentent les items centraux.

Figure 8

Diagramme à barres des fréquences terminologiques (mots centraux hachurés)



Il en est de même pour la distribution des rangs d'un même mot (voir Tableaux 6 et 7 ainsi que les Figures 9 et 10). Ici, nous avons choisi le mot « ordinateur », supposé appartenir au noyau central, et le mot « tablette », supposé appartenir à la première périphérie. Dans un cas comme dans l'autre, aucune distribution n'est de forme normale, rendant la moyenne un estimé inadéquat pour le type de distribution examiné.

Enfin, on remarque dans le Tableau 8 et la Figure 11 que la distribution des rangs moyens de l'ensemble des mots ne suit pas non plus la loi normale.

Tableau 6*Distribution des rangs d'importance du mot « ordinateur »*

Mot	Position supposée	Nombre de participants	1 ^{re} place	2 ^e place	3 ^e place	4 ^e place	Rang moyen du mot
Ordinateur	Centrale	100	90	7	1	2	1,06

Tableau 7*Distribution des rangs d'importance du mot « tablette »*

Mot	Position supposée	Nombre de participants	1 ^{re} place	2 ^e place	3 ^e place	4 ^e place	Rang moyen du mot
Tablette	Périphérique 1	60	8	7	20	25	3,0

Figure 9

Diagramme à barres de la distribution des rangs du mot « ordinateur »

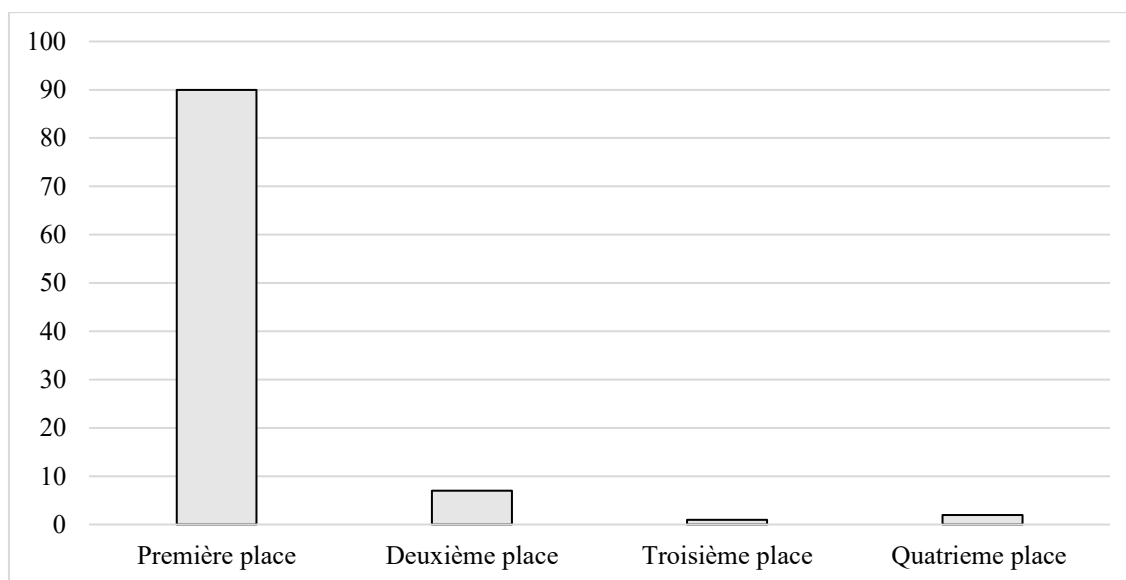
**Figure 10**

Diagramme à barres de la distribution des rangs du mot « tablette »

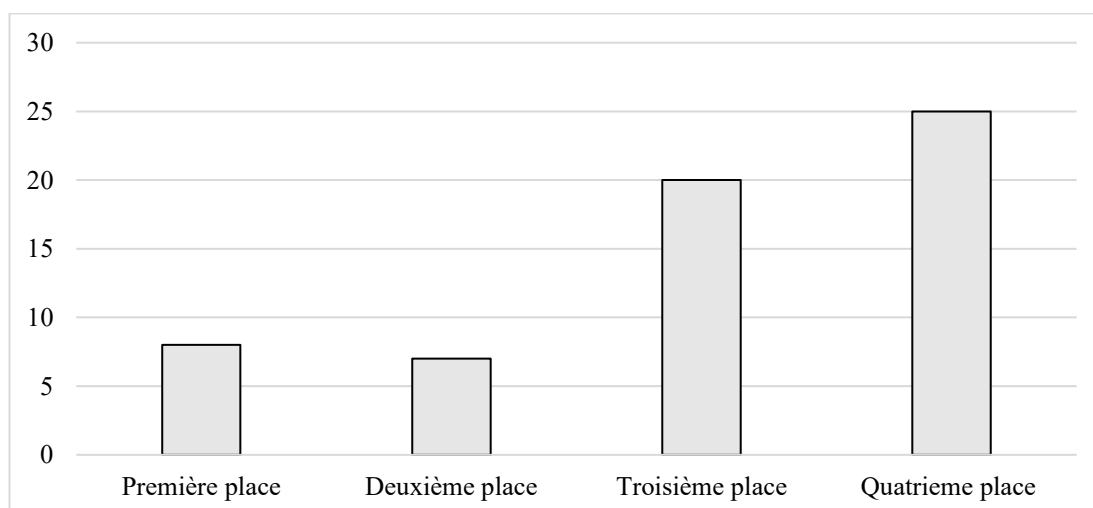
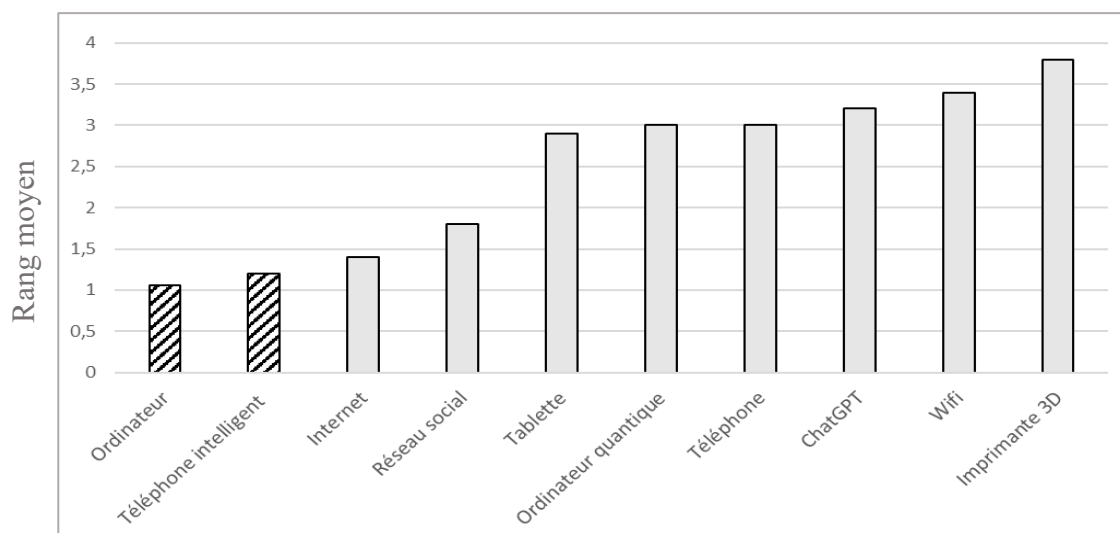


Tableau 8*Importance des mots*

Mot	Position supposée	Rang moyen
Ordinateur	Centrale	1,06
Téléphone intelligent	Centrale	1,20
Internet	Contrastée	1,40
Réseau social	Contrastée	1,80
Tablette	Périphérique n°1	2,90
Ordinateur quantique	Périphérique n°1	3,00
Téléphone	Périphérique n°1	3,00
ChatGPT	Périphérique n°2	3,20
Wifi	Périphérique n°2	3,40
Imprimante 3D	Périphérique n°2	3,80
Moyenne		2,48

Note. Les données en gras dans le tableau représentent les items centraux.

Figure 11*Diagramme à barres de l'importance*

La distribution de la fréquence des mots et celle de leur importance ne suivant pas une loi normale, la moyenne ne pouvait pas constituer un bon indicateur de tendance centrale et par conséquent, celle-ci ne pouvait pas non plus représenter un bon seuil pour classer les éléments dans l'une des quatre cases de la représentation. De plus, l'absence de ligne directrice dans les écrits de référence pour choisir un autre indicateur qui soit plus adapté à des distributions non normales (Abric, 2003; Lo Monaco, Delouée et Rateau, 2016; Vergès, 1992) invite à considérer les autres indicateurs de position comme la médiane ou les percentiles, ceux-ci étant plus adaptés pour résumer les données d'une distribution non normale.

Le choix de ce nouvel indicateur a donc été guidé par nos données. Il résulte d'un compromis entre la variabilité du nombre de termes dans le noyau central selon l'indicateur choisi et le respect des caractéristiques des éléments centraux. Par exemple, nous avons choisi initialement la médiane comme seuil, mais avec cette configuration, les mots cités plus de trois fois et dont le rang médian était inférieur à 2,5 pouvaient être candidats au noyau central. Attendu qu'un terme avec une fréquence supérieure à 3 et ayant un rang médian inférieur à 2,5 ne pouvait pas représenter un mot souvent cité et dont l'importance était partagée par plusieurs participants, nous n'avons pas retenu la médiane comme seuil. Nous avons donc exploré la piste des percentiles.

Après plusieurs essais¹, le choix concernant les fréquences des mots s'est porté sur le percentile 85 de l'ensemble des fréquences des mots. En d'autres termes, le critère de fréquence pour être candidat au noyau central était que le mot présente une fréquence équivalente à 15 % des fréquences les plus élevées de l'échantillon. Dans la présente étude, cela concerne un mot cité dix fois ou plus sur un total de 1096 mots cités.

Concernant l'importance attribuée, nous avons préalablement calculé le rang médian comme indicateur de l'importance globale d'un terme. Ensuite, nous avons proposé de calculer le percentile 20 de l'ensemble des rangs médians. En d'autres termes, le critère

¹ En guise d'illustration, voici deux exemples avec les résultats les plus extrêmes. Le premier avec le percentile 80 pour les fréquences et le percentile 15 pour les rangs médians. Cette configuration aboutissait à un seul et unique candidat au noyau central. En revanche en conservant le percentile 80 pour les fréquences, mais en augmentant au percentile 20 pour les rangs médians, 21 candidats au noyau central émergeaient; ce qui était beaucoup trop.

d'importance pour être candidat au noyau central est que le mot présente un rang médian qui se situe dans les 20 % de rangs médians les plus importants de l'ensemble des mots. Dans notre cas, cela représente un mot dont le rang médian était égal ou inférieur à 2. Dans la présente étude, cela concerne un mot qui est placé par les participants en première ou seconde position dans 50 % des cas au minimum.

En résumé, les seuils retenus dans notre étude pour qu'un item soit candidat au noyau central étaient une fréquence supérieure ou égale à 10 et un rang médian d'importance égale ou inférieure à 2.

Question du seuil d'appartenance au noyau central et confirmation de l'appartenance au noyau central dans une recherche sur les représentations sociales

Dans le courant structuraliste des représentations sociales, l'objectif de recherche vise à établir la structure de la représentation et particulièrement son noyau central. Pour cela, il est nécessaire de tester la structure préalablement établie grâce à l'analyse prototypique. Cette structure est initialement « supposée », car l'analyse prototypique est une opération de réagencement des mots selon leurs fréquences et leurs importances et non pas un test statistique. Il reste donc aux chercheurs à éprouver de façon empirique la structure de la représentation. Ainsi, ils vont tester la distribution des mots candidats au noyau pour déterminer si celle-ci correspond à une distribution d'un item suffisamment important et fréquent pour appartenir réellement au noyau central. Pour résumer, les chercheurs tentent de savoir si les items font consensus au sein de la population étudiée à travers leur fréquence et leur importance. Bien que des questionnaires aient été créés spécifiquement

pour cette tâche, comme le *Questionnaire de mise en cause* ou le *Test d'indépendance au contexte*, la question de savoir à partir de quand un item fait consensus n'est pas résolue de manière définitive. L'historique des méthodes de détermination du consensus, et par extension de la centralité, permettra aux lecteurs de mieux comprendre pourquoi il est plus pertinent de tester la forme générale de la distribution empirique plutôt que de procéder par seuil arbitraire.

Dans les études initiales sur le sujet (Flament, 1999; Moliner, Rateau et Cohen-Scali, 2002), les chercheurs avaient retenu un seuil arbitraire minimal de 75 % de réponse pour déclarer qu'un item appartenait au noyau central (Baggio et Renaud, 2015). Le premier problème était de justifier le seuil pour établir un consensus. En effet, quel critère permet de justifier le choix d'un seuil à 75 % plutôt qu'un seuil à 50 % de réponse pour s'assurer du consensus? Un second problème était qu'il n'y avait pas de test d'inférence statistique qui était utilisé pour pouvoir tester les résultats de manière rigoureuse.

Dans les recherches ultérieures (Abric, 2003; Flament et Rouquette, 2003; Lheureux et Lo Monaco, 2011; Lo Monaco, Lheureux et Hamili-Falkowicz, 2008), les chercheurs ont donc mis en place des tests statistiques afin d'étayer davantage leurs conclusions (Baggio et Renaud, 2015). Le premier test retenu pour départager les items appartenant au noyau central de ceux hors noyau fut celui du Khi deux d'ajustement (Baggio et Renaud, 2015; Pearson, 1900). Or, le Khi deux d'ajustement ne règle pas la question du seuil

arbitraire, puisqu'il faut déterminer une fréquence d'ajustement préalablement au test afin d'examiner si les fréquences observées diffèrent de la fréquence théorique.

Les chercheurs (Abric, 2003; Lo Monaco, Lheureux et Hamili-Falkowicz, 2008; Moliner, Rateau et Cohen-Scali, 2002) ont donc essayé de contourner le problème en utilisant le *Test de Kolmogorov-Smirnov* (Kolmogorov, 1933). Pour bien comprendre l'argumentaire concernant l'utilisation du Kolmogorov-Smirnov, un exemple fictif sera utilisé pour faciliter la compréhension des points techniques.

Un chercheur étudie la représentation de la vie privée et plusieurs mots sont suspectés d'être centraux, dont le mot « couple ». Le *Test d'indépendance au contexte* est choisi pour confirmer la centralité. Dans le cas du mot « couple », la question posée sera : « La "Vie privée" concerne-t-elle le terme "couple"? ». Tout comme dans l'article original (Lo Monaco, Lheureux et Hamili-Falkowicz, 2008), quatre modalités de réponse étaient disponibles pour cet exemple : « oui », « plutôt oui », « plutôt non » et « non ». Le chercheur recrute 100 participants afin de confirmer la centralité de l'item. Pour analyser les résultats, il choisit d'appliquer le *Test de Kolmogorov-Smirnov*.

Dans cette situation, nous allons suivre pas à pas les étapes du *Test de Kolmogorov-Smirnov* (Kolmogorov, 1933), puis nous examinerons les divergences avec l'utilisation faite par Abric (2003) de ce même test. Malgré l'utilisation commune du test, des divergences de résultats émergent et seront l'objet d'explications.

Le *Test de Kolmogorov-Smirnov* (Kolmogorov, 1933) suit plusieurs étapes; la lecture du Tableau 9 dans les pages suivantes peut faciliter la compréhension de l'explication. La première étape consiste à choisir la loi mathématique ou la distribution théorique avec laquelle comparer la distribution observée. Dans l'exemple, nous avons choisi une distribution théorique où 100 % des participants ont répondu « oui » à la question pour confirmer que l'item est central. Une fois ce choix effectué, il est nécessaire de calculer les fréquences relatives cumulatives de chacune des deux distributions¹ afin de les comparer (voir les colonnes $F(x)$ et $Sn(x)$ du Tableau 9). Pour rappel, la fréquence relative cumulative est la fréquence de chaque intervalle de réponse divisée par le nombre total d'observations. Attendu qu'il y avait quatre modalités dans notre exemple, il y avait donc quatre fréquences relatives cumulées à calculer pour chacune des distributions (observée et théorique). Une fois les fréquences cumulées calculées, il suffit d'effectuer une soustraction entre chacune des valeurs obtenues (observées et théoriques)². Ainsi, après avoir effectué l'ensemble des soustractions (voir colonne D du Tableau 9), on conserve la valeur absolue de la différence la plus grande (en gras dans la colonne D). Enfin, la différence retenue est comparée avec la valeur maximale³ attendue de la table du *Test de Kolmogorov-Smirnov* afin de savoir si le test s'avère significatif ou non (Kanji, 1999).

¹ Dans les citations ci-après, l'auteur (Kanji, 1999) nomme $F(x)$ la fréquence relative cumulative observée et $Sn(x)$ la fréquence relative cumulative prédite par la loi mathématique ou distribution théorique choisie.

² Dans les citations ci-après les auteurs (Abriç, 2003; Kanji, 1999) nomment D la différence qui résulte de la soustraction entre les deux distributions.

³ Noté $D(\alpha)$ par Kanji (1999), mais D_{\max} par Abriç (2003).

Tableau 9

Exemple illustratif

Modalité de réponse à la question : « Dans un contexte d'usage des technologies de l'information et des communications, la "Vie privée" concerne-t-elle le terme "couple" »?	Nombre de participants	Fréquence observée cumulée	Nombre attendu de participants pour 100 % de réponses centrales (H_0)	Fréquence cumulative prédite (théorique)	Fréquence relative cumulative observée $Sn(x)$	Fréquence relative cumulative prédite $F(x)$	$D = F - Sn $ Si $D > D_{max}$ (0,136), alors le test est significatif
Réponses originales							
Non	5	10 %	0	0 %	0,1	0	$D = 0,05$
Plutôt non	0	10 %	0	0 %	0,1	0	$D = 0,05$
Plutôt oui	40	90 %	0	0 %	0,90	0	$D = 0,90$
Oui	5	100 %	100	100 %	1,00	1	$D = 0$
Total	50	100 %	100	100 %	1,00	1	$D = 0$
$D_{max} =$	0,136						
Réponses dichotomisées							
Non	5	10 %	0	0 %	0,10	0	$D = 0,10$
Oui	45	100 %	100	100 %	1,00	1	$D = 0$
Total	100	100 %	100	100 %	1,00	1	
$D_{max} =$	0,136						

Note. $N = 100$. D_{max} correspond à la différence maximale acceptable entre la distribution observée et celle théorique (forme normale) pour conclure au rejet de l'hypothèse.

Les données en gras dans le tableau représentent les données auxquelles nous faisons référence dans le paragraphe précédent.

Cependant, dans le cas où n est supérieur à 35 (ici $n = 100$), cette valeur maximale attendue se calcule comme suit : $D_{\max} = \frac{1,36}{\sqrt{n}}$ pour $p = 0,05$ (Kanji, 1999), soit pour notre exemple $\frac{1,36}{\sqrt{100}} = 0,136$. Ainsi, si le D est plus grand que le D_{\max} , alors les deux distributions sont considérées différentes et, par conséquent, l'item ne peut pas être déclaré comme faisant partie du noyau central. Inversement, si le D ne dépasse pas la valeur du D_{\max} , alors les deux distributions sont considérées similaires et, par conséquent, l'item est attribué au noyau central. En résumé, effectuer un test d'ajustement de Kolmogorov-Smirnov revient à effectuer l'opération suivante :

À partir de l'échantillon, la distribution cumulative $S_n(x)$ est déterminée et tracée sous forme de fonction échelonnée. La distribution cumulative $F(x)$ de la population supposée est également tracée sur le même diagramme. La différence maximale entre ces deux distributions $D = |F - S_n|$ fournit la statistique du test et celle-ci est comparée à la valeur $D(\alpha)$. Si $D > D(\alpha)$, l'hypothèse nulle selon laquelle l'échantillon provient de la population supposée est rejetée. (Kanji, 1999, p. 76) [traduction libre]

Abric (2003) propose un calcul différent lorsqu'il utilise le *Test de Kolmogorov-Smirnov* :

Il s'agit d'un test de comparaison de distribution qui permet de définir le seuil au-delà duquel on peut considérer que l'item est rejeté [considéré comme appartenant à la zone périphérique]. Cet indice dépend de la taille de l'échantillon et du seuil de significativité choisi. Il est calculé par la formule $D = 1 - \frac{1,36}{\sqrt{n}}$ x 100, où n représente le nombre de sujets. Cette formule est appliquée pour des effectifs supérieurs à trente-cinq personnes. Si l'effectif est moindre, on trouve le seuil dans des tables proposées par Kanji.

En fait, ce que Abric (2003) propose sans l'explicitier, c'est de reprendre la formule du *Test de Kolmogorov-Smirnov* $D = |F - S_n|$ et de résoudre une équation à une inconnue.

Ici, rappelons que F constitue la fréquence relative cumulative observée et Sn la fréquence relative cumulative prédite par la loi ou la distribution choisie. Si le chercheur sait que le D ne doit pas dépasser le D_{max} , pour que l'item soit considéré comme appartenant au noyau central et que pour un échantillon de plus de 35 personnes, celui-ci se calcule par $\frac{1,36}{\sqrt{n}}$ et qu'enfin, le chercheur connaît la taille de son échantillon (n), alors on peut écrire l'équation suivante : $\frac{1,36}{\sqrt{n}} = |F - Sn|$.

Mais cette équation ne peut être résolue telle quelle. Pour la résoudre, Abric (2003) est parti du présupposé que la comparaison entre la distribution de l'échantillon s'effectuait avec une distribution théorique où 100 % de l'échantillon qualifie de central l'élément testé, ici « couple ». Ce présupposé revient à dire que la fréquence cumulée attendue sera égale à 1 (voir colonne $Sn(x)$ du Tableau 9, colonne 7 de la modalité « oui »). Ainsi, dans notre exemple, l'équation suit les étapes suivantes : $\frac{1,36}{\sqrt{100}} = 1 - Sn$, soit $Sn = 1 - \frac{1,36}{\sqrt{100}}$; ce qui fait $Sn = 1 - 0,136 = 0,864$. Enfin, en multipliant le résultat par 100, le chercheur retombe sur la formule de Abric et connaît le pourcentage de réponses positives (« oui ») nécessaires pour que le test confirme l'appartenance de l'item au noyau central.

Cependant, cette approche pose deux problèmes majeurs. Le premier est que l'enjeu du seuil précédemment évoqué n'est pas résolu. Encore une fois, qu'est-ce qui justifie plus de choisir une distribution théorique où 100 % de l'échantillon déclare l'item central

plutôt que 90 % de l'échantillon? Un second problème est que la mise en place de ce procédé de calcul nécessite obligatoirement de dichotomiser l'ensemble des quatre modalités de réponse proposées initialement. Cela impacte les résultats comme le montre le Tableau 9 (voir « réponses dichotomisées »).

Concrètement, dès que nous effectuons le test en conservant l'ensemble des quatre modalités, nous nous apercevons que la différence entre les fréquences cumulatives observées et théoriques de la modalité « plutôt oui » devient significative (en gras dans le Tableau 9). Bien qu'une très grande majorité de personnes, c'est-à-dire 90 personnes sur 100, donnent des réponses positives (« plutôt oui ») pour l'item « couple », l'application classique du test aboutit à rejeter l'appartenance de l'item au noyau central, car le $D(0,95)$ est supérieur au $D_{\max}(0,136)$. Or, en dichotomisant les variables, le test n'est plus significatif, puisque le D ne dépasse plus le D_{\max} comme l'indique le Tableau 9.

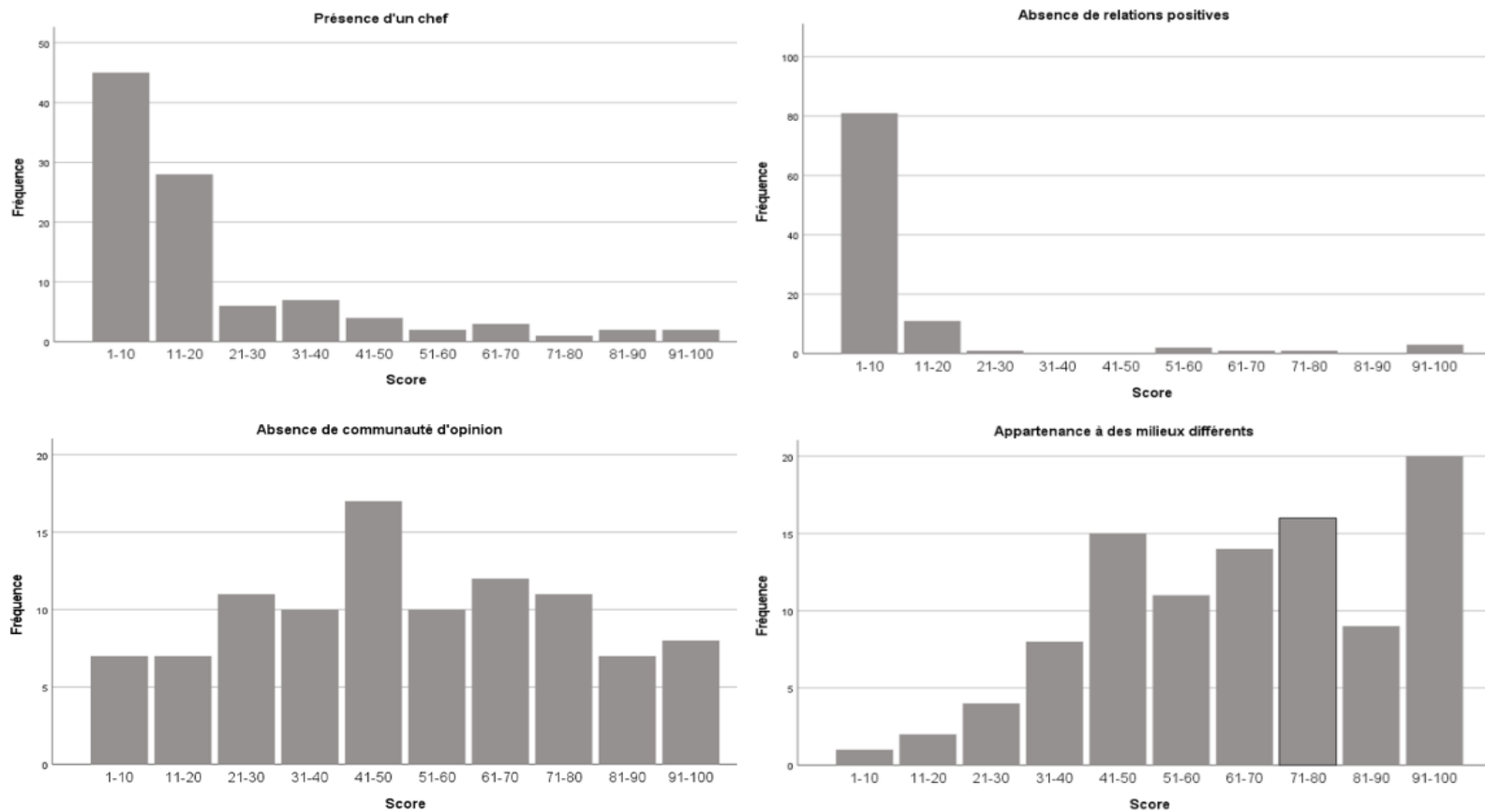
En conclusion, peu importe le test utilisé, la question du seuil demeure à trancher. Qu'ils s'agissent du test du Khi deux ou du Kolmogorov-Smirnov, il est nécessaire que les chercheurs choisissent subjectivement un seuil permettant de rejeter l'hypothèse nulle. D'ailleurs, la subjectivité lors de la détermination du seuil d'appartenance au noyau central s'observe concrètement dans la littérature, puisque nous retrouvons globalement trois seuils récurrents. Le premier seuil est de 50 % et il a été retrouvé dans les études de Tafani (2001) (Tafani, Souchet, Codaccioni et Mugny (2003), de Dany et Apostolidis (2007), de Lheureux, Rateau et Guimelli (2008), de Lheureux et Lo Monaco (2011) et de

Zbinden, Souchet, Girandola et Bourg (2011). Le second seuil choisi par les chercheurs est de 75 %. Nous pouvons notamment le retrouver dans les études de Flament (1999), Tafani (2002), Tafani et Bellon (2001), Lo Monaco, Lheureux et Hamili-Falkowicz (2008); Lo Monaco, Rateau et Guimelli (2007) et Wachelke (2013). Un troisième seuil souvent retrouvé est celui de 100 % comme dans les études de Lheureux, Rateau et Guimelli (2008), Lo Monaco, Lheureux et Hamili-Falkowicz (2008) ou encore Methivier (2010). D'autres seuils sont retrouvés, mais plus « exotiques » comme dans l'article de Salesses (2005a, 2005b) où le seuil est de 85 %, ou encore l'article de Apostolidis, Lionel, Cress et Pecly Wolter (2013) où le seuil est de 66,6 %.

Pour éviter le problème de l'arbitraire, une solution a été proposée par Baggio et Renaud (2015). Premièrement, ces auteurs proposent de passer d'une échelle de réponse ordinale à une échelle continue pour pouvoir utiliser le Test de Cramer-von Mises qui teste la distribution de l'échantillon avec la distribution théorique choisie (Baggio et Renaud, 2015; Cramér, 1928). Deuxièmement, ils proposent de tester si la distribution observée diffère d'une loi exponentielle (avec $\lambda = 1 / \text{moyenne à l'item}$) au lieu d'une distribution théorique où 100 % de la population considère l'item comme central, car selon eux, la loi exponentielle est un marqueur plus approprié du consensus. En effet, la distribution d'un item appartenant au noyau devrait avoir une forte asymétrie à droite, signe qu'une très forte majorité de personnes aurait déclaré l'item comme appartenant au noyau. En exemple, voici la distribution des items de l'article de Baggio et Renaud (2015), avec les items appartenant au noyau en haut de la Figure 12 et périphérique en bas.

Figure 12

Histogrammes des réponses pour les 4 items relatifs à la représentation du groupe d'amis adapté de Baggio et Renaud (2015, p. 43)



Note. Où 1, l'item représente totalement le groupe d'amis et 100, il ne le représente pas du tout. Les deux premiers du haut présentent un effet plancher très net; ce qui caractérise des items centraux.

Ainsi, la forme de la loi exponentielle explique le choix des auteurs (Baggio et Renaud, 2015). En effet, plus la distribution d'un item est consensuelle, plus sa distribution est asymétrique et, par conséquent, sa fonction de répartition présentera un fort effet de plancher, par rapport aux scores bruts, comme le montre la Figure 13.

Ainsi, pour Baggio et Renaud (2015), les caractéristiques de la loi exponentielle rendraient donc adéquatement compte du phénomène de consensus autour d'un item. Pour ces auteurs, il suffit de paramétrer la loi exponentielle comme suit : $\lambda = \frac{1}{\bar{x}}$, où \bar{x} représente la moyenne de la distribution étudiée. Ainsi, l'utilisation de ce test vient résoudre la question de la perte d'informations due à la dichotomisation et règle le problème du seuil d'appartenance au noyau central en se focalisant sur la distribution de l'échantillon.

Analyse de similitude

L'analyse de similitude est une technique multivariée issue de la théorie des graphes. Le but de cette analyse est d'étudier la connexité des éléments de la représentation (Moliner et Lo Monaco, 2017). Plus exactement, l'analyse de similitude permet de mettre en exergue l'organisation de l'ensemble des items en étudiant les relations qu'entretiennent les items entre eux; ces relations pouvant être similaires ou antagonistes dans un graphe valué ou pondéré. Une très brève introduction à la théorie des graphes nous semble nécessaire. Si le lecteur n'a aucune connaissance de cette théorie, nous lui conseillons de lire le paragraphe qui suit tout en se reportant simultanément à la Figure 14.

Figure 13

Distributions des 4 items précédents (en noir) et distributions théoriques de la loi exponentielle pour chacun des 4 items précédents (en rouge) adapté de Baggio et Renaud (2015, p. 45)

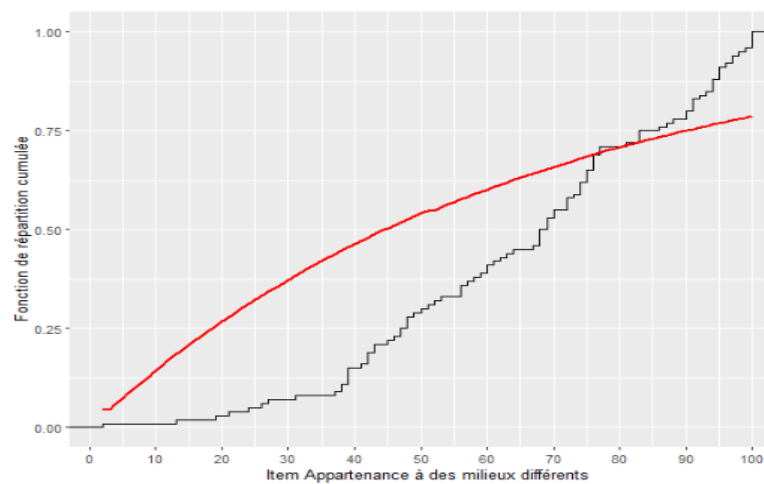
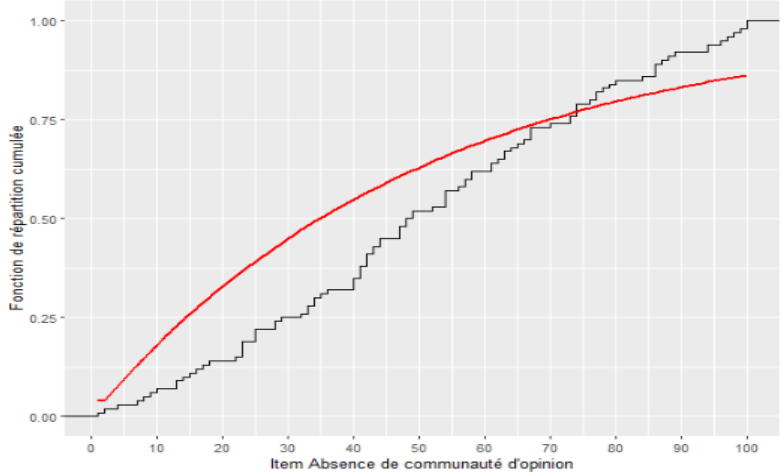
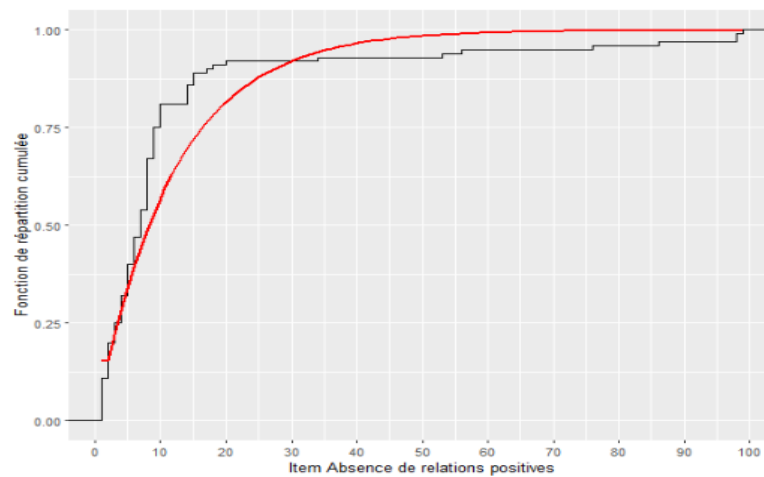
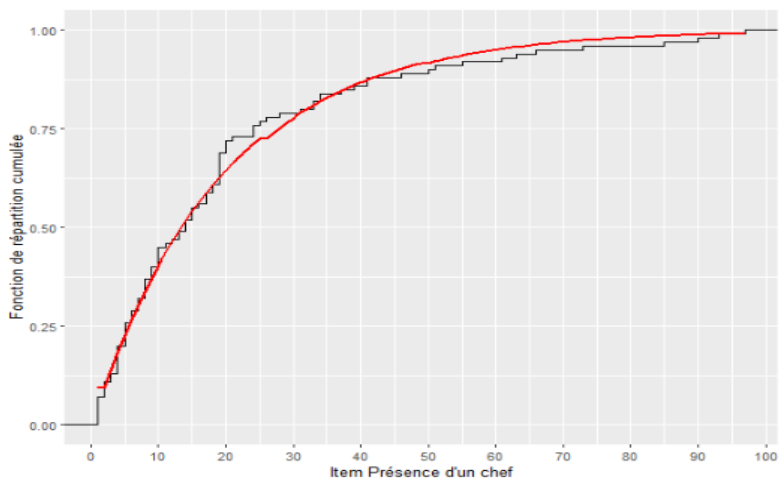
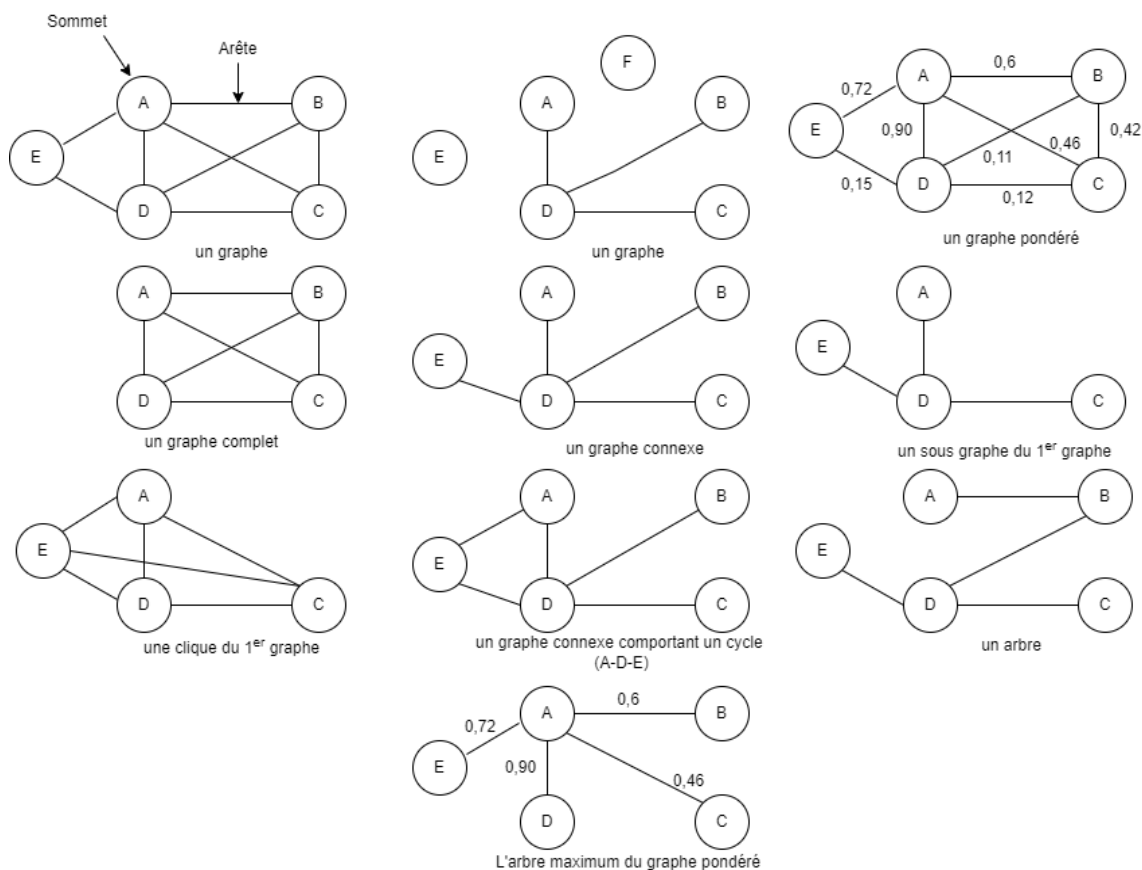


Figure 14

Typologie des graphes



Un graphe est composé de deux types d'éléments : des sommets et des arêtes. Dans le cadre des représentations sociales, le graphe illustre à travers les sommets, les items que le chercheur souhaite étudier. Le graphe permet aussi de représenter, grâce aux arêtes, les relations que les items entretiennent entre eux (Flament, 1965; Flament, 1981; Vergès et Bouriche, 2001). Il est possible de qualifier la force des relations entre les items grâce à une pondération, transformant le graphe en graphe pondéré. Cette pondération est calculée grâce à un indice de similitude choisi par le chercheur pour créer une matrice de similitude

interitems. La qualification des graphes ne s'arrête pas aux graphes pondérés, puisqu'il existe toute une typologie de graphes qu'il est utile de connaître.

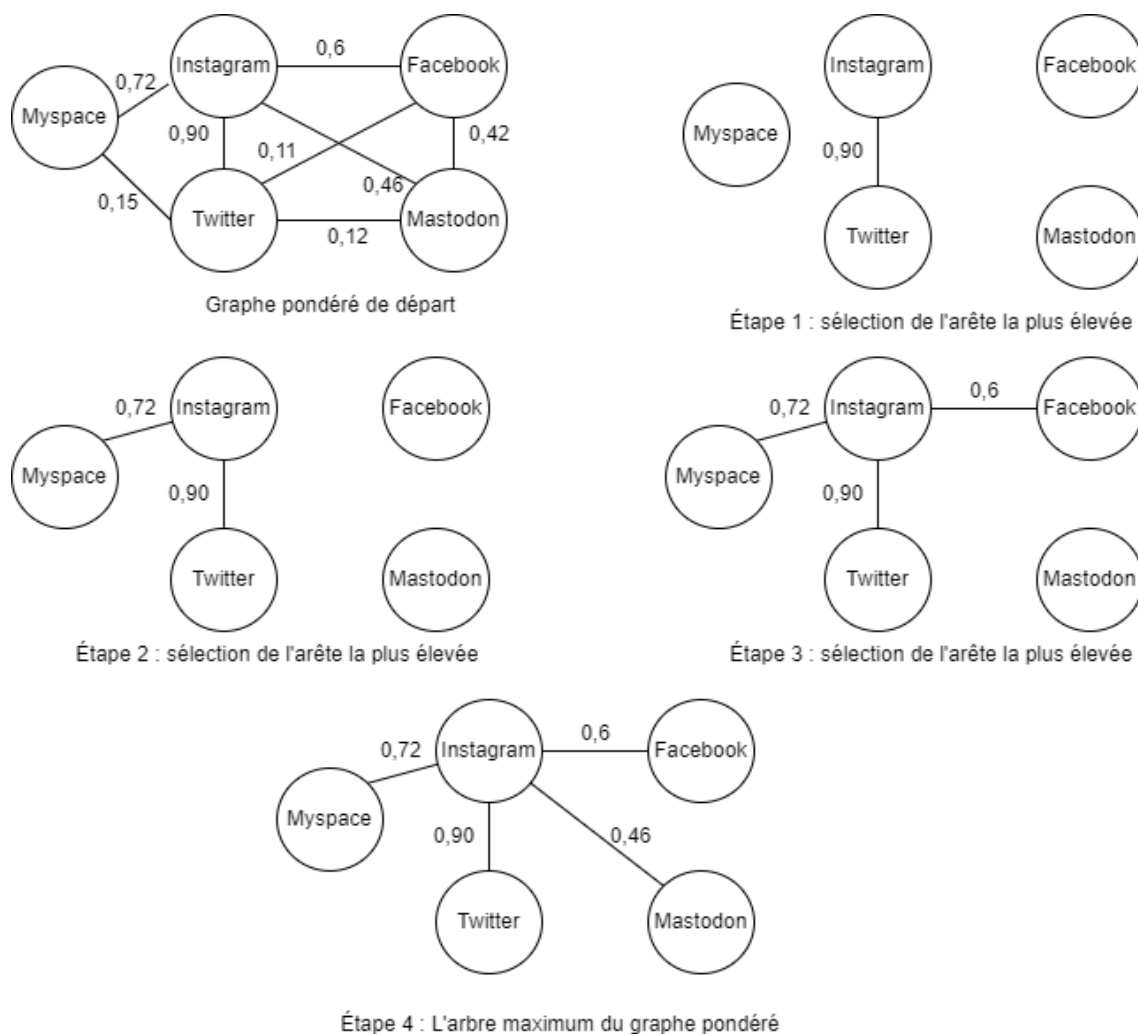
Certains graphes ont par exemple des caractéristiques particulières, comme le graphe complet dont chacun des sommets est relié avec la totalité des autres sommets. Un graphe peut aussi être connexe. Dans ce cas, il s'agit d'un graphe dont tous les sommets sont reliés par au moins une arête. On peut aussi obtenir un sous-graphe lorsqu'on supprime certaines arêtes et certains sommets d'un graphe initial. De plus, un sous-graphe peut être qualifié de « cliques » lorsque celui-ci prend la forme d'un sous-graphe complet. En outre, on peut rencontrer des graphes présentant des cycles, c'est-à-dire qu'il existe un ensemble d'arêtes qui relie un certain nombre de sommets et permet de revenir au point de départ sans passer deux fois par la même arête. Ces caractéristiques peuvent se cumuler pour former un graphe connexe avec un ou des cycles. Cependant, il existe aussi des graphes connexes sans cycle qu'on nomme « arbres ». Enfin, on peut obtenir ce qu'on appelle un arbre des maximums si l'on crée un arbre pondéré faisant apparaître uniquement les arêtes dont la pondération est la plus grande.

La première étape de l'analyse de similitude est la transformation des données initiales en matrice de similitude (Vergès et Bouriche, 2001). Pour notre étude, l'indice de similitude conseillé pour construire la matrice de similitude est le Tau de Kendall (Kendall, 1938; Vergès et Bouriche, 2001), car nos données étaient ordinales. Cette statistique mesure la corrélation de rang entre deux variables.

La seconde étape est l'exploration de la matrice de similitudes pour créer les premiers graphes (Vergès et Bouriche, 2001). Pour cela, on effectue l'algorithme de l'arbre maximum qui consiste à retenir l'arête ayant la pondération la plus forte. La valeur la plus élevée dans la matrice de similitude sera la première arête de notre graphe. On continue l'algorithme en retenant les arêtes les plus fortes et on s'assure qu'elles ne créent pas un cycle. Si le choix de l'arête la plus forte mène à la création d'un cycle, on choisit l'arête suivante plus forte. L'algorithme s'arrête une fois tous les sommets reliés. La Figure 15 représente un exemple fictif où nous avons appliqué l'algorithme sur la représentation des réseaux sociaux.

La troisième étape consiste simplement à tracer l'arbre des maximums et les graphes seuils. Les graphes seuils sont le dessin de graphes ayant des arêtes ne dépassant pas un certain seuil de pondération. Ils permettent de mieux comprendre comment le graphe global s'est construit au fil des seuils successifs (Bouriche, 2003).

Rappelons aussi aux lecteurs que cette analyse n'a pas pour but de tester la centralité des items, mais leur connexité. On peut donc incorporer des items d'autres zones pour comprendre les relations qui existent entre les différents items. De plus, il ne s'agit pas d'un test statistique, il n'y a donc pas de but inférentiel. Cette analyse vise la description et l'exploration de la structure de la représentation ainsi que l'intensité des relations qu'entretiennent les mots entre eux.

Figure 15*Algorithme de l'arbre des maximums***Analyses descriptives**

Pour ne pas rendre cette partie redondante, nous ne présenterons pas les données sociodémographiques de l'échantillon qui a permis d'effectuer l'analyse prototypique, puisqu'elles sont identiques à celles du premier temps de collectes présentées dans la partie Méthode, section Participants. Précisions faites, nous commencerons par les

données descriptives des questionnaires *Privacy Orientation Scale* (Baruh et Cemalcilar, 2014) et le questionnaire *Security Behavior Intentions Scale* (Egelman et Peer, 2015b), pour conclure avec les données descriptives du *Test d'indépendance au contexte* et du *Questionnaire de caractérisation*.

Analyse descriptive du questionnaire Privacy Orientation Scale

Ce questionnaire (Baruh et Cemalcilar, 2014) comporte 16 items que nous avons traduits. Il permet de mesurer l'attitude concernant la vie privée lors de l'utilisation des réseaux sociaux des participants. Le recueil de données en vue de sa validation s'est effectué durant la première et la seconde collecte de données, comme indiqué dans le schéma récapitulatif (voir Figure 6), et comportait en tout 408 participants. Nous n'avons conservé que les participants ayant rempli la totalité du questionnaire. Les quelques données manquantes sont issues de l'échantillon de la seconde collecte et s'expliquent par le positionnement des *Questionnaires sociodémographiques* en fin d'expérience (attrition due à la longueur de la tâche, environ une quinzaine de minutes).

L'échantillon était différent de l'échantillon de l'article original (Baruh et Cemalcilar, 2014), car les moyennes des sous-échelles de « la vie privée comme un droit », la « préoccupation personnelle à propos de la vie privée » et la « préoccupation concernant la vie privée d'autrui » étaient supérieures à celle de l'article. Cependant, notre échantillon avait une moyenne plus faible concernant la sous-échelle de « la vie privée conditionnelle aux autres » (voir Tableau 10).

Tableau 10

Test t sur échantillon unique des différentes sous-échelles de la Privacy Orientation Scale de l'étude avec les moyennes dans l'article de validation originale

Sous-échelle	Moyenne de l'étude	Intervalle de confiance à 95%		Moyenne de l'article original	<i>p</i>	<i>d</i> de Cohen
		Borne inférieure	Borne supérieure			
La vie privée comme un droit	4,41	4,32	4,51	4,09	< 0,001	0,58
La préoccupation personnelle à propos de la vie privée	3,69	3,55	3,82	3,37	< 0,001	0,38
La vie privée conditionnelle aux autres	3,27	3,14	3,41	3,48	0,003	-0,24
Préoccupation concernant la vie privée d'autrui	4,27	4,18	4,35	4,08	< 0,001	0,36

Concernant les rares données comparables entre notre échantillon et l'échantillon de Baruh et Cemalcılar (2014), notre échantillon comportait 408 participants majoritairement jeunes ($M = 22,21$ ans, $ÉT = 4,71$) alors que celui de Baruh et Cemalcılar (2014) comportait 1 244 individus et sa moyenne d'âge était de 45,7 ans ($ÉT = 14,39$). Notre échantillon était majoritairement composé des femmes (75,2 %) alors que l'échantillon de Baruh et Cemalcılar (2014) était plus équilibré, avec 50,2 % de femmes. Les étudiants de notre échantillon étaient majoritairement détenteurs d'une licence équivalent au baccalauréat québécois ($n = 159$), suivis de près par 149 bacheliers (équivalent au DEC), 69 étudiants avaient déjà un master complété (équivalent d'une maîtrise) et 3 participants avaient un doctorat complété. Les individus constituant l'échantillon de Baruh et Cemalcılar (2014) étaient moins diplômés, car 61 % d'entre eux avaient le niveau lycée (équivalent du Cégep au Québec). Enfin, les deux dernières informations sur l'échantillon de Baruh et Cemalcılar (2014) étaient l'usage d'Internet dans les trois dernières années (98,4 %) et des données concernant l'ethnicité de l'échantillon. Bien que la première donnée ne fût pas collectée directement dans notre étude, on peut l'estimer à 100 % de l'échantillon, puisque le questionnaire était administré en ligne. Concernant l'ethnicité, cette donnée ne fut pas demandée, car il est illégal de demander des statistiques ethniques en France (Conseil Constitutionnel, 2007). En effet, depuis la loi informatique et liberté du 6 janvier 1978, il est interdit de recueillir des données ethniques dans ce pays. En 2007, le Conseil constitutionnel a rappelé cette interdiction concernant une nouvelle loi sur l'immigration. Sa décision fut motivée au titre de l'article premier de la constitution. Celui-ci indique que la France « assure l'égalité

devant la loi de tous les citoyens sans distinction d'origine, de race ou de religion ». Les statistiques ethniques pourraient contrevenir à cet article en étant utilisées pour faire des différences entre les citoyens. Les seules exceptions tolérées sont les projets d'études ayant comme objet la diversité même des ethnies avec l'accréditation préalable de la CNIL.

Concernant le reste des données sociodémographiques, les individus de notre échantillon étaient principalement célibataires ($n = 290$) et de nationalité française ($n = 346$). Les étudiants en psychologie ($n = 72$) étaient les plus nombreux ainsi que ceux en droit ($n = 49$). Les autres domaines d'études ayant moins de 15 représentants n'ont pas été affichés. Concernant les questions politiques, les participants rapportaient principalement être assez intéressés ($n = 162$) ou peu intéressés ($n = 140$).

Lorsqu'on demandait aux participants de se positionner sur une échelle visuelle analogique où 0 représente un engagement politique à gauche et 10 à droite, on obtenait une tendance plutôt à gauche. La valeur 3 était la modalité de réponse la plus fréquente ($n = 66$) pour cette question. La seconde modalité la plus fréquente était 4 ($n = 55$) et indique un échantillon plutôt à gauche sur le spectre politique. La moyenne sur cette échelle était de 3,74 ($ET = 2,38$). Ce résultat était cohérent avec le parti politique des participants qui était majoritairement à gauche : Europe Écologie Les Verts (EELV) avec 92 participants. Cependant la première modalité était « Je ne souhaite pas répondre » avec 131 participants ayant choisi de ne pas divulguer cette information. Pour la catégorie

socioprofessionnelle du père, mais aussi de la mère, la modalité principale était « Employés ». On retrouvait 76 personnes dont le père était employé et 122 dont la mère était employée. Enfin, le revenu des participants était majoritairement compris entre 0-5000€ avec 188 participants. La quasi-totalité de l'échantillon ne faisait partie d'aucune association défendant la vie privée ($n = 394$, soit 96,57 %). Les participants connaissaient la CNIL ($n = 158$, soit 38,73 %) et 164 participants (40,20 %) appliquaient quelques-uns de ces conseils. Pour terminer, les participants connaissaient le VPN ($n = 223$, soit 54,66 %), mais de nombreuses personnes n'en utilisaient jamais ($n = 191$, soit 46,81 %).

Il nous semble important de rapporter les variables ayant obtenu plus de 5 % de modalités « Je ne souhaite pas répondre » que nous avons noté « Privé » dans le Tableau 11 (voir Appendice F) : le statut marital (5,15 %), l'orientation politique (9,80 %), le parti politique (32,11 %), la catégorie socioprofessionnelle du père (9,80 %), la catégorie socioprofessionnelle de la mère (6,37 %), les ressources financières annuelles (11,76 %), la mise en place des conseils de la CNIL (12,55 %) et l'utilisation d'un VPN (5,64 %). L'ensemble des données est regroupé dans un tableau disponible en Appendice F (voir Tableau 11).

Analyse descriptive du questionnaire Security Behavior Intentions Scale

Le questionnaire *Security Behavior Intentions Scale* (Egelman et Peer, 2015a, 2015b) comporte 16 items que nous avons traduits. Cet instrument vise à évaluer les connaissances des participants en matière de sécurité informatique. Le recueil de données

en vue de sa validation s'est effectué durant le premier temps et le second temps de l'expérience comme indiqué dans le schéma récapitulatif (voir Figure 6) et comportait en tout 404 participants.

L'échantillon était différent de l'échantillon de l'article original, car la moyenne de la sous-échelle de sécurité de l'outil était plus élevée. Le score moyen de la sous-échelle concernant la génération de mot de passe n'était pas différent de l'article. Enfin, les scores moyens des sous-échelles concernant la conscience proactive et les mises à jour étaient plus faibles que dans l'article original (voir Tableau 12).

Concernant les données comparables entre notre échantillon et l'échantillon de Egelman et Peer (2015a), notre échantillon comportait 404 participants majoritairement jeunes ($M = 22,22$ ans, $ÉT = 4,73$) alors que celui de Egelman et Peer (2015a) comportait 457 individus avec un âge moyen de 32,91 ans ($ÉT = 11,19$). Notre échantillon était majoritairement féminin ($n = 307$), alors que l'échantillon d'Egelman et Peer (2015a) était légèrement plus masculin avec 58,2 % d'hommes (contrairement au nôtre qui en comportait 22,3 %). Les étudiants de notre échantillon étaient majoritairement détenteurs d'une licence équivalente au baccalauréat québécois ($n = 159$), suivis de près par 147 bacheliers (équivalent au DEC), 69 étudiants avaient déjà un master complété (équivalent d'une maîtrise) et 3 participants avaient un doctorat complété. Les individus constituant l'échantillon de Egelman et Peer (2015a) avaient au moins terminé le lycée

(équivalent du Cégep) pour 33 % d'entre eux et 33,3 % avaient une licence (équivalent au baccalauréat).

Tableau 12

Test t sur échantillon unique des différentes sous-échelles de la Security Behavior Intentions Scale de l'étude avec les moyennes dans l'article de validation originale

Sous-échelle	Moyenne de l'étude	Intervalle de confiance à 95%		Moyenne de l'article original	<i>p</i>	<i>d</i> de Cohen
		Borne inférieure	Borne supérieure			
Sécurité de l'outil	4,34	4,22	4,46	3,21	< 0,001	1,56
Génération de mots de passe	3,31	3,23	3,40	3,25	0,18	0,11
Conscience proactive	3,10	2,98	3,24	3,73	< 0,001	-0,75
Mises à jour	3,19	3,04	3,34	3,47	< 0,001	-0,29

Concernant les autres données sociodémographiques, les individus de notre échantillon étaient principalement célibataires ($n = 289$) et de nationalité française ($n = 344$). Les étudiants en psychologie ($n = 72$) étaient les plus nombreux ainsi que ceux en droit ($n = 49$). Les autres domaines d'études ayant moins de 15 représentants n'ont pas été affichés. Concernant les questions politiques, les participants rapportaient principalement être assez intéressés ($n = 162$) ou peu intéressés ($n = 139$). Lorsqu'on demandait aux participants de se positionner sur une échelle visuelle analogique où 0 représente un engagement politique à gauche et 10 à droite, on obtenait une tendance plutôt à gauche. La valeur 3 était la modalité de réponse la plus fréquente ($n = 66$) pour cette question. La seconde modalité la plus fréquente était 4 ($n = 55$) et indique un échantillon plutôt à gauche sur le spectre politique. La moyenne sur cette échelle était de 3,74 ($ÉT = 2,38$). Ce résultat était cohérent avec le parti politique des participants qui était majoritairement à gauche : Europe Écologie Les Verts (EELV) avec 91 participants. Cependant, la première modalité était « Privé » avec 130 participants ayant choisi de ne pas divulguer cette information. Pour la catégorie socioprofessionnelle du père, mais aussi de la mère, la modalité principale était « Employés ». On retrouvait 76 personnes dont le père était employé et 122 dont la mère était employée. Le revenu des participants était majoritairement compris entre 0-5000 € avec 187 participants. La majorité de l'échantillon ne faisait partie d'aucune association défendant la vie privée ($n = 394$, soit 97,52 %). Les participants connaissaient la CNIL ($n = 158$, soit 39,11 %) et 164 participants (40,59 %) appliquaient quelques-uns de ces conseils. Pour terminer, les

participants connaissaient le VPN ($n = 223$ soit 55,20 %), mais plusieurs personnes n'en utilisaient jamais ($n = 191$, soit 47,28 %).

Comme précédemment, il nous semble important de rapporter les variables ayant obtenu plus de 5 % de modalités « Je ne souhaite pas répondre » que nous avons noté « Privé » dans le Tableau 13 (voir Appendice G) : le statut marital (4,95 %), l'orientation politique (9,90 %), le parti politique (32,18 %), la catégorie socioprofessionnelle du père (9,65 %), la catégorie socioprofessionnelle de la mère (6,44 %), les ressources financières annuelles (11,63 %), la mise en place des conseils de la CNIL (13,55 %) et l'utilisation d'un VPN (6,23 %). L'ensemble des données est disponible dans un tableau en Appendice G (voir Tableau 13).

Analyse descriptive du Test d'indépendance au contexte

Le *Test d'indépendance au contexte* (Lo Monaco, Lheureux et Hamili-Falkowicz, 2008) comportait 11 items qui étaient supposés centraux et exploitables pour le questionnaire conformément à ce qui a été indiqué plus haut. Le recueil de données a été effectué durant le second temps de l'expérience comme indiqué dans le schéma récapitulatif (voir Figure 6) et comportait en tout 163 participants. Il n'y a aucune donnée manquante concernant le test des items candidats au noyau central et quelques données manquantes concernant les variables sociodémographiques.

Concernant notre échantillon, les 163 participants ($M = 22,24$ ans, $ÉT = 4,48$) étaient majoritairement des femmes ($n = 112$; 68,71 % des participants), principalement célibataires ($n = 111$) et de nationalité française ($n = 123$). Les étudiants de l'échantillon étaient majoritairement détenteurs d'une licence équivalente au baccalauréat québécois ($n = 62$), suivis de près par 54 bacheliers (équivalent au DEC), 29 étudiants avaient déjà un master complété (équivalent d'une maîtrise) et un seul avait un doctorat complété. La plupart des étudiants étudiaient en psychologie ($n = 25$) et en droit ($n = 19$).

Concernant les questions politiques, les participants rapportaient principalement être assez intéressés ($n = 56$) ou peu intéressés ($n = 52$). Lorsqu'on demandait aux participants de se positionner sur une échelle visuelle analogique où zéro représente un engagement politique à gauche et 10 à droite, on obtenait une tendance plutôt à gauche. La modalité « Je ne souhaite pas répondre » était la réponse la plus fréquente ($n = 24$) pour cette question. La seconde modalité la plus fréquente était 4 ($n = 20$) et indique un échantillon plutôt à gauche sur le spectre politique. La moyenne sur cette échelle était de 4,02 ($ÉT = 2,33$). Ce résultat était cohérent avec l'affiliation politique des participants qui se positionnait majoritairement à gauche : Europe Écologie Les Verts (EELV) avec 29 participants. Cependant, la première modalité était « Privé » avec 72 participants ayant choisi de ne pas divulguer cette information. Pour la catégorie socioprofessionnelle du père, mais aussi de la mère, la modalité principale était « Employés »; on retrouvait 33 personnes dont le père était employé et 42 dont la mère était employée. Le revenu des participants était principalement compris entre 0 et 5000 € avec 62 participants. Une

grande partie de l'échantillon ne faisait partie d'aucune association défendant la vie privée ($n = 147$, soit 90,18 %). Les participants connaissaient la CNIL ($n = 60$, soit 36,81 %) et 59 participants (36,20 %) appliquaient quelques-uns de ces conseils. Pour terminer, les participants connaissaient le VPN ($n = 91$, soit 55,83 %), mais plusieurs personnes n'en utilisaient jamais ($n = 74$, soit 45,40 %).

Comme précédemment, nous souhaitons mettre en exergue les variables ayant obtenu plus de 5 % de modalités « Je ne souhaite pas répondre » : le statut marital (6,13 %), l'orientation politique (14,72 %), le parti politique (44,17 %), la catégorie socioprofessionnelle du père (12,88 %), la catégorie socioprofessionnelle de la mère (11,66 %), les ressources financières annuelles (19,02 %) ainsi que l'application des conseils de la CNIL (12,88 %). L'ensemble des données est disponible dans un tableau en Appendice H (voir Tableau 14).

L'échelle visuelle analogique utilisée pour le *Test d'indépendance au contexte* allait de 0 (« Toujours, dans tous les cas, l'item fait partie de la vie privée en contexte d'utilisation des TIC ») à 10 (« Jamais, dans aucun cas, l'item ne fait pas partie de la vie privée en contexte d'utilisation des TIC »).

Concernant les statistiques descriptives des items testés avec le *Test d'indépendance au contexte*, on remarque que les moyennes sont relativement basses et sont comprises entre 1,55 pour la plus faible et 3,77 pour la plus haute (voir Tableau 15). Bien que la

moyenne seule ne puisse statuer sur la centralité d'un item, elle indique la tendance de l'item à être central. Conformément au paragraphe précédent, plus la moyenne de l'item sera proche de 0, plus les participants l'auront considérée comme centrale et plus celle-ci sera proche de 10, moins les participants l'auront considérée comme centrale. Les médianes sont comprises entre 1 et 3. Les écarts-types sont compris entre 1,94 et 2,72.

Tableau 15

Caractéristiques des scores de chaque item au Test d'indépendance au contexte

Variable	<i>n</i>	<i>M</i>	<i>ÉT</i>	Médiane	Minimum	Maximum	Asymétrie	Aplatissement
Donnée personnelle	163	1,55	2,11	1	0	10	1,70	3,06
Confidentialité	163	1,61	1,94	1	0	9,00	1,50	2,10
Intimité	163	1,66	2,04	1	0	10	1,46	2,45
Respect	163	2,26	2,51	2	0	9	1,00	-0,03
Protection	163	2,33	2,42	2	0	10	1,07	0,59
Droit	163	2,52	2,39	2	0	10	0,81	-0,02
Sécurité	163	2,52	2,63	2	0	10	0,95	0,48
Liberté	163	3,15	2,72	3	0	10	0,55	-0,66
Choix	163	3,25	2,56	3	0	10,00	0,41	-0,71
Secret	163	3,61	2,63	3	0	10	0,39	-0,73
Anonymat	163	3,77	2,62	3	0	10,00	0,30	-0,74

On peut noter que les items ayant la moyenne la plus faible (« confidentialité », « donnée personnelle » et « intimité ») présentent aussi les écarts-types les plus faibles. Les items « confidentialité » et « donnée personnelle » sont les seuls ayant une médiane aussi basse que 1.

Les items ont tous un minimum égal à 0, mais certains n'ont pas atteint le maximum de l'échelle qui était 10. Par exemple aucun des 163 participants ayant effectué le *Test d'indépendance au contexte* n'a choisi la borne « Jamais, dans aucun cas » pour l'item « confidentialité », tout comme pour l'item « respect ». Concernant le coefficient d'asymétrie (égal à 0 pour une distribution parfaitement normale), on remarque que certains items présentent une asymétrie plus forte que les autres comme « confidentialité » (1,50), « donnée personnelle » (1,70) et « intimité » (1,46). À l'inverse, certains items ont une asymétrie faible comme, « anonymat » (0,30) ou « choix » (0,41). D'autres mots sont eux caractérisés par une asymétrie se situant entre ces deux extrêmes comme « protection » (1,07) ou « sécurité » (0,95). Enfin, on peut remarquer une tendance identique concernant l'aplatissement. En effet, des mots comme « confidentialité » (2,10), « donnée personnelle » (3,06) et « intimité » (2,45) ont un aplatissement bien plus fort qu'« anonymat » (-0,74) ou « choix » (-0,71).

Afin que le lecteur ait une idée générale des distributions de chacune des trois tendances qui se dégagent du Tableau 15, nous proposons trois diagrammes en points (*dotplots*) (voir Figures 16 à 18). Rappelons que chaque point du diagramme représente un individu ayant choisi la valeur indiquée sur l'échelle visuelle analogique (0 = « Toujours, dans tous les cas, l'item fait partie de la vie privée en contexte d'utilisation des TIC (10 = « Jamais, dans aucun cas, l'item ne fait pas partie de la vie privée en contexte d'utilisation des TIC »)).

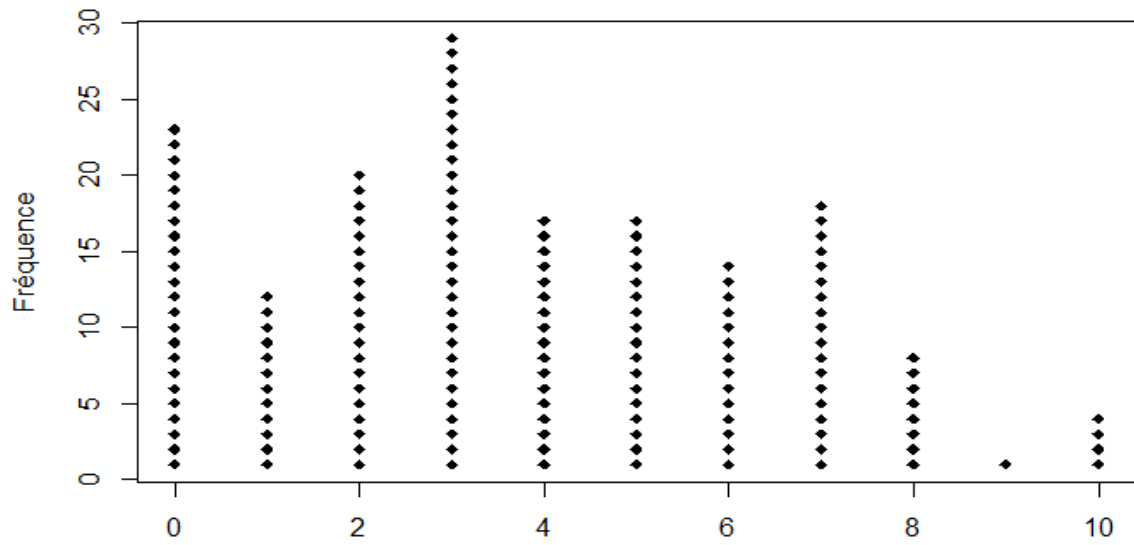
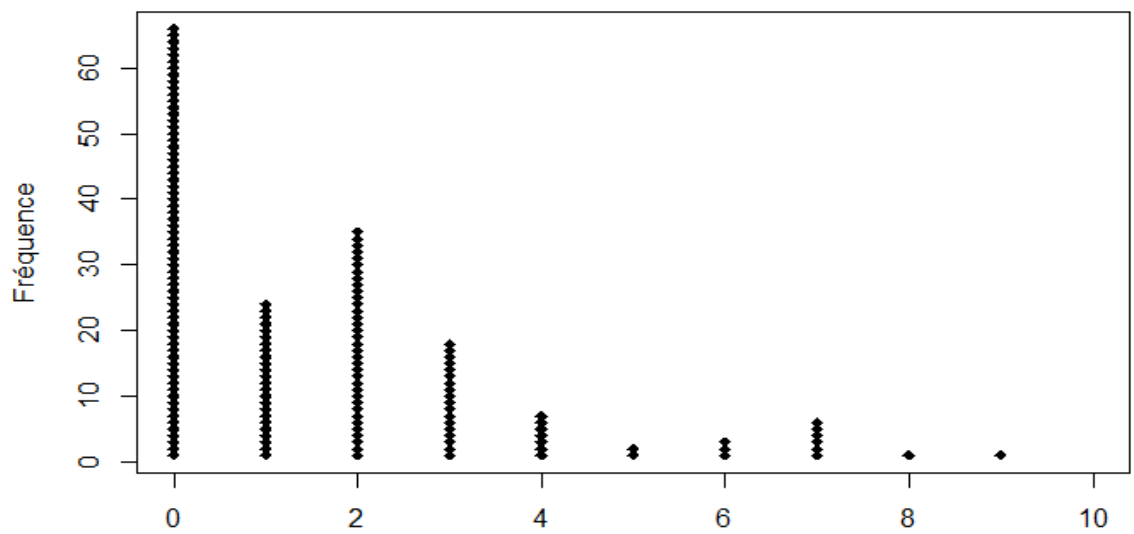
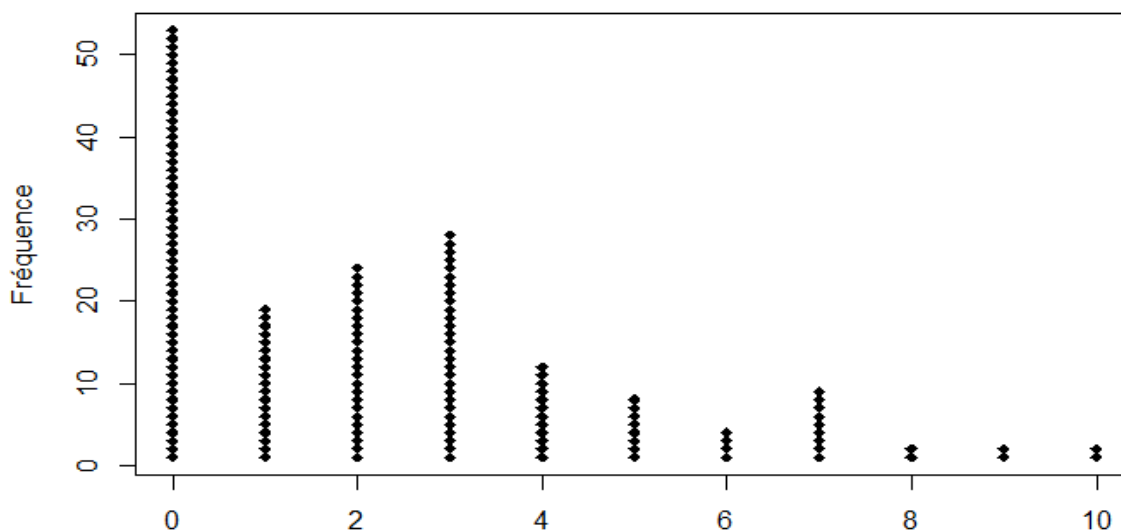
Figure 16*Diagramme à points de l'item « anonymat »***Figure 17***Diagramme à points de l'item « confidentialité »*

Figure 18

Diagramme à points de l'item « protection »



Analyse descriptive du Questionnaire de caractérisation

Le *Questionnaire de caractérisation* (Abric, 2003) comporte 21 items. Parmi ceux-ci, 10 items n'étaient pas centraux et 11 étaient supposés centraux et exploitables pour le questionnaire conformément à ce qui a été indiqué à la Figure 6. Tout comme pour le *Test d'indépendance au contexte*, la passation du questionnaire a été effectuée lors de la seconde collecte. Au total, nous avons 172 participants qui avaient rempli le *Questionnaire de caractérisation*. Il n'y a aucune donnée manquante concernant le test des items candidats au noyau central et quelques données manquantes ($n = 12$) concernant les variables sociodémographiques.

L'échantillon comportait 172 participants principalement jeunes ($M = 22,24$ ans, $ÉT = 4,48$), majoritairement des femmes ($n = 112$), qui étaient principalement célibataires ($n = 111$) et de nationalité française ($n = 123$).

Les étudiants de notre échantillon étaient majoritairement détenteurs d'une licence équivalente au baccalauréat québécois ($n = 62$), suivis de près par 54 bacheliers (équivalent au DEC, 29 étudiants avaient déjà un master complété (équivalent d'une maîtrise) et un seul avait un doctorat complété. La plupart des étudiants étudiaient en psychologie ($n = 25$) et en droit ($n = 19$). Les autres domaines d'études ayant moins de 7 représentants n'ont pas été affichés.

Concernant les questions politiques, les participants rapportaient principalement être assez intéressés ($n = 56$) ou peu intéressés ($n = 52$). Lorsqu'on demandait aux participants de se positionner sur une échelle visuelle analogique où 0 représente un engagement politique à gauche et 10 à droite, on obtenait une tendance plutôt à gauche. La modalité « Je ne souhaite pas répondre » était la réponse la plus fréquente ($n = 24$) pour cette question. La seconde modalité la plus fréquente était 4 ($n = 20$) et indique un échantillon plutôt à gauche sur le spectre politique. La moyenne sur cette échelle était de 4,02 ($ÉT = 2,33$). Ce résultat était cohérent avec le parti politique des participants qui était majoritairement à gauche : Europe Écologie Les Verts (EELV) avec 29 participants. Cependant, la première modalité était « Privé » avec 72 participants ayant choisi de ne pas divulguer cette information. Pour la catégorie socioprofessionnelle du père, mais aussi de

la mère, la modalité principale était « Employés ». On retrouvait 33 personnes dont le père était employé et 42 dont la mère était employée. Le revenu des participants était majoritairement compris entre 0 et 5000€ avec 62 participants. La plupart des participants ne faisaient partie d'aucune association défendant la vie privée ($n = 147$, soit 85,47 %). Les participants connaissaient la CNIL ($n = 60$, 34,88 %) et 59 participants (34,30 %) appliquaient quelques-uns de ces conseils. Pour terminer, les participants connaissaient le VPN ($n = 91$, soit 52,91 %), mais plusieurs personnes n'en utilisaient jamais ($n = 74$, soit 43,02 %).

Comme précédemment, nous souhaitons mettre en évidence les variables ayant obtenu plus de 5 % de modalités « Je ne souhaite pas répondre » : le statut marital (5,81 %), l'orientation politique (13,95 %), le parti politique (41,86 %), la catégorie socioprofessionnelle du père (12,21 %), la catégorie socioprofessionnelle de la mère (11,05 %), les ressources financières annuelles (18,02 %) et l'application des conseils de la CNIL (12,21 %). L'ensemble des données est disponible dans un tableau en Appendice I (voir Tableau 16).

Concernant les données descriptives du *Questionnaire de caractérisation*, on peut remarquer que ce sont les items « confidentialité » et « donnée personnelle » qui sont considérés respectivement par 77 % et 74 % de l'échantillon comme étant les plus caractéristiques de la vie privée en contexte d'usage des TIC (voir Tableau 17).

Tableau 17*Données descriptives concernant le Questionnaire de caractérisation*

Item	Le plus caractéristique	%	Ni plus ni moins caractéristique	%	Le moins caractéristique	%	Moyenne
Confidentialité	133	77	27	16	12	7	1,30
Donnée personnelle	128	74	30	17	14	8	1,34
Protection	96	56	56	33	20	12	1,56
Intimité	95	55	53	31	24	14	1,59
Sécurité	92	53	58	34	22	13	1,59
Anonymat	87	51	43	25	42	24	1,74
Consentement	85	49	62	36	25	15	1,65
Respect	64	37	66	38	42	24	1,87
Droit	59	34	67	39	46	27	1,92
Liberté	51	30	69	40	52	30	2,01
RGPD	46	27	84	49	42	24	1,98
Secret	41	24	72	42	59	34	2,10
Choix	40	23	71	41	61	35	2,12
Mot de passe	38	22	75	44	59	34	2,12
Réseaux sociaux	33	19	44	26	95	55	2,36
GAFAM	30	17	53	31	89	52	2,34
Cookies	24	14	48	28	100	58	2,44
Famille	22	13	60	35	90	52	2,40
Photo	17	10	65	38	90	52	2,42
Ami	14	8	47	27	111	65	2,56
Relation	9	5	54	31	109	63	2,58

Ensuite, on retrouve des items qui ont un pourcentage plus élevé de modalité « le plus caractéristique » sans qu'il se démarque aussi clairement que pour les deux items cités

précédemment. C'est notamment le cas pour : « anonymat » (51 %), « consentement » (49 %), « intimité » (55 %), « protection » (56 %), et « sécurité » (53 %).

D'autres items ont comme mode « ni plus ni moins caractéristique », comme les mots : « choix » (41 %), « droit » (39 %), « liberté » (40 %), « mot de passe » (44 %), « respect » (38 %), « RGPD » c'est-à-dire le règlement général sur la protection des données (49 %) et « secret » (42 %).

Enfin, on retrouve des items dont le mode est « le moins caractéristique » comme « ami » (65 %), « cookies » (58 %), « famille » (52 %), « GAFAM » c'est-à-dire Google, Amazon, Facebook, Apple, Microsoft (52 %), « photo » (52 %), « relation » (63 %) et « réseaux sociaux » (55 %).

Validation des questionnaires et comparaison de moyennes

Nous présentons l'ACP avec une rotation Varimax pour pouvoir effectuer la comparaison avec les articles originaux, ainsi que l'analyse factorielle exploratoire (AFE) avec extraction en axes principaux et rotation Varimax qui est plus appropriée pour une analyse factorielle.

Concernant le questionnaire *Privacy Orientation Scale* (Baruh et Cemalcılar, 2014), l'indice KMO est de 0,82; ce qui est adéquat. En comparant les structures factorielles du questionnaire, on s'aperçoit qu'elles sont semblables, puisqu'il ressort à chaque fois une

structure à 4 facteurs (voir Tableau 18) dont la traduction est : « la vie privée en tant que droit » pour le premier, « la préoccupation concernant la protection de la vie privée » pour le deuxième, « la vie privée conditionnelle aux actions des autres personnes » pour le troisième et « la préoccupation pour la vie privée d'autrui » pour la quatrième. Ces quatre facteurs expliquaient 70 % de la variance dans l'article original (Baruh et Cemalcılar, 2014). Dans notre étude, ils expliquent un peu plus de 60 % de la variance avec l'ACP. L'analyse avec l'extraction en axes principaux obtient une variance de 47,5 % pour la structure factorielle du questionnaire. On remarque que les items 2, 15 et 16 présentent une moins bonne saturation factorielle que dans l'article original. Enfin, et bien que la cohérence interne de la sous-échelle de la vie privée en tant que droit s'avère plus faible (0,65), le reste des sous-échelles présente un alpha de Cronbach convenable compris entre 0,77 et 0,81 (voir Tableau 19). Ces éléments suggèrent que la traduction du questionnaire *Privacy Orientation Scale* présente une structure similaire à celle de l'article original (Baruh et Cemalcılar, 2014).

Tableau 18*Saturation des items de la Privacy Orientation Scale selon l'analyse*

Item	ACP* de l'article original				ACP* sur les données				AFE** sur les données			
	1	2	3	4	1	2	3	4	1	2	3	4
1. Les lois sur la vie privée devraient être renforcées pour protéger la vie privée des personnes	0,79				0,73				0,60			
2. Les gens ont besoin de protection légale contre l'usage inapproprié des données personnelles	0,83				0,74				0,60			
3. Si j'avais à rédiger une loi constitutionnelle aujourd'hui, j'inclurais probablement la vie...	0,74			0,32	0,75				0,60			
4. Quand je partage des informations sur ma vie personnelle avec quelqu'un, je m'inquiète...		0,80				0,80				0,70		
5. Je me préoccupe du fait que les gens autour de moi savent trop de choses à propos de moi		0,83				0,83				0,78		
6. Je me préoccupe des conséquences du partage d'informations reliées à mon identité		0,67				0,71				0,61		
7. Je m'inquiète de partager des informations avec plus de personnes que je ne le voudrais		0,78				0,74				0,65		
8. Si quelqu'un n'est pas prudent concernant la protection de sa propre vie privée, je ne peux pas ...			0,76				0,76				0,66	
9. Pour que je puisse apprécier les aspects privés de ma vie, j'ai besoin que mes amis prennent soin...			0,77				0,80				0,74	
10. Je ne pourrais jamais considérer une personne digne de confiance, si elle distribue des détails sur sa ...			0,80				0,79				0,90	

Tableau 18*Saturation des items de la Privacy Orientation Scale selon l'analyse (suite)*

Item	ACP* de l'article original				ACP* sur les données				AFE** sur les données			
	1	2	3	4	1	2	3	4	1	2	3	4
11. Le niveau de vie privée dont je peux profiter dépend du degré auquel les personnes autour de...			0,78				0,66				0,54	
12. Il est important pour moi de respecter la vie privée des individus, même s'ils ne font pas attention...				0,72				0,82				0,76
13. Je valorise la vie privée des autres personnes autant que la mienne				0,81				0,78				0,68
14. Même si quelqu'un n'est pas prudent à propos de sa vie privée, je fais de mon mieux pour ...				0,83				0,83				0,77
15. Je fais toujours du mieux que je peux pour ne pas envahir la vie privée des autres personnes				0,73				0,56				0,46
16. Le respect de la vie privée d'autrui devrait être une priorité importante dans les relations sociales				0,79				0,56				0,46

Notes. Ne figurent dans le tableau que les scores supérieurs à 0,30. Voir Appendice E pour visualiser le questionnaire au complet.

* ACP = Analyse en composante principale avec extraction en composante principale et rotation Varimax.

** AFE = Analyse factorielle exploratoire avec extraction en axes principaux et rotation Varimax.

Tableau 19*Alpha de Cronbach des sous-échelles de la Privacy Orientation Scale*

Sous-échelle	Article original	Présente étude
La vie privée en tant que droit	0,83	0,65
Préoccupations concernant la protection de la vie privée	0,83	0,81
La vie privée conditionnelle aux autres	0,85	0,77
Préoccupation pour la vie privée d'autrui	0,86	0,78

Enfin, les analyses exploratoires ont abouti à un seul résultat significatif comme l'attestent les tableaux suivants (voir Tableaux 20 à 23). Plus précisément, il n'a pas été observé de différences de moyennes des scores des sous-échelles entre les sexes. Il n'a pas non plus été observé de différences de moyennes dans les sous-échelles selon que les participants aient révélé ou garder privée leur orientation politique, de même qu'aucun lien n'a été observé entre l'orientation politique et les scores aux sous-échelles. Enfin, un seul et unique lien est apparu comme significatif entre le nombre de réponses gardées privées et le score de la sous-échelles préoccupation personnelle.

Tableau 20

Comparaison selon le genre pour les différents scores des sous-échelles de la Privacy Orientation Scale avec un test t pour échantillons indépendants

Sous-échelle	Sexe	n	Moyenne	Intervalle de confiance à 95%		p	d de Cohen
				Borne inférieure	Borne supérieure		
La vie privée comme un droit	Homme	35	4,36	4,07	4,59	0,66	-0,10
	Femme	112	4,42	4,33	4,51		
Préoccupation personnelle à propos de la vie privée	Homme	35	3,54	3,25	3,81	0,33	-0,22
	Femme	112	3,71	3,55	3,85		
La vie privée conditionnelle aux autres	Homme	35	3,20	2,88	3,54	0,56	-0,13
	Femme	112	3,31	3,16	3,44		
Préoccupation concernant la vie privée d'autrui	Homme	35	4,14	3,95	4,33	0,17	-0,30
	Femme	112	4,30	4,21	4,39		

Tableau 21

Comparaison selon la déclaration ou non de l'orientation politique¹ et les différents scores des sous-échelles de la Privacy Orientation Scale avec un test t pour échantillons indépendants

Sous-échelle	Orientation politique	n	Moyenne	Intervalle de confiance à 95%		p	d de Cohen
				Borne inférieure	Borne supérieure		
La vie privée comme un droit	Orientation révélée	130	4,45	4,35	4,55	0,06	0,41
	Privé	24	4,22	4,01	4,42		
Préoccupation personnelle à propos de la vie privée	Orientation révélée	130	3,69	3,53	3,82	0,82	0,04
	Privé	24	3,66	3,39	3,91		
La vie privée conditionnelle aux autres	Orientation révélée	130	3,28	3,14	3,43	0,72	0,08
	Privé	24	3,22	2,90	3,55		
Préoccupation concernant la vie privée d'autrui	Orientation révélée	130	4,29	4,19	4,40	0,09	0,31
	Privé	24	4,13	3,98	4,28		

¹ L'évaluation de la politique se faisait comme suit ; plus le score est proche de zéro et plus la personne se considérait à gauche. Plus la personne se rapprochait de 10 et plus elle se considérait comme à droite.

Tableau 22

Corrélation entre la tendance politique¹ et les différents scores des sous-échelles de la Privacy Orientation Scale

Sous-échelle	Corrélation	Intervalle de confiance à 95%		p
		Borne inférieure	Borne supérieure	
La vie privée comme un droit	0,04	-0,129	0,21	0,71
Préoccupation personnelle à propos de la vie privée	0,06	-0,140	0,26	0,51
La vie privée conditionnelle aux autres	0,13	0,050	0,25	0,16
Préoccupation concernant la vie privée d'autrui	-0,17	-0,350	0,04	0,06

Tableau 23

Corrélation entre le nombre de réponse « Je ne souhaite pas répondre » et les différents scores des sous-échelles de la Privacy Orientation Scale

Sous-échelle	Corrélation	Intervalle de confiance à 95%		p
		Borne inférieure	Borne supérieure	
La vie privée comme un droit	-0,16	-0,306	-0,03	0,05
Préoccupation personnelle à propos de la vie privée	-0,18	-0,310	-0,03	0,03
La vie privée conditionnelle aux autres	-0,05	-0,200	0,14	0,58
Préoccupation concernant la vie privée d'autrui	-0,09	-0,250	0,08	0,28

¹ L'évaluation de la politique se faisait comme suit ; plus le score est proche de zéro et plus la personne se considérait à gauche. Plus la personne se rapprochait de 10 et plus elle se considérait comme à droite.

Concernant le questionnaire *Security Behavior Intentions Scale* (Egelman et Peer, 2015b), l'indice KMO est convenable et s'élève à 0,72. En comparant, les structures factorielles du questionnaire, on s'aperçoit qu'elles sont semblables. En effet, pour chacune d'entre elles, une structure à quatre facteurs émerge (voir Tableau 24). Les facteurs étaient « sécurité de l'outil », « génération de mot de passe », « conscience proactive » et « mises à jour ». Ces quatre dimensions expliquaient 55,6 % de la variance dans l'article original (Egelman et Peer, 2015b). Dans notre étude, ils expliquent 50,6 % de la variance avec l'ACP. L'analyse avec l'extraction en axes principaux génère une variance de 36,2 % pour la structure factorielle du questionnaire.

On remarque que plusieurs items ont une moins bonne saturation factorielle, comme les numéros 3, 5, 6, 12 et 13, comparativement à l'article original. Enfin, la cohérence interne n'est pas aussi bonne que dans l'article original comme en témoigne la sous-échelle concernant la sécurité et les mots de passe (voir Tableau 25). La dimension conscience, bien que faible, reste acceptable avec un alpha de 0,62. Enfin, la dernière échelle concernant les mises à jour est semblable à la cohérence de l'article original, 0,71 pour notre étude contre 0,72 pour l'article original. Globalement, la structure factorielle du questionnaire *Security Behavior Intentions Scale* traduit présente une structure similaire à celle de l'article original (Egelman et Peer, 2015b).

Tableau 24*Saturation des items du questionnaire Security Behavior Intentions Scale*

Item	ACP* de l'article original				ACP* de l'étude				AFE** de l'étude			
	1	2	3	4	1	2	3	4	1	2	3	4
1. J'ai prévu un verrouillage automatique de mon écran lorsque je ne l'utilise pas durant une période de temps...	0,76				0,74				0,61			
2. J'utilise un mot de passe pour verrouiller mon ordinateur ou ma tablette	0,72				0,76				0,65			
3. Je verrouille manuellement mon écran d'ordinateur quand je m'en éloigne	0,71				0,54				0,42			
4. J'utilise un code d'identification personnel ou un mot de passe pour déverrouiller mon téléphone intelligent	0,72				0,79				0,70			
5. Je ne change pas mon mot de passe, à moins d'y être contraint		0,69				0,52				0,32		
6. J'utilise différents mots de passe pour chacun des comptes que j'ai		0,74				0,58				0,39		
7. Quand je crée un nouveau compte, j'essaye d'utiliser un mot de passe qui dépasse les exigences minimums du site		0,74				0,75				0,72		
8. Je n'utilise pas de caractères spéciaux dans mon mot de passe si ce n'est pas obligatoire		0,74				0,72				0,56		
9. Quand quelqu'un m'envoie un lien, je l'ouvre sans vérifier auparavant où il me dirige			0,70				0,67				0,56	

Tableau 24*Saturation des items du questionnaire Security Behavior Intentions Scale (suite)*

Item	ACP* de l'article original				ACP* de l'étude				AFE** de l'étude			
	1	2	3	4	1	2	3	4	1	2	3	4
10. Je reconnais le site que je visite en me basant sur son apparence, plutôt que de regarder la barre d'adresse			0,55				0,76				0,66	
11. J'envoie des informations à des sites internet sans vérifier premièrement qu'elles sont envoyées de façon sécurisée...			0,73									
12. Quand je parcours des sites internet, je survole les liens avec la souris pour voir où ils me dirigent, avant de ...			0,54									
13. Si je découvre un problème de sécurité, je continue à faire ce que je faisais parce que je considère que quelqu'un ...			0,63				0,52				0,37	
14. Lorsque je suis sollicité pour une mise à jour de logiciel, je l'installe tout de suite				0,80				0,80				0,57
15. J'essaye de m'assurer que les logiciels que j'utilise sont à jour				0,77				0,83				0,85
16. Je vérifie si mon antivirus se met régulièrement à jour par lui-même				0,73				0,67				0,54

Note. Ne figurent dans le tableau que les scores supérieurs à 0,30. Voir Appendice E pour visualiser le questionnaire au complet

* ACP = Analyse en composante principale avec extraction en composante principale et rotation Varimax; ** AFE = Analyse factorielle exploratoire avec extraction en axes principaux et rotation Varimax.

Tableau 25

*Alpha de Cronbach des sous-échelles du questionnaire
Security Behavior Intentions Scale*

Sous-échelle	Article original	Étude
Sécurité de l'outil	0,76	0,67
Génération de mot de passe	0,73	0,60
Conscience proactive	0,67	0,62
Mises à jour	0,72	0,71

Pour terminer, les analyses exploratoires ont abouti à peu de résultats significatifs (voir Tableaux 26 à 29). Parmi ces résultats, on peut observer qu'il existe une faible corrélation négative entre l'orientation politique et la sous-échelle concernant la sécurité de l'outil informatique, $r_s = -0,29$, $p < 0,001$. Il existe aussi une faible corrélation entre l'orientation politique et la sous-échelle concernant la sécurité de l'outil informatique, $r_s = -0,21$, $p = 0,02$.

Tableau 26

Comparaison selon le genre pour les différents scores des sous-échelles de la Security Behavior Intentions Scale avec un test t pour échantillons indépendants

Sous-échelle	Sexe	n	Moyenne	Intervalle de confiance à 95%		p	d de Cohen
				Borne inférieure	Borne supérieure		
Sécurité de l'outil	Homme	35	4,37	4,12	4,59	0,74	0,06
	Femme	112	4,33	4,18	4,45		
Génération de mots de passe	Homme	35	3,43	3,28	3,61	0,08	0,31
	Femme	112	3,25	3,15	3,36		
Conscience proactive	Homme	35	3,31	3,06	3,59	0,11	0,31
	Femme	112	3,04	2,9	3,21		
Mises à jour	Homme	35	3,15	2,83	3,45	0,75	-0,06
	Femme	112	3,21	3,03	3,40		

Tableau 27

Comparaison selon la déclaration ou non de l'orientation politique¹ et les différents scores des sous-échelles de la Security Behavior Intentions Scale avec un test t pour échantillons indépendants

Sous-échelle	Sexe	n	Moyenne	Intervalle de confiance à 95%		p	d de Cohen
				Borne inférieure	Borne supérieure		
Sécurité de l'outil	Orientation révélée	128	4,35	4,22	4,47	0,67	0,11
	Privé	24	4,27	3,92	4,58		
Génération de mots de passe	Orientation révélée	128	3,28	3,19	3,37	0,24	-0,31
	Privé	24	3,45	3,21	3,72		
Conscience proactive	Orientation révélée	128	3,14	3,00	3,27	0,32	0,25
	Privé	24	2,93	2,61	3,32		
Mises à jour	Orientation révélée	128	3,16	2,99	3,35	0,45	-0,16
	Privé	24	3,32	2,92	3,69		

¹ L'évaluation de la politique se faisait comme suit ; plus le score est proche de zéro et plus la personne se considérait à gauche. Plus la personne se rapprochait de 10 et plus elle se considérait comme à droite.

Tableau 28

Corrélation de Spearman entre la tendance politique¹ et les différents scores des sous-échelles de la Security Behavior Intentions Scale

Sous-échelle	Corrélation	Intervalle de confiance à 95%		<i>p</i>
		Borne inférieure	Borne supérieure	
Sécurité de l'outil	-0,29	0,10	0,47	< 0,001
Génération de mots de passe	0,10	-0,10	0,28	0,30
Conscience proactive	0,01	-0,18	0,21	0,91
Mises à jour	0,21	0,01	0,38	0,02

Tableau 29

Corrélation de Spearman entre le nombre de réponses « Je ne souhaite pas répondre » et les différents scores des sous-échelles de la Security Behavior Intentions Scale

Sous-échelle	Corrélation	Intervalle de confiance à 95%		<i>p</i>
		Borne inférieure	Borne supérieure	
Sécurité de l'outil	0,32	-0,14	0,21	0,69
Génération de mots de passe	0,08	-0,08	0,23	0,35
Conscience proactive	-0,08	-0,24	0,07	0,30
Mises à jour	0,06	-0,10	0,22	0,47

¹ L'évaluation de la politique se faisait comme suit ; plus le score est proche de zéro et plus la personne se considérait à gauche. Plus la personne se rapprochait de 10 et plus elle se considérait comme à droite.

Analyse

Durant la première partie de cette section, les résultats de la première collecte nécessaire à l'élaboration de la structure supposée la représentation seront présentés. Dans une seconde partie seront présentés les éléments du noyau central qui auront été confirmés par nos analyses ainsi que la présentation des relations entre les items de la relation.

Étape préliminaire à la découverte du noyau central : l'établissement des hypothèses de structure dégagées par l'analyse prototypique et thèmes émergents

Avant d'accomplir l'analyse prototypique, il est pertinent d'utiliser le logiciel Sémato pour lemmatiser les mots (Plante, 2022). La lemmatisation consiste à récupérer les mots sujets à variation, comme les noms (masculin/féminin, singulier/pluriel), les verbes (temps, personne, mode), etc., pour les transformer en une forme canonique (masculin, singulier et infinitif). La forme est la façon dont les mots sont enregistrés dans le dictionnaire, par exemple les verbes sont enregistrés dans le dictionnaire à l'infinitif et le reste des mots est enregistré au masculin singulier. Cette procédure amène une grande réduction de la taille du dictionnaire où ses formes naturelles (lexèmes) sont substituées par leurs lemmes.

Une fois cette transformation effectuée sur les données par le logiciel Sémato, celui-ci fait apparaître 143 formes après l'opération de lemmatisation. Le logiciel propose la proportion de chaque zone de la représentation. Dans la présente étude, la zone du noyau central représente 11,19 % des formes totales avec 16 mots différents. La zone de la première périphérie représente quant à elle 5,59 % des formes totales avec huit mots

différents. La zone mixte représente 32,87 % des formes totales, soit 47 mots différents et enfin, la seconde périphérie avec 50,35 % des formes totales, soit 72 mots différents. Comme indiqué précédemment, le seuil concernant les fréquences était de 10 citations et celui des rangs était de 2; grâce à ceux-ci, nous avons pu tracer la Figure 19 qui représente la structure projetée de la représentation.

Figure 19*Représentation des hypothèses structurales de la représentation*

	Rang médian inférieur au centile 15 (≤ 2)	Rang médian supérieur au centile 15 (> 2)
Fréquence supérieure au centile 85 (≥ 10)	Personnel, intimité, confidentialité, secret, respect, sécurité, protection, donnée, droit, privé, information, donnée personnelle, liberté, confidentiel, anonymat, choix	Famille, social, réseau social, cookies, intime, mot de passe, danger, ami
Fréquence supérieure au centile 85 (< 10)	Information personnelle, RGPD, consentement, photo, important, image, amour, partagé, sexuel, contrôle, protégé, vie personnelle, sexualité, bancaire, protéger, CNIL, oublié, ciblé, fondamental, restreint, sécurisé, devoir, propriété, donnée privée, GAFAM, vie, donné, Big Brother, navigation, inexistant, précaution, article, vie sexuelle, carte bancaire, divulguer, donne, publicité ciblée, exister, cercle restreint, droit fondamental, calme, téléphone, vie familiale, quotidien, donnée sensible, coordonnées, vie privée	Public, passe, relation, partage, caché, Internet, identité, couple, familial, publicité, individuel, surveillance, VPN, Facebook, intrusion, cookie, cacher, tranquillité, discrétion, sensible, utilisation, piratage, vente, confiance, limite, revente, loisir, jardin secret, espionnage, silence, propre, problème, maison, savoir, foyer, politique, argent, GAFA, expression, professionnel, réseau, condition, unique, intérêt, opinion personnelle, distance, localisation, anonyme, voyeurisme, accord, vie sentimentale, information privée, vigilance, risque, accès, règlement, contexte, insécurité, violation, sentimental, historique, perso, privacy, responsabilité, cryptage, google, nécessaire, regarder, vendu, difficile, traçabilité, commercialisation

Nous présentons aussi les tableaux détaillés de chaque zone, avec sa fréquence et le rang médian de chaque mot (voir Tableaux 30 à 33).

Tableau 30*Zone n°1 candidat au noyau central (fréquence élevée et rang faible)*

Mot	Fréquence	Rang
Personnel	106	2
Intimité	69	2
Confidentialité	60	1
Secret	50	2
Respect	44	2
Sécurité	42	2
Protection	42	2
Donnée	27	2
Droit	21	2
Privé	20	2
Information	20	2
Donnée personnelle	18	2
Liberté	15	2
Confidentiel	13	2
Anonymat	10	2
Choix	10	2

Tableau 31*Zone n°2 première périphérie (fréquence élevée et rang élevé)*

Mot	Fréquence	Rang
Famille	36	3
Social	23	3
Réseau social	17	3
Cookies	13	3
Intime	12	2,5
Mot de passe	11	3
Danger	11	3
Ami	10	3,5

Tableau 32*Zone n°3 zone mixte (fréquence faible et rang faible)*

Mot	Fréquence	Rang	Mot	Fréquence	Rang	Mot	Fréquence	Rang
Information personnelle	8	2	Navigation	2	2	Restreint	3	1
RGPD	8	2	Inexistant	2	2	Sécurisé	3	1
Consentement	8	2	Précaution	2	2	Devoir	3	1
Photo	7	2	Article	2	2	Propriété	3	2
Important	7	2	Vie sexuelle	2	1	Donnée privée	3	2
Image	5	2	Carte bancaire	2	1,5	GAFAM	2	1,5
Amour	5	2	Divulguer	2	1	Vie	2	2
Partagé	4	1,5	Donne	2	2	Big Brother	2	2
Sexuel	4	1	Publicité ciblée	2	2	Donnée	2	1,5
Contrôle	4	2	Exister	2	2			
Protégé	4	2	Cercle restreint	2	1,5			
Vie personnelle	4	1,5	Droit fondamental	2	1,5			
Sexualité	3	2	Calme	2	2			
Bancaire	3	1	Téléphone	2	1,5			
Protéger	3	1	Vie familiale	2	1			
CNIL	3	2	Quotidien	2	1,5			
Oubli	3	2	Donnée sensible	2	1,5			
Ciblé	3	2	Coordonnées	2	2			
Fondamental	3	1	Vie privée	2	1			

Tableau 33*Zone n°4 seconde périphérie (fréquence faible et rang élevé)*

Mot	Fréquence	Rang	Mot	Fréquence	Rang	Mot	Fréquence	Rang
Public	8	3	Loisir	4	4	Information privée	2	2,5
Passé	7	3	Jardin secret	3	4	Vigilance	2	2,5
Relation	7	3	Espionnage	3	3	Risque	2	3,5
Partage	7	3	Silence	3	3	Accès	2	2,5
Caché	7	3	Propre	3	4	Règlement	2	3,5
Internet	7	3	Problème	3	4	Contexte	2	4
Identité	6	3	Maison	3	4	Insécurité	2	3
Couple	6	3	Savoir	3	4	Violation	2	4
Familial	6	2,5	Foyer	3	3	Sentimental	2	3,5
Publicité	6	3,5	Politique	3	3	Historique	2	3,5
Individuel	6	3	Argent	3	3	Perso	2	3
Surveillance	6	3,5	GAFa	2	3,5	Privacy	2	3
VPN	5	3	Expression	2	3,5	Responsabilité	2	3
Facebook	5	4	Professionnel	2	3	Cryptage	2	3,5
Intrusion	5	4	Réseau	2	4	Google	2	2,5
Cookie	5	3	Condition	2	3	Nécessaire	2	3
Cacher	5	3	Unique	2	3,5	Regarder	2	3
Tranquillité	4	3,5	Intérêt	2	3	Vendu	2	2,5

Tableau 33*Zone n°4 seconde périphérie (fréquence faible et rang élevé) (suite)*

Mot	Fréquence	Rang	Mot	Fréquence	Rang	Mot	Fréquence	Rang
Discrétion	4	3	Opinion personnelle	2	3,5	Difficile	2	2,5
Sensible	4	3	Confiance	2	2,5	Traçabilité	2	2,5
Public	4	4	Limite	2	2,5	Commercialisation	2	3
Passe	4	3,5	Revente	2	3			
Relation	4	2,5	Loisir	2	4			
Utilisation	4	3	Distance	2	4			
Piratage	4	3,5	Localisation	2	3			
Vente	4	3	Anonyme	2	3			
Confiance	8	3	Voyeurisme	2	4			
Limite	7	3	Accord	2	4			
Revente	7	3	Vie sentimentale	2	4			

À la lecture de l'ensemble des termes composant la représentation, on peut noter que certains se recoupent dans des thèmes. À travers une analyse thématique exploratoire réalisée manuellement en rassemblant les mots ayant le même champ sémantique, quatre thèmes émergent : la technologie, la protection, l'intimité et le danger. L'un de ces quatre thèmes se retrouve exclusivement dans des termes du système périphérique (zones hors du noyau central). Les trois autres, quant à eux, traversent l'ensemble des zones de la représentation.

La technologie est le thème avec le plus de diversité dans ces termes. On suppose que notre objet d'étude aide à avoir un nombre conséquent de termes dans ce thème. On retrouve dans celui-ci 25 termes de la zone périphérique : « réseau social », « réseau », « cookies », « mot de passe », « photo », « image », « donnée », « bancaire », « donnée privée », « GAFAM », « navigation », « carte bancaire », « publicité ciblée », « téléphone », « données sensibles », « coordonnées », « Internet », « VPN », « Facebook », « GAFA », « localisation », « information privée », « historique », « cryptage » et « Google ». À noter que cela représente 18,88 % des différents items de l'ensemble de la représentation et 21,26 % des différents items du système périphérique.

Concernant les termes candidats au noyau central, trois termes ont été retenus. Seuls les termes « information », « donnée » et « donnée personnelle » ont été considérés comme faisant partie du thème technologie. Cela représente 2,10 % des différents items de la représentation et 18,75 % des différents items candidats au noyau central.

De plus, on peut noter la présence de sous-thèmes de la technologie dans le thème global de la technologie. Par exemple, on pourrait faire émerger un sous-thème sur le type de données. Autre point intéressant, certains sous-thèmes technologiques sont exclusifs aux éléments périphériques comme les médiums technologiques utilisés ou encore les entreprises liées à la technologie.

Tout comme le thème de la technologie, le thème de la protection fait à la fois partie des zones périphériques et du noyau central. Concernant la protection, nous avons considéré les neuf items du système périphérique suivants : « RGPD », « consentement », « contrôle », « protéger », « CNIL », « oubli », « sécurisé », « précaution » et « règlement ». Cela représente 6,99 % des différents items que l'on peut trouver dans la représentation et 7,87 % de ceux du système périphérique.

À noter que les termes, comme « VPN »¹ ou « cryptage », n'ont pas été inclus dans ce thème. Statuer sur l'inclusion du « VPN » fut difficile, car il protège partiellement la vie privée. De plus, le VPN peut aussi permettre de rejoindre des sites uniquement accessibles depuis certains pays. Attendu que l'intention derrière ce mot n'est pas complètement claire, nous avons préféré ne pas l'inclure dans ce thème. Pour des raisons

¹ Les néophytes pensent qu'un VPN rend anonyme. Malheureusement, cela est faux. En fait, le VPN ne fait que déplacer le problème. Ce n'est plus votre fournisseur Internet qui a accès aux données de navigation, mais votre fournisseur de VPN. De plus, si le VPN change effectivement la localisation de l'ordinateur, il ne change pas son adresse MAC. Celle-ci est une série de chiffres unique au monde reliée à un et unique appareil.

semblables, le cryptage n'a pas été inclus, car il ne cible pas directement la protection de la vie privée, mais plutôt celle des données.

Concernant les termes candidats au noyau central, sept mots ont été considérés comme appartenant au thème de la protection : « confidentialité », « secret », « sécurité », « protection », « droit », « confidentiel » et « anonymat ». Ces termes représentent 4,9 % des différents items de la représentation et 43,75 % des différents items candidats au noyau central.

Comme les thèmes précédents, le thème de l'intimité fait à la fois partie des éléments périphériques et des éléments candidats au noyau central. Ce thème concerne 18 termes périphériques qui sont : « famille », « amour », « sexuel », « vie personnelle », « sexualité », « vie sexuelle », « cercle restreint », « vie familiale », « quotidien », « relation », « couple », « familiale », « foyer », « vie sentimentale », « sentimental », « maison », « intime » et « ami ». Ces items constitutifs de ce thème représentent 12,59 % des différents items de la représentation totale et 14,17 % des différents items du système périphérique.

Concernant les termes candidats au noyau central, les termes « personnel » et « intimité » sont les seuls qui pourraient y être intégrés; ce qui représente 1,40 % des différents items de la représentation et 12,5 % des différents items candidats au noyau central.

Enfin, un dernier thème apparaît, soit le danger. Il a comme particularité de n'avoir aucun terme candidat au noyau central. Par conséquent, la représentation sociale de la vie privée étudiée auprès de cet échantillon universitaire français a la particularité de ne comporter aucun élément central propre au thème du danger. Dans ce thème, on retrouve les 13 termes suivants : « cookies », « danger », « Big Brother », « publicité ciblée », « surveillance », « intrusion », « piratage », « espionnage », « problème », « voyeurisme », « risque », « insécurité » et « violation ». Cela représente 9,09 % des différents items pour la totalité de la représentation et 10,24 % des différents items du système périphérique.

Validation du noyau central avec le Test d'indépendance au contexte

Nous rappelons aux lecteurs que tous les mots candidats au noyau central n'ont pas été retenus, conformément à la Figure 6. Onze mots candidats ont été conservés : « anonymat », « choix », « confidentialité », « donnée personnelle », « droites », « intimité », « liberté », « protection », « respect », « secret » et « sécurité ». Afin de déterminer si un item appartient au noyau central, il est nécessaire d'observer la valeur p et le graphique où sont réunies la distribution théorique de loi exponentielle et empirique de notre échantillon (Baggio et Renaud, 2015). Ainsi, trois cas de figure sont possibles.

Le premier cas, la valeur p est non significative. Dans ce cas, on peut conclure que l'item appartient bel et bien au noyau central, car la distribution de l'item n'est pas

différente de la distribution exponentielle. La lecture graphique est optionnelle et ne fait que confirmer que les distributions se superposent.

Le second cas survient lorsque la valeur p est inférieure à 0,05. Celle-ci nécessite obligatoirement que le chercheur se rapporte au graphique où sont représentées les distributions empirique et théorique. Si la distribution de l'item se situe au-dessus de la distribution exponentielle, alors l'item fait partie du noyau central, car sa distribution est plus asymétrique que la distribution exponentielle. En revanche, si la distribution de l'item se situe en-dessous de la distribution exponentielle, alors l'item n'appartient pas au noyau central, car sa distribution est moins asymétrique que la distribution exponentielle. De plus, dans cette étape, le coefficient d'asymétrie et l'oméga carré maximal¹ permettent d'étayer la lecture des distributions par une valeur numérique. Conjointement à la valeur p , l'allure de la courbe, la forme observée sur le diagramme en points superposés, le coefficient d'asymétrie et l'oméga carré maximal, nous proposons d'autres statistiques descriptives afin d'étayer les résultats dans le Tableau 34.

¹ L'oméga carré maximal est la distance maximale que le test calcule entre la distribution empirique et théorique. Donc, plus elle est grande, plus la différence entre les courbes sera marquée.

Tableau 34*Résumé des scores de chaque item du Test d'indépendance au contexte*

Variable	0*	1	2	3	4	5	6	7	8	9	10	Coefficient d'asymétrie	Oméga ² max	Valeur <i>p</i>
Anonymat	23	12	20	29	17	17	14	18	8	1	4	0,31	3,43	< 0,001
Choix	36	9	24	25	18	16	15	11	4	4	1	0,41	3,1	< 0,001
Confidentialité	66	24	35	18	7	2	3	6	1	1	0	1,53	1,49	< 0,001
Donnée personnelle	81	16	23	19	13	1	3	2	3	0	2	1,73	1,45	< 0,001
Droit	49	15	26	25	16	11	8	7	3	2	1	0,83	1,72	< 0,001
Intimité	75	12	28	26	7	5	5	3	0	0	2	1,49	2,32	< 0,001
Liberté	40	16	17	26	16	12	10	16	4	3	3	0,56	1,79	< 0,001
Protection	53	19	24	28	12	8	4	9	2	2	2	1,09	1,66	< 0,001
Respect	60	20	24	15	17	3	6	11	4	3	0	1,02	1,89	< 0,001
Secret	25	13	24	25	20	15	12	14	10	2	3	0,40	3,35	< 0,001
Sécurité	45	17	27	29	18	7	9	3	3	4	1	0,97	1,46	< 0,001

Note. 0 = « Toujours dans tous les cas, l'item est considéré comme faisant partie de la vie privée » et 10 = « Jamais dans aucun cas, l'item est considéré comme faisant partie de la vie privée ».

Dans le Tableau 34, tous les items présentent des tests significatifs ; ce qui signifie que tous les items sont différents de la courbe exponentielle théorique de référence. On se retrouve donc dans la deuxième possibilité décrite préalablement. La lecture des valeurs des coefficients d'asymétrie permet de remarquer que certains items se différencient nettement des autres, notamment par leur fort coefficient d'asymétrie. C'est le cas de « confidentialité » avec un coefficient d'asymétrie de 1,53, « donnée personnelle » avec un coefficient de 1,73 et enfin « intimité » avec un coefficient de 1,49. Au vu de ces fortes valeurs, ainsi que de la distribution de leurs scores et de la significativité du test, on peut s'attendre à ce que ces items fassent partie du noyau central. Cela est confirmé par les Figures 20 à 25.

Figure 20

Graphique de la distribution de l'item « confidentialité »

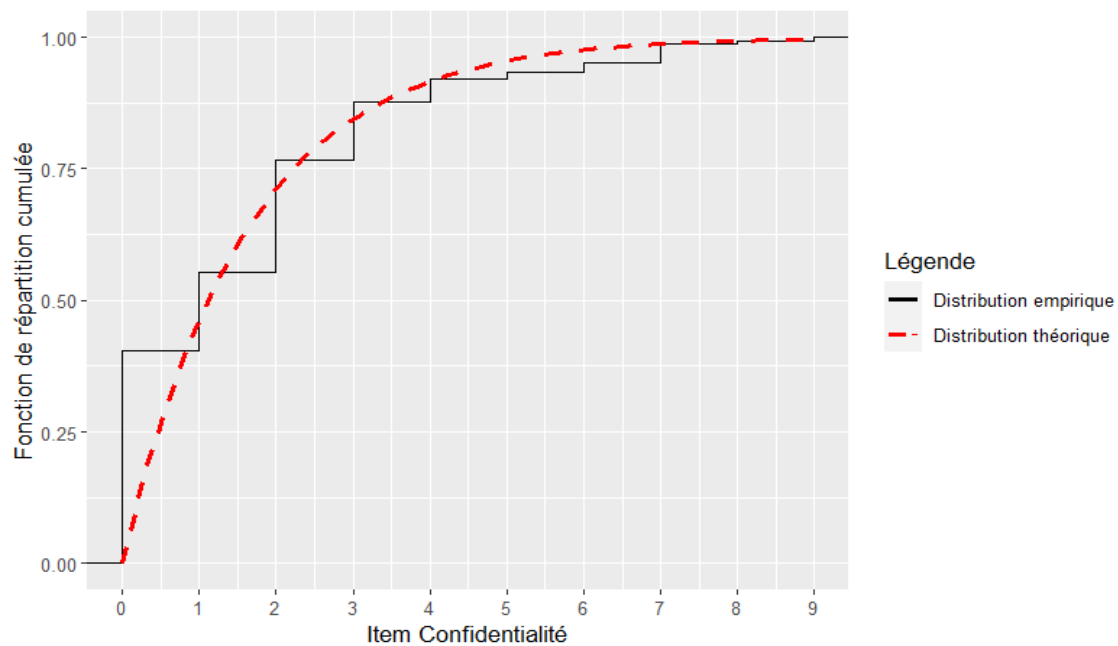
**Figure 21**

Diagramme à points de la distribution de l'item « confidentialité »

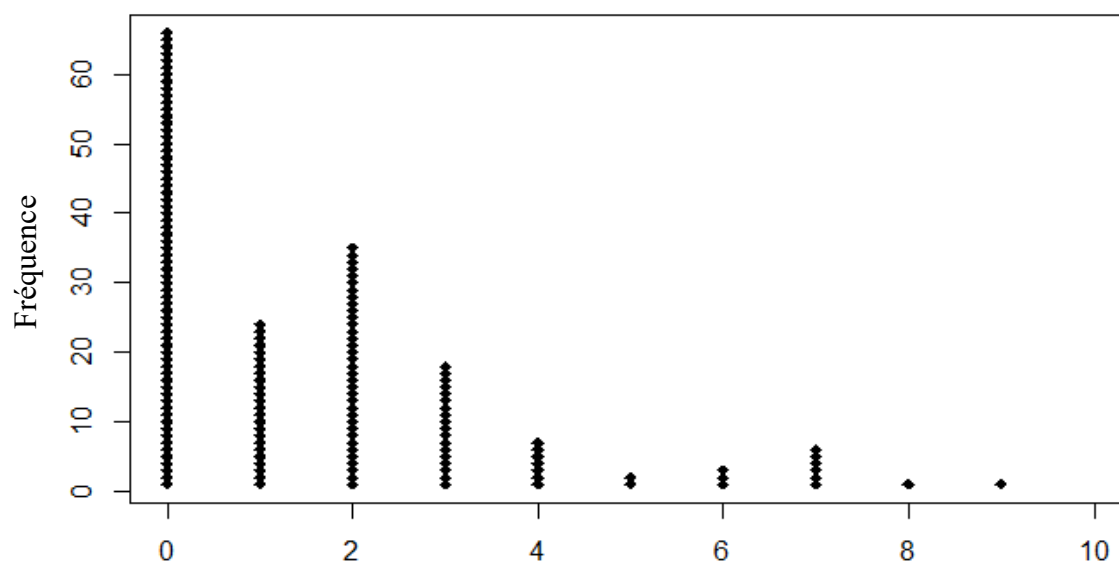


Figure 22

Graphique de la distribution de l'item « donnée personnelle »

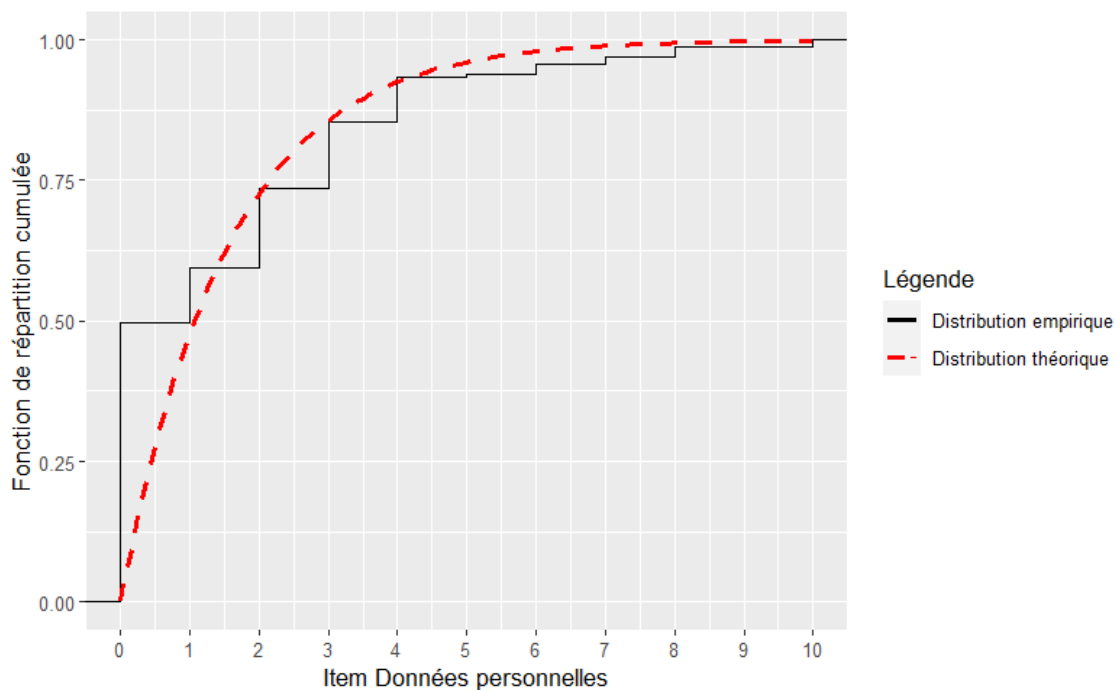
**Figure 23**

Diagramme à points de la distribution de l'item « donnée personnelle »

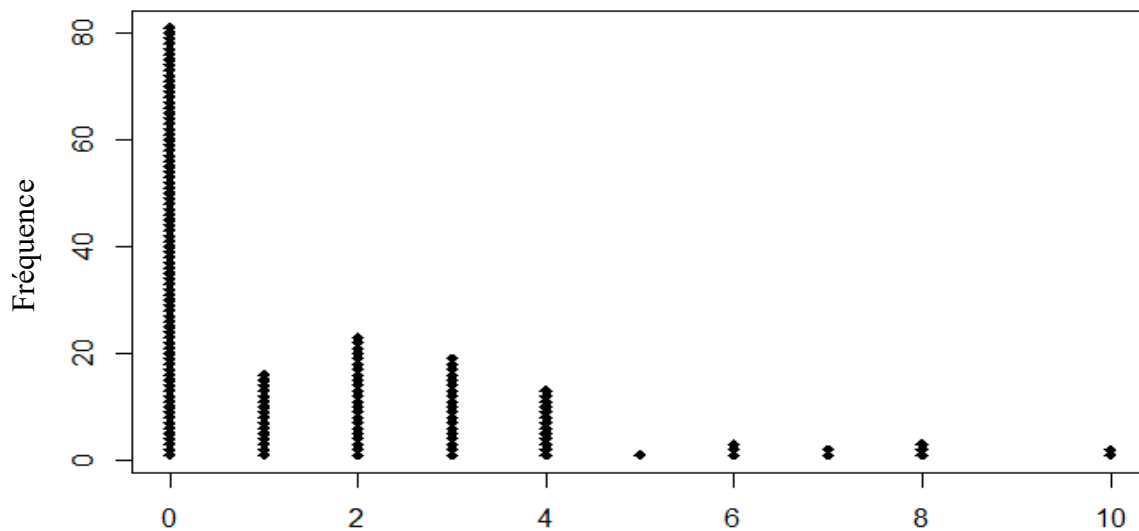
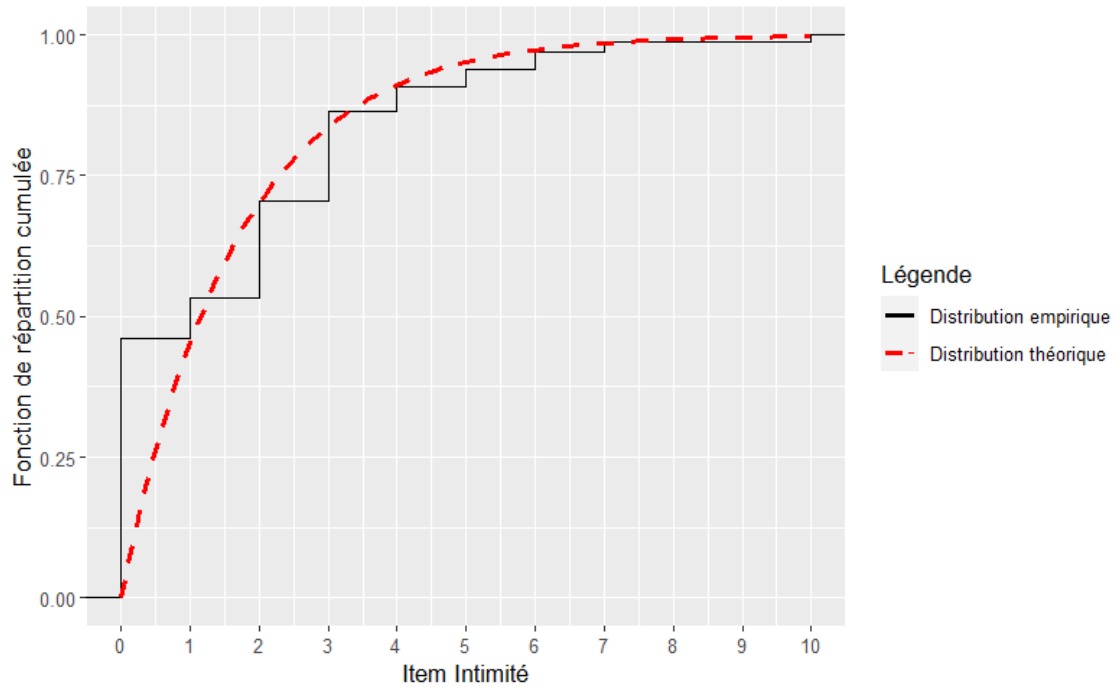
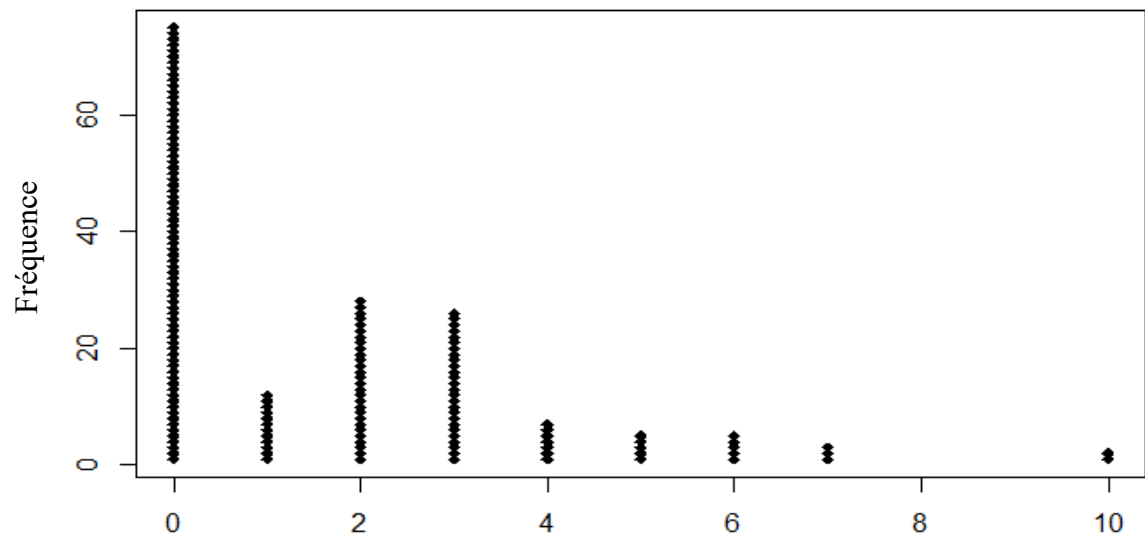


Figure 24*Graphique de la distribution de l'item « intimité »***Figure 25***Diagramme à points de l'item « intimité »*

Il est possible de constater que chaque marche de la distribution empirique est au-dessus de la courbe théorique exponentielle et qu'un fort effet de plancher se manifeste dès le début de la courbe. Pour l'item « intimité », pratiquement la moitié des participants cotent qu'il est toujours dans tous les cas, considéré comme faisant partie de la vie privée (soit la catégorie 0).

Un autre groupe de mots se détache dans le Tableau 34 : « droit », « protection », « respect » et « sécurité ». Ces mots présentent un coefficient d'asymétrie supérieur à 0,80 et leur distribution dans les modalités de réponse reste relativement élevée. Cependant, la valeur du coefficient pour ces items est plus difficile à interpréter, car l'effet de plancher, bien que présent, n'est pas aussi important que pour les items précédents, comme cela apparaît dans les figures suivantes (voir Figures 26 à 33).

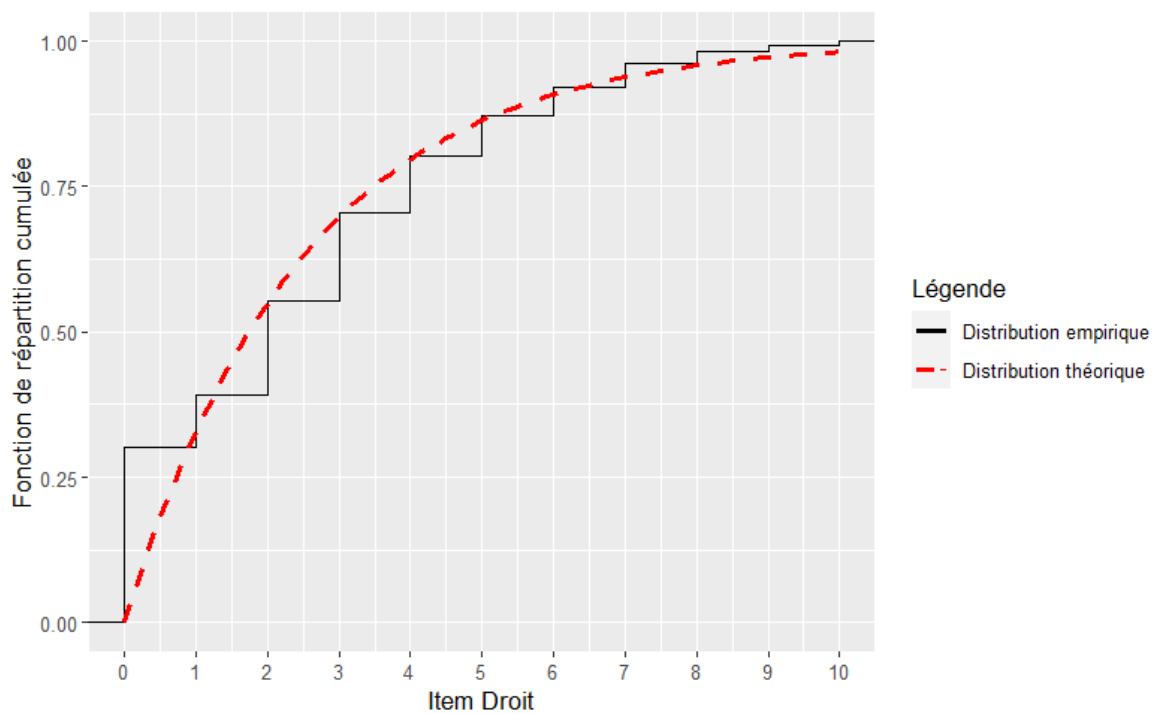
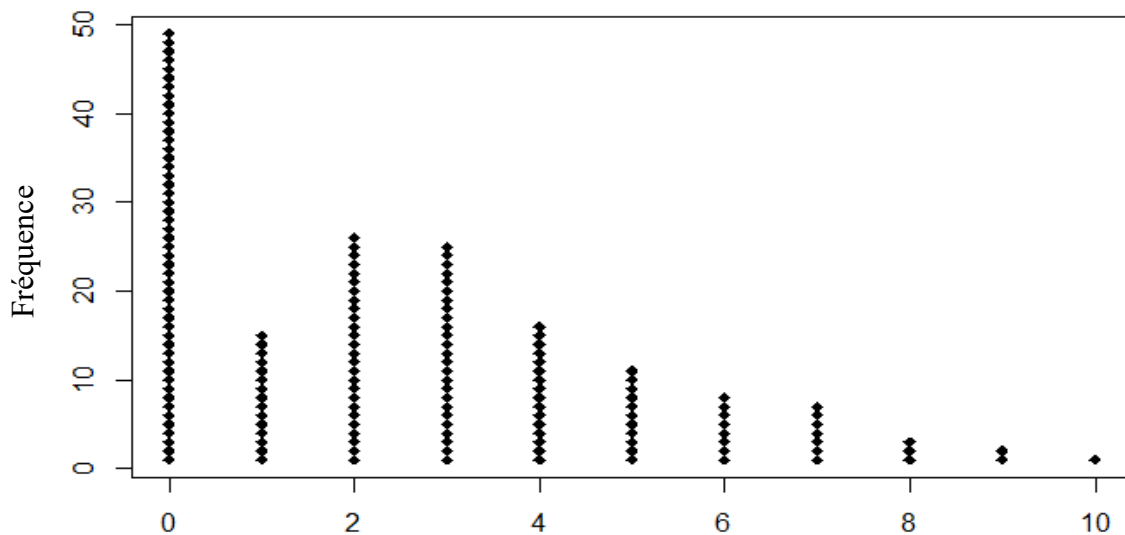
Figure 26*Graphique de la distribution de l'item « droit »***Figure 27***Diagramme à points de l'item « intimité »*

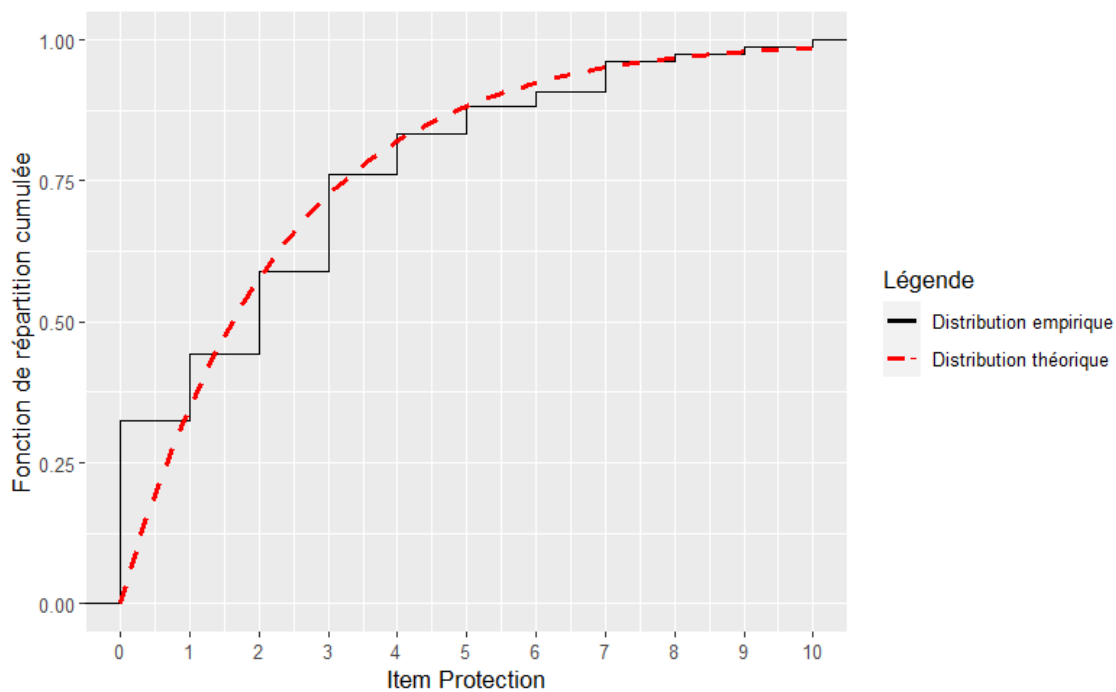
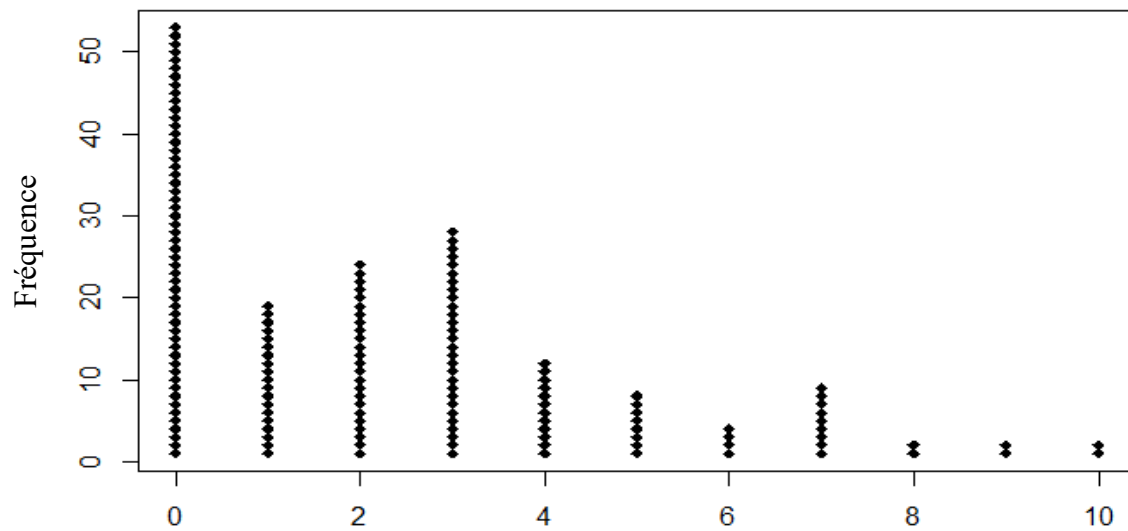
Figure 28*Graphique de la distribution de l'item « protection »***Figure 29***Diagramme à points de la distribution de l'item « protection »*

Figure 30

Graphique de la distribution de l'item « respect »

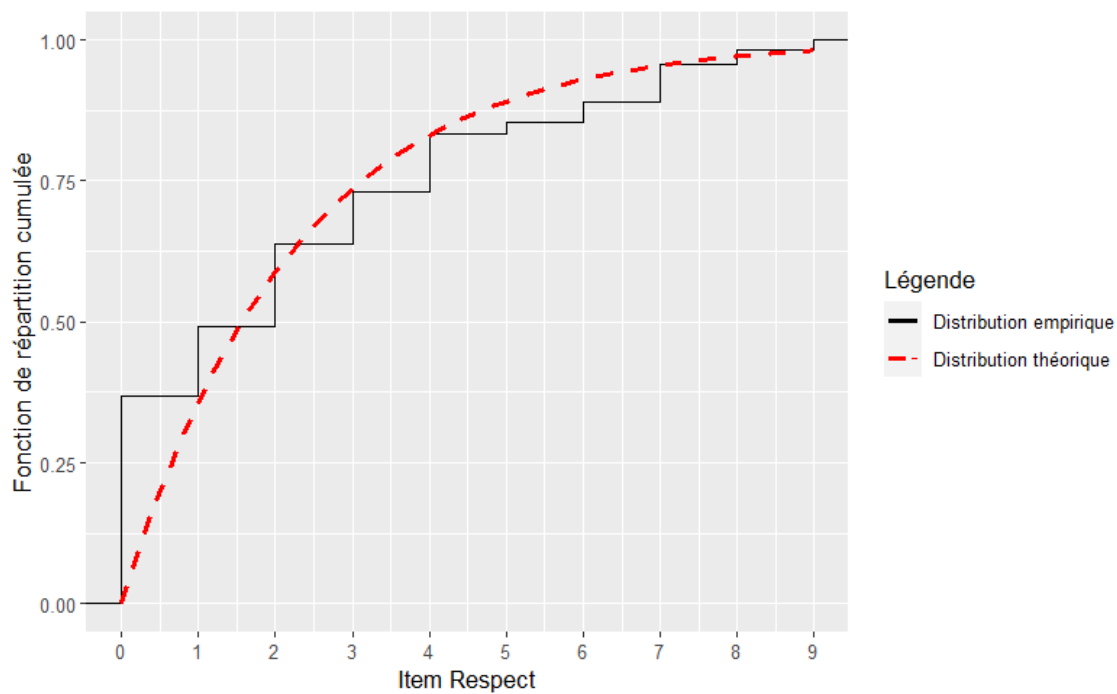
**Figure 31**

Diagramme à points de la distribution de l'item « respect »

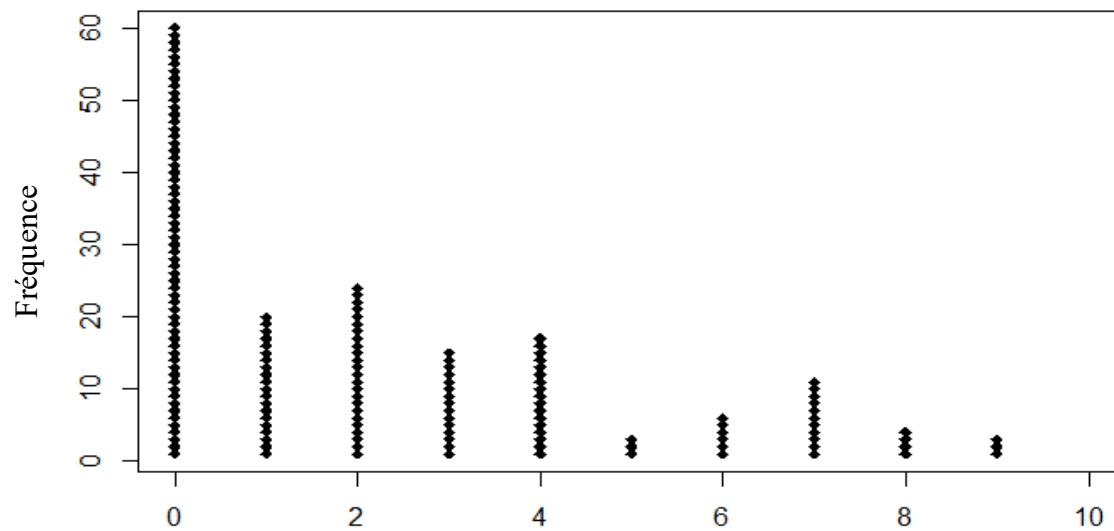


Figure 32

Graphique de la distribution de l'item « sécurité »

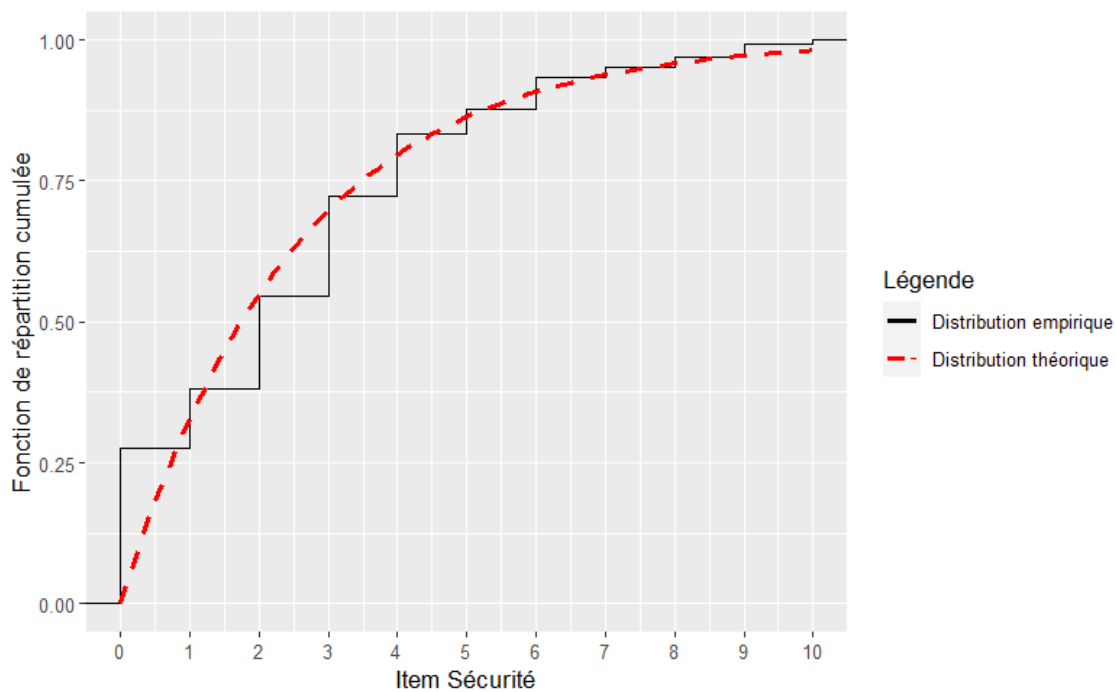
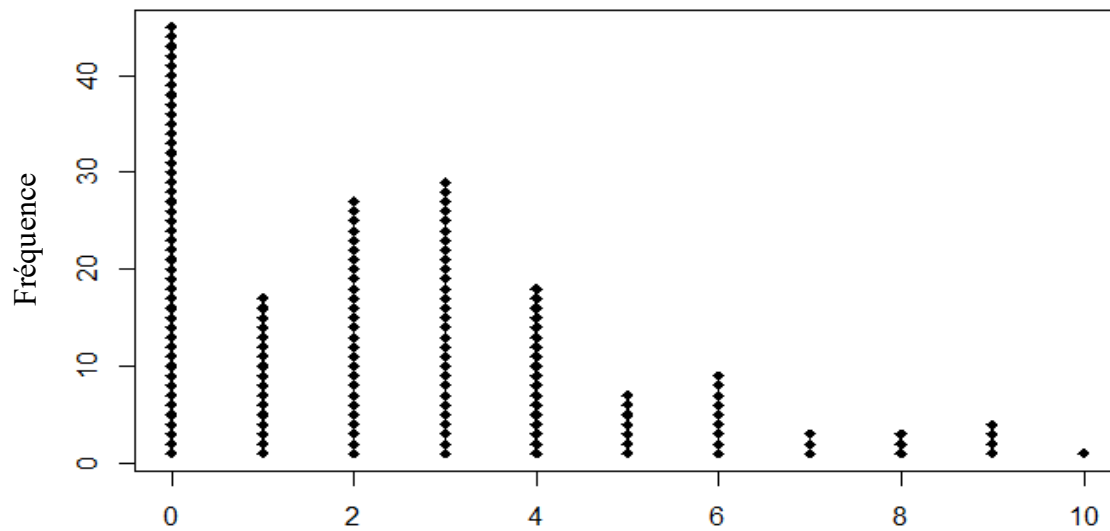
**Figure 33**

Diagramme à points de la distribution de l'item « sécurité »



Malgré tout, ces items font donc partie du noyau central de la représentation du fait de la significativité de leur test, de l'allure de leur courbe et de leur coefficient d'asymétrie élevé.

Enfin, le troisième et dernier groupe, avec « anonymat », « choix », « liberté » et « secret », comporte un test significatif, mais un faible coefficient d'asymétrie et leurs courbes de distribution se situent en-dessous de la distribution exponentielle (voir Figures 34 à 41).

Figure 34

Graphique de la distribution de l'item « anonymat »

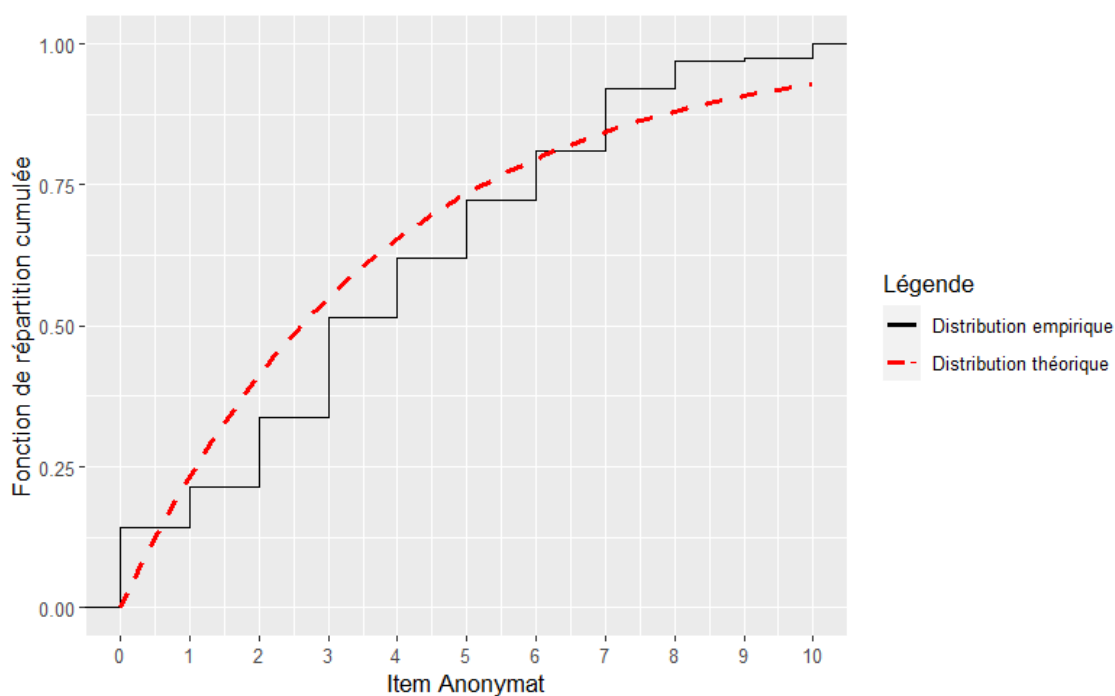
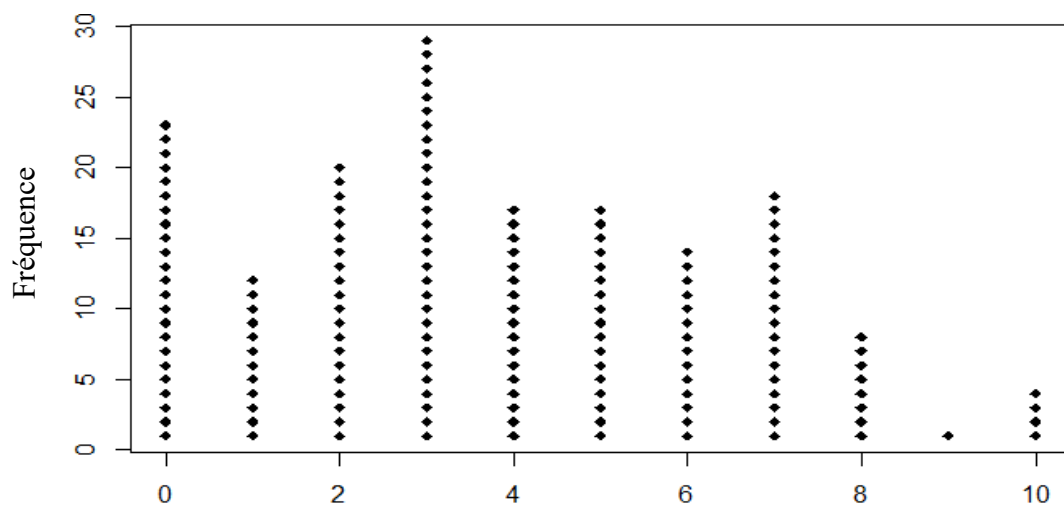


Figure 35

Diagramme à points de la distribution de l'item « anonymat »

**Figure 36**

Graphique de la distribution de l'item « choix »

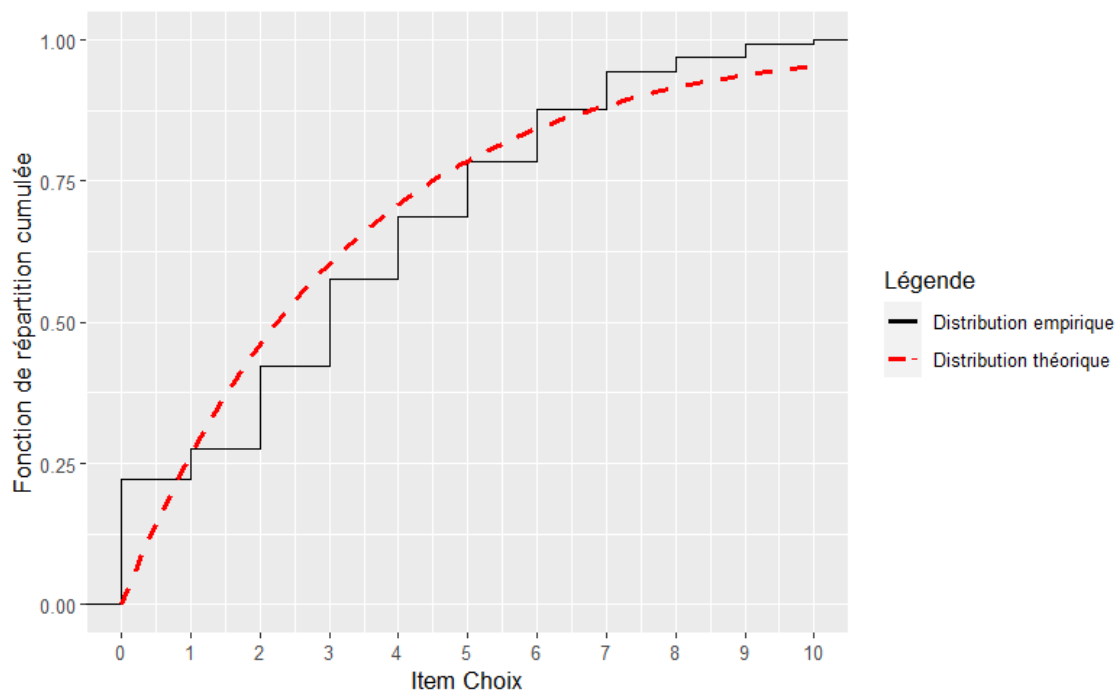
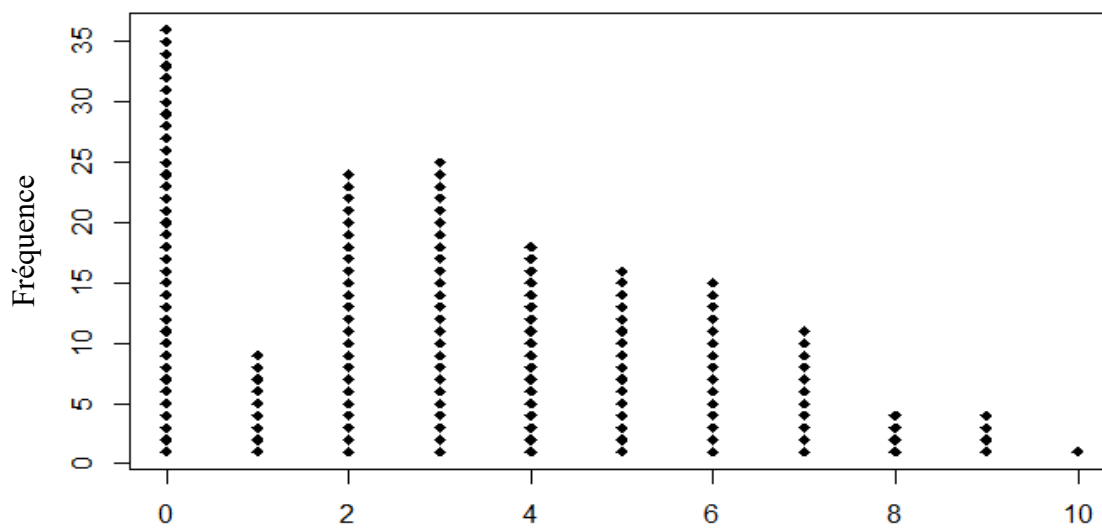


Figure 37

Diagramme à points de la distribution de l'item « choix »

**Figure 38**

Graphique de la distribution de l'item « liberté »

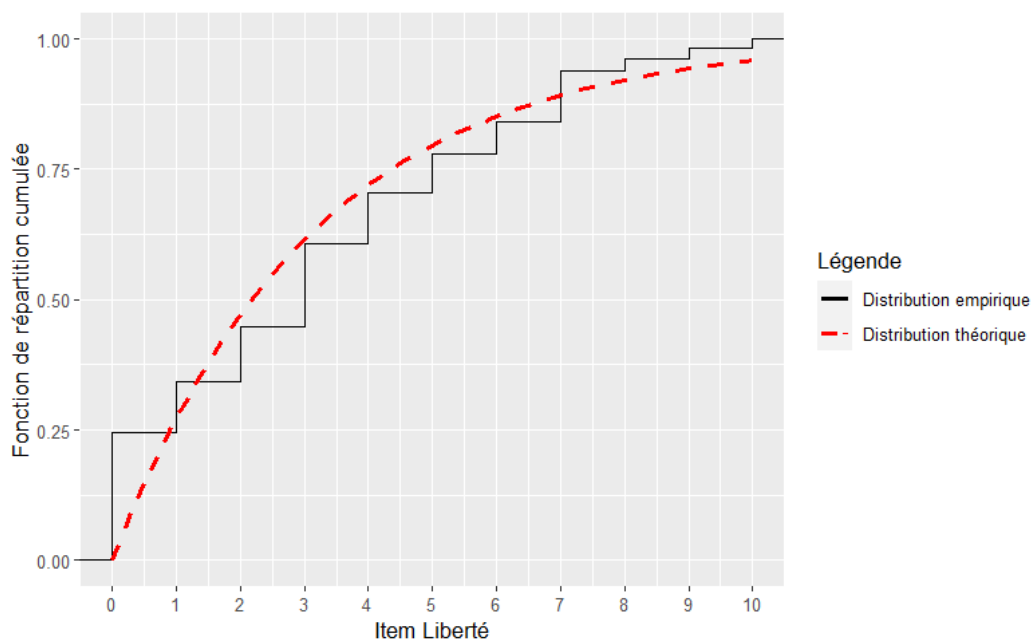


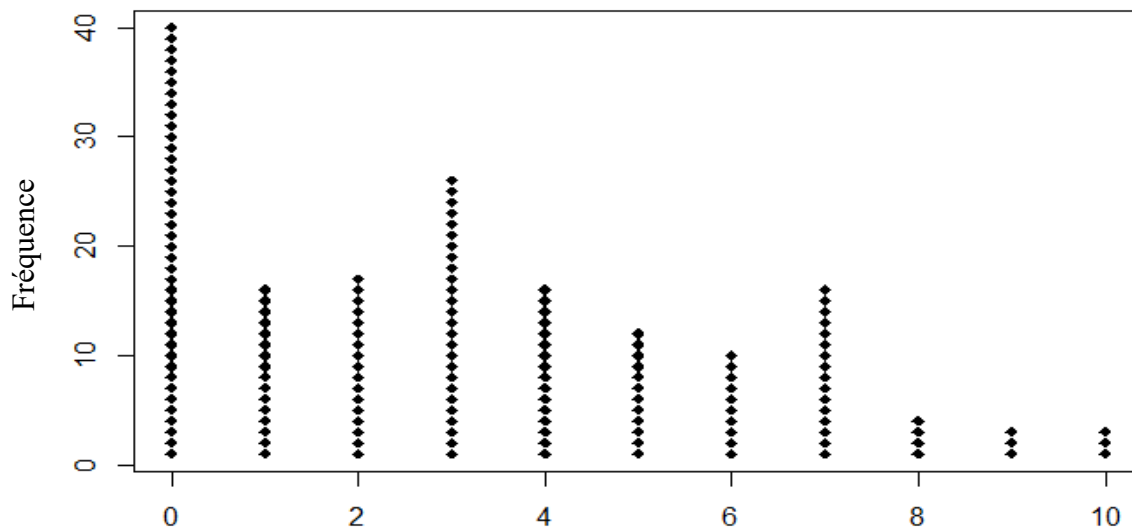
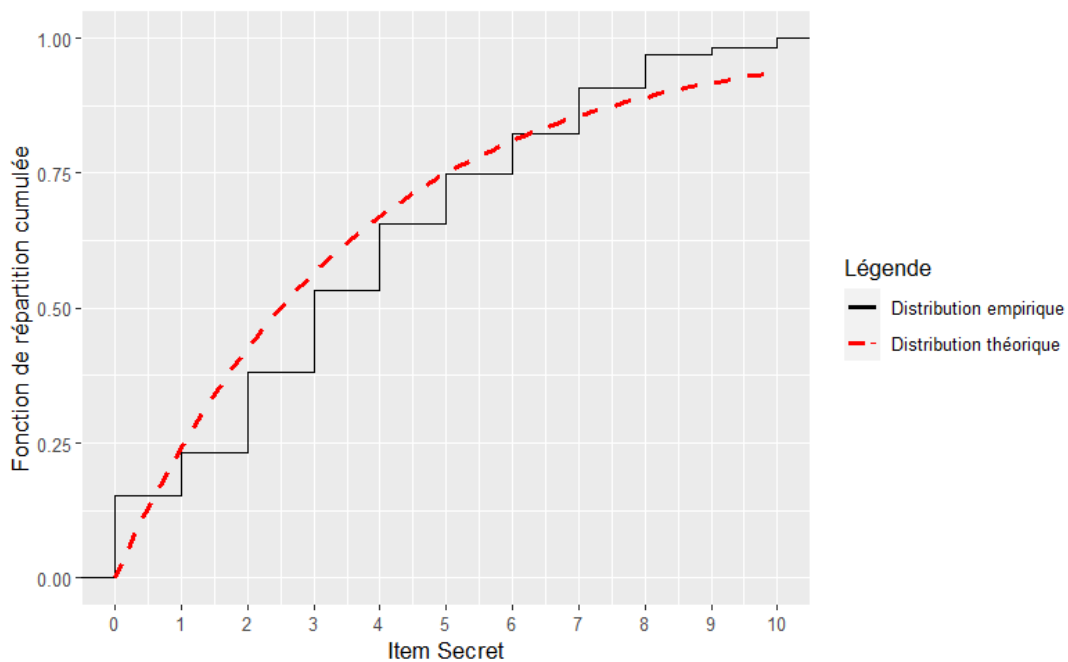
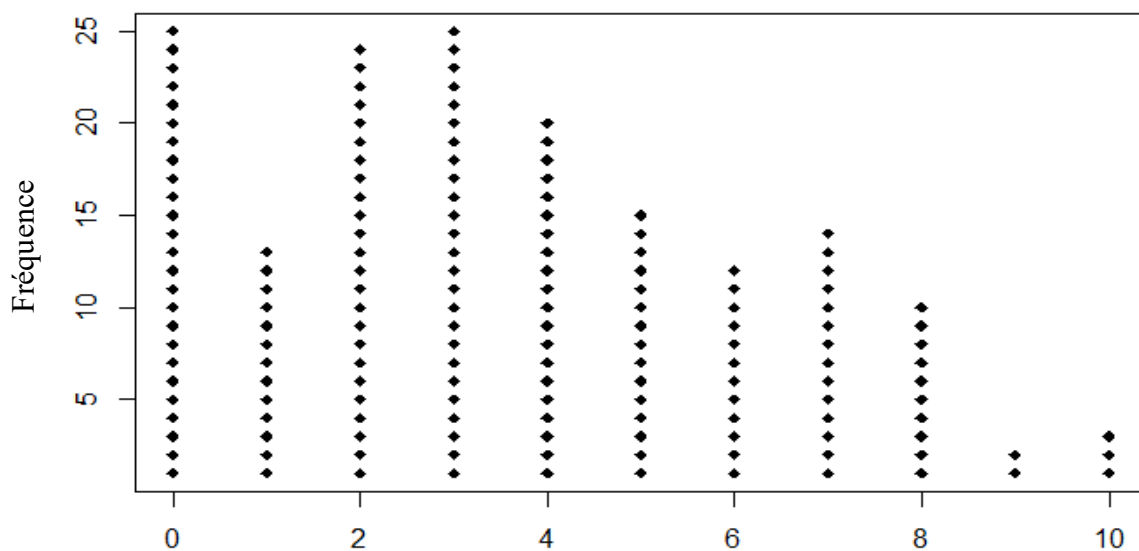
Figure 39*Diagramme à points de la distribution de l'item « liberté »***Figure 40***Graphique de la distribution de l'item « secret »*

Figure 41

Diagramme à points de la distribution de l'item « secret »



Dans ce groupe, on remarque dès la modalité de réponse numéro 2 que les courbes s'affaissent et se retrouvent en-dessous de la distribution exponentielle, signe que ces éléments n'appartiennent pas au noyau central de la représentation. Voici un tableau récapitulatif avec les résultats ainsi que les données ayant mené à celui-ci (voir Tableau 35).

Tableau 35

Résumé des scores de chaque item du Test d'indépendance au contexte

Variable	< 0,05	Position par rapport à la courbe exponentielle	Coefficient d'asymétrie	Oméga ² max	Conclusion
Donnée personnelle	Oui	Au-dessus	1,73	1,45	Fait partie du noyau
Confidentialité	Oui	Au-dessus	1,53	1,49	Fait partie du noyau
Intimité	Oui	Au-dessus	1,49	2,32	Fait partie du noyau
Protection	Oui	Au-dessus	1,09	1,66	Fait partie du noyau
Respect	Oui	Au-dessus	1,02	1,89	Fait partie du noyau
Sécurité	Oui	Au-dessus	0,97	1,46	Fait partie du noyau
Droit	Oui	Au-dessus	0,83	1,72	Fait partie du noyau
Secret	Oui	En-dessous	0,40	3,35	Hors noyau
Liberté	Oui	En-dessous	0,56	1,79	Hors noyau
Choix	Oui	En-dessous	0,41	3,09	Hors noyau
Anonymat	Oui	En-dessous	0,31	3,42	Hors noyau

Relation entre les items de la représentation avec l'analyse de similitude

L'analyse de similitude consiste à représenter visuellement et de manière planaire la matrice de similitude construite à partir des données du *Questionnaire de caractérisation*. Rappelons que cette matrice fut construite via la mesure d'association ordinaire du Tau de Kendall (Kendall, 1938).

Il est important de rappeler qu'il y avait des mots candidats autres que ceux candidats au noyau central, entre autres ceux issus de la zone mixte ou de la première et de la seconde périphérie. À la lecture du corrélogramme, plusieurs aspects émergent. Le premier est que

les corrélations s'étendent de -0,28 à 0,46 (voir Figure 42). On remarque aussi que les corrélations sont principalement négatives, ce qui indique que les items unis par ce type de corrélation sont classés de manières opposées. Par exemple, il est possible de présumer qu'une grande partie de notre échantillon ait choisi de classer l'item « sécurité » comme caractéristique de la vie privée en contexte technologique et en même temps, l'item « cookies » comme non caractéristique de la vie privée en contexte technologique. Deuxièmement, les mots ayant les corrélations les plus fortes entre eux sont des mots qui n'appartiennent pas au noyau central de la représentation, comme « GAFAM », « réseaux sociaux » et « cookies ». D'autres mots n'ont pas de lien évident avec la technologie, mais plus avec la vie privée de manière générale comme : « famille », « ami » et « relation ». De plus, les éléments ayant les plus hautes corrélations positives, et donc classées dans les mêmes catégories, sont majoritairement des items non caractéristiques et qui n'appartiennent pas au noyau central.

Nous proposons deux arbres maximums pour illustrer les résultats, le premier élaboré avec les valeurs algébriques des corrélations, le second avec les valeurs absolues des corrélations (voir Figures 43 et 44). Dans le second cas, nous avons rajouté a posteriori les signes des corrélations afin que le lecteur puisse avoir le sens de la relation.

Figure 42

Matrice de similitude de l'analyse de similitude comprenant les 21 items et leurs Tau C de Kendall

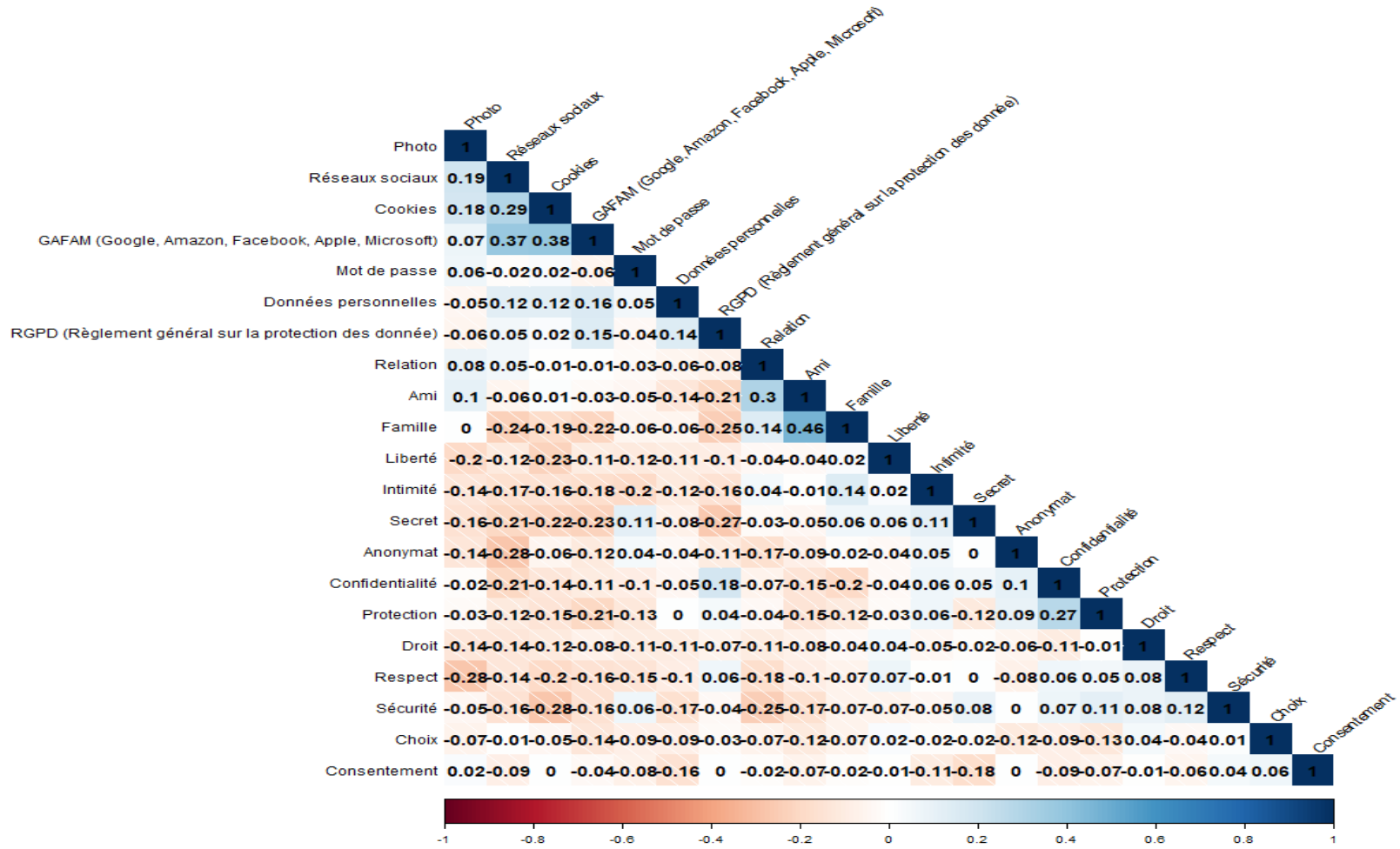


Figure 43

Arbre maximum avec les valeurs algébriques

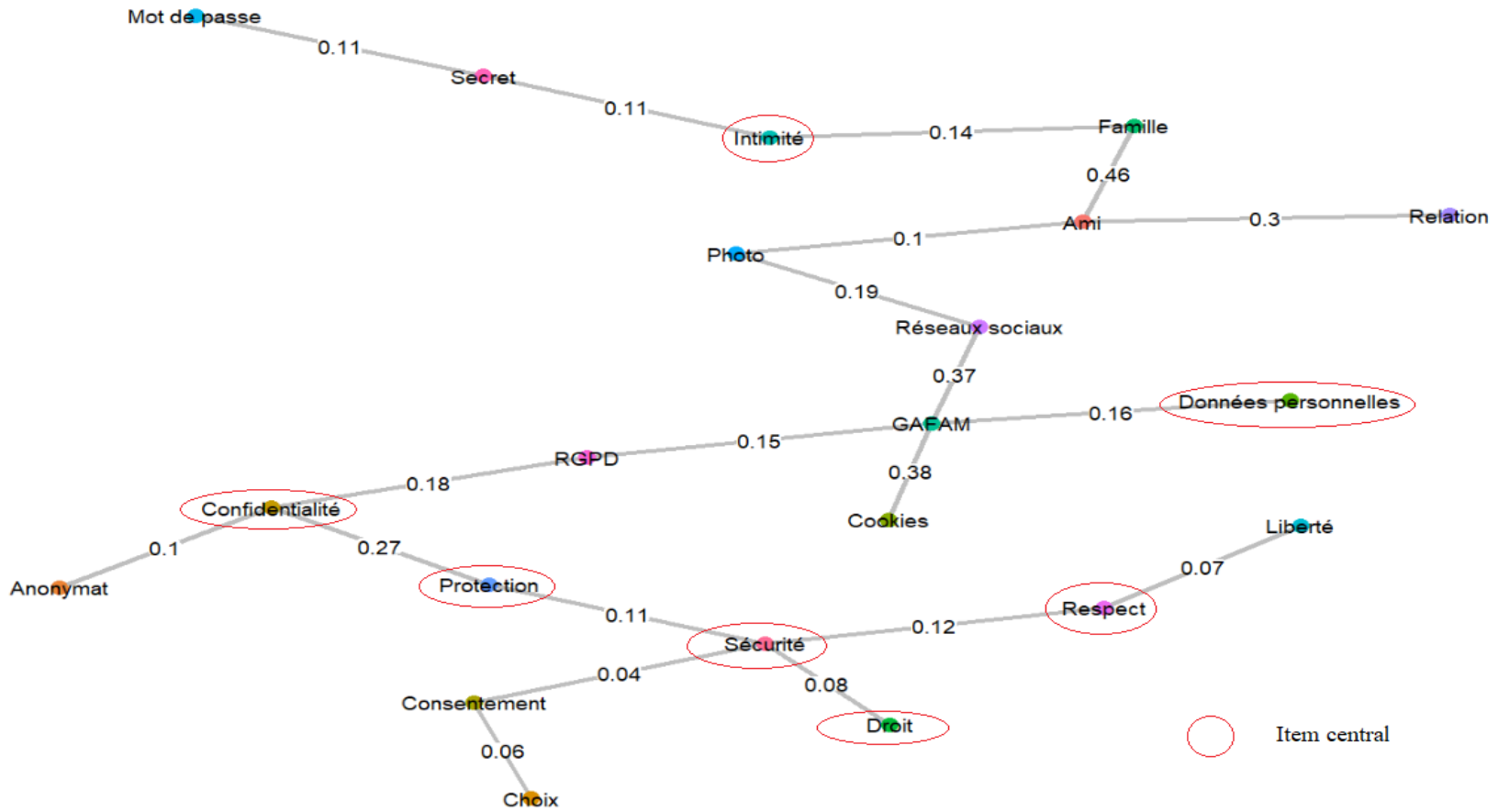
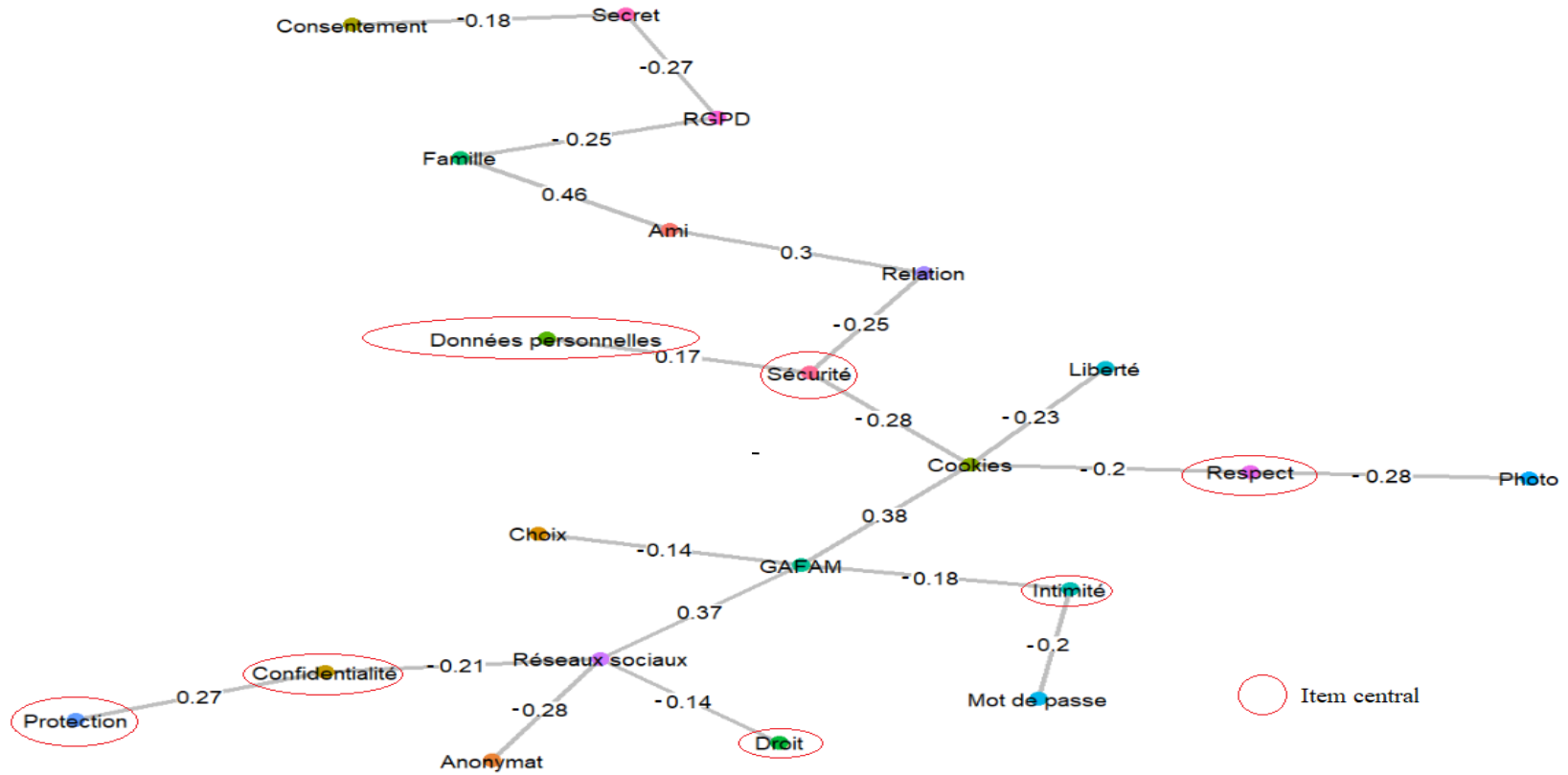


Figure 44

Arbre maximum avec les valeurs absolues



L'arbre maximum issu de la matrice permet de représenter les 21 items sélectionnés (voir Figure 43). Pour l'approche structuraliste, il s'agit d'obtenir l'armature de la matrice de similitude en conservant les similitudes les plus fortes entre les items à travers le graphe le plus simple possible. L'arbre maximum est constitué de 20 arêtes dont les valeurs varient entre 0,46 pour l'arête famille-ami et 0,06 pour l'arête anonymat-confidentialité. Ainsi, cet arbre qui ne prend en considération que la valeur algébrique, met au premier plan les items classés dans les mêmes catégories : soit le plus caractéristique de la vie privée, soit ni plus ni moins caractéristique de la vie privée, ou encore le moins caractéristique de la vie privée.

On remarque que les items ayant le plus de connexité (de voisins) sont « GAFAM » et « sécurité » avec tous deux quatre voisins directs. On constate que les items considérés comme centraux par notre analyse précédente sont divisés par des éléments n'appartenant pas au noyau. Un premier bloc en bas du graphe est constitué par certains items du noyau comme « confidentialité », « protection », « sécurité », « droit », « respect » et d'autres items hors noyau comme « consentement », « anonymat », « choix », « liberté ». Le thème du bloc gravite autour de la protection de la vie privée.

Au centre, un autre thème apparaît et est constitué du bloc « RGPD », « GAFAM », « cookies », « réseaux sociaux » et « donnée personnelle ». L'item « donnée personnelle » appartenant au noyau central se retrouve isolé.

Plus haut, le terme « photo » se retrouve à l'articulation entre deux blocs, celui des enjeux technologiques et celui de l'intimité, constitué par les items « famille », « ami », « relation » et « intimité ». Enfin, en haut à gauche se retrouvent les items « mot de passe » et « secret » qui ne dégagent pas réellement de thème.

Il est intéressant de noter que les items présentant les plus fortes corrélations sont les items ayant été classifiés le plus souvent comme n'étant pas caractéristiques de la vie privée. À la lumière des résultats de la précédente analyse, les items possédant les plus hautes corrélations ne font pas partie du noyau central.

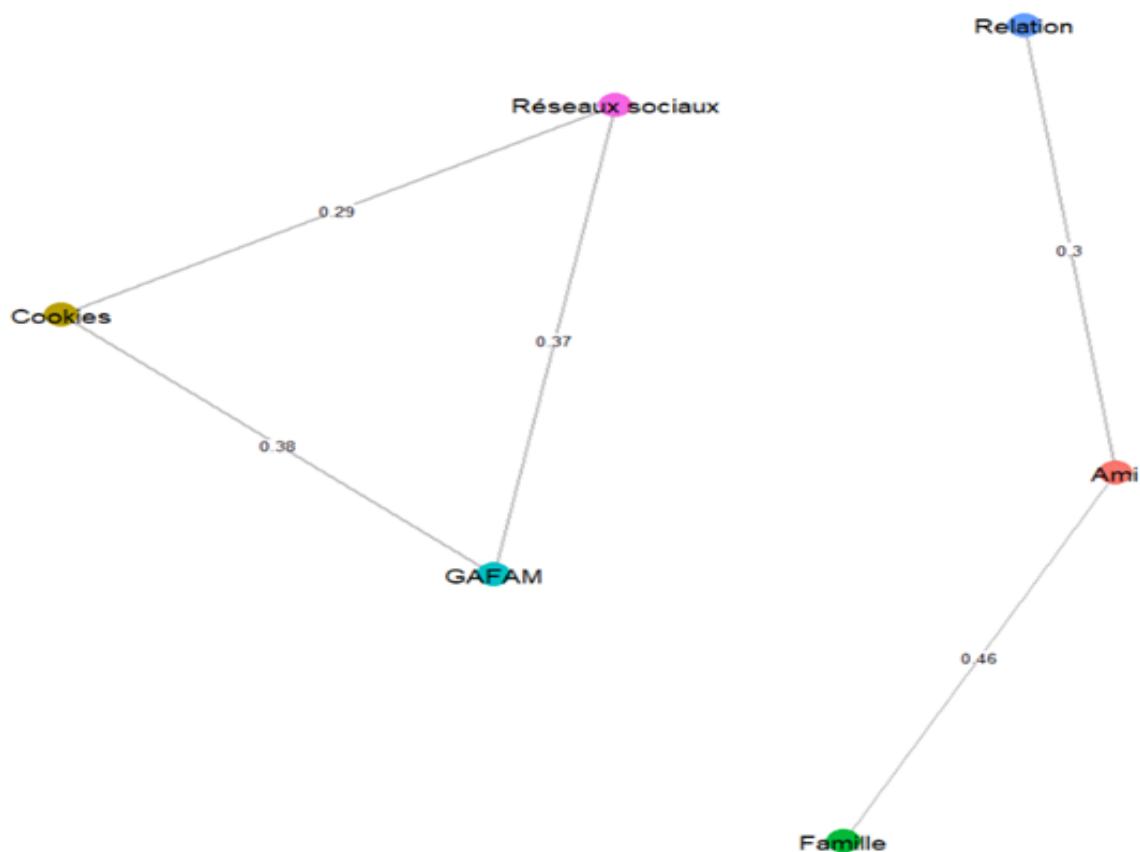
Le prochain graphe est l'arbre des maximums issu de la matrice des similitudes des valeurs absolues (voir Figure 44). Cet arbre fait apparaître les items qui sont classés de la même façon et non pas dans la même catégorie. Ainsi, on peut voir quel type de lien entretiennent les items, s'il s'agit d'opposition ou de similitude. L'arbre qui se dégage de cette manière de procéder présente des caractéristiques différentes. L'arbre maximum est constitué de 20 arêtes dont les valeurs varient entre $|0,46|$ pour l'arête famille-ami et $|0,14|$ pour les arêtes choix-GAFAM et réseaux sociaux-droit.

Comme précédemment, l'item « GAFAM » a le plus de connexité ainsi que l'item « cookies », avec tous deux quatre voisins directs. La division précédemment constatée est encore plus flagrante, car tous les items considérés comme centraux se retrouvent

isolés. Deux blocs restent cependant inchangés, le bloc « GAFAM », « cookies », « réseaux sociaux » et celui « famille », « relation », « ami ».

À la suite des arbres, nous proposons de présenter les graphes seuils à travers quatre seuils différents (Tau C de Kendall) : supérieur à 0,27, supérieur à 0,18, supérieur à 0,11, supérieur à 0,08. Rappelons qu'un graphe seuil est un graphe qui fait apparaître l'ensemble des arêtes qui dépassent un certain seuil. Attendu que les graphes seuils sont descriptifs, nous avons choisi les précédents seuils, car ils permettent au lecteur de voir à partir de quels moments (seuils) les éléments commencent à s'agglomérer entre-deux. Ces graphes avec des seuils spécifiques permettent aussi de comprendre comment les éléments classés dans la même classe s'agglomèrent pour constituer la représentation sociale de la vie privée en contexte technologique.

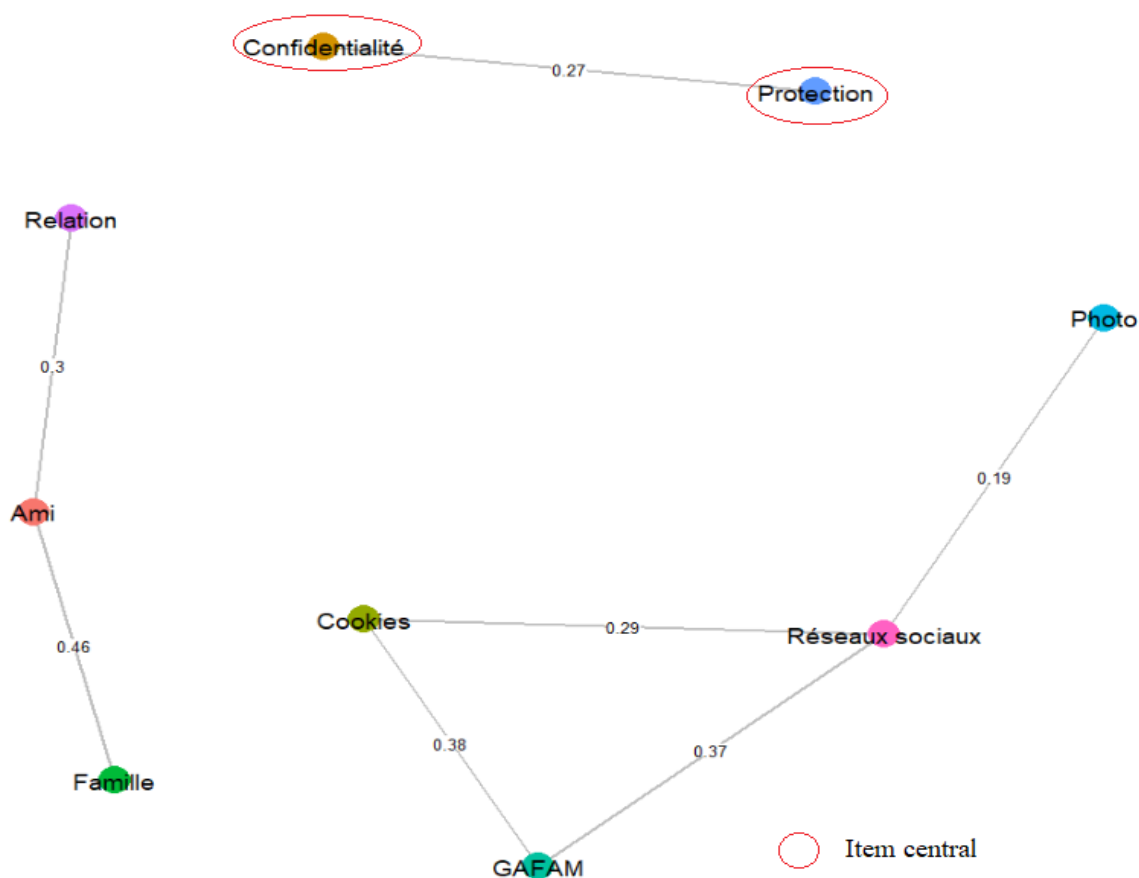
Le graphe seuil à 0,27 met en évidence deux dimensions, la première portant sur Internet avec les réseaux sociaux, les GAFAM et les cookies et la seconde portant sur une sphère plus intime avec la famille, les amis, les relations (voir Figure 45). Il est intéressant de constater que la totalité de ces items a été majoritairement classée comme n'étant pas caractéristique de la vie privée en contexte technologique.

Figure 45*Grappe seuil à 0,27*

Dans le graphe seuil à 0,18, les deux dimensions précédemment décrites restent stables, si ce n'est l'ajout de l'item « photo » à l'ensemble portant sur Internet (voir Figure 46). Les premiers éléments appartenant au noyau central apparaissent : il s'agit de « confidentialité » et de « protection ».

Figure 46

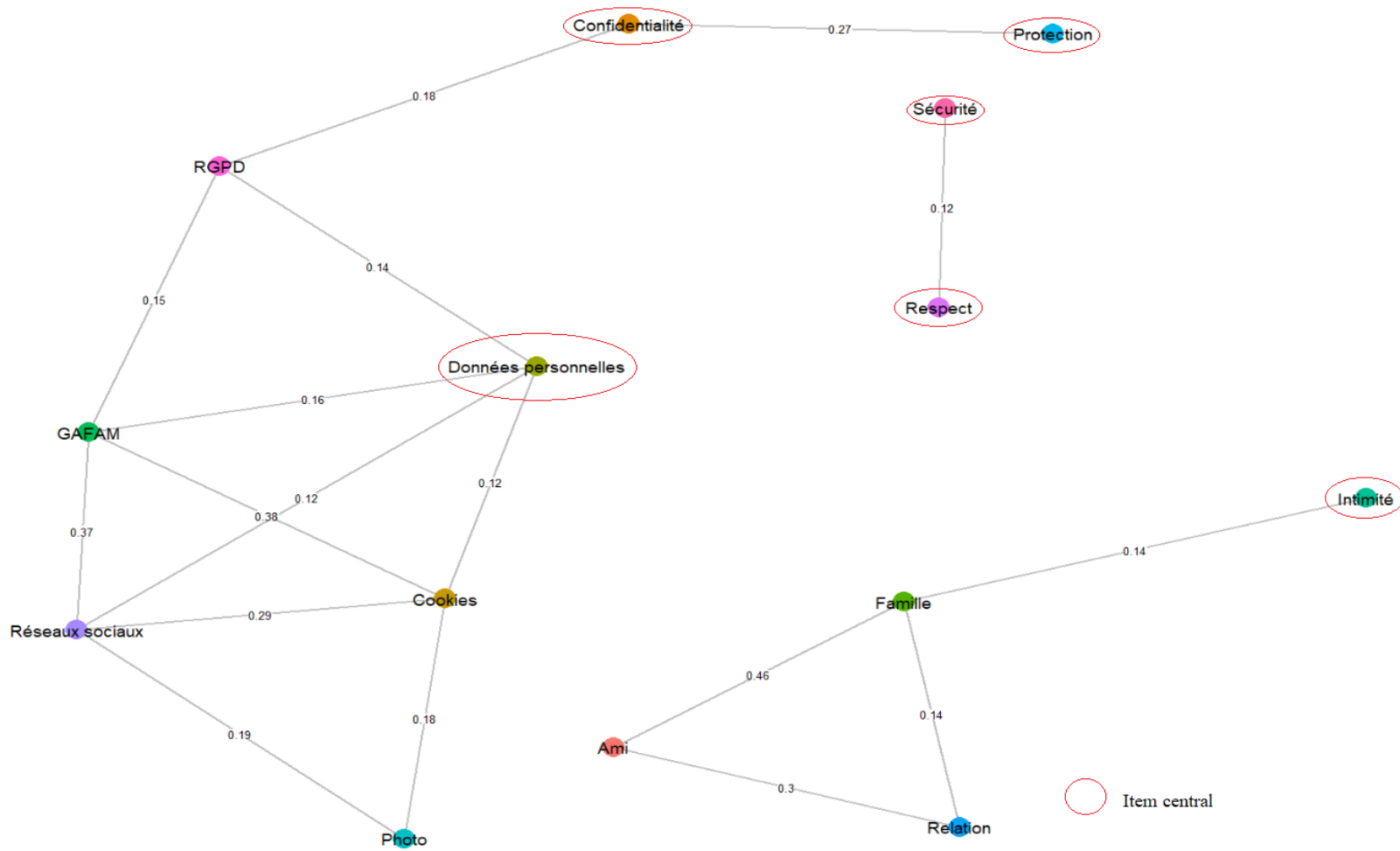
Graphe seuil à 0,18



Dans le graphe seuil à 0,11, le bloc concernant Internet s'agrandit et intègre les éléments du noyau central précédent (« confidentialité » et « protection »), mais aussi « donnée personnelle » (voir Figure 47). Cet amalgame est rendu possible grâce à l'item « RGPD ». On remarque deux thèmes à l'intérieur de ce bloc, celui de la protection des données d'un côté et des menaces pour la vie privée de l'autre. Le bloc renvoyant à la vie privée de façon générale est toujours présent et s'enrichit de l'item « intimité » qui fait partie du noyau central. Enfin, un nouveau groupement apparaît, constitué de « sécurité » et de « respect ».

Figure 47

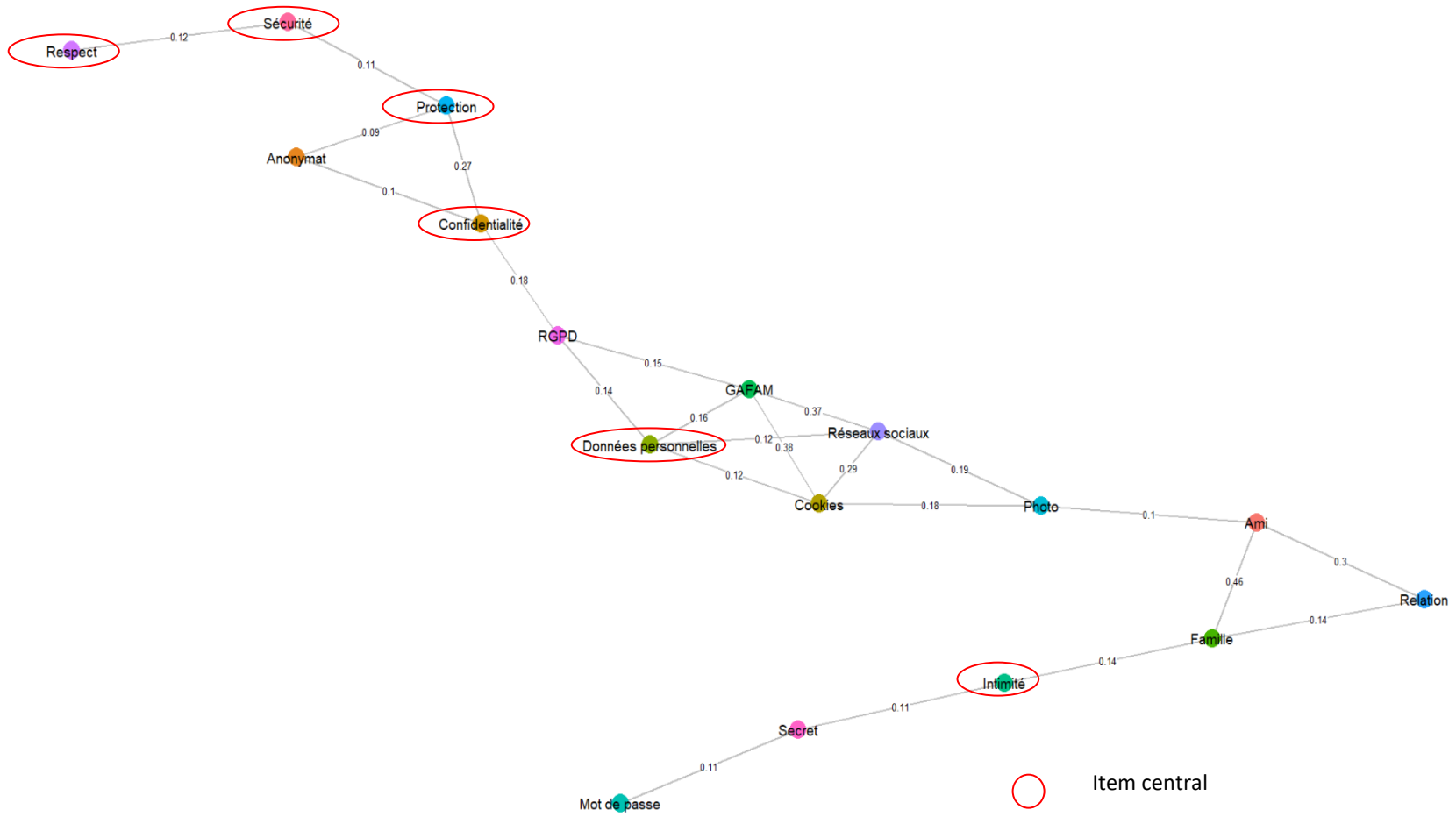
Grappe seuil à 0,11



Dans le graphe seuil à 0,08 l'ensemble des 21 termes ne sont pas encore apparus sur celui-ci (voir Figure 48). Cependant, c'est avec ce seuil de 0,08 que l'on observe les trois blocs précédents se regrouper sous la forme d'un graphe connexe. Les trois thèmes précédents sont toujours présents. Le bloc sur les risques envers la vie privée est enchâssé entre la protection des données personnelles d'un côté et la sphère intime de l'autre. Le bloc concernant la protection des données s'est aggloméré avec la partie « sécurité » et « respect ». Il s'enrichit aussi de l'item « anonymat ». Le bloc portant sur Internet n'a pas changé si ce n'est les nouvelles connexions à travers le duo « photo » et « ami ». Enfin, la sphère de l'intimité s'est vue agrandie par l'item « secret ».

Figure 48

Graphe seuil à 0,08



Discussion

La discussion abordera les résultats selon la séquence suivante. En premier lieu, nous discuterons des résultats au vu des objectifs, questions et hypothèses de recherches proposés dans la partie problématique ainsi qu’au regard des recherches actuelles. Dans un second temps, nous aborderons les applications pratiques que peuvent avoir nos résultats ainsi que les orientations futures que pourraient prendre les recherches. Enfin, nous évoquerons les forces et les limites de cette recherche.

Le premier résultat de notre thèse est la structure de la représentation sociale de la vie privée en contexte technologique. Nous avons choisi de commencer par l’interprétation de la structure par la discussion des résultats du système périphérique¹. En effet, les résultats portant sur le noyau central doivent être discutés à la lumière des deux phases de l’étude alors que ceux des zones périphériques ne peuvent l’être qu’à la lumière de la première, celle-ci n’ayant pas de second temps de validation. Nous entamerons donc l’interprétation des résultats par une analyse qualitative et exploratoire du système périphérique afin d’avoir une première idée de la représentation qui permettra de discuter les résultats avec les données actuelles de la recherche.

¹ Le système périphérique est constitué de la 1^{re} et 2^e périphérie ainsi que de la zone mixte.

Un des thèmes saillants de la périphérie est la technologie. En effet, rappelons que ce thème représente 21,26 % des différents items dans la périphérie, soit 21,93 % des items totaux générés par les participants dans la périphérie. Outre la prépondérance du thème de la technologie, nous pensons que celui-ci occupe ce que Abric (1994) nomme la « fonction de concrétisation ». Celle-ci permet à la représentation d'être facilement appréhendable et partageable par les différents individus qui sont confrontés à l'objet de la représentation, dans notre cas la vie privée en contexte technologique.

Selon l'auteur, cette fonction est contextuelle, c'est-à-dire qu'elle est relative à l'environnement dans lequel évolue le groupe social qui possède la représentation étudiée. Cette dimension contextuelle permet l'ancrage dans le réel de la représentation en s'adaptant au changement d'environnement. Dans notre étude, il semble que les éléments périphériques ayant pour thème la technologie corroborent la présence de cette fonction de concrétisation. Nous pensons notamment que des termes comme « Facebook », « GAFAM », « Internet », « téléphone », « données sensibles », « localisation », et « historique » donnent une idée tangible du contexte dans lequel prend place la représentation et illustrent par conséquent cette fonction.

Ces termes périphériques pourraient aussi suggérer que les participants ne sont pas confrontés aux mêmes situations technologiques (p. ex., utilisation de l'ordinateur c. utilisation de la tablette) et qu'elles ne seraient pas toutes associées à la vie privée. En effet, s'il y avait eu un consensus autour du thème de la technologie, la plupart des

candidats au noyau central auraient gravité autour du thème technologique, or ce n'est pas le cas. On peut donc se poser la question suivante : est-ce que les individus pensent à la vie privée uniquement lorsqu'ils sont en contact avec certaines technologies? Si tel était le cas, cela suggérerait que certains individus seraient plus sensibles aux enjeux concernant la vie privée lorsqu'ils surfent sur Facebook alors que d'autres le seraient uniquement lorsqu'ils initialiseront une opération de géolocalisation. Cependant, la technologie n'est pas l'unique thème qui pourrait se comporter de la sorte.

Tout aussi fréquent dans le système périphérique (14,17 % des différents items, soit 20,98 % des items totaux générés par les participants de la périphérie), le thème de l'intimité semble lui aussi illustrer la fonction de concrétisation. En effet, nous pensons que les termes comme « famille », « vie personnelle », « quotidien », « foyer » et « cercle restreint » permettent un ancrage concret du concept de vie privée. Tout comme le thème de la technologie, nous nous demandons si la diversité des termes périphériques concernant l'intimité ne pourrait pas être le signe d'une association entre certaines situations sociales et la vie privée. La pluralité des items et leur place dans la périphérie pourraient-elles faire en sorte qu'uniquement certaines relations sociales pourraient activer le lien avec la vie privée chez les individus? Par exemple, on peut facilement imaginer que l'activation des thèmes de la vie privée soit différente lors d'une interaction avec un membre de la famille par rapport à un inconnu ou un collègue de travail.

Autre thème important qui émerge de l'analyse prototypique : le thème du danger. Celui-ci contient des menaces concrètes comme la publicité ciblée, le piratage, la surveillance ou encore la violation, qui corroborent, par leur aspect concret, la fonction de concrétisation de la périphérie. Mais au-delà d'appuyer la fonction de concrétisation, le thème du danger vient éclairer d'une façon inattendue le paradoxe de la vie privée, car les résultats vont à l'encontre de celui-ci.

En effet, nos résultats suggèrent que les dangers envers la vie privée ne semblent pas être un enjeu préoccupant pour nos participants, et ce, pour plusieurs raisons. La première de ces raisons est le type de risque qu'évoquent les participants dans leurs représentations de la vie privée. Gerber, Reinheimer et Volkamer (2019) expliquent que selon le type de risque, les individus estiment différemment la probabilité d'être confrontés à un risque ainsi qu'aux préjudices qui en découlent. Par exemple, lorsque le risque est diffus (p. ex., des entreprises collectent et analysent vos données), celui-ci est perçu comme plus probable et moins préjudiciable qu'un risque qualifié de « spécifique » (p. ex., des entreprises collectent et analysent vos données et votre assureur pourrait les obtenir et revoir votre prime). Dans notre cas, nos résultats corroborent les perceptions des participants de l'étude de Gerber, Reinheimer et Volkamer (2019), puisqu'on note que sur les 13 risques répertoriés¹, un seul peut être considéré comme spécifique, soit « publicité ciblée ». On peut faire donc l'hypothèse que l'échantillon entretient une conception du

¹ Insécurité, violation, risque, voyeurisme, problème, espionnage, piratage, intrusion, surveillance, Big Brother, danger, cookies, publicité ciblée.

risque majoritairement diffuse; ce qui impliquerait une idée du risque encouru moins grave, mais plus probable pour la vie privée.

La seconde raison vient préciser les probabilités envisagées par les participants concernant le danger autour de la vie privée. En effet, les faibles fréquences des items rattachés au danger (4,29 % des items totaux générés) suggèrent que les risques restent une préoccupation peu fréquente des participants, d'autant plus que les items sur le risque sont localisés exclusivement en périphérie. Cela suggère une absence de consensus entre les participants et par conséquent, qu'ils ne considèrent pas les risques concernant la vie privée en contexte technologique comme essentiels à la définition de celle-ci.

Enfin, ces résultats nous invitent à nous questionner sur les méthodes utilisées pour recueillir nos données. Manifestement, lorsqu'on utilise une association libre, le risque apparaît de façon très relative (peu de mots, localisation périphérique et type de risque énoncé faiblement préjudiciable) et cela nous interroge sur la présence réelle des préoccupations chez les participants concernant la vie privée. Il semble plus approprié de considérer cette vision relative qu'ont les participants du risque comme un faible niveau de préoccupation. Ces faibles préoccupations pourraient expliquer le manque de protection de la vie privée qu'observent plusieurs auteurs et qu'ils décrivent comme « paradoxale » (Brown, 2001; Hargittai et Marwick, 2016; Lutz et Strathoff, 2014; Staddon, Acquisti et LeFevre, 2013).

Contrairement aux auteurs qui défendent l'existence du paradoxe de la vie privée, nos résultats soutiennent l'idée d'un comportement plus plausible que paradoxal. En effet, selon nos résultats, on pourrait se poser les questions suivantes : pourquoi un individu se préoccuperait-il des dangers concernant sa vie privée si celui-ci les perçoit comme très restreints et peu dommageables? De plus, pourquoi mettre en place des moyens de protection au vu de ces mêmes dangers qui paraissent fort diffus? En caricaturant la situation, ce serait comme trouver paradoxal que les individus ne mettent pas une tenue d'apiculteur afin d'éviter les piqûres d'insectes lorsqu'ils sortent dans leur jardin.

Une explication additionnelle en lien avec des mécanismes macrosociaux peut être évoquée, notamment le contexte médiatique dans lequel évoluent les participants. Plus précisément, il est possible que la presse, à travers un certain cadrage médiatique, favorise l'apparition ou le maintien de certains éléments de la représentation. À ce sujet, la recherche de Mols et Janssen (2017) portant sur la presse des Pays-Bas relèvent six discours sur la vie privée à la suite de l'affaire Snowden. Le cadrage médiatique qui met de l'avant l'idée que la personne n'a « rien à cacher » (7 % des corpus analysés) et donc ne risque rien, pourrait amener les personnes à minimiser certains dangers pour la vie privée. De même, le cadrage « la fin justifie les moyens » (18 % des corpus analysés) pourrait expliquer que l'individu légitimise et minimise certains dangers comme la surveillance des citoyens par l'État. Une possibilité pour valider cette piste d'explication serait d'utiliser le paradigme sociogénétique des représentations sociales et d'effectuer des entretiens semi-directifs avec les participants.

Au vu de nos résultats et interprétations, nous souscrivons à la perspective de Dienlin et Trepte (2015) qui proposaient que le paradoxe de la vie privée soit un artefact méthodologique. Cependant et contrairement à eux, nous pensons que cet artefact n'est pas uniquement attribuable à des modalités de réponse restreintes qui impliquaient un manque de puissance statistique, mais plutôt qu'il est attribuable à une question de type de collecte qui ne laisse pas spontanément le participant exprimer sa vision du phénomène. Cela nous questionne sur la qualité des réponses obtenues. Premièrement, nous pensons que celles-ci sont tributaires des questions et préconceptions du chercheur. Ainsi, certains chercheurs demanderont si les participants ont bien paramétré leur compte Facebook alors que d'autres demanderont s'ils naviguent sur Tor (Dingledine et Mathewson, 2002) et communiquent avec Signal (Signal Messenger, 2018). Le participant répond donc aux risques que les chercheurs ont estimés importants. Deuxièmement, proposer des préoccupations « clé en mains » pour interroger des participants pourrait créer un effet d'amorce qui les conditionne à imaginer leurs réponses à des risques auxquels ils ne penseraient pas spontanément. La vraie question est donc de savoir quelle préoccupation concernant la vie privée anime réellement les individus en contexte technologique, voire même si ces préoccupations existent. Une possibilité serait d'observer le phénomène à travers le paradigme sociodynamique des représentations sociales.

Si avec le thème du danger nous avons pu explorer une facette du paradoxe de la vie privée, le thème portant sur la protection permet d'apporter un autre éclairage sur le phénomène. Comme précédemment, les items périphériques constituant ce thème

semblent corroborer la fonction de concrétisation. Nous pensons notamment que celle-ci s'exprime à travers la citation de moyens institutionnels de protection comme la CNIL, ou légaux comme le RGPD, le consentement ou encore l'item « règlement ». À notre connaissance, nous n'avons pas recensé de questionnaire interrogeant l'utilisation et la connaissance des moyens de protection légaux à disposition du citoyen pour protéger sa vie privée. Or, si les moyens concrets de protection qui viennent à l'esprit des participants sont des moyens légaux ou institutionnels et que ceux-ci ne semblent pas être interrogés par les recherches actuelles, comment la recherche peut-elle statuer sur l'absence de comportement de protection? Ne faut-il pas plutôt y voir un mécanisme supplémentaire qui alimenterait le paradoxe de la vie privée en diminuant les moyens de protection que mettent en place les individus et par conséquent, qui soutiendrait l'hypothèse de l'artefact méthodologique?

Précisément, la plupart des études actuelles semblent n'interroger que des comportements portant sur des paramétrages techniques que les participants mettent en place (Barth, de Jong, Junger, Hartel et Roppelt, 2019; Dienlin et Trepte, 2015; Young et Quan-Haase, 2013).

Selon nous, cet angle mort de la recherche est dû à deux raisons. La première est une question de temporalité, la seconde une question de géographie. En effet, le RGPD n'est appliqué que depuis 2018 et ne concerne que les pays de l'Union européenne ou les entreprises ayant une succursale sur le continent ou commerçant avec des habitants d'un

des pays de l'Union européenne. Nous pensons qu'il s'agit donc d'une question de temps avant que les auteurs d'autres pays ne s'intéressent à l'aspect légal de la protection de la vie privée dans les recherches sur le paradoxe de la vie privée. D'ores et déjà, le continent nord-américain emboîte le pas à l'Union européenne et se dote de loi semblable au RGPD, comme le Canada et sa loi sur la protection des renseignements personnels et les documents électroniques (Deschênes et Archambault, 2019) ou encore l'état de Californie et sa loi sur la protection de la vie privée des consommateurs (Darnault, 2021). Pour autant, les aspects temporels et géographiques n'expliquent pas tout. Le design de recherche nous semble aussi important pour expliquer la faible référence à la dimension légale et institutionnelle dans les recherches sur le paradoxe de la vie privée.

En effet, les réponses portant sur les institutions et les mécanismes légaux n'indiquent pas uniquement les manières de se protéger existantes dans la représentation sociale des individus, elles indiquent aussi leurs dimensions. Dans notre cas, de telles réponses mettent en exergue une dimension qui dépasse le cadre individuel. Précisément, nous pensons que ces résultats révèlent l'aspect macrosocial que revêt la protection des données dans la représentation sociale des individus, car les lois et les institutions exercent leurs pouvoirs sur l'ensemble de la population. Or, nous pensons que trop peu de recherches sont conçues pour capter la partie macrosociale du phénomène. En effet, les revues de la littérature exposent la prépondérance des modèles cognitifs expliquant le paradoxe de la vie privée (Bélanger et Crossler, 2011; Gerber, Gerber et Volkamer, 2018; Kokolakis, 2017; Li, 2012) et ces modèles relèvent du paradigme de l'individualisme méthodologique

(Valade, 2012). Celui-ci nous semble peu adapté à recueillir les mécanismes macrosociaux qui traversent pourtant la représentation de la vie privée en contexte technologique. En effet, le questionnement ouvert que propose l'association libre permet de recueillir tout type de réponse, y compris celles auxquelles le chercheur ne s'attend pas. Mais l'analyse prototypique permet d'aller plus loin. Premièrement, elle permet d'avoir accès aux éléments consensuels de la représentation. Deuxièmement, certaines de ces zones proposent des pistes de discussion intéressantes comme la zone des éléments contrastés (faiblement cités, mais jugés très importants). En effet, selon Abric (2003), ces éléments pourraient être une partie d'un noyau central appartenant à des sous-groupes et qui aurait une représentation différente. Cependant, il ajoute que ces éléments pourraient être tout aussi bien complémentaires aux zones périphériques. Dans le cas où il s'avérait que les éléments contrastés soient bien le noyau central d'un sous-groupe, l'analyse prototypique pourrait, en plus de renseigner sur les consensus qui traversent les groupes, documenter les dynamiques internes des groupe étudiés en indiquant l'existence de sous-groupes. Cependant, à ce jour, aucune méthodologie n'a été initiée pour départager la nature de ces éléments (un noyau d'un sous-groupe ou un complément aux zones périphériques).

En résumé, la discussion des résultats de la zone périphérique a permis d'observer que les participants semblaient être confrontés à des contextes technologiques différents (l'ordinateur, le téléphone intelligent, etc.), mais aussi des situations intimes différentes (la famille, les amis, etc.). Aussi, la discussion des thèmes observés, tels que le danger et

la protection, nous a permis de questionner le paradoxe de la vie privée, notamment d'alimenter la thèse du paradoxe de la vie privée comme artefact méthodologique en montrant que la présence de dangers sérieux dans la représentation était questionnable et que les aspects de la protection étaient partiellement interrogés dans les recherches antérieures. Cependant, la réponse à la question de recherche sur la structure de la représentation sociale de la vie privée en contexte technologique reste inachevée sans la discussion des résultats des termes candidats au noyau central ainsi que leur validation.

Les items candidats au noyau central se rattachent à trois thèmes principaux : la technologie, la protection et l'intimité. Les items candidats au noyau central ayant comme thème la technologie ne sont pas aussi fréquents que dans la périphérie. En effet, dans la périphérie, ce thème représente presque 10 % des items totaux générés par les participants de la représentation alors que chez les candidats au noyau central, il ne représente qu'environ 6 % des items totaux générés. Concernant le thème de l'intimité et de la protection, le pourcentage de citations augmente substantiellement, doublant pour le thème de l'intimité et septuplant pour le thème de la protection. Le fait que ces items soient considérés comme importants dans le noyau central et que leur fréquence augmente de manière substantielle dans cette partie de la représentation peut être interprété comme la manifestation de l'aspect consensuel propre au noyau central. Cependant, pour finaliser l'analyse de la structure de la représentation, il est nécessaire de discuter des éléments confirmés comme centraux par le *Test d'indépendance au contexte*.

Dans les 11 mots testés¹, seuls 7 font réellement partie du noyau central : « confidentialité », « donnée personnelle », « droit », « intimité », « protection », « respect » et « sécurité ». On remarque que la présence des trois grands thèmes, à savoir la technologie, l'intimité et la protection, est confirmée dans la structure de la représentation. Aussi, il est intéressant de mentionner que la nature des items centraux est en accord avec la théorie. En effet, comme le rappelle Moliner (2016), les éléments centraux ont tendance à être abstraits, définitionnels de la représentation et transcontextuel alors que les éléments périphériques relèvent d'expériences singulières et contextuelles. Ces caractéristiques sont bien retrouvées dans les sept items centraux.

Toutefois, ces résultats ne permettent pas uniquement de confirmer des points théoriques ou d'étayer les pistes exploratoires précédemment abordées : elles permettent de statuer sur les hypothèses concernant la structure de la représentation. Il apparaît que le thème de la protection se retrouve de façon prépondérante dans les éléments centraux. Factuellement, plus de la moitié des sept items centraux sont rattachés au thème de la protection, à savoir : « confidentialité », « droit », « protection » et « sécurité ».

La confirmation de nos deux hypothèses de recherche, à savoir que le noyau central comportera des items liés au thème de la protection et que les items centraux ayant un lien avec le thème de la protection seront majoritaires, continue d'alimenter nos interrogations

¹ Pour rappel, les mots testés étaient : « anonymat », « choix », « confidentialité », « donnée personnelle », « droit », « intimité », « liberté », « protection », « respect », « secret » et « sécurité ».

concernant le paradoxe de la vie privée. En effet, nous émettons l'hypothèse que la proportion importante d'items gravitant autour du champ lexical de la protection suggère une représentation sociale de la vie privée en contexte technologique intrinsèquement sécuritaire, du moins pour l'échantillon d'universitaires français inclus dans l'étude. Dans ce contexte, comment est-il possible de conclure à un paradoxe dans le comportement des utilisateurs? En effet, l'idée de protection est majoritaire dans le noyau central. De plus, les travaux faisant le lien entre engagement et représentation confirment qu'une intention se manifeste plus souvent ou qu'une action est entreprise plus fréquemment dès lors qu'elle est activée par un item central (Eyssartier, Joule et Guimelli, 2007; Zbinden *et al.*, 2011). Or, sans perception de danger dans le noyau central, quelle motivation peut-il y avoir à contrer un possible danger?

Le dévoilement des différents liens entre items grâce à l'analyse de similitude¹ apporte des nuances supplémentaires, tout en répondant à la dernière question de recherche sur les liens qu'entretiennent les différents éléments. Une des premières nuances est que dans l'arbre maximum en valeurs algébriques (voir Figure 43), ce sont les mots périphériques qui sont associés le plus souvent dans les mêmes catégories, en l'occurrence celle nommée « non caractéristique ». Les items suivants se retrouvent dans cet arbre avec une corrélation qui est égale ou supérieure à 0,30 : « GAFAM », « cookies », « réseaux sociaux », « famille », « ami » et « relation ». Cela signifie que les participants ont eu plus

¹ Si le lecteur souhaite se remémorer en profondeur, la théorie des graphes et l'analyse de similitude, nous le renvoyons à la Figure 14.

de facilité à associer ces mots (périphériques), en l'occurrence dans la catégorie « non caractéristique », que les éléments centraux. De façon globale, l'arbre suggère qu'il est donc plus aisé pour les participants de créer du lien entre des notions qui n'ont pas de rapport avec la vie privée en contexte technologique plutôt que des notions qui en auraient. Ceci revient également à considérer que le noyau central présente un sens diffus, lequel pourrait d'ailleurs contribuer à la faible mobilisation des participants envers la protection de la vie privée.

Pour étayer cette interprétation, revenons au graphe seuil¹ à 0,27 (voir Figure 45), car celui-ci fait apparaître la totalité des corrélations pour ce seuil. Là encore, ce sont les relations entre les éléments périphériques qui portent sur la technologie et les relations proches qui apparaissent en premier.

Enfin, le second arbre maximum qui est calculé avec des valeurs absolues fait apparaître (voir Figure 44) les liens qui sont le plus souvent classés de la même façon, mais pas forcément dans la même catégorie. En mettant de l'avant les valeurs absolues les plus fortes, cet arbre met en exergue les associations les plus fortes lorsque celles-ci sont positives et les oppositions les plus fortes quand celles-ci sont négatives. Comme précédemment observé, les trios réseaux sociaux-GAFAM-cookies et famille-ami-relation sont toujours associés. Cependant, ces trios sont en même temps isolés par un

¹ Pour rappel, un graphe seuil est un graphe qui fait apparaître toutes les corrélations égales ou supérieures à un seuil préalablement déterminé par le chercheur.

ensemble de corrélations négatives. Concernant le trio réseaux sociaux-GAFAM-cookies, le fait qu'il ne soit pas caractéristique de la vie privée en contexte technologique pourrait être considéré comme un signe que ces items font partie des dangers de la vie privée. Cette piste explicative devrait faire l'objet de travaux supplémentaires afin qu'elle puisse être validée ou infirmée. Concernant le trio famille-ami-relation, il est possible que l'ensemble des participants s'accordent sur le fait que ce trio soit considéré comme non caractéristique vis-à-vis du contexte technologique que la plupart des participants les classent de façon identique et qu'il en résulte de fortes corrélations entre les items. Enfin, l'arbre maximum en valeur absolue fait apparaître les items centraux isolés et dispersés.

Il est possible d'émettre l'hypothèse que cette « éclatement » des liens entre items centraux est dû à ce que Flament (1996) nomme des éléments centraux « excentriques ». Selon lui, l'excentricité d'un élément est la distance maximale qu'il peut exister entre un item donné et un autre. Ainsi, plus l'item serait aux limites du graphe, plus il serait susceptible d'être central, mais de façon « excentrique ». Or, cette hypothèse nous semble difficilement défendable, car elle va à l'encontre de la théorie de la représentation sociale et en particulier la caractéristique de connexité des éléments centraux (Moliner et Guimelli, 1994). De surcroît, les résultats du *Test d'indépendance au contexte* viennent invalider cette possibilité en mettant aussi à l'extrémité du graphe des éléments non centraux comme « anonymat », « liberté » ou encore « choix ».

Cependant, Flament (1994) a aussi proposé la notion d'autonomie qui nous semble mieux expliquer notre situation que celle des éléments centraux excentriques. Pour ce chercheur, il existe des représentations autonomes d'un côté, ayant un noyau central et une organisation claire. D'un autre côté, il existe des représentations non autonomes avec un faible niveau d'organisation et aucun noyau central. Cette représentation non autonome serait traversée par des éléments qui pourraient être regroupés sous plusieurs grands thèmes ayant plus ou moins de liens avec la représentation, sans arriver à constituer un noyau central.

La seconde nuance portera donc sur l'autonomie de la représentation, car sans partager une vision quelque peu dichotomique, nous nous demandons jusqu'à quel degré la représentation de la vie privée en contexte technologique s'avère autonome. En effet, deux éléments vont à l'encontre de l'idée d'une organisation claire caractéristique de l'autonomie.

Le premier que nous retrouvons dans nos données est la présence de corrélations faiblement positives entre les items centraux dans l'arbre en valeur algébrique. Celle-ci indique que pour les participants, il n'y a pas de façon unanime de classer les items centraux dans une catégorie. Ce premier élément soulève la question de la clarté de la représentation. En effet, si les items ne sont pas catégorisés de façon identique par la plupart des individus, on peut se demander jusqu'à quel point la conceptualisation de la représentation est partagée par l'ensemble des individus.

Le second élément est que ces faibles corrélations donnent lieu à un phénomène observable uniquement si l'on compare l'arbre maximum en valeur algébrique avec celui en valeur absolue et qui va lui aussi à l'encontre de l'évidence attendue pour une représentation autonome. Dans nos données, on observe un regroupement d'items centraux dans l'arbre maximum en valeur algébrique qui se morcelle dans l'arbre maximum en valeur absolue, faisant apparaître les items centraux comme isolés dans ce dernier. Cela semble soutenir la faible clarté de notre représentation. En effet, si l'organisation de la représentation était évidente, les corrélations les plus élevées dans les deux arbres devraient concorder pour les items centraux, puisqu'il s'agit des éléments les plus stables de la représentation, or ce n'est pas le cas.

Cependant, un seul lien stable entre éléments centraux perdure dans les deux arbres, celui concernant le lien entre « confidentialité » et « protection ». Par conséquent, la conservation de ce lien à travers les deux arbres suggère qu'il s'agit du seul lien suffisamment consensuel chez les participants pour subsister dans les deux arbres. Il s'agit d'un élément supplémentaire qui étaye l'idée d'une représentation sociale de la vie privée conçue comme sécurisée ou sécuritaire.

L'ensemble des résultats discutés précédemment invite aussi à penser de nouvelles applications, notamment sur la sensibilisation aux enjeux de la vie privée. Par exemple, les résultats des trois analyses déployées (prototypique, similitude et *Test d'indépendance au contexte*) suggèrent qu'il serait pertinent de communiquer sur des liens forts comme

ceux reliant la confidentialité et la protection lors de campagnes de sensibilisation. Mais les résultats indiquent aussi la nécessité d'apporter des éléments de définition de la vie privée lors de ces campagnes. En effet, la faible présence d'éléments dans le noyau central définissant la vie privée interroge sur la capacité réelle de définir ce qui relève de la vie privée en contexte technologique. Ainsi, on peut supposer que les individus ne savent pas quand ils sont à risque de compromettre leur vie privée, car n'étant pas capables de la définir concrètement. La question de la capacité à se protéger se pose donc.

Le terme « donnée personnelle » semble échapper à ce constat de composantes diffuses sur la vie privée dans le noyau. Il représente l'item le plus concret du noyau. Toutefois, cette recherche ne peut pas expliciter ce que le participant conçoit derrière cet item central. Lorsque les individus font référence à ces termes, pensent-ils à leur adresse, à leur âge ou à leur statut marital? De façon similaire, à quoi font référence concrètement les participants lorsqu'ils écrivent le mot « droit »? Au droit à l'image, au RGPD, au fait que la vie privée est un droit constitutionnel? La mise en place d'étude qualitative pourrait répondre à ces interrogations.

Aussi, les travaux de Piermattéo, Lo Monaco, Moreau, Girandola et Tavani (2014) reposent sur une méthodologie prometteuse pour répondre à ce genre de limite. Ces auteurs proposent d'ajouter une contextualisation sémantique afin d'affiner les résultats de l'association libre (Piermattéo *et al.*, 2014; Piermattéo, Tavani et Lo Monaco, 2018). Cette phase additionnelle consiste simplement à demander une phrase utilisant chaque

mot que le participant a donné. On pourrait par exemple demander à un participant de faire une phrase avec le mot « droit » qu'il aurait donné lors de la tâche associative. Ainsi, le participant pourrait donner la réponse suivante : « La vie privée est un droit », alors qu'un autre participant aurait pu écrire : « le droit à l'oubli est fondamental pour la vie privée en contexte technologique ». Outre diminuer la polysémie des mots (Piermattéo, Tavani et Lo Monaco, 2018), cette procédure permettrait d'effectuer une typologie des contextes. Nous pensons même pouvoir aller plus loin en créant un corpus qui rassemblerait chaque phrase illustrative afin de mener une analyse de correspondance sur tableau textuel (ex. : le tableau textuel du mot droit aurait comme première observation : « La vie privée est un droit »). Celle-ci pourrait résumer l'ensemble de ces contextes à travers les axes factoriels qui représenteraient les dimensions latentes de la conceptualisation et ainsi offrir une vision plus précise de ceux-ci.

Aussi, dans la continuité de ces travaux (Piermattéo *et al.*, 2014; Piermattéo, Tavani et Lo Monaco, 2018), il nous semble important de mener plus de recherches reposant sur un paradigme qualitatif sur le sujet. Celles-ci restent encore trop peu nombreuses (Kokolakis, 2017). De plus, les sujets connexes comme l'intimité et la technologie, qui sont des thèmes récurrents de la représentation sociale, doivent aussi être interrogés lors de ces recherches pour pouvoir identifier les aires communes de ces sujets. Comme avec l'analyse prototypique, de telles recherches permettraient elles aussi de connaître de manière spontanée et approfondie si des risques et des moyens de protection sont réellement présents dans les discours des participants. Ainsi, il serait possible de mieux

cerner dans quel contexte technologique évoluent les individus (p. ex., l'utilisation du téléphone c. l'utilisation de l'ordinateur, etc.).

Si l'aspect qualitatif offre des pistes prometteuses, nous pensons qu'il est aussi possible d'améliorer les recherches quantitatives, notamment pour qu'elles puissent permettre un recueil de données plus précis qui rendra accessible le recueil direct de comportements de défense de la vie privée qui font souvent défaut dans la recherche. Deux raisons sont avancées pour expliquer cet état de fait. La première proposée par Kokolakis (2017) est que le recueil s'intéresse souvent à des comportements trop peu fréquents (p. ex., paramétrage du compte, qui intervient habituellement une seule fois) pour représenter un comportement de protection. L'autre raison que les auteurs relèvent (Jensen, Potts et Jensen, 2005; Kokolakis, 2017; Staddon, Acquisti et LeFevre, 2013) est que le recueil se fait majoritairement à travers des mesures autorapportées. Pour Kokolakis (2017) ces limites suggèrent un manque de validité écologique des études sur la vie privée en contexte technologique.

Malheureusement, Kokolakis (2017) n'a pas détaillé de marche à suivre pour ressourdre concrètement ces problèmes. Nous proposons que l'utilisation d'une modalité de réponse du type « Je ne souhaite pas répondre » puisse pallier la limite énoncée par Kokolakis (2017) et par conséquent, améliorer la qualité des études sur la vie privée. Ainsi, nous faisons l'hypothèse que la réponse « Je ne souhaite pas répondre » pourrait être considérée comme la manifestation d'une défense de la vie privée. Cette interprétation se

base sur les travaux de Tourangeau (2018) sur les questions dites sensibles. Selon ces travaux (Tourangeau, 2018), on peut considérer le refus délibéré de répondre d'un participant comme l'expression de sa préoccupation concernant sa vie privée.

Si tel est le cas, quelle explication donner à la corrélation négative entre la proportion de réponses « Privé » et la sous-échelle de préoccupation personnelle de la *Privacy Orientation Scale*? En effet, de prime abord, ce résultat peut paraître contre-intuitif, car il indique que ces deux variables varient dans un sens contraire. Concrètement, cela implique que dans notre échantillon, les proportions de réponse « Privé » les plus élevées apparaissent avec un score faible de préoccupations personnelles concernant la vie privée. À la lecture des quatre questions¹ qui composent cette sous-échelle, une explication possible pourrait être que l'individu anticipant déjà ces inquiétudes ferait de la rétention d'information préventive de manière systématique. La conséquence serait que l'individu serait rarement inquiet, car faisant déjà le nécessaire en tout temps pour qu'aucune de ces informations ne soit révélée. Il assurerait ainsi sa tranquillité. Nous nous demandons si le même phénomène ne serait pas à l'œuvre pour expliquer l'absence de résultat concernant la variable politique.

¹ Pour rappel, les questions étaient : « Quand je partage des informations sur ma vie personnelle avec quelqu'un, je m'inquiète qu'il pourrait les divulguer à d'autres personnes », « Je me préoccupe du fait que les gens autour de moi savent trop de choses à propos de moi », « Je me préoccupe des conséquences du partage d'informations reliées à mon identité », « Je m'inquiète de partager des informations avec plus de personnes que je ne le voudrais ».

En effet, nos résultats ne montrent aucune différence de moyenne dans les scores des sous-échelles de la *Privacy Orientation Scale* des personnes ayant révélé leur appartenance politique et celles ne l'ayant pas fait. Pour autant, les variables portant sur la politique avaient un taux de réponse « Privé¹ » supérieur à 8 % minimalement et pouvaient atteindre jusqu'à 44 % selon la phase de notre étude. D'ailleurs, il semble que nos données ne soient pas différentes d'autres échantillons, car dans sa revue de littérature, Krumpal (2013) trouve des chiffres compris entre 14,4 % et 22,7 % concernant le refus de révéler l'intention de vote. Serait-il possible que certains participants n'aient pas souhaité partager leurs opinions politiques, car elles auraient pu être interprétées comme socialement répréhensibles (p. ex., positionnement à l'extrême gauche ou à l'extrême droite du spectre de la vie politique) ? Par exemple, certains participants qui auraient voté pour le Rassemblement National auraient pu souhaiter conserver cette information pour eux, car ne voulant pas être vus comme d'extrême droite. Ainsi, l'intention de vote serait considérée comme une donnée sensible. Leon, Aizpurua et van der Valk (2021) définissent une donnée sensible comme étant une information considérée comme non acceptable socialement, soit parce qu'elle déroge à une loi, soit parce qu'elle déroge à une règle socialement admise.

Cependant, il nous semble difficile de présumer que 44 % de l'échantillon auraient voulu dissimuler son vote pour le Rassemblement National. Une piste supplémentaire

¹ La modalité « Privé » a été utilisée pour remplacer le nom de la modalité complète « Je ne souhaite pas répondre », signe d'une préoccupation pour la vie privée.

serait que cette autocensure ait été augmentée par la passation en ligne. À ce sujet, Neubaum et Krämer (2018) avancent dans leur étude que les individus préfèrent s'abstenir de s'exprimer en ligne lorsqu'ils entretiennent des opinions marginales. Cette hypothèse met en exergue que la modalité « Je ne souhaite pas répondre » ne peut être à elle seule une modalité suffisante pour détecter un comportement de défense de la vie privée. Il serait nécessaire de lui adjoindre un champ pour connaître les raisons du refus et être en mesure de départager les données privées des données sensibles.

Une dernière piste possible provient des travaux de Krämer et Schäwel (2020) où les auteurs expliquent que les données sensibles sont plus difficilement partagées. Cependant, les auteurs ne précisent pas la cause intrinsèque présumée du phénomène. Enfin, une dernière piste est celle des travaux de Neubaum et Krämer (2018). Ils proposent une explication des taux de refus plus élevés en mettant en lumière le rôle de l'évaluation du risque de dévoilement et de l'anticipation des conséquences négatives de la révélation d'informations sensibles. Neubaum et Krämer (2018) font remarquer que lors de cette évaluation, l'individu anticipe les conséquences négatives de l'expression de ses opinions, notamment lorsqu'il sent un climat hostile à celles-ci. Ainsi, l'individu voudrait éviter les jugements, les rejets ou même les attaques personnelles comme des insultes qu'il pourrait encourir s'il exprimait son opinion. Par exemple, un individu sur un forum dédié à l'évolution des espèces préférera sans doute garder le silence sur ses croyances religieuses associées au créationnisme. Même si durant la passation en ligne, il n'y avait pas d'interactions avec le chercheur ou avec des groupes, il est possible que les participants

n'aient pas souhaité partager les informations politiques supposant que l'environnement ne leur serait pas favorable. Ainsi, il nous semble que l'ajout d'une modalité « Je ne souhaite pas répondre » permet de pallier l'absence de recueil direct de comportements de défense de la vie privée, mais elle pourrait aussi permettre de dresser une classification des données sensibles.

D'autres variables présentent des proportions de refus importantes. Les ressources financières du participant sont l'une d'entre elles, avec des taux de refus de réponses supérieurs à 5 % et les catégories socioprofessionnelles des parents en sont d'autres. Cependant, leurs refus restent plus faibles que ceux atteints par les variables politiques. Concernant les explications possibles à ces refus de réponses, il est possible qu'elles soient identiques à celles avancées précédemment. Les catégories socioprofessionnelles peuvent constituer une donnée sensible. Par exemple, les travaux sur la stigmatisation des ouvriers ou des employés font apparaître que certaines catégories sociales peuvent être stigmatisantes pour la personne (Avenel, 2009; Bosmans, Mousaid, De Cuyper, Hardonk, Louckx et Vanroelen, 2016), voire pour leurs enfants (Latchoumanin, 2013). Les travaux sur la stigmatisation des chômeurs (Bourguignon et Herman, 2006; Herman, 2007) corroborent eux aussi cette hypothèse.

Pour conclure, les interprétations et discussions des résultats nous ont permis de proposer de nouvelles pistes d'explications, mais nous pensons qu'elles peuvent aussi suggérer de nouvelles possibilités de recherche. Pour donner suite aux explications

proposées, il nous semble qu'au moins trois pistes de recherche peuvent être établies afin d'approfondir les connaissances sur les problématiques de la vie privée. La première, inspirée par l'article de Neubaum et Krämer (2018), porte sur l'impact du mode de passation (passation en ligne c. passation avec un examinateur). La seconde est l'insertion délibérée de questions sur les données sensibles afin de répertorier les comportements de défense de la vie privée. Plusieurs devis de recherche nous semblent intéressants dans ce cas. L'un d'eux consisterait à faire varier le nombre de questions sensibles afin de savoir si un grand nombre de questions amène l'individu à un refus systématique ou au contraire à un renoncement à toute protection. Un autre serait, comme nous avons essayé, de tenter de constituer une échelle de mesure de la défense de la protection de la vie privée à travers le nombre total de refus de réponses des participants sur leur *Questionnaire sociodémographique*. Cependant, l'absence de résultat concernant cet essai pousserait à ajouter à ce genre de question un champ libre afin de pouvoir connaître la raison des refus.

Bien que l'ajout d'une modalité mesurant la protection de la vie privée soit pertinent pour l'amélioration de la recherche, le manque d'outil de recueil des comportements de défense de la vie privée demeure un enjeu réel. En effet, malgré leur manque de fiabilité, les mesures autorapportées (Jensen, Potts et Jensen, 2005; Kokolakis, 2017; Staddon, Acquisti et LeFevre, 2013) continuent à être utilisées pour l'étude du phénomène (Kokolakis, 2017). Comme l'affirmaient déjà Jensen et ses collègues en 2005, il est nécessaire de plonger les participants dans des expériences plus réalistes (p. ex., simulation de navigation interne) qu'un questionnaire autorapporté. Il nous semble donc

primordial de continuer la réflexion autour des méthodologies de recueil des comportements observables directement de défense de la vie privée.

Mais ce travail de réflexion sur le recueil de données ne doit pas se cantonner aux comportements directement observables de défense de la vie privée durant la collecte, mais bien interroger l'ensemble du recueil des données. À ce titre, nous aimerions revenir sur les limites de notre recherche. En effet, les variations des proportions des différents sous-groupes (homme c. femme pour le genre, le domaine d'études du droit ou de la psychologie pour les domaines d'études) qui sont apparues durant notre recueil de données pourraient être considérées comme influençant les résultats de l'analyse prototypique, attendu que cette analyse est principalement une opération de réagencement des données. Ainsi, la proportion inégale de certains groupes de l'échantillon pourrait être considérée comme une erreur d'échantillonnage qui influence l'émergence de certains candidats au noyau central plutôt que d'autres, entraînant une limite sérieuse.

Cependant, soutenir l'hypothèse d'une erreur d'échantillonnage semble difficile pour plusieurs raisons. En effet, si l'on prend l'exemple de la variable sexe, on peut mieux saisir les nuances importantes qui s'imposent. Dans cette étude, les femmes ont répondu de manière prépondérante. Il serait possible de pallier aux problèmes en équilibrant les échantillons conformément aux données populationnelles de l'INSEE¹, soit 55 % de femmes et 45 % d'hommes pour l'année 2019 (Insée, 2021). Cependant, il nous semble,

¹ Institut national de la statistique et des études économiques.

au regard de la recherche, que les proportions sans intervention de notre part pour les influencer (p. ex., viser spécifiquement les hommes) pourraient être en soi une information. En effet, nous faisons l'hypothèse que ces proportions inégales pourraient représenter le niveau d'enjeux que l'objet social suscite chez les différents sous-groupes, car la littérature note que les femmes sont plus préoccupées par les questions de vie privée que les hommes (Cho et Hung, 2011; Park et Kim, 2020; Tifferet, 2019).

Contrairement à nos résultats qui ne montraient aucune différence de score entre hommes et femmes sur les préoccupations évaluées, Lewis, Kaufman et Christakis (2008) constatent même que les femmes ont plus de comportements sécuritaires que les hommes concernant leur vie privée, en paramétrant leur profil sur les réseaux sociaux pour qu'il soit plus confidentiel que celui des hommes (Lewis, Kaufman et Christakis, 2008). Pour autant, nous pensons qu'il serait possible d'améliorer les recherches en délimitant mieux les groupes sociaux comme le proposait Rouquette (1994) à travers les niveaux d'analyse psychosociaux de Doise (1982).

Dans notre cas, nous pouvons imaginer que l'analyse prototypique soit menée de façon générale sur l'échantillon étudiant, puis sur des sous-groupes plus spécifiques et déterminés à travers les quatre niveaux d'analyse psychosociaux de Doise (1982). Si l'on suit les recommandations de Rouquette (1994), il faudrait que les sous-groupes soient définis sur trois des quatre niveaux d'analyse psychosociaux. Ainsi, nous pourrions voir jusqu'à quel point les différents sous-groupes d'une population ont les mêmes candidats

à la centralité. Amener ces précisions dans la détermination des groupes nous semble une avenue de recherche importante, car en conservant ce même niveau de précision pour la validation des items centraux, nous pourrions voir s'il existe des différences entre les différents mots faisant partie du noyau central selon les sous-groupes d'étudiants. Cependant, en mettant en place ces modifications, il nous semble important d'arriver à recruter des échantillons suffisamment conséquents pour avoir un certain équilibre entre chaque niveau d'analyse.

Il nous semble aussi important de revenir sur les seuils choisis pour les fréquences et les rangs lors de l'analyse prototypique. Pour rappel, nous pensions que l'indicateur qu'est la moyenne pour découpage de la représentation en quatre zones n'était pas pertinent comme nous en faisons la proposition (section résultat, partie analyse prototypique). Ce problème perdure depuis la création de l'analyse prototypique, soit depuis 30 ans (Vergès, 1992). C'est d'ailleurs pour ces raisons que nous avons préféré utiliser les centiles qui sont plus adaptés à ce genre de cas.

Cependant, cette solution n'est pas exempte de défaut. Bien qu'elle soit plus juste statistiquement, le choix du chercheur pour la détermination du seuil constitue une limite qu'il est important d'interroger. En effet, il est impossible pour le chercheur de justifier que le centile 80 représente mieux un mot fortement cité que le centile 85, autant pour les fréquences que pour les rangs, car cette décision demeure guidée par les données que le

chercheur a dans sa collecte de données¹. Ainsi, on peut légitimement s'interroger à partir de quels chiffres la fréquence et le rang d'un mot sont suffisamment élevés pour être considérés comme signifiants. La solution pourrait être de soumettre la distribution des fréquences des items à un test binomial comme le proposent Boussoco et al. (2016). Ainsi, un premier tri pourrait être opéré afin d'enlever les items dont la probabilité d'être cités ne battrait pas le hasard, contrairement aux items dont la distribution le battrait. Attendu que les auteurs (Boussoco et al., 2016) veulent dépasser une distribution due au hasard, la probabilité du mot attendue doit dépasser les 50 % d'être cité. Cependant, la réponse à cette limite demeure partielle, car il ne faut pas oublier qu'un test binomial ne peut être effectué que si les modalités de la variable sont dichotomiques.

Or, dans le cas de l'analyse prototypique, uniquement la fréquence d'un mot peut être traitée comme tel, avec une modalité « cité » ou « non cité ». Le rang des mots quant à lui ne peut être traité de la sorte, car étant une variable ordinale. Comme nous pouvons le constater, aucune méthode n'est pleinement satisfaisante, car aucune d'entre elles ne règle entièrement la question des délimitations entre les quatre zones. À ce jour, remplacer la moyenne par des centiles nous a semblé être la solution la moins problématique. Pour autant, il nous paraît capital pour les futurs travaux de continuer la recherche fondamentale sur la cohérence entre la théorie et l'aspect mathématiques de celle-ci.

¹ La partie résultat, section analyse prototypique résume bien la façon dont le chercheur est guidé par la réalité de ces données.

Aussi, nous pensons que l'association libre de l'analyse prototypique peut être parfois la source d'une polysémie des termes recueillis. En effet, il est parfois difficile de préciser le sens du terme évoqué. Cette polysémie limite l'analyse qualitative qui peut être faite des termes recueillis. Si nous voulons tirer un maximum de résultats qualitatifs issus de l'analyse prototypique, nous pensons qu'il est nécessaire de mettre en place de façon systématique la contextualisation sémantique proposée par Piermattéo et al. (2014). Ainsi, comme nous le proposons précédemment, celle-ci pourrait limiter fortement la polysémie de certains termes en précisant leur contexte d'usage.

De plus, nous aimerions revenir sur l'analyse de cluster de la *Privacy Orientation Scale* de Baruh et Cemalcılar (2014). Celle-ci n'a pas abouti au même regroupement de participants que dans l'article original (défenseur de la vie privée, individualiste de la vie privée et indifférents de la vie privée). Attendu l'inadéquation des résultats avec ceux de l'article et la pauvreté de ceux-ci (deux groupes qui se divisaient entre ceux étant préoccupés et étant indifférents), nous avons fait le choix de ne pas mettre des analyses en clusters pour ne pas alourdir la thèse avec des résultats non concluants et de surcroît annexes.

Aussi, nous estimons que l'utilisation de Facebook pour le recrutement représente une limite à considérer attendu notre objet d'étude. Ce choix fut guidé par l'ergonomie cognitive du processus. En effet, choisir des réseaux comme TikTok ou Snapchat aurait probablement donné lieu à beaucoup plus de perte de participants, attendu l'expérience

qu'ils proposent (vidéo courte et dynamique). En visant, des groupes Facebook spécifiquement étudiants, nous avons supposé que nous bénéficierons d'un certain effet d'engagement. De plus, bien que le recours à Facebook eut été justifié par le contexte de la pandémie du COVID-19 qui rendait impossible un recrutement en personne, les études ultérieures devraient diversifier les sources dont sont issus les participants. Le fait de recruter des participants en personne permettrait d'obtenir des participants ayant des rapports à l'objet différent, par exemple les gens n'utilisant pas Facebook.

Enfin, nous aimerions revenir sur notre revue de littérature. Premièrement, la thèse s'insérant dans la théorie des représentations sociales, cela implique impliquait d'utiliser un paradigme plus ayant une composante inductifinductive forte. Cette position dernier explique aussi l'absence d'une proposition d'une de définition de la vie privée de notre part. Le but était d'obtenir la « définition » des participants, et non pas de la comparer avec la définition d'un expert. Deuxièmement, mis à part l'aspect historique de la vie privée, l'ensemble de la revue a été menée dans une perspective psychologique, car il s'agissait d'une thèse en psychologie. Cependant, il nous semble pertinent pour des travaux ultérieurs d'ajouter les éclairages de la criminologie (Powell et al., 2018), de la sociologie (Lyon, 2018 ; Nippert-Eng, 2019) ou encore la communication sociale (Brandtzaeg et Lüders, 2018). De plus, cette recherche est basée sur une revue de la littérature effectuée de 2015 à 2018; depuis, de nouvelles recherches proposant de nouvelles pistes (Waldman, 2020) ou continuant d'alimenter la critique envers le paradoxe de la vie privée (Solove, 2021) sont parues. Par ailleurs, le contexte juridique et politique

concernant la vie privée en France a beaucoup changé. Ces changements ont beaucoup de conséquences, comme le Projet de loi d'orientation et de programmation du ministère de la justice pour les années 2023-2027. Cette dernière permet par exemple à la police d'activer du matériel électronique à distance (Quadrature du net, 2023).

Pour finir notre discussion, nous aimerions revenir sur les forces de notre étude. L'une des premières est d'avoir effectué un travail conséquent de vulgarisation du courant structuraliste des représentations sociales, notamment en explicitant les présupposés statistiques du choix des indicateurs de l'analyse prototypique permettant de déterminer les candidats au noyau central. Ce travail de vulgarisation fut poursuivi en expliquant la formule du Kolmogorov-Smirnov proposée par Abric (2003) et censée résoudre le problème du seuil permettant de confirmer l'appartenance d'un candidat au noyau central. Outre les explications, nous avons pu constater que l'utilisation du Kolmogorov-Smirnov, tel que le faisait Abric (2003), ne résolvait pas le problème de détermination du seuil. En expliquant et appliquant les solutions proposées par Baggio et Renaud (2015), nous pensons avoir confirmé l'importance de changer de méthode de calcul pour la détermination des éléments centraux.

Bien que les explications et rectifications méthodologiques soient importantes sur le plan de la recherche fondamentale, elles ne sont pas l'unique apport de notre recherche. L'apport majeur de notre recherche se situe dans le recueil de la représentation lui-même. En effet, la méthodologie de l'association libre a permis de se décentrer de nos

préconceptions de chercheur, tout en permettant d'être plus proche de la vision que pouvaient avoir nos participants du phénomène de la vie privée. Ainsi, en adoptant un point de vue psychosocial, nous avons pu voir la place que pouvait occuper l'intime, la technologie et la notion de protection dans la représentation des participants tout en questionnant le paradoxe de la vie privée.

Conclusion

Le premier but de cette thèse visait à identifier la structure de la représentation sociale de la vie privée en contexte technologique des étudiants français. Dans un second temps, elle visait à confirmer une partie importante de la représentation : le noyau central. Plus précisément, le second temps a rendu possible la confirmation de nos hypothèses de recherche, à savoir la présence d'éléments faisant référence à la protection de façon majoritaire. De plus, ce second temps a permis d'observer les liens qui reliaient les différents items. Enfin, ces deux collectes ont été l'occasion de proposer un nouvel indicateur pour l'analyse prototypique. Ces deux collectes ont aussi permis l'expérimentation des recommandations de la méthode de Baggio et Renaud (2015) pour la confirmation du noyau central. Autant le nouvel indicateur que la méthodologie de Baggio et Renaud (2015) ont permis de gagner en précision et en cohérence mathématique. À travers ce dernier point, nous avons pu réaliser un travail d'éclaircissement méthodologique.

De façon plus précise, l'analyse prototypique menée dans notre première expérience a permis de générer les premières hypothèses concernant la répartition des mots à l'intérieur de la représentation sociale de la vie privée en contexte technologique. Préalablement au déploiement de l'analyse, des éclaircissements méthodologiques et mathématiques sur l'indicateur traditionnel qu'est la moyenne ont permis sa remise en cause. En effet, lors de nos explications, nous avons montré en quoi les distributions non

normales des fréquences et des rangs des mots rendaient impropre l'usage de la moyenne. À la suite de ce constat, nous avons choisi le centile comme indicateur plus approprié aux diverses distributions.

Une fois la méthodologie éclaircie et déterminée, les premiers résultats furent l'obtention des 11 candidats au noyau central, à savoir : « anonymat », « choix », « confidentialité », « donnée personnelle », « droit », « intimité », « liberté », « protection », « respect », « secret » et « sécurité ». En outre, en ayant l'ensemble des items de la représentation, nous avons pu mener une rapide analyse thématique qui a permis de noter trois grands thèmes traversant l'ensemble de la représentation, soit (1) la technologie; (2) l'intimité; et (3) la protection. Ces premiers résultats ont révélé que les items ayant pour thème la protection et l'intimité étaient majoritaires dans le noyau central. L'analyse a aussi permis de révéler un autre thème majeur, celui du danger. Cependant, et contrairement aux autres thèmes cités précédemment, celui-ci était cantonné aux zones périphériques. La position périphérique du danger indiquait l'aspect non consensuel qu'occupait ce thème dans la représentation de la vie privée. De surcroît, une analyse plus fine des risques et dangers évoqués par les participants a abouti à des précisions importantes. En effet, la majorité des risques cités par les participants étaient diffus. Selon la littérature, ce genre de risque est plus probable que des risques spécifiques dans la perception des individus, mais est généralement perçu comme moins dommageable pour leur vie privée.

De fait, ces premiers résultats questionnent la théorie du paradoxe de la vie privée. D'une part, le danger n'est pas une caractéristique nécessaire pour représenter la vie privée comme l'atteste l'absence totale de mots rattachés au risque chez les candidats au noyau central. En outre, leur faible dangerosité perçue appuie l'absence ou la faiblesse des protections mises en place et répertoriées dans de nombreuses études. D'autre part, la présence majoritaire de candidats au noyau central ayant pour thème la protection a alimenté cette remise en question. Cependant, les candidats au noyau central n'étant pas confirmés à cette étape de la recherche, il a fallu préalablement les valider pour confirmer la présence majoritaire d'items reliés à la protection.

En effet, ce second temps a permis de tester l'appartenance des candidats au noyau central. Un second travail de familiarisation portant sur la méthodologie a été nécessaire. Celui-ci a permis d'expliquer et de critiquer la formule du Kolmogorov-Smirnov proposée par Abric (2003). Ainsi, nous avons pu étayer les raisons de son échec pour résoudre l'arbitraire du seuil de confirmation d'appartenance au noyau central des items candidats à celui-ci. Par la suite, nous avons repris les solutions proposées par Baggio et Renaud (2015) pour tester les 11 items candidats.

Sur la base de nos analyses, sept items sur 11 ont été validés comme centraux dont 4 étaient reliés au thème de la protection. Ainsi, nous avons pu vérifier nos hypothèses opérationnelles de recherche, à savoir qu'il y avait effectivement des éléments centraux qui étaient rattachés au thème de la protection et qu'ils étaient majoritaires à l'intérieur du

noyau. Ces résultats nous ont amenés à formuler une hypothèse qui voudrait que la représentation sociale de la vie privée en contexte technologique soit perçue comme intrinsèquement sécurisée pour l'échantillon utilisé. Attendu que la représentation et en particulier les éléments du noyau central sont des guides pour l'action, il se peut que les individus ne sentent pas la nécessité de protéger leur vie privée attendu qu'il la représente comme déjà sécurisée. Ainsi, il n'y aurait plus de raison de constater comme paradoxale l'absence de protection de la vie privée chez les participants, attendu la faible présence du danger et la présence majoritaire d'items renvoyant à la protection. Cependant, la validation de cette hypothèse nécessiterait des études ultérieures afin d'être plus amplement étayée.

Enfin, l'analyse de similitude qui visait à quantifier les liens similaires et les oppositions des différents items lors du second temps a abouti sur une question intéressante au vu de l'instabilité des liens entre les arbres : celle de l'autonomie de la représentation. En effet, la faible ressemblance entre les deux arbres suggérait un faible degré d'organisation de la représentation, et par conséquent, une représentation faiblement autonome. Ainsi, il est apparu que les seuls liens stables se trouvaient entre des éléments périphériques et cotés comme non caractéristiques de la représentation. Le seul lien stable entre items centraux dans les deux arbres était celui reliant confidentialité et protection. Ce dernier résultat soulignait une fois de plus l'idée d'une représentation sécuritaire de la vie privée en contexte technologique.

Enfin, l'exploitation de résultats annexes et qui ne faisait pas l'objet d'une question de recherche s'est révélée intéressante. Ainsi, nous avons mis en œuvre la possibilité de recueillir des comportements observables directement de défense de la vie privée à travers la modalité de réponse « Je ne souhaite pas répondre ». Il nous a été possible de constater que l'idéologie politique, les catégories socioprofessionnelles des parents, ainsi que les revenus étaient des données sociodémographiques particulièrement sensibles au vu du haut taux de refus de réponse qu'ils enregistraient. Ces résultats nous ont paru prometteurs, car ils permettent de dresser une typologie de données sensibles en matière de vie privée. De plus, ces résultats permettent le rapprochement avec les recherches sur les données sensibles qui apportent de nouvelles pistes d'explication et d'interprétation. Par exemple, les travaux sur les données sensibles nous ont permis de soutenir que les individus voulaient garder certaines données secrètes, car l'information en question ne pourrait pas être admise socialement ou encore qu'elle pourrait faire l'objet d'un jugement social négatif.

En somme, les résultats des deux temps de collecte ont permis la découverte de la structure de la représentation, son noyau central ainsi que les liens qu'entraînaient certains items entre eux. À travers ces résultats, nous avons pu mettre à jour la faible présence du danger et des risques à l'intérieur de la représentation à travers leurs localisations exclusivement périphériques dans la représentation. De plus, leur typologie les caractérisait comme faiblement dommageables. Nous avons pu aussi constater la présence importante des thèmes rattachés à la protection et la sécurité, peu importe la zone de la

représentation. Ce constat nous a amenés à formuler l'hypothèse que la représentation de la vie privée est intrinsèquement sécuritaire pour les individus.

Pour terminer, nous aimerions remettre en perspective les résultats de nos travaux. La découverte de la représentation sociale de la vie privée en contexte technologique d'étudiants français remet en question la théorie du paradoxe de la vie privée et alimente l'hypothèse qu'il soit un possible artefact méthodologique. Il s'agit là d'un des principaux apports scientifiques de notre thèse sur la question de la vie privée en contexte technologique. Mais qu'en est-il de la contribution sociale de notre thèse? En réalité, cette question s'est posée dès la fin de l'analyse de nos résultats, lorsque nous envisagions des applications pratiques à notre travail de recherche. Au début de notre réflexion, celles-ci prenaient principalement la forme de campagnes de prévention où nous aurions ciblé des éléments centraux afin d'assurer une communication efficace pour que les individus adoptent plus facilement des comportements sécuritaires concernant leur vie privée. Bien que ces campagnes soient nécessaires pour améliorer la sécurité des données personnelles de la population, nous pensons que cette solution est parcellaire.

Premièrement, sur le plan technique, une récente revue de la littérature en informatique (El-Mhamdi et al., sous-presse) permet de saisir que la solution d'une campagne d'accès sur les comportements sécuritaires ne peut pas représenter l'unique solution. En effet, cette revue portant sur la sécurité des nouveaux algorithmes d'apprentissage utilisés par les moteurs de recherche, suggère qu'il est fortement

improbable d'obtenir à la fois d'excellentes prédictions et à la fois protéger la confidentialité des données des utilisateurs (El-Mhamdi et al., sous-presse).

Deuxièmement, outre la dimension technique qui vient limiter l'apport d'une campagne portant uniquement sur la responsabilité individuelle de protéger sa vie privée, l'ensemble de notre démarche intellectuelle ainsi qu'une partie de nos résultats incitent à proposer une réponse plus macro aux problématiques de la vie privée. En effet, une des visées de notre étude était d'élargir l'angle de recherche traditionnellement adopté dans les travaux sur la vie privée, en incorporant une perspective sociale à notre problématique. Grâce à celle-ci, nous avons pu saisir que pour les participants de notre recherche, les problématiques de vie privée dépassaient elles aussi le cadre individuel. Les réponses plus institutionnelles comme « CNIL » ou réglementaires comme « RGPD » suggèrent une protection reposant plus sur des mécanismes macros qu'individuels. Au vu de ces résultats, il nous semble donc pertinent d'alimenter notre réflexion avec les travaux de Zuboff (2020) pour mieux comprendre la pertinence d'une campagne de sensibilisation aux lois et institutions défendant la vie privée ainsi qu'au problème de vie privée, notamment la récolte des données personnelles.

L'intérêt des travaux de Zuboff (2020) est de retracer l'histoire des premières entreprises d'Internet. L'auteur explique que l'intérêt pour les données s'est manifesté lorsqu'il a fallu trouver un modèle économique viable pour ces entreprises (Zuboff, 2020). En effet, la plupart d'entre elles faisaient faillite faute de modèle économique efficient.

Cependant, l'une d'elles, nommée Google, parvint à trouver un modèle rentable en adoptant celui de la régie publicitaire. C'est à partir de ce moment que les données personnelles sont devenues un enjeu économique. Grâce aux données, les entreprises allaient pouvoir prévoir et modifier le comportement de leurs utilisateurs et vendre ces solutions clés en main à n'importe quelle entreprise (Zuboff, 2020). L'établissement du modèle de la régie basé sur les données pourrait entraîner une conséquence sur les représentations sociales des individus. En effet, les intérêts économiques de ces acteurs pourraient les motiver à maintenir chez les utilisateurs des représentations sociales de la vie privée perçues comme protégées. Par conséquent, les utilisateurs maintenant de telles représentations seraient moins enclins à remettre en question l'utilisation des données et les entreprises pourraient continuer à prélever les données personnelles sans trop de difficultés.

De plus, l'implantation de ce modèle d'activité n'est possible qu'à travers une collecte massive et constante des données personnelles des utilisateurs pour assurer la compétitivité et la rentabilité de ce dernier, car Google n'est plus le seul à s'être lancé dans la régie publicitaire. Si l'on fait le lien avec la revue de littérature de El-Mhamdi et al. (2022) précédemment citée, on comprend que la mise en place de ces algorithmes risque très probablement de se généraliser, puisqu'elle permettra aux entreprises sur Internet d'obtenir des prédictions de meilleures qualités et par conséquent, de garantir leurs revenus.

De surcroît, le monde des affaires ne serait pas le seul à profiter de meilleures prédictions à travers l'implantation de ces algorithmes. En effet, les états pourraient en tirer parti pour surveiller les citoyens dans un contexte d'insécurité globale, comme celui qui caractérise la possibilité d'attentats terroristes (Bigo, 2008). Dans ses travaux, Bigo (2008) propose le concept de ban-opticon, un système de gouvernance caractérisé par la surveillance. Ce système abuserait de son pouvoir à travers l'institutionnalisation de lois d'exceptions pour surveiller et ostraciser certains groupes de citoyens, notamment par le profilage en lien avec de possibles comportements prédits (Bigo, 2008). Or, ce profilage n'est possible que par la récupération de données personnelles sur les citoyens. À ce titre, le programme de surveillance mise en place à la suite du 11 septembre par les agences gouvernementales américaines (NSA) et dénoncé par Edward Snowden en 2013 peut rappeler le fonctionnement du ban-opticon.

Ainsi on se rend compte que le futur de la vie privée est autant menacé par des intérêts économiques que politiques. Au vu de la puissance des acteurs en question (GAFAM, États), la protection de la vie privée ne peut pas se faire uniquement à travers une réponse technique, elle doit aussi incorporer un volet institutionnel et réglementaire.

À la lumière de ces recherches et de nos résultats, nous pensons qu'il serait pertinent de mettre en place à la fois des campagnes de sensibilisation sur les enjeux de société qu'amène la collecte massive de données personnelles et à la fois d'enseigner des moyens de protection techniques et légaux de la vie privée à la population. Enfin, nous suggérons

la création de pôles pluridisciplinaires qui intègrent des disciplines telles que l'informatique, le droit, l'économie, les sciences politiques, la sociologie et la psychologie afin d'avoir une vision complète du phénomène et de créer des solutions innovantes aux problématiques de vie privée résultant de l'interdisciplinarité de ces pôles.

Références

- Abric, J.-C. (1976). *Jeux, conflits et représentations sociales* [Thèse de doctorat d'État]. Université de Provence, Aix en Provence, France.
- Abric, J.-C. (1994). *Pratiques sociales et représentations*. Presses universitaires de France.
- Abric, J.-C. (2003). *Méthodes d'étude des représentations sociales*. Erès.
- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221), 509-514. <https://doi.org/10.1126/science.aaa1465>
- Aim, M.-A., Decarsin, T., Bovina, I., & Dany, L. (2018). Health and social representations: A structural approach. *Papers on Social Representations*, 27(1), 3.1-3.21.
- AlSudiari, M. A., & Vasista, T. (2012). Cloud computing and privacy regulations: An exploratory study on issues and implications. *Advanced Computing*, 3(2), 159-169. <https://doi.org/10.5121/acij.2012.3216>
- Altman, I. (1975). *The environment and social behavior*. Brooks.
- Apostolidis, T., Lionel, D., Cress, P., & Pecly Wolter, R. (2013). Efeito das modalidades de resposta sobre a estruturação de uma representação social: o exemplo da representação dos estudos. *Interamerican Journal of Psychology*, 45(211-222).
- Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, 41(3), 258-290. <https://doi.org/10.1027/1864-9335/a000179>
- Avenel, C. (2009). La construction du « problème des banlieues » entre ségrégation et stigmatisation. *Journal français de psychiatrie*, 34(3), 36-44. <https://doi.org/10.3917/jfp.034.0036>
- Ayalon, O., & Toch, E. (2016). Not even past: Information aging and temporal privacy in online social networks. *Human-Computer Interaction*, 32(2), 73-102. <https://doi.org/10.1080/07370024.2016.1203791>
- Baggio, S., & Renaud, O. (2015). Détecter la centralité des éléments d'une représentation sociale : un indice alternatif de mesure. *Les cahiers internationaux de psychologie sociale*, 105(1), 33-52. <https://doi.org/10.3917/cips.105.0033>

- Bagola, B. (2004). L'américanisation de la langue française sur Internet? Quelques aspects de la terminologie officielle et de l'usage des internautes. *GLOBE: Revue internationale d'études québécoises*, 7(2), 101-124. <https://doi.org/10.7202/1000863ar>
- Barth, S., & de Jong, M. D. (2017). The privacy paradox—Investigating discrepancies between expressed privacy concerns and actual online behavior—A systematic literature review. *Telematics and Informatics*, 34(7), 1038-1058. <https://doi.org/10.1016/j.tele.2017.04.013>
- Barth, S., de Jong, M. D., Junger, M., Hartel, P. H., & Roppelt, J. C. (2019). Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources. *Telematics and Informatics*, 41(1), 55-69. <https://doi.org/10.1016/j.tele.2019.03.003>
- Baruh, L., & Cemalcilar, Z. (2014). It is more than personal: Development and validation of a multidimensional Privacy Orientation Scale. *Personality and Individual Differences*, 70(1), 165-170. <https://doi.org/10.1016/j.paid.2014.06.042>
- Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review. *Journal of Communication*, 67(1), 26-53. <https://doi.org/10.1111/jcom.12276>
- Béché, E. (2017). Étudier l'appropriation des TIC à l'école en combinant l'examen des usages et des représentations sociales des utilisateurs. Une analyse à partir du contexte d'intégration de l'ordinateur et l'Internet dans quatre lycées de Yaoundé (Cameroun). *Tic&société*, 10(2-3). <https://doi.org/10.4000/ticetsociete.2108>
- Bélanger, F., & Crossler, R. E. (2011). Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly*, 35(4), 1017-1042. <https://doi.org/10.2307/41409971>
- Bergström, A. (2015). Online privacy concerns: A broad approach to understanding the concerns of different groups for different uses. *Computers in Human Behavior*, 53(1), 419-426. <https://doi.org/10.1016/j.chb.2015.07.025>
- Billingham, P., & Parr, T. (2020). Online public shaming: Virtues and vices. *Journal of Social Philosophy*, 51(3), 371-390. <https://doi.org/10.1111/josp.12308>
- Bosmans, K., Mousaid, S., De Cuyper, N., Hardonk, S., Louckx, F., & Vanroelen, C. (2016). Dirty work, dirty worker? Stigmatisation and coping strategies among domestic workers. *Journal of Vocational Behavior*, 92(1), 54-67. <https://doi.org/10.1016/j.jvb.2015.11.008>

- Bourguignon, D., & Herman, G. (2006, Novembre). *Je suis chômeur (se), je suis stigmatisé (e). Des conséquences de la stigmatisation aux stratégies de défense de soi*. Communication présentée à Institut National de l'Orientation et de l'Insertion Professionnelles. Actes du Colloque AFPA/INOIP « Exclusions et discriminations : comprendre et agir », Lille, France.
- Bouriche, B. (2003). L'analyse de similitude. Dans J.-C. Abric (Éd.), *Méthodes d'étude des représentations sociales* (pp. 221-252). Erès.
- Boussoco, J., Dany, L., Giboreau, A., & Urdapilleta, I. (2016). « Faire la cuisine » : approche socio-représentationnelle et distance à l'objet. *Les cahiers internationaux de psychologie sociale*, III(3), 367-395. <https://doi.org/10.3917/cips.111.0367>
- Brandtzaeg, P. B., & Lüders, M. (2018). Time collapse in social media: extending the context collapse. *Social Media + Society*, 4(1). <https://doi.org/10.1177/2056305118763349>
- Brown, B. (2001). *Studying the Internet experience*. <https://www.hpl.hp.com/techreports/2001/HPL-2001-49.pdf>
- Burgoon, J. K., Parrott, R., Le Poire, B. A., Kelley, D. L., Walther, J. B., & Perry, D. (1989). Maintaining and restoring privacy through communication in different types of relationships. *Journal of Social and Personal Relationships*, 6(2), 131-158. <https://doi.org/10.1177/026540758900600201>
- Casilli, A. A. (2013). Contre l'hypothèse de la « fin de la vie privée ». La négociation de la privacy dans les médias sociaux. *Revue française des sciences de l'information et de la communication*, (3). <https://doi.org/10.4000/rfsic.630>
- Cho, V., & Hung, H. (2011). The effectiveness of short message service for communication with concerns of privacy protection and conflict avoidance. *Journal of Computer-Mediated Communication*, 16(2), 250-270. <https://doi.org/10.1111/j.1083-6101.2011.01538.x>
- Christofides, E., & O'Doherty, K. (2016). Company disclosure and consumer perceptions of the privacy implications of direct-to-consumer genetic testing. *New Genetics and Society*, 35(2), 101-123. <https://doi.org/10.1080/14636778.2016.1162092>
- Cisco. (2017). *Cisco visual networking index: Global mobile data traffic forecast update, 2016-2021*. https://www.ris.org/uploadi/editor/1341575327white_paper_c11-520862.pdf
- CNIL. (2016). 10 conseils pour rester net sur le web. Récupéré le 17/01/2023 de <https://www.cnil.fr/fr/10-conseils-pour-rester-net-sur-le-web>

- CNIL. (2018a). *Les Français et la protection des données personnelles*. https://www.cnil.fr/sites/default/files/atoms/files/barometre_ifop_rgpd-2018.pdf
- CNIL. (2018b). *Le règlement général sur la protection des données - RGPD*. Récupéré le 18/01/2022 de <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>
- CNRTL. (2017). Étymologie et histoire du mot « Privée ». Récupéré le 04/04/2017 de <http://www.cnrtl.fr/etymologie/priv%C3%A9>
- CNRTL. (2018). Définition du mot « Séclusion ». Récupéré le 28/09/2021 à <https://www.cnrtl.fr/lexicographie/seclusion>
- Conseil Constitutionnel. (2007). *Décision n° 2007-557 DC du 15 novembre 2007*. Récupéré le 14/02/2022 de <https://www.conseil-constitutionnel.fr/decision/2007/2007557DC.htm>
- Cramér, H. (1928). On the composition of elementary errors: First paper: Mathematical deductions. *Scandinavian Actuarial Journal*, 1928(1), 13-74. <https://doi.org/10.1080/03461238.1928.10416862>
- Credoc. (2018). Baromètre du numérique 2018 : 18^e éd. [Rapport]. Récupéré le 5/05/2019 de <https://www.credoc.fr/publications/barometre-du-numerique-2018>
- Culnan, M. J. (1993). "How Did They Get My Name?": An exploratory investigation of consumer attitudes toward secondary information use. *MIS Quarterly*, 17(3), 341-363. <https://doi.org/10.2307/249775>
- Dany, L., & Apostolidis, T. (2007). Approche structurale de la représentation sociale de la drogue : interrogations autour de la technique de mise en cause. *Les cahiers internationaux de psychologie sociale*, 73(1), 11-26. <https://doi.org/10.3917/cips.073.0011>
- Darnault, C. (2021). La protection des données personnelles : présentation des approches européennes et américaines. *Cahiers Tocqueville des Jeunes Chercheurs*, 3(1), 153-176.
- Debatin, B., Lovejoy, J. P., Horn, A.-K., & Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication*, 15(1), 83-108. <https://doi.org/10.1111/j.1083-6101.2009.01494.x>
- De Rosa, A. S. (2003). Le « réseau d'associations » : une technique pour détecter la structure, les contenus, les indices de polarité, de neutralité et de stéréotypie du champ sémantique liés aux représentations sociales. Dans J.-C. Abric (Éd.), *Méthodes*

- d'étude des représentations sociales* (pp. 81-117). Érès. <https://doi.org/10.3917/eres.abric.2003.01.0081>
- Deschênes, C., & Archambault, P. (2019). *Le RGPD et le régime canadien de protection des renseignements personnels : quelles sont les différences?* Récupéré le 19/02/2021 de <https://www.soulier-avocats.com/wp-content/uploads/prev/2019/03/Langlois-avocats-Article-LPRPSP-vs-RGPD-Bulletin-Soulier-avocats-Franc%CC%A7ais.pdf>
- Dienlin, T., & Trepte, S. (2015). Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors. *European Journal of Social Psychology*, 45(3), 285-297. <https://doi.org/10.1002/ejsp.2049>
- Dingledine, R., & Mathewson, N. (2002) Tor. (Version alpha). Récupéré du site <https://www.torproject.org/download/>
- Doise, W. (1982). *L'explication en psychologie sociale*. Presses universitaires de France.
- Doise, W. (1986). Les représentations sociales : définition d'un concept. Dans W. Doise & A. Palmonari (Éds), *L'étude des représentations sociales* (pp. 81-94). Delachaux & Niestlé.
- Domo. (2020). *Data never sleeps 8.0*. <https://www.blogdumoderateur.com/une-minute-internet-2020/>
- Donoghue, B. (2017). *Risks: Hacking, identity theft and burglary that are associated with students (aged 18-24) in sharing personal data through Social Media Applications* [Thèse de doctorat inédite]. Cardiff Metropolitan University, Royaume-Uni.
- Dropbox. (2021). *À quoi équivaut 1 To de stockage?* Récupéré le 28/10/2021 de <https://www.dropbox.com/fr/features/cloud-storage/how-much-is-1tb>
- Eckhoff, D., & Wagner, I. (2017). Privacy in the smart city—applications, technologies, challenges, and solutions. *IEEE Communications Surveys & Tutorials*, 20(1), 489-516. <https://doi.org/10.1109/COMST.2017.2748998>
- Egelman, S., & Peer, E. (2015a, Avril). *Scaling the security wall: Developing a Security Behavior Intentions Scale (Sebis)*. Communication présentée à Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Séoul, République de Corée. <https://doi.org/10.1145/2702123.2702249>
- Egelman, S., & Peer, E. (2015b). Predicting privacy and security attitudes. *ACM SIGCAS Computers and Society*, 45(1), 22-28. <https://doi.org/10.1145/2738210.2738215>

- El-Mhamdi, E.-M., Farhadkhani, S., Guerraoui, R., Gupta, N., Hoang, L.-N., Pinot, R., & Stephan, J. (2022). *SoK: On the Impossible Security of Very Large Foundation Models*. <https://doi.org/10.48550/arXiv.2209.15259>
- Eyssartier, C., Joule, R.-V., & Guimelli, C. (2007). Effets comportementaux et cognitifs de l'engagement dans un acte préparatoire activant un élément central versus périphérique de la représentation du don d'organes. *Psychologie française*, 52(4), 499-517. <https://doi.org/10.1016/j.psfr.2007.01.004>
- Flament, C. (1965). *Théorie des graphes et structure sociale*. Mouton-Gauthier-Villar.
- Flament, C. (1981). L'analyse de similitude : une technique pour les recherches sur les représentations sociales. *Cahiers de psychologie cognitive/Current Psychology of Cognition*, 1(4), 375-395.
- Flament, C. (1994). Aspects périphériques des représentations sociales. Dans C. Guimelli (Éd.), *Structures et transformations des représentations sociales* (pp. 85-118). Delachaux et Niestlé.
- Flament, C. (1996). Quand les Elements centraux d'une représentation sont excentriques- Note Technique. *Papers on Social Representations*, 5(1), 145-156.
- Flament, C. (1999). La représentation sociale comme système normatif. *Psychologie et société*, 1(1), 29-53.
- Flament, C., & Rouquette, M.-L. (2003). *Anatomie des idées ordinaires : comment étudier les représentations sociales*. Armand Colin.
- Freud, S. (1908). *On the sexual theories of children*. Read Books Ltd.
- Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, 77(1), 226-261. <https://doi.org/10.1016/j.cose.2018.04.002>
- Gerber, N., Reinheimer, B., & Volkamer, M. (2019). Investigating people's privacy risk perception. *Proceedings on Privacy Enhancing Technologies*, 2019(3), 267-288. <https://doi.org/10.2478/popets-2019-0047>
- Gross, R., & Acquisti, A. (2005, Novembre). *Information revelation and privacy in online social networks (The Facebook case)*. Communication présentée Workshop on Privacy in the Electronic Society (WPES), Virginie, États-Unis. <https://doi.org/10.1145/1102199.1102214>

- Haigh, T., & Priestley, M. (2015). Where code comes from: Architectures of automatic control from Babbage to algol. *Communications of the ACM*, 59(1), 39-44. <https://doi.org/10.1145/2846088>
- Halpérin, J.-L. (2005). L'essor de la « privacy » et l'usage des concepts juridiques. *Droit et société*, 61(3), 765-782. <https://doi.org/10.3917/drs.061.0765>
- Halpérin, J.-L. (2015). Protection de la vie privée et privacy : deux traditions juridiques différentes? *Les nouveaux cahiers du conseil constitutionnel*, 48(3), 59-68. <https://doi.org/10.3917/nccc1.048.0059>
- Hargittai, E., & Marwick, A. (2016). "What can I really do?" Explaining the privacy paradox with online apathy. *International Journal of Communication*, 10(1), 3737-3757.
- Heider, F. (1925). Ding und medium. *Symposion*, 1, 109-158.
- Hendricks, L., Lumadue, R., & Waller, L. R. (2012, Janvier). *The evolution of bullying to cyber bullying: An overview of the best methods for implementing a cyber bullying preventive program*. <http://www.nationalforum.com/Electronic%20Journal%20Volumes/Hendricks,%20LaVelle%20Cyber%20Bullying%20NFJCA%20V1%20N1%202012.pdf>
- Herman, G. (2007). *Travail, chômage et stigmatisation : une analyse psychosociale*. De Boeck
- Hutuleac, A. (2021). *#dogpoopgirl*. Roumanie: DaKINO. Récupéré de <https://www.imdb.com/title/tt14403262/>
- Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299-314. <https://doi.org/10.2307/1390807>
- Insée. (2021). *D'Admission post-bac à Parcoursup : quels effets sur la répartition des néo-bacheliers dans les formations d'enseignement supérieur?* Récupéré le 28/10/2022 de <https://www.insee.fr/fr/statistiques/5432519?sommaire=5435421&q=%C3%A9tudiant+universitaire>
- Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, 51(8), 56-59. <https://doi.org/10.1109/MC.2018.3191268>
- Jakobson, R. (1963). *Essais de linguistique générale*. Éditions de Minuit.

- Jensen, C., Potts, C., & Jensen, C. (2005). Privacy practices of Internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies*, 63(1-2), 203-227. <https://doi.org/10.1016/j.ijhcs.2005.04.019>
- Jeong, E. (2020). *Troll feminism: The rise of popular feminism in South Korea* [Thèse de doctorat inédite]. University of York, Angleterre.
- Jeong, Y., & Kim, Y. (2017). Privacy concerns on social networking sites: Interplay among posting types, content, and audiences. *Computers in Human Behavior*, 69(1), 302-310. <https://doi.org/10.1016/j.chb.2016.12.042>
- Jodelet, D. (1984). Représentation sociale : phénomène, concept et théorie. Dans S. Moscovici (Éd.), *Psychologie sociale* (pp. 357-378). Presses universitaires de France.
- Jodelet, D. (1989). *Les représentations sociales* (7^e éd.). Presses universitaires de France.
- Jodelet, D. (Éd.) (1991). *Grand dictionnaire de la psychologie*. Larousse.
- Kanji, G. K. (1999). *100 statistical tests*. Sage.
- Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1/2), 81-93. <https://doi.org/10.2307/2332226>
- Khan, R., Kumar, P., Jayakody, D. N. K., & Liyanage, M. (2019). A survey on security and privacy of 5G technologies: Potential solutions, recent advancements, and future directions. *IEEE Communications Surveys & Tutorials*, 22(1), 196-248. <https://doi.org/10.1109/COMST.2019.2933899>
- Kim, S. (2015). *Being surplus in the age of new media: Ing-yeo subjectivity and youth culture in South Korea* [Thèse de doctorat inédite]. George Mason University, États-Unis.
- Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, 64(1), 122-134. <https://doi.org/10.1016/j.cose.2015.07.002>
- Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Inst. Ital. Attuari, Giorn.*, 4(1), 83-91.
- Krämer, N. C., & Schäwel, J. (2020). Mastering the challenge of balancing self-disclosure and privacy in social media. *Current Opinion in Psychology*, 31(1), 67-71. <https://doi.org/10.1016/j.copsyc.2019.08.003>

- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, 47(4), 2025-2047. <https://doi.org/10.1007/s11135-011-9640-9>
- Kuisel, R. (1990). L'américanisation de la France (1945-1970). *Les cahiers du centre de recherches historiques. Archives*(5). <https://doi.org/10.4000/ccrh.2889>
- Larousse. (2017). *Définition du mot privé*. Récupéré le 04/04/2017 de <https://www.larousse.fr/dictionnaires/francais/priv%C3%A9/64013>
- Larousse. (2021). *Définition du mot synapsie*. Récupéré le 25/04/2021 de <https://www.larousse.fr/dictionnaires/francais/synapsie/76111>
- Latchoumanin, M. (2013). Quand l'orientation scolaire rime avec stigmatisation et exclusion. *Travaux & documents*, 2013(45), 69-85. <https://hal.univ-reunion.fr/hal-02170530/document>
- Laufer, R. S., & Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues*, 33(3), 22-42. <https://doi.org/10.1111/j.1540-4560.1977.tb01880.x>
- Lee, H., Park, H.m & Kim, J. (2013). Why do people share their context information on Social Network Services? A qualitative study and an experimental study on users' behavior of balancing perceived benefit and risk. *International Journal of Human-Computer Studies*, 71(9), 862-877. <https://doi.org/10.1016/j.ijhcs.2013.01.005>
- Leon, C. M., Aizpurua, E., & van der Valk, S. (2021). The impact of confidentiality assurances on participants' responses to sensitive questions. *International Journal of Public Opinion Research*, 33(4), 1024-1038. <https://doi.org/10.1093/ijpor/edaa039>
- Le Robert. (2017). *Définition du mot privé*. Récupéré le 04/04/2017 de <https://dictionnaire.lerobert.com/definition/priv%C3%A9>
- Levin, A., & Nicholson, M. J. (2005). Privacy law in the United States, the EU and Canada: The allure of the middle ground. *University of Ottawa Law & Technology*, 2(1), 357-395.
- Lewis, K., Kaufman, J., & Christakis, N. (2008). The taste for privacy: An analysis of college student privacy settings in an online social network. *Journal of Computer-Mediated Communication*, 14(1), 79-100. <https://doi.org/10.1111/j.1083-6101.2008.01432.x>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median.

Journal of Experimental Social Psychology, 49(4), 764-766. <https://doi.org/10.1016/j.jesp.2013.03.013>

- Lheureux, F., & Guimelli, C. (2009). Organisation interne et affectivité des représentations sociales : une étude exploratoire. *Anuario de Psicología*, 40(3), 357-376.
- Lheureux, F., & Lo Monaco, G. (2011). Hiérarchie intra-noyau et négociabilité des éléments centraux d'une représentation sociale? *Les cahiers internationaux de psychologie sociale*, 91(3), 213-230. <https://doi.org/10.3917/cips.091.0211>
- Lheureux, F., Rateau, P., & Guimelli, C. (2008). Hiérarchie structurale, conditionnalité et normativité des représentations sociales. *Les cahiers internationaux de psychologie sociale* 77(1), 41-55. <https://doi.org/10.3917/cips.077.0041>
- Li, Y. (2012). Theories in online information privacy research: A critical review and an integrated framework. *Decision Support Systems*, 54(1), 471-481. <https://doi.org/10.1016/j.dss.2012.06.010>
- Lo Monaco, G., Delouée, S., & Rateau, P. (2016). *Les représentations sociales : théories, méthodes et applications*. De Boeck Supérieur.
- Lo Monaco, G., Lheureux, F., & Hamili-Falkowicz, S. (2008). Test d'indépendance au contexte (TIC) et structure des représentations sociales. *Swiss Journal of Psychology*, 67(2), 119-123. <https://doi.org/10.1024/1421-0185.67.2.119>
- Lo Monaco, G., Piermattéo, A., Rateau, P., & Tavani, J. L. (2016). Methods for studying the structure of social representations: A critical review and agenda for future research. *Journal for the Theory of Social Behaviour*, 47(3), 306-331. <https://doi.org/10.1111/jtsb.12124>
- Lo Monaco, G., Rateau, P., & Guimelli, C. (2007). Nexus, représentations sociales et masquage des divergences intra et intergroupes. *Bulletin de psychologie*, 492(6), 581-592. <https://doi.org/10.3917/bupsy.492.0581>
- Lutz, C., & Strathoff, P. (2014). *Privacy concerns and online behavior—Not so paradoxical after all? Viewing the privacy paradox through different theoretical lenses. Viewing the Privacy Paradox Through Different Theoretical Lenses (April 15, 2014)*. <https://doi.org/10.2139/ssrn.2425132>
- Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society*, 1(2). <https://doi.org/10.1177/2053951714541861>

- Lyon, D. (2018). *The culture of surveillance: Watching as a way of life*. John Wiley & Sons.
- Mao, B., & Saintourens, T. (2016). *Cyber fragiles : enquête sur les dangers de nos vies connectées*. Éditions Tallandier.
- Margulis, S. T. (1977). Conceptions of privacy: Current status and next steps. *Journal of Social Issues*, 33(3), 5-21. <https://doi.org/10.1111/j.1540-4560.1977.tb01879.x>
- Margulis, S. T. (2003). On the status and contribution of Westin's and Altman's theories of privacy. *Journal of Social Issues*, 59(2), 411-429. <https://doi.org/10.1111/1540-4560.00071>
- Mariotti, F. (2003). Tous les objets sociaux sont ils des objets de représentation sociale? Questions autour de la pertinence. *Journal international sur les représentations sociales*, 1(1), 2-18.
- Markovà, I. (1996). Towards an epistemology of social representation. *Journal for the Theory of Social Behaviour*, 26(2), 177-196. <https://doi.org/10.1111/j.1468-5914.1996.tb00528.x>
- Markovà, I. (2016). *The dialogical mind: Common sense and ethics*. Cambridge University Press.
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11(3), 320-341. https://doi.org/10.1207/s15328007sem1103_2
- Martin, C. R., & Savage-McGlynn, E. (2013). A 'good practice' guide for the reporting of design and analysis for psychometric evaluation. *Journal of Reproductive and Infant Psychology*, 31(5), 449-455. <https://doi.org/10.1080/02646838.2013.835036>
- Mayer-Schönberger, V. (2011). *Delete: The virtue of forgetting in the digital age*. Princeton University Press.
- Methivier, J. (2010). Impact de la peur sur les représentations sociales du travail et du chômage, chez de jeunes adultes en recherche d'emploi. *Bulletin de psychologie*, 507(3), 183-189. <https://doi.org/10.3917/bupsy.507.0183>
- Micheau, F. (2017). *La relation des Français à leur téléphone portable*. Récupéré le 24/06/2021 de <https://www.opinion-way.com/fr/sondage-d-opinion/sondages-publies/etudes-publiees/opinionway-pour-volpy-la-relation-des-francais-a-leur-telephone-portable-avril-2017.html>

- Moliner, P. (1988). *La représentation sociale comme grille de lecture : étude expérimentale de sa structure et aperçu sur ses processus de transformation*. [Thèse de doctorat inédite]. Université d'Aix-Marseille 1, France.
- Moliner, P. (1993). Cinq questions à propos des représentations sociales. *Les cahiers internationaux de psychologie sociale*, 20(1), 5-14.
- Moliner, P. (1994). Les méthodes de repérage et d'identification du noyau des représentations sociales. Dans C. Guimelli (Éd.), *Structures et transformations des représentations sociales* (pp. 199-232): Delachaux & Niestlé.
- Moliner, P. (1996). *Images et représentations sociales : de la théorie des représentations à l'étude des images sociales*. Presse universitaire de Grenoble.
- Moliner, P. (2015). Objectivation et ancrage du message iconique. Propositions théoriques et pistes de recherche. *Sociétés*, 130(4), 81-94. <https://doi.org/10.3917/soc.130.0081>
- Moliner, P. (2016). De la théorie du noyau central à la théorie du noyau Matrice. *Papers on Social Representations*, 26(2), 1-13.
- Moliner, P., & Lo Monaco, G. (2017). *Méthodes d'association verbale pour les sciences humaines et sociales : fondements conceptuels et aspects pratiques*. Presses universitaires de Grenoble.
- Moliner, P., Rateau, P., & Cohen-Scali, V. (2002). *Les représentations sociales : pratique des études de terrain*. Presses universitaires de Rennes.
- Mols, A., & Janssen, S. (2017). Not interesting enough to be followed by the NSA. *Digital Journalism*, 5(3), 277-298. <https://doi.org/10.1080/21670811.2016.1234938>
- Moore, A. D. (Éd.) (2021). 2. Defining privacy. Dans *Privacy Rights* (pp. 11-32). Penn State University Press.
- Moscovici, S. (1961). *La psychanalyse, son image et son public : étude sur la représentation sociale de la psychanalyse*. Presses universitaires de France.
- Mulligan, D. K., Koopman, C., & Doty, N. (2016). Privacy is an essentially contested concept: A multi-dimensional analytic for mapping privacy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083), 20160118. <https://doi.org/10.1098/rsta.2016.0118>
- Naker, S., & Greenbaum, D. (2017). Now you see me: Now you still do: Facial recognition technology and the growing lack of privacy. *Boston University Journal of Science & Technology Law*, 23(1), 88-122.

- Nippert-Eng, C. E. (2019). *Islands of privacy*. University of Chicago Press
- Neubaum, G., & Krämer, N. C. (2018). What do we fear? Expected sanctions for expressing minority opinions in offline and online communication. *Communication Research*, 45(2), 139-164. <https://doi.org/10.1177/0093650215623837>
- Online Etymology Dictionary. (2017). Definition of privacy. Récupéré le 10/12/2017 de <https://www.etymonline.com/word/privacy>
- Oxford. (2017). *Definition of privacy*. Récupéré le 04/04/2017 de <https://www.oxforddictionaries.com/>
- Paine, C., Reips, U.-D., Stieger, S., Joinson, A., & Buchanan, T. (2007). Internet users' perceptions of 'privacy concerns' and 'privacy actions'. *International Journal of Human-Computer Studies*, 65(6), 526-536. <https://doi.org/10.1016/j.ijhcs.2006.12.001>
- Parent, W. A. (1983). Privacy, morality, and the law. *Philosophy & Public Affairs*, 12(4), 269-288. <http://www.jstor.org/stable/2265374>
- Park, N., & Kim, Y. (2020). The Impact of social networks and privacy on electronic word-of-mouth in Facebook: Exploring gender differences. *International Journal of Communication*, 14(1), 176-199.
- Paul-Dauphin, A., Guillemin, F., Virion, J.-M., & Briançon, S. (1999). Bias and precision in visual analogue scales: A randomized controlled trial. *American Journal of Epidemiology*, 150(10), 1117-1127. <https://doi.org/10.1093/oxfordjournals.aje.a009937>
- Pawlowski, S. D., & Jung, Y. (2015). Social representations of cybersecurity by university students and implications for instructional design. *Journal of Information Systems Educatio*, 26(4), 281-294.
- Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157-175. <https://doi.org/10.1080/14786440009463897>
- Perbal, B. (2018). *Les données personnelles et la propriété du soi* [Thèse de doctorat inédite]. Université Côte d'Azur, France. <https://theses.hal.science/tel-02506246/document>
- Piaget, J. (1932). *The moral development of the child*. Routledge.

- Piermattéo, A., Lo Monaco, G., & Girandola, F. (2016). When commitment can be overturned. *Environment and Behavior*, 48(10), 1270-1291. <https://doi.org/10.1177/0013916515597511>
- Piermattéo, A., Lo Monaco, G., Moreau, L., Girandola, F., & Tavani, J. L. (2014). Context variations and pluri-methodological issues concerning the expression of a social representation: The example of the Gypsy community. *The Spanish Journal of Psychology*, 17(1), E85. <https://doi.org/10.1017/sjp.2014.84>
- Piermattéo, A., Tavani, J.-L., & Lo Monaco, G. (2018). Improving the study of social representations through word associations: Validation of semantic contextualization. *Field Methods*, 30(4), 329-344. <https://doi.org/10.1177/1525822X18781766>
- Pierre, J. (2014). *Le cadre privatif : des données aux contextes : approche interdimensionnelle des enjeux de médiation de la vie privée* [Thèse de doctorat inédite]. Université Grenoble Alpes, Grenoble, France.
- Plante, P. (2022). *Analyse prototypique depuis une saisie de données en format matriciel -- Mode d'emploi*. Récupéré le 28/03/2022 de <https://semato.uqam.ca/guidexpert-ato/analyse-prototypique-mode-emploi.asp>
- Powell, A., Stratton, G., & Cameron, R. (2018). *Digital criminology: Crime and justice in digital society*. Routledge.
- Quadrature du net (2023, 17 octobre). *L'Assemblée adopte l'activation à distance des appareils électroniques*. La quadrature du net. <https://www.laquadrature.net/2023/07/25/lassemblee-adopte-lactivation-a-distance-des-appareils-electroniques/>
- Pucheral, P., Rallet, A., Rochelandet, F., & Zolynski, C. (2016). La Privacy by design : une fausse bonne solution aux problèmes de protection des données personnelles soulevés par l'Open data et les objets connectés? *Legicom*, 56(1), 89-99. <https://doi.org/10.3917/legi.056.0089>
- Rallet, A., Rochelandet, F., & Zolynski, C. (2015). De la Privacy by Design à la Privacy by Using. *Reseaux*, 189(1), 15-46. <https://doi.org/10.3917/res.189.0015>
- Rateau, P. (1995). Dimensions descriptive, fonctionnelle et évaluative des représentations sociales-une étude exploratoire. *Papers on Social Representations*, 4(1), 133-146.
- Rey, B. (2012). La privacy à l'ère du numérique. *Terminal. Technologie de l'information, culture & société*, (110), 91-103. <https://doi.org/10.4000/terminal.1242>

- Robert, L. (2021). *Définition du mot « artefact »*. Récupéré le 17/09/2021 de <https://www.lerobert.com/google-dictionnaire-fr?param=art%C3%A9fact>
- Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, 4(3), 328-350. [https://doi.org/10.1016/0010-0285\(73\)90017-0](https://doi.org/10.1016/0010-0285(73)90017-0)
- Rouquette, M.-L. (1994). *Sur la connaissance des masses : essai de psychologie politique*. Presses universitaires de Grenoble.
- Salesses, L. (2005a). Rôle du niveau de connaissance dans le processus de structuration d'une représentation sociale. *Les cahiers internationaux de psychologie sociale*, 66(2), 25-42. <https://doi.org/10.3917/cips.066.0025>
- Salesses, L. (2005b). Effet d'attitude dans le processus de structuration d'une représentation sociale. *Psychologie française*, 50(4), 471-485. <https://doi.org/10.1016/j.psfr.2005.06.002>
- Sanchez, M. A., Velut, G., Nivoix, P., Mayet, A., Dany, L., Meynard, J. B., Deparis, X., & Berger, F. (2016). Social representations of public health among healthcare workers within the French Armed Forces. *Revue d'épidémiologie et de santé publique*, 64(1), 7-14. <https://doi.org/10.1016/j.respe.2015.10.001>
- Schmitz, C., & LimeSurvey Project Team. (2012). *LimeSurvey: An Open Source survey tool (Version Cloud)*. LimeSurvey Project. <https://www.limesurvey.org/>
- Schoeman, F. D. (1984). *Philosophical dimensions of privacy: An anthology*. Cambridge University Press.
- Scopus. (2022). *Recherche Scopus "privacy"*. Récupéré le 04/01/2023 de <https://www-scopus-com>
- Signal Messenger. (2018). *Signal (Version 6.11.0)*. <https://signal.org/fr/>
- Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 35(4), 989-1016. <https://doi.org/10.2307/41409970>
- Solove, D. J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review*, 15(3) 477-564. <https://doi.org/10.2307/40041279>
- Solove, D. J. (2007). *The future of reputation: Gossip, rumor, and privacy on the Internet*. Yale University Press.

- Solove, D. J. (2021). The myth of the privacy paradox. *The George Washington Law Review*, 89(1), 1-51. <http://dx.doi.org/10.2139/ssrn.3536265>
- Staddon, J., Acquisti, A., & LeFevre, K. (2013, Septembre). *Self-reported social network behavior: Accuracy predictors and implications for the privacy paradox*. Communication présentée à 2013 International Conference on Social Computing Alexandria, Virginie, États-Unis. <https://doi.org/10.1109/SocialCom.2013.48>
- Steinfeld, N. (2016). "I agree to the terms and conditions":(How) do users read privacy policies online? An eye-tracking experiment. *Computers in Human Behavior*, 55(1), 992-1000. <https://doi.org/10.1016/j.chb.2015.09.038>
- Stephenson, W. (1953). *The study of behavior; Q-technique and its methodology*. University of Chicago Press
- Stuart, A., & Levine, M. (2017). Beyond 'nothing to hide': When identity is key to privacy threat under surveillance. *European Journal of Social Psychology*, 47(6), 694-707. <https://doi.org/10.1002/ejsp.2270>
- Tafani, E. (2001). Attitudes, engagement et dynamique des représentations sociales : études expérimentales. *Revue internationale de psychologie sociale*, 14(1), 7-30.
- Tafani, E., & Bellon, S. (2001). Principe d'homologie structurale et dynamique représentationnelle. Dans P. Moliner (Éd.), *La dynamique des représentations sociales* (pp. 163-194). Presses universitaires de Grenoble.
- Tafani, E., Bellon, S., & Apostolidis, T. (2002). Théorie des champs sociaux et dynamique représentationnelle : approche expérimentale des effets des asymétries positionnelles sur la structure d'une représentation sociale. *Revue internationale de psychologie sociale*, 15(2), 57-90.
- Tafani, E., Souchet, L., Codaccioni, C., & Mugny, G. (2003). Influences majoritaire et minoritaire sur la représentation sociale de la drogue. *Nouvelle revue de psychologie sociale*, 2(3), 343-354.
- Takabi, H., Joshi, J. B., & Ahn, G.-J. (2010). Security and privacy challenges in cloud computing environments. *IEEE Security & Privacy*, 8(6), 24-31. <https://doi.org/10.1109/MSP.2010.186>
- Thomson, R., Yuki, M., & Ito, N. (2015). A socio-ecological approach to national differences in online privacy concern: The role of relational mobility and trust. *Computers in Human Behavior*, 51(1), 285-292. <https://doi.org/10.1016/j.chb.2015.04.068>

- Tifferet, S. (2019). Gender differences in privacy tendencies on social network sites: A meta-analysis. *Computers in Human Behavior*, 93(1), 1-12. <https://doi.org/10.1016/j.chb.2018.11.046>
- Tourangeau, R. (2018). Confidentiality, privacy, and anonymity. Dans D. Vannette & J. Krosnick (Éds), *The Palgrave handbook of survey research* (pp. 501-507). Springer.
- Valade, B. (2012). De l'explication dans les sciences sociales : holisme et individualisme. Dans J.-M. Berthelot (Éd.), *Épistémologie des sciences sociales* (pp. 357-405). Presses universitaires de France.
- Vasilomanolakis, E., Daubert, J., Luthra, M., Gazis, V., Wiesmaier, A., & Kikiras, P. (2015). *On the security and privacy of Internet of Things architectures and systems*. Communication présentée à 2015 International Workshop on Secure Internet of Things (SIoT), Vienne, Autriche. <https://doi.org/10.1109/SIOT.2015.9>
- Vergès, P. (1992). L'évocation de l'argent : une méthode pour la définition du noyau central d'une représentation. *Bulletin de Psychologie*, 45(405), 203-209.
- Vergès, P., & Bouriche, B. (2001). *L'analyse des données par les graphes de similitude*. Sciences Humaines.
- Wachelke, J. (2013). Black and white or shades of grey? Comparing social representations centrality models. *Spanish Journal of Psychology*, 1(1), E33. <https://doi.org/10.1017/sjp.2013.16>
- Wagner, I., & Eckhoff, D. (2018). Technical privacy metrics: A systematic survey. *ACM Computing Surveys (CSUR)*, 51(3), 1-38. <https://doi.org/10.1145/3168389>
- Wagner, W., Duveen, G., Far, R., Jovchelovitch, S. L.-C., Markovà, I., & Rose, D. (1999). Theory and method of social representation. *Asian Journal of Social Psychology*, 2(1), 95-125. <https://doi.org/10.1111/1467-839X.00028>
- Waldman, A. E. (2020). Cognitive biases, dark patterns, and the 'privacy paradox'. *Current opinion in psychology*, 31, 105-109. <https://doi.org/10.1016/j.copsyc.2019.08.025>
- Weber, R. H. (2010). Internet of Things—New security and privacy challenges. *Computer Law & Security Review*, 26(1), 23-30. <https://doi.org/10.1016/j.clsr.2009.11.008>
- Weinstein, W. L. (1971). *The private and the free: A conceptual inquiry*. Atherton Press.
- Westin, A. F. (1967). *Privacy and freedom*. Atheneum.

- Whitman, J. Q. (2004). The two western cultures of privacy: Dignity versus liberty. *Yale Law Journal*, 113(6), 1151-1221. <https://doi.org/10.2307/4135723>
- Wolter, R. P. (2011). Effet du contexte dans l'évaluation d'un objet social : le cas d'une manifestation collective en France. *Temas em Psicologia*, 19(1), 139-146.
- Xanthos, A. (2020). 1980 : quand l'utilisateur rencontre l'ordinateur. *Études de lettres*(312), 55-58. <https://doi.org/10.4000/edl.2346>
- Xu, H., Dinev, T., Smith, J., & Hart, P. (2011). Information privacy concerns: Linking individual perceptions with institutional privacy assurances. *Journal of the Association for Information Systems*, 12(12). <https://doi.org/10.17705/1jais.00281>
- Xuan, Z., Salehi, N., Naranjit, S., Alwaalan, S., Voids, S., & Cosley, D. (2013, Mai). *The many faces of Facebook: Experiencing social media as performance, exhibition, and personal archive*. Communication présentée CHI 2013: Changing Perspectives, Paris, France.
- Young, A. L., & Quan-Haase, A. (2013). Privacy protection strategies on Facebook. *Information, Communication & Society*, 16(4), 479-500. <https://doi.org/10.1080/1369118x.2013.777757>
- Zbinden, A., Souchet, L., Girandola, F., & Bourg, G. (2011). Communication engageante et représentations sociales : une application en faveur de la protection de l'environnement et du recyclage. *Pratiques psychologiques*, 17(3), 285-299. <https://doi.org/10.1016/j.prps.2010.10.002>
- Zuboff, S. (2020). *L'âge du capitalisme de surveillance*. Zulma.
- Zureik, E., Harling Stalker, L. H., Smith, E., Lyon, D., & Chan, Y. E. (2010). *Surveillance, privacy, and the globalization of personal information: International comparisons*. McGill-Queen's Press-MQUP.

Appendice A
Message de recrutement

[Une participation de quelques minutes = 1 chance de gagner 1 des 5 cartes-cadeaux Amazon (Valeur : 20€)]

Bonjour,

Dans le cadre de mon doctorat, je réalise une étude qui vise le recueil des perceptions des étudiants sur les informations personnelles. Auriez-vous quelques minutes pour le remplir?

Lien : <https://privacyuqtr.limequery.com/793827?lang=fr>

En le remplissant, vous aurez peut-être la chance de gagner une des 5 cartes-cadeaux Amazon Valeur : 20€).

N.B. IL FAUT ABSOLUMENT ÊTRE ÉTUDIANT POUR LE REMPLIR!

Je vous remercie d'avance

Appendice B

Tableau 1. Résumé des variables sociodémographiques de la première
collecte de données

Tableau 1*Résumé des variables sociodémographiques de la première collecte de données*

Variable	Modalité	n	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Âge	Âge	366	99,19	22,46	5,61	21	21
	Manquant	1	0,27				
	Privé ¹	2	0,54				
Sexe	Femme	280	75,88				Femme
	Homme	83	22,49				
	Autre	3	0,81				
	Privé	3	0,81				
Statut marital	Célibataire	266	72,09				Célibataire
	Concubin	70	18,97				
	Marié	12	3,25				
	Pacsé	8	2,17				
	Privé	13	3,52				
Nationalité	Française	320	86,72				Française
	Étranger	24	6,50				
	Binationaux	23	6,23				
	Privé	2	0,54				
Niveau d'études	BAC	123	33,33				Licence
	Licence	143	38,75				
	Master	60	16,26				
	BTS/DUT	25	6,78				
	Privé	13	3,52				
	Doctorat	3	0,81				
Domaine d'études (n > 5 rapportés)	Psychologie	70	18,97				Psychologie
	Droit	38	10,30				
	Santé	28	7,59				
	Langue étrangère	24	6,50				
	Informatique	15	4,07				
	Biologie	14	3,79				
	Enseignement	13	3,52				
	Histoire	12	3,25				
	Comptabilité gestion	11	2,98				
	Lettre	11	2,98				
	Ingénierie	10	2,71				
	Mathématiques	8	2,17				
	Sociologie	8	2,17				
	Économie	8	2,17				
	Marketing	6	1,63				
	Manquant	2	0,54				

¹ Nous avons remplacé la modalité de réponse « Je ne souhaite pas répondre » par « Privé » afin d'être plus explicite.

Tableau 1*Résumé des variables sociodémographiques de la première collecte de données (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Intéressement politique	Peu intéressé	142	38,48				Peu intéressé
	Assez intéressé	136	36,86				
	Très intéressé	51	13,82				
	Pas du tout intéressé	36	9,76				
	Privé	4	1,08				
Orientation politique	Orientation	310	84,01	3,59	2,34	3	3
	Manquant	29	7,86				
	Privé	30	8,13				
Parti politique (<i>n</i> > 5)	Privé	98	26,56				Privé
	L'Europe Écologie Les Verts (EELV)	88	23,85				
	Autre	45	12,20				
	Parti Socialiste	43	11,65				
	La République En Marche	29	7,86				
	La France Insoumise	23	6,23				
	Les Républicains	17	4,61				
	Le Rassemblement national	8	2,17				
	Le Parti Communiste Français	7	1,90				
CSP. du père	Employés	70	18,97				Employés
	Cadres	69	18,7				
	Ouvriers	41	11,11				
	Professions intellectuelles supérieures	40	10,84				
	Retraités	37	10,03				
	Privé	27	7,32				
	Professions intermédiaires	23	6,23				
	Chefs d'entreprise	21	5,69				
	Artisans	15	4,07				
	Agriculteurs	12	3,25				
	Commerçants	7	1,9				
	Autres personnes sans activité professionnelle	7	1,9				

Tableau 1*Résumé des variables sociodémographiques de la première collecte de données (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
CSP. de la mère	Employés	115	31,17				Employés
	Cadres	51	13,82				
	Autres personnes sans activité professionnelle	49	13,28				
	Professions intermédiaires	48	13,01				
	Professions intellectuelles supérieures	40	10,84				
	Retraités	18	4,88				
	Ouvriers	12	3,25				
	Privé	12	3,25				
	Commerçants	11	2,98				
	Chefs d'entreprise	6	1,63				
	Agriculteurs	4	1,08				
	Artisans	3	0,81				
	Ressources financières annuelles	0-5000€	176	47,70			
5001-10000€		88	23,85				
Privé		32	8,67				
10001-15000€		28	7,59				
15001-20000€		14	3,79				
20001-25000€		13	3,52				
25001-30000€		9	2,44				
30001€ et plus		9	2,44				
Membre d'une association défendant la vie privée	Non	247	66,94				Non
	Manquant	118	31,98				
	Oui	2	0,54				
	Privé	2	0,54				
Connaissance de la CNIL	Manquant	118	31,98				Manquant
	Oui	98	26,56				
	Plutôt oui	63	17,07				
	Non	60	16,26				
	Plutôt non	22	5,96				
	Ni oui, ni non	8	2,17				
	Privé	0	0,00				

Tableau 1*Résumé des variables sociodémographiques de la première collecte de données (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Application des conseils de la CNIL	Manquant	118	31,98				Manquant
	Quelques-uns	105	28,46				
	La plupart	76	20,6				
	Aucun	35	9,49				
	Privé	29	7,86				
	Tous	6	1,63				
Connaissance du VPN	Oui	132	35,77				Oui
	Manquant	118	31,98				
	Plutôt oui	43	11,65				
	Non	40	10,84				
	Plutôt non	28	7,59				
	Ni oui, ni non	8	2,17				
Utilisation du VPN	Privé	0	0,00				Manquant
	Manquant	118	31,98				
	Jamais	117	31,71				
	Parfois	58	15,72				
	Rarement	41	11,11				
	Toujours	20	5,42				
Privé	15	4,07					

Appendice C

Tableau 2. Résumé des variables sociodémographiques de la seconde
collecte de données

Tableau 2*Résumé des variables sociodémographiques de la seconde collecte de données*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Âge	Âge	149	91,41	22,24	4,48	21	21
	Privé	2	1,23				
Sexe	Femme	112	68,71				Femme
	Homme	35	21,47				
	Autre	2	1,23				
	Privé	2	1,23				
Statut marital	Célibataire	111	68,10				Célibataire
	Concubin	28	17,18				
	Marié	2	1,23				
	Privé	10	6,13				
	Pacsé	0	0				
Nationalité	Française	123	75,46				Française
	Binationaux	15	9,20				
	Étranger	8	4,91				
	Privé	5	3,07				
Niveau d'études	BAC	54	33,13				Licence
	Licence	62	38,04				
	Master	29	17,79				
	Privé	3	1,84				
	BTS/DUT	2	1,23				
	Doctorat	1	0,61				
Domaine d'études (<i>n</i> > 5 rapportés)	Psychologie	70	18,97				Psychologie
	Droit	38	10,30				
	Santé	28	7,59				
	Langue étrangère	24	6,50				
	Informatique	15	4,07				
	Biologie	14	3,79				
	Enseignement	13	3,52				
	Histoire	12	3,25				
	Comptabilité	11	2,98				
	gestion						
	Lettre	11	2,98				
	Ingénierie	10	2,71				
	Mathématiques	8	2,17				
	Sociologie	8	2,17				
	Économie	8	2,17				
	Marketing	6	1,63				
Manquant	2	0,54					
Privé	2	0,54					

Tableau 2*Résumé des variables sociodémographiques de la seconde collecte de données (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Intéressement politique	Assez intéressé	56	34,36				Assez intéressé
	Peu intéressé	52	31,9				
	Très intéressé	23	14,11				
	Pas du tout intéressé	16	9,82				
	Manquant	12	7,36				
	Privé	4	2,45				
Orientation politique	Orientation	118	72,39	4,02	2,53	4	4
	Privé	24	14,72				
	Manquant	21	12,88				
Parti politique (<i>n</i> > 5)	Privé	72	44,17				Privé
	L'Europe Écologie Les Verts (EELV)	29	17,79				
	La République En Marche	13	7,98				
	Manquant	12	7,36				
	Autre	11	6,75				
	La France Insoumise	8	4,91				
	Parti Socialiste	7	4,29				
CSP. du père	Employés	33	20,25				Employés
	Privé	22	13,5				
	Cadres	21	12,88				
	Ouvriers	21	12,88				
	Professions intellectuelles supérieures	12	7,36				
	Professions intermédiaires	12	7,36				
	Manquant	12	7,36				
	Chefs d'entreprise	11	6,75				
	Artisans	6	3,68				
	Commerçants	6	3,68				
	Agriculteurs	4	2,45				
	Autres personnes sans activité professionnelle	3	1,84				

Tableau 2*Résumé des variables sociodémographiques de la seconde collecte de données (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
CSP. de la mère	Employés	42	25,77				Employés
	Professions intermédiaires	22	13,5				
	Privé	19	11,66				
	Cadres	18	11,04				
	Autres personnes sans activité professionnelle	16	9,82				
	Professions intellectuelles supérieures	15	9,2				
	Manquant	12	7,36				
	Commerçants	5	3,07				
	Ouvriers	5	3,07				
	Agriculteurs	5	3,07				
	Artisans	2	1,23				
	Chefs d'entreprise	2	1,23				
Ressources financières annuelles	0-5000€	62	38,04				0-5000€
	5001-10000€	32	19,63				
	Privé	31	19,02				
	Manquant	12	7,36				
	10001-15000€	10	6,13				
	15001-20000€	10	6,13				
	20001-25000€	2	1,23				
	25001-30000€	2	1,23				
30001€ et plus	2	1,23					
Membre d'une association défendant la vie privée	Non	147	90,18				Non
	Manquant	12	6,75				
	Oui	3	1,84				
	Privé	2	1,23				
Connaissance de la CNIL	Oui	60	36,81				Oui
	Plutôt oui	38	23,31				
	Ni oui, ni non	10	6,13				
	Plutôt non	15	9,20				
	Non	29	17,79				
	Manquant	11	6,75				
Privé	0	0,00					

Tableau 2*Résumé des variables sociodémographiques de la seconde collecte de données (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Application des conseils de la CNIL	La plupart	56	34,36				Quelques-uns
	Quelques-uns	59	36,20				
	Privé	21	12,88				
	Aucun	13	7,98				
	Manquant	11	6,75				
	Tous	3	1,84				
Connaissance du VPN	Oui	91	55,83				Oui
	Plutôt oui	31	19,02				
	Non	17	10,43				
	Manquant	11	6,75				
	Plutôt non	10	6,13				
	Ni oui, ni non	3	1,84				
	Privé	0	0				
Utilisation du VPN	Jamais	74	45,40				Jamais
	Parfois	35	21,47				
	Rarement	25	15,34				
	Manquant	11	6,75				
	Toujours	10	6,13				
	Privé	8	4,91				

Appendice D

Précaution à l'utilisation de l'analyse en composante principale

Précaution à l'utilisation de l'analyse en composante principale

Dans les articles originaux, les auteurs de la *Security Behavior Intentions Scale* (Egelman et Peer, 2015b) et les auteurs de la *Privacy Orientation Scale* (Baruh et Cemalcılar, 2014) utilisent l'ACP avec une rotation Varimax pour effectuer « l'extraction » des construits latents. Rappelons que dans ces deux articles, l'objectif poursuivi par les chercheurs était de déterminer la structure de leur questionnaire.

Cependant, la littérature déconseille l'utilisation de l'ACP dans ce cas (Martin et Savage-McGlynn, 2013), car elle est une méthode de réduction des données et non d'extraction des construits latents. Il est préférable d'effectuer une AFE pour l'extraction des construits. N'ayant pas accès aux données originales des auteurs, nous avons voulu malgré tout comparer nos résultats avec les leurs afin de valider les questionnaires. Pour ce faire, nous avons donc utilisé la même technique que les auteurs pour savoir si la structure factorielle des questionnaires était identique avec notre échantillon. Les résultats de l'AFE avec une méthode d'extraction en axe principale et une rotation Varimax sont également fournis.

Enfin, il nous semble important de conclure en mentionnant que des « diminutions » de pourcentage de variance expliquée peuvent s'observer entre l'ACP et l'AFE. Cela s'explique par le type de variance étudiée par chacune des analyses. En effet, l'AFE se base sur le calcul de la variance commune alors que l'ACP se base sur la variance totale qui comprend la variance commune, la variance spécifique ainsi que l'erreur de mesure.

Appendice E
Questionnaires des collectes de données

QUESTIONNAIRE DE LA PREMIÈRE COLLECTE DE DONNÉES

Étude sur les représentations sociales des étudiants universitaires dans un contexte d'usage des technologies de l'information et des communications

Votre participation à la recherche vise à mieux comprendre les représentations sociales de la vie privée des étudiants universitaires. Votre collaboration serait grandement appréciée.

Objectifs

Le but de cette lettre d'information est de vous aider à comprendre exactement ce qu'implique votre éventuelle participation à la recherche de sorte que vous puissiez prendre une décision éclairée à ce sujet. Prenez donc le temps de la lire attentivement et n'hésitez pas à poser toute question que vous jugerez utile.

Tâche

Votre participation consiste à répondre à plusieurs questions concernant la vie privée, de façon exhaustive. La durée de votre participation est d'environ une vingtaine de minutes et celle-ci s'effectuera en ligne.

Risques, inconvénients, inconforts

Aucun risque spécifique n'est associé à votre participation. Le temps consacré au projet, soit environ une vingtaine de minutes, demeure le seul inconvénient.

Bénéfices

En guise de compensation, les participants peuvent participer à un tirage au sort de 5 cartes-cadeaux d'une valeur de 20€ sur Amazon.fr.

De plus, vous contribuez à l'avancement des connaissances au sujet de la vie privée. Ceci représente le principal bénéfice associé à votre participation. Par ailleurs, le fait de participer à cette recherche vous offre également une occasion de réfléchir sur le concept de vie privée.

Confidentialité

Votre participation au projet est confidentielle et les résultats de la recherche, qui pourront être diffusés sous forme d'articles et de communications scientifiques, ne permettront pas d'identifier les participants.

Les données recueillies seront conservées sur un lecteur encrypté par mot de passe au local MS0084 de l'Université du Québec à Trois-Rivières (3351, boulevard des Forges, Trois-Rivières, Québec, G8Z 4M3 Canada). L'ordinateur concerné est protégé par un mot de passe et se trouve dans un local fermé à clé. Les seules personnes qui auront accès aux coordonnées des participants seront le chercheur et les superviseurs de la recherche. Toutes ces personnes ont signé un engagement à la confidentialité. Les noms et coordonnées des participants seront détruits 5 ans après la publication des résultats sous forme d'article scientifique. Les données anonymisées du

projet seront toutefois conservées pour servir dans le cadre de projets ultérieurs pour étudier le concept de vie privée.

Participation volontaire

Votre participation à cette étude se fait sur une base volontaire. Vous êtes entièrement libre de participer ou non, de refuser de répondre à certaines questions ou de vous retirer en tout temps sans préjudice et sans avoir à fournir d'explications.

Remerciement

Votre collaboration est précieuse. Nous l'apprécions et vous en remercions.

Informations additionnelles sur la recherche

Pour obtenir de plus amples renseignements ou pour toute question concernant ce projet de recherche, vous pouvez communiquer avec le chercheur responsable à l'adresse eudes.loubet@uqtr.ca ou contacter un des superviseurs de la recherche, soit Michael Cantinotti au (+1) 819-376-5011 poste 3503, soit Élias Rizkallah au (+1) 514-987-3000 poste 5266.

Question ou plainte concernant l'éthique de la recherche

Cette recherche est approuvée par le comité d'éthique de la recherche avec des êtres humains de l'Université du Québec à Trois-Rivières et un certificat portant le numéro CER-19-258-07.29 a été émis le 26 juin 2019.

Pour toute question ou plainte d'ordre éthique concernant cette recherche, vous devez communiquer avec la secrétaire du comité d'éthique de la recherche de l'Université du Québec à Trois-Rivières, par téléphone (+1) 819 376-5011, poste 2129 ou par courrier électronique CEREH@uqtr.ca.

CONSENTEMENT À PARTICIPER À LA RECHERCHE

Engagement de la chercheuse ou du chercheur

Nous, Eudes Loubet, Michael Cantinotti et Élias Rizkallah, nous engageons à procéder à cette étude conformément à toutes les normes éthiques qui s'appliquent aux projets comportant la participation de sujets humains.

***Obligatoire**

Consentement du participant*

Je confirme avoir lu et compris la lettre d'information au sujet du projet « Étude sur la perception de la vie privée chez des étudiants universitaires ». J'ai bien saisi les conditions, les risques et les bienfaits éventuels de ma participation. On a répondu à toutes mes questions à mon entière satisfaction. J'ai disposé de suffisamment de temps pour réfléchir à ma décision de participer ou non à cette recherche. Je comprends que ma participation est entièrement volontaire et que je peux décider de me retirer en tout temps, sans avoir à subir aucun préjudice.

- Je consens à participer à l'étude
- Je refuse de participer à l'étude

Si vous souhaitez participer au tirage de 5 cartes-cadeaux Amazon d'une valeur de 20€, veuillez indiquer un courrier électronique où vous contacter, dans l'éventualité où vous seriez parmi les gagnants :

QUESTIONNAIRE SOCIODEMOGRAPHIQUE

Quel est votre genre?

- Homme
- Femme
- Je ne souhaite pas répondre
- Autre : _____

Quel est votre âge?

- Je ne souhaite pas répondre

Quel est votre statut marital?

- Marié
- Concubinage
- Pacsé
- Célibataire
- Veuf
- Divorcé
- Je ne souhaite pas répondre

Quelle est votre nationalité?

- J'ai uniquement la nationalité française
- J'ai la nationalité française et une ou plusieurs autres nationalités
- J'ai une ou plusieurs nationalités différentes de la nationalité française
- Je ne souhaite pas répondre

Avez-vous déjà résidé (travail, étude, ...) dans un pays étranger plus de 6 mois?

- Oui
- Non
- Je ne souhaite pas répondre

Dans quel autre pays avez-vous vécu?

- Je ne souhaite pas répondre

Quel est le diplôme le plus élevé que vous avez obtenu?

- BTS/DUT
- Licence
- Master
- Doctorat
- Je ne souhaite pas répondre
- Autre

Dans quelle discipline étudiez-vous? (p. ex., Psychologie, Mathématique, Histoire, etc...)

-
- Je ne souhaite pas répondre

Dans quelle mesure diriez-vous que vous vous intéressez à la politique?

- Pas du tout intéressé
- Peu intéressé
- Assez intéressé
- Très intéressé
- Je ne souhaite pas répondre

En politique, les gens parlent parfois de « gauche » ou de « droite ». En utilisant la règle ci-dessous, où vous placeriez-vous sachant que 0 signifie la « gauche » et 10 signifie la « droite »?

0 ————— 10

- Je ne souhaite pas répondre

Quel est le parti politique qui correspond le plus à vos idées politiques?

- Le Parti Communiste Français (PCF)
- La France Insoumise (FI)
- L'Europe Écologie Les Verts (EELV)
- La République En Marche (LREM)
- Le Mouvement Démocrate (MODEM)
- Le Mouvement radical (MR)
- Les Républicains (LR)
- L'Union des Démocrates et Indépendants (UDI)

- Agir, la droite constructive (AGIR)
- Debout la France (DLF)
- Le Rassemblement national (RN)
- Les Patriotes (LP)
- Je ne souhaite pas répondre
- Autre : _____

Quelle est la catégorie socioprofessionnelle de votre père?

- Agriculteurs
- Artisans
- Commerçants
- Chefs d'entreprise
- Cadres
- Professions intellectuelles supérieures
- Professions intermédiaires
- Employés
- Ouvriers
- Retraités
- Autres personnes sans activité professionnelle
- Je ne souhaite pas répondre

Quelle est la catégorie socioprofessionnelle de votre mère?

- Agriculteurs
- Artisans
- Commerçants
- Chefs d'entreprise
- Cadres
- Professions intellectuelles supérieures
- Professions intermédiaires
- Employés
- Ouvriers
- Retraités
- Autres personnes sans activité professionnelle
- Je ne souhaite pas répondre

Dans quelle fourchette estimeriez-vous vos ressources financières annuelles? (salaire, aide parentale, bourse, etc.)

- 0-5000€
- 5001-10000€
- 10001-15000€
- 15001-20000€
- 20001-25000€
- 25001-30000€
- 30001€ et plus
- Je ne souhaite pas répondre

ASSOCIATION LIBRE

Dans un contexte d'usage de technologies de l'information et des communications, quels sont les 4 termes et/ou expressions qu'évoque pour vous « VIE PRIVÉE »?

TERME 1 : _____

TERME 2 : _____

TERME 3 : _____

TERME 4 : _____

ASSOCIATION HIERACHISEE

Veillez classer les termes et/ou expressions que vous venez d'écrire selon l'IMPORTANT que vous leur accordez pour caractériser « VIE PRIVÉE » dans un contexte d'usage de technologies de l'information et des communications.

Veillez donc placer en 1^{re} position le terme qui, selon votre expérience, est le plus important pour caractériser « VIE PRIVÉE » dans un contexte d'usage de technologies de l'information et des communications, et ainsi de suite jusqu'à la 4^e position où se trouvera le terme le moins important selon votre expérience pour caractériser « VIE PRIVÉE » dans un contexte d'usage de technologies de l'information et des communications.

TERME 1 (LE PLUS IMPORTANT) : _____

TERME 2 : _____

TERME 3 : _____

TERME 4 (LE PLUS IMPORTANT) : _____

ORIENTATION PRIVACY SCALE (TRADUCTION)

S'il vous plait, veuillez indiquer votre réponse à chacun des éléments ci-dessous en cochant la case de votre choix.

Les lois sur la vie privée devraient être renforcées pour protéger la vie privée des personnes

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Les gens ont besoin de protection légale contre l'usage inapproprié des données personnelles

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Si j'avais à rédiger une loi constitutionnelle aujourd'hui, j'inclurais probablement la vie privée comme droit fondamental

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quand je partage des informations sur ma vie personnelle avec quelqu'un, je m'inquiète qu'il pourrait les divulguer à d'autres personnes

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je me préoccupe du fait que les gens autour de moi savent trop de choses à propos de moi

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je me préoccupe des conséquences du partage d'informations liées à mon identité

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je m'inquiète de partager des informations avec plus de personnes que je ne le voudrais

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Si quelqu'un n'est pas prudent concernant la protection de sa propre vie privée, je ne peux pas lui faire confiance au sujet du respect de la mienne

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Pour que je puisse apprécier les aspects privés de ma vie, j'ai besoin que mes amis prennent soin de protéger leur vie privée aussi

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je ne pourrais jamais considérer une personne digne de confiance si elle distribue des détails sur sa propre vie privée sans discernement

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Le niveau de vie privée dont je peux profiter dépend du degré auquel les personnes autour de moi protègent leur propre vie privée

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Il est important pour moi de respecter la vie privée des individus, même s'ils ne font pas attention à la protection de leur propre vie privée

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je valorise la vie privée des autres personnes autant que la mienne

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Même si quelqu'un n'est pas prudent à propos de sa vie privée, je fais de mon mieux pour respecter la vie privée de cette personne

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je fais toujours du mieux que je peux pour ne pas envahir la vie privée des autres personnes

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Le respect de la vie privée d'autrui devrait être une priorité importante dans les relations sociales

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

SECURITY BEHAVIOR INTENTIONS SCALE (TRADUCTION)

S'il vous plait, veuillez indiquer votre réponse à chacun des éléments ci-dessous en cochant la case de votre choix.

J'ai prévu un verrouillage automatique de mon écran lorsque je ne l'utilise pas durant une période de temps prolongée

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'utilise un mot de passe pour verrouiller mon ordinateur ou ma tablette

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je verrouille manuellement mon écran d'ordinateur quand je m'en éloigne

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'utilise un code d'identification personnel ou un mot de passe pour déverrouiller mon téléphone intelligent

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je ne change pas mon mot de passe, à moins d'y être contraint

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'utilise différents mots de passe pour chacun des comptes que j'ai

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quand je crée un nouveau compte, j'essaye d'utiliser un mot de passe qui dépasse les exigences minimums du site

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je n'utilise pas de caractères spéciaux dans mon mot de passe si ce n'est pas obligatoire

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quand quelqu'un m'envoie un lien, je l'ouvre sans vérifier auparavant où il me dirige

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je reconnais le site que je visite en me basant sur son apparence plutôt que de regarder la barre d'adresse

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'envoie des informations à des sites Internet sans vérifier premièrement qu'elles sont envoyées de façon sécurisée (p. ex., SSL, « https:// », une icône d'un cadenas)

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quand je parcours des sites Internet, je survole les liens avec la souris pour voir où ils me dirigent avant de cliquer sur eux.

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Si je découvre un problème de sécurité, je continue à faire ce que je faisais parce que je considère que quelqu'un d'autre règlera le problème

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Lorsque je suis sollicité pour une mise à jour de logiciel, je l'installe tout de suite

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'essaie de m'assurer que les logiciels que j'utilise sont à jour

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je vérifie si mon antivirus se met régulièrement à jour par lui-même

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

AUTRES QUESTIONS

Faites-vous partie d'une association ayant dans ses missions la défense du droit à la vie privée des citoyens?

- Oui
- Non
- Je ne souhaite pas répondre

À quelle fréquence votre entourage vous a-t-il déjà fait des remarques à propos du fait que vous vous préoccupez de votre vie privée sur Internet?

- Jamais
- Rarement
- Parfois
- Souvent
- Tout le temps
- Je ne souhaite pas répondre

Connaissez-vous la CNIL (Commission nationale de l'informatique et des libertés)?

- Non
- Plutôt non
- Ni oui, ni non
- Plutôt oui
- Oui
- Je ne souhaite pas répondre

Appliquez-vous les conseils qu'elle délivre en matière de protection de la vie privée?

- Aucun
- Quelques-uns
- La plupart
- Tous
- Je ne souhaite pas répondre

Savez-vous ce qu'est un VPN (Virtual Personal Network)?

- Non
- Plutôt non
- Ni oui, ni non
- Plutôt oui
- Oui
- Je ne souhaite pas répondre

Utilisez-vous un VPN (Virtual Private Network)?

- Jamais
- Rarement
- Parfois
- Toujours
- Je ne souhaite pas répondre

Merci de votre participation. Si vous souhaitez faire un commentaire ou signaler une erreur, nous vous invitons à remplir l'espace suivant :

QUESTIONNAIRE DE LA SECONDE COLLECTE DE DONNÉES

Étude sur les représentations sociales des étudiants universitaires dans un contexte d'usage des technologies de l'information et des communications

Votre participation à la recherche vise à mieux comprendre les représentations sociales de la vie privée des étudiants universitaires. Votre collaboration serait grandement appréciée.

Objectifs

Le but de cette lettre d'information est de vous aider à comprendre exactement ce qu'implique votre éventuelle participation à la recherche de sorte que vous puissiez prendre une décision éclairée à ce sujet. Prenez donc le temps de la lire attentivement et n'hésitez pas à poser toute question que vous jugerez utile.

Tâche

Votre participation consiste à répondre à plusieurs questions concernant la vie privée, de façon exhaustive. La durée de votre participation est d'environ une vingtaine de minutes et celle-ci s'effectuera en ligne.

Risques, inconvénients, inconforts

Aucun risque spécifique n'est associé à votre participation. Le temps consacré au projet, soit environ une vingtaine de minutes, demeure le seul inconvénient.

Bénéfices

En guise de compensation, les participants peuvent participer à un tirage au sort de 5 cartes-cadeaux d'une valeur de 20€ sur Amazon.fr.

De plus, vous contribuez à l'avancement des connaissances au sujet de la vie privée. Ceci représente le principal bénéfice associé à votre participation. Par ailleurs, le fait de participer à cette recherche vous offre également une occasion de réfléchir sur le concept de vie privée.

Confidentialité

Votre participation au projet est confidentielle et les résultats de la recherche, qui pourront être diffusés sous forme d'articles et de communications scientifiques, ne permettront pas d'identifier les participants.

Les données recueillies seront conservées sur un lecteur encrypté par mot de passe au local MS0084 de l'Université du Québec à Trois-Rivières (3351, boulevard des Forges, Trois-Rivières, Québec, G8Z 4M3 Canada). L'ordinateur concerné est protégé par un mot de passe et se trouve dans un local fermé à clé. Les seules personnes qui auront accès aux coordonnées des participants seront le chercheur et les superviseurs de la recherche. Toutes ces personnes ont signé un engagement à la confidentialité. Les noms et coordonnées des participants seront détruits 5 ans après la publication des résultats sous forme d'article scientifique. Les données anonymisées du

projet seront toutefois conservées pour servir dans le cadre de projets ultérieurs pour étudier le concept de vie privée.

Participation volontaire

Votre participation à cette étude se fait sur une base volontaire. Vous êtes entièrement libre de participer ou non, de refuser de répondre à certaines questions ou de vous retirer en tout temps sans préjudice et sans avoir à fournir d'explications.

Remerciement

Votre collaboration est précieuse. Nous l'apprécions et vous en remercions.

Informations additionnelles sur la recherche

Pour obtenir de plus amples renseignements ou pour toute question concernant ce projet de recherche, vous pouvez communiquer avec le chercheur responsable à l'adresse eudes.loubet@uqtr.ca ou contacter un des superviseurs de la recherche, soit Michael Cantinotti au (+1) 819-376-5011 poste 3503, soit Élias Rizkallah au (+1) 514-987-3000 poste 5266.

Question ou plainte concernant l'éthique de la recherche

Cette recherche est approuvée par le comité d'éthique de la recherche avec des êtres humains de l'Université du Québec à Trois-Rivières et un certificat portant le numéro CER-19-258-07.29 a été émis le 26 juin 2019.

Pour toute question ou plainte d'ordre éthique concernant cette recherche, vous devez communiquer avec la secrétaire du comité d'éthique de la recherche de l'Université du Québec à Trois-Rivières, par téléphone (+1) 819 376-5011, poste 2129 ou par courrier électronique CEREH@uqtr.ca.

CONSENTEMENT À PARTICIPER À LA RECHERCHE

Engagement de la chercheuse ou du chercheur

Nous, Eudes Loubet, Michael Cantinotti et Élias Rizkallah, nous engageons à procéder à cette étude conformément à toutes les normes éthiques qui s'appliquent aux projets comportant la participation de sujets humains.

***Obligatoire**

Consentement du participant*

Je confirme avoir lu et compris la lettre d'information au sujet du projet « Étude sur la perception de la vie privée chez des étudiants universitaires ». J'ai bien saisi les conditions, les risques et les bienfaits éventuels de ma participation. On a répondu à toutes mes questions à mon entière satisfaction. J'ai disposé de suffisamment de temps pour réfléchir à ma décision de participer ou non à cette recherche. Je comprends que ma participation est entièrement volontaire et que je peux décider de me retirer en tout temps, sans avoir à subir aucun préjudice.

- Je consens à participer à l'étude
- Je refuse de participer à l'étude

Si vous souhaitez participer au tirage de 5 cartes-cadeaux Amazon d'une valeur de 20€, veuillez indiquer un courrier électronique où vous contacter, dans l'éventualité où vous seriez parmi les gagnants :

ASSOCIATION LIBRE

Dans un contexte d'usage de technologies de l'information et des communications, quels sont les 4 termes et/ou expressions qu'évoque pour vous « VIE PRIVÉE »?

TERME 1 : _____

TERME 2 : _____

TERME 3 : _____

TERME 4 : _____

ASSOCIATION HIÉRACHISÉE

Veillez classer les termes et/ou expressions que vous venez d'écrire selon l'IMPORTANT que vous leur accordez pour caractériser « VIE PRIVÉE » dans un contexte d'usage de technologies de l'information et des communications.

Veillez donc placer en 1^{re} position le terme qui, selon votre expérience, est le plus important pour caractériser « VIE PRIVÉE » dans un contexte d'usage de technologies de l'information et des communications, et ainsi de suite jusqu'à la 4^e position où se trouvera le terme le moins important selon votre expérience pour caractériser « VIE PRIVÉE » dans un contexte d'usage de technologies de l'information et des communications.

TERME 1 (LE PLUS IMPORTANT) : _____

TERME 2 : _____

TERME 3 : _____

TERME 4 (LE PLUS IMPORTANT) : _____

QUESTIONNAIRE DE CARACTÉRISATION

Voici une liste de 21 termes :

Ami, anonymat, choix, confidentialité, consentement, cookies, donnée personnelle, droit, famille, GAFAM (Google, Amazon, Facebook, Apple, Microsoft), intimité, liberté, mot de passe, photo, protection, relation, réseaux sociaux, respect, RGPD (règlement général sur la protection des données), secret, sécurité.

Effectuez les tâches suivantes dans l'ordre indiqué.

Maintenant et avant de passer à l'étape suivante, relisez les termes que vous avez choisis pour votre classement. Si vous le désirez, modifiez l'ordre dans lequel vous les avez agencés en gardant à l'esprit que plus un terme est haut dans le classement, plus il est caractéristique de la vie privée dans un contexte d'usage de technologies de l'information et des communications.

Tirez et glissez les 7 termes les plus caractéristiques de la vie privée dans un contexte d'usage de technologies de l'information et des communications et ordonnez-les de façon à ce que le 1^{er} terme le plus caractéristique se retrouve tout en haut de la liste, suivi par le 2^e terme le plus caractéristique et ainsi de suite. Enfin, terminez avec le 7^e terme le plus caractéristique qui se retrouvera tout en bas de la liste.

Le plus caractéristique	Ni plus ni moins	Le moins caractéristique
1.	1.	1.
2.	2.	2.
3.	3.	3.
4.	4.	4.
5.	5.	5.
6.	6.	6.
7.	7.	7.

TEST D'INDÉPENDANCE AU CONTEXTE

Nous allons vous soumettre plusieurs questions. Pour chacune d'entre elles, nous vous demandons de positionner le curseur sur le continuum qui correspond le mieux à votre pensée. Ce continuum va de : « Jamais, dans aucun cas » à « Toujours, dans tous les cas ».

Dans un contexte d'usage des technologies de l'information et des communications, ...

... la « Vie privée » concerne-t-elle le terme « Anonymat »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Choix »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Confidentialité »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Donnée personnelle »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Droit »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Intimité »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Liberté »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Respect »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Protection »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Secret »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

... la « Vie privée » concerne-t-elle le terme « Sécurité »?

Jamais, _____
dans aucun cas

Toujours,
dans tous les cas

ORIENTATION PRIVACY SCALE (TRADUCTION)

S'il vous plait, veuillez indiquer votre réponse à chacun des éléments ci-dessous en cochant la case de votre choix.

Les lois sur la vie privée devraient être renforcées pour protéger la vie privée des personnes

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Les gens ont besoin de protection légale contre l'usage inapproprié des données personnelles

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Si j'avais à rédiger une loi constitutionnelle aujourd'hui, j'inclurais probablement la vie privée comme droit fondamental

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quand je partage des informations sur ma vie personnelle avec quelqu'un, je m'inquiète qu'il pourrait les divulguer à d'autres personnes

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je me préoccupe du fait que les gens autour de moi savent trop de choses à propos de moi

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je me préoccupe des conséquences du partage d'informations liées à mon identité

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je m'inquiète de partager des informations avec plus de personnes que je ne le voudrais

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Si quelqu'un n'est pas prudent concernant la protection de sa propre vie privée, je ne peux pas lui faire confiance au sujet du respect de la mienne

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Pour que je puisse apprécier les aspects privés de ma vie, j'ai besoin que mes amis prennent soin de protéger leur vie privée aussi

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je ne pourrais jamais considérer une personne digne de confiance si elle distribue des détails sur sa propre vie privée sans discernement

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Le niveau de vie privée dont je peux profiter dépend du degré auquel les personnes autour de moi protègent leur propre vie privée

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Il est important pour moi de respecter la vie privée des individus, même s'ils ne font pas attention à la protection de leur propre vie privée

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je valorise la vie privée des autres personnes autant que la mienne

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Même si quelqu'un n'est pas prudent à propos de sa vie privée, je fais de mon mieux pour respecter la vie privée de cette personne

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je fais toujours du mieux que je peux pour ne pas envahir la vie privée des autres personnes

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Le respect de la vie privée d'autrui devrait être une priorité importante dans les relations sociales

Fortement en désaccord	En désaccord	Ni en accord, ni en désaccord	En accord	Fortement en accord
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

SECURITY BEHAVIOR INTENTIONS SCALE (TRADUCTION)

S'il vous plait, veuillez indiquer votre réponse à chacun des éléments ci-dessous en cochant la case de votre choix.

J'ai prévu un verrouillage automatique de mon écran lorsque je ne l'utilise pas durant une période de temps prolongée

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'utilise un mot de passe pour verrouiller mon ordinateur ou ma tablette

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je verrouille manuellement mon écran d'ordinateur quand je m'en éloigne

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'utilise un code d'identification personnel ou un mot de passe pour déverrouiller mon téléphone intelligent

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je ne change pas mon mot de passe, à moins d'y être contraint

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'utilise différents mots de passe pour chacun des comptes que j'ai

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quand je crée un nouveau compte, j'essaie d'utiliser un mot de passe qui dépasse les exigences minimums du site

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je n'utilise pas de caractères spéciaux dans mon mot de passe si ce n'est pas obligatoire

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quand quelqu'un m'envoie un lien, je l'ouvre sans vérifier auparavant où il me dirige

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je reconnais le site que je visite en me basant sur son apparence plutôt que de regarder la barre d'adresse

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'envoie des informations à des sites Internet sans vérifier premièrement qu'elles sont envoyées de façon sécurisée (p. ex., SSL, « https:// », une icône d'un cadenas)

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Quand je parcours des sites Internet, je survole les liens avec la souris pour voir où ils me dirigent avant de cliquer sur eux.

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Si je découvre un problème de sécurité, je continue à faire ce que je faisais parce que je considère que quelqu'un d'autre règlera le problème

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Lorsque je suis sollicité pour une mise à jour de logiciel, je l'installe tout de suite

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

J'essaye de m'assurer que les logiciels que j'utilise sont à jour

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Je vérifie si mon antivirus se met régulièrement à jour par lui-même

Jamais	Rarement	Parfois	Souvent	Toujours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

AUTRES QUESTIONS

Faites-vous partie d'une association ayant dans ses missions la défense du droit à la vie privée des citoyens?

- Oui
- Non
- Je ne souhaite pas répondre

À quelle fréquence votre entourage vous a-t-il déjà fait des remarques à propos du fait que vous vous préoccupez de votre vie privée sur Internet?

- Jamais
- Rarement
- Parfois
- Souvent
- Tout le temps
- Je ne souhaite pas répondre

Connaissez-vous la CNIL (Commission nationale de l'informatique et des libertés)?

- Non
- Plutôt non
- Ni oui, ni non
- Plutôt oui
- Oui
- Je ne souhaite pas répondre

Appliquez-vous les conseils qu'elle délivre en matière de protection de la vie privée?

- Aucun
- Quelques-uns
- La plupart
- Tous
- Je ne souhaite pas répondre

Savez-vous ce qu'est un VPN (Virtual Personal Network)?

- Non
- Plutôt non
- Ni oui, ni non
- Plutôt oui
- Oui
- Je ne souhaite pas répondre

Utilisez-vous un VPN (Virtual Private Network)?

- Jamais
- Rarement
- Parfois
- Toujours
- Je ne souhaite pas répondre

QUESTIONNAIRE SOCIODÉMOGRAPHIQUE

Quel est votre genre?

- Homme
- Femme
- Je ne souhaite pas répondre
- Autre : _____

Quel est votre âge?

- _____
- Je ne souhaite pas répondre

Quel est votre statut marital?

- Marié
- Concubinage
- Pacsé
- Célibataire
- Veuf
- Divorcé
- Je ne souhaite pas répondre

Quelle est votre nationalité?

- J'ai uniquement la nationalité française
- J'ai la nationalité française et une ou plusieurs autres nationalités
- J'ai une ou plusieurs nationalités différentes de la nationalité française
- Je ne souhaite pas répondre

Avez-vous déjà résidé (travail, étude, ...) dans un pays étranger plus de 6 mois?

- Oui
- Non
- Je ne souhaite pas répondre

Dans quel autre pays avez-vous vécu?

- _____
- Je ne souhaite pas répondre

Quel est le diplôme le plus élevé que vous avez obtenu?

- BTS/DUT
- Licence
- Master
- Doctorat
- Je ne souhaite pas répondre
- Autre

Dans quelle discipline étudiez-vous? (p. ex., Psychologie, Mathématique, Histoire, etc...)

-
- Je ne souhaite pas répondre

Dans quelle mesure diriez-vous que vous vous intéressez à la politique?

- Pas du tout intéressé
- Peu intéressé
- Assez intéressé
- Très intéressé
- Je ne souhaite pas répondre

En politique, les gens parlent parfois de « gauche » ou de « droite ». En utilisant la règle ci-dessous, où vous placeriez-vous sachant que 0 signifie la « gauche » et 10 signifie la « droite »?

0 ————— 10

- Je ne souhaite pas répondre

Quel est le parti politique qui correspond le plus à vos idées politiques?

- Le Parti Communiste Français (PCF)
- La France Insoumise (FI)
- L'Europe Écologie Les Verts (EELV)
- La République En Marche (LREM)
- Le Mouvement Démocrate (MODEM)
- Le Mouvement radical (MR)
- Les Républicains (LR)
- L'Union des Démocrates et Indépendants (UDI)

- Agir, la droite constructive (AGIR)
- Debout la France (DLF)
- Le Rassemblement national (RN)
- Les Patriotes (LP)
- Je ne souhaite pas répondre
- Autre : _____

Quelle est la catégorie socioprofessionnelle de votre père?

- Agriculteurs
- Artisans
- Commerçants
- Chefs d'entreprise
- Cadres
- Professions intellectuelles supérieures
- Professions intermédiaires
- Employés
- Ouvriers
- Retraités
- Autres personnes sans activité professionnelle
- Je ne souhaite pas répondre

Quelle est la catégorie socioprofessionnelle de votre mère?

- Agriculteurs
- Artisans
- Commerçants
- Chefs d'entreprise
- Cadres
- Professions intellectuelles supérieures
- Professions intermédiaires
- Employés
- Ouvriers
- Retraités
- Autres personnes sans activité professionnelle
- Je ne souhaite pas répondre

Dans quelle fourchette estimeriez-vous vos ressources financières annuelles? (salaire, aide parentale, bourse, etc.)

- 0-5000€
- 5001-10000€
- 10001-15000€
- 15001-20000€
- 20001-25000€
- 25001-30000€
- 30001€ et plus
- Je ne souhaite pas répondre

Merci de votre participation. Si vous souhaitez faire un commentaire ou signaler une erreur, nous vous invitons à remplir l'espace suivant :

Appendice F

Tableau 11. Résumé des variables sociodémographiques du questionnaire
Privacy Orientation Scale

Tableau 11*Résumé des variables sociodémographiques du questionnaire
Privacy Orientation Scale*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Âge	Âge	399	97,79	22,21	4,71	21	21
	Manquant	4	0,98				
	Privé	3	0,74				
Sexe	Femme	307	75,25				Femme
	Homme	90	22,06				
	Autre	5	1,23				
	Manquant	3	0,74				
	Privé	3	0,74				
Statut marital	Célibataire	290	71,08				Célibataire
	Concubin	81	19,85				
	Privé	21	5,15				
	Marié	9	2,21				
	Pacsé	4	0,98				
	Manquant	3	0,74				
Nationalité	Française	346	84,80				Française
	Binationaux	32	7,84				
	Étranger	22	5,39				
	Privé	5	1,23				
	Manquant	3	0,74				
Niveau d'études	BAC	149	36,52				Licence
	Licence	159	38,97				
	Master	69	16,91				
	BTS/DUT	17	4,17				
	Privé	8	1,96				
	Doctorat	3	0,74				
Domaine spécifique (<i>n</i> > 15)	Psychologie	72	17,65				Psychologie
	Droit	49	12,01				
	Santé	32	7,84				
	Biologie	24	5,88				
	Langue étrangère	22	5,39				
	Informatique	16	3,92				
	Manquant	3	0,74				
	Privé	4	0,98				

Tableau 11

*Résumé des variables sociodémographiques du questionnaire
Privacy Orientation Scale (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Intéressement politique	Peu intéressé	140	34,31				Assez intéressé
	Assez intéressé	162	39,71				
	Très intéressé	58	14,22				
	Pas du tout intéressé	39	9,56				
	Manquant	3	0,74				
	Privé	6	1,47				
Orientation politique (un score de 0 représente un positionnement à « gauche » et 10 à « droite »)	Orientation	335	82,11	3,74	2,38	4	3
	Privé	40	9,80				
	Manquant	33	8,09				
Parti politique (<i>n</i> > 5)	Privé	131	32,11				Privé
	L'Europe Écologie Les Verts (EELV)	92	22,55				
	Autre	44	10,78				
	La République En Marche	37	9,07				
	Parti Socialiste	36	8,82				
	La France Insoumise	26	6,37				
	Manquant	3	0,74				
CSP. de la mère	Employés	122	29,9				Employés
	Professions intermédiaires	59	14,46				
	Cadres	52	12,75				
	Autres personnes sans activité professionnelle	49	12,01				
	Professions intellectuelles supérieures	43	10,54				
	Privé	26	6,37				
	Ouvriers	13	3,19				
	Commerçants	11	2,7				
	Retraités	10	2,45				
	Agriculteurs	9	2,21				
	Chefs d'entreprise	6	1,47				
	Artisans	5	1,23				
	Manquant	3	0,74				

Tableau 11

*Résumé des variables sociodémographiques du questionnaire
Privacy Orientation Scale (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
CSP. du père	Employés	76	18,63				Employés
	Cadres	71	17,4				
	Ouvriers	51	12,5				
	Privé	40	9,8				
	Professions intellectuelles supérieures	36	8,82				
	Professions intermédiaires	31	7,6				
	Chefs d'entreprise	25	6,13				
	Retraités	25	6,13				
	Agriculteurs	16	3,92				
	Artisans	15	3,68				
	Commerçants	11	2,7				
	Autres personnes sans activité professionnelle	8	1,96				
	Manquant	3	0,74				
Finance annuelle	0-5000€	188	46,08				0-5000€
	5001-10000€	94	23,04				
	10001-15000€	29	7,11				
	15001-20000€	19	4,66				
	20001-25000€	11	2,7				
	Privé	8	11,76				
	30001€ et plus	9	2,21				
	25001-30000€	7	1,72				
	Manquant	3	0,74				
Membre d'une association défendant la vie privée	Non	394	96,57				Non
	Oui	5	1,23				
	Manquant	5	1,23				
	Privé	4	0,98				
Connaissance de la CNIL	Oui	158	38,73				Oui
	Plutôt oui	101	24,75				
	Non	89	21,81				
	Plutôt non	37	9,07				
	Ni oui, ni non	18	4,41				
	Manquant	5	1,23				
Privé	0	0,00					

Tableau 11

*Résumé des variables sociodémographiques du questionnaire
Privacy Orientation Scale (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Application des conseils de la CNIL	Quelques-uns	164	40,2				Quelques-uns
	La plupart	132	32,35				
	Aucun	48	11,76				
	Privé	50	12,25				
	Tous	9	2,21				
	Manquant	5	1,23				
Connaissance du VPN	Oui	223	54,66				Oui
	Plutôt oui	74	18,14				
	Non	57	13,97				
	Plutôt non	38	9,31				
	Ni oui, ni non	11	2,7				
	Manquant	5	1,23				
	Privé	0	0				
Utilisation du VPN	Jamais	191	46,81				Jamais
	Parfois	93	22,79				
	Rarement	66	16,18				
	Toujours	30	7,35				
	Privé	23	5,64				
	Manquant	5	1,23				

Appendice G

Tableau 13. Résumé des variables sociodémographiques du questionnaire
Security Behavior Intentions Scale

Tableau 13*Résumé des variables sociodémographiques du questionnaire
Security Behavior Intentions Scale*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Âge	Âge	399	98,76	22,21	4,71	21	21
	Privé	3	0,74				
	Manquant	2	0,50				
Sexe	Femme	305	75,50				Femme
	Homme	90	22,28				
	Autre	5	1,24				
	Privé	3	0,74				
	Manquant	1	0,25				
Statut marital	Célibataire	289	71,53				Célibataire
	Concubin	81	20,05				
	Privé	20	4,95				
	Marié	9	2,23				
	Pacsé	4	0,99				
	Manquant	1	0,25				
Nationalité	Française	344	85,15				Française
	Binationaux	32	7,92				
	Étranger	22	5,45				
	Privé	5	1,24				
	Manquant	1	0,25				
Niveau d'études	BAC	147	36,39				Licence
	Licence	159	39,36				
	Master	69	17,08				
	BTS/DUT	17	4,21				
	Privé	8	1,98				
	Doctorat	3	0,74				
Domaine spécifique (<i>n</i> > 15)	Psychologie	72	17,82				Psychologie
	Droit	49	12,13				
	Santé	32	7,92				
	Biologie	24	5,94				
	Langue étrangère	22	5,45				
	Informatique	14	3,47				
	Privé	3	0,74				
	Manquant	1	0,25				

Tableau 13

*Résumé des variables sociodémographiques du questionnaire
Security Behavior Intentions Scale (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Intéressement politique	Assez intéressé	162	40,1				Assez intéressé
	Peu intéressé	139	34,41				
	Très intéressé	58	14,36				
	Pas du tout intéressé	38	9,41				
	Privé	6	1,49				
	Manquant	1	0,25				
Orientation politique (un score de 0 représente un positionnement à « gauche » et 10 à « droite »)	Orientation	331	81,93	3,74	2,38	4	3
	Privé	40	9,9				
	Manquant	33	8,17				
Parti politique (<i>n</i> > 5)	Privé	130	32,18				Privé
	L'Europe Écologie Les Verts (EELV)	91	22,52				
	Autre	44	10,89				
	Parti Socialiste	36	8,91				
	La République En Marche	37	9,16				
	La France Insoumise	26	6,44				
	Manquant	1	0,25				
CSP. de la mère	Employés	122	30,2				Employés
	Professions intermédiaires	59	14,6				
	Cadres	51	12,62				
	Autres personnes sans activité professionnelle	48	11,88				
	Professions intellectuelles supérieures	43	10,64				
	Privé	26	6,44				
	Ouvriers	13	3,22				
	Commerçants	11	2,72				
	Retraités	10	2,48				
	Agriculteurs	9	2,23				
	Chefs d'entreprise	6	1,49				
	Artisans	5	1,24				
	Manquant	1	0,25				

Tableau 13

*Résumé des variables sociodémographiques du questionnaire
Security Behavior Intentions Scale (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
CSP. du père	Employés	76	18,81				Employés
	Cadres	71	17,57				
	Ouvriers	51	12,62				
	Privé	39	9,65				
	Professions intellectuelles supérieures	35	8,66				
	Professions intermédiaires	31	7,67				
	Retraités	25	6,19				
	Chefs d'entreprise	25	6,19				
	Agriculteurs	16	3,96				
	Artisans	15	3,71				
	Commerçants	11	2,72				
	Autres personnes sans activité professionnelle	8	1,98				
	Manquant	1	0,25				
	Finance annuelle	0-5000€	187	46,29			
5001-10000€		94	23,27				
Privé		47	11,63				
10001-15000€		29	7,18				
15001-20000€		19	4,7				
20001-25000€		11	2,72				
30001€ et plus		9	2,23				
25001-30000€		7	1,73				
Manquant		1	0,25				
Membre d'une association défendant la vie privée	Non	394	97,52				Non
	Oui	5	1,24				
	Privé	4	0,99				
	Manquant	1	0,25				
Connaissance de la CNIL	Oui	158	39,11				Oui
	Plutôt oui	101	25				
	Non	89	22,03				
	Plutôt non	37	9,16				
	Ni oui, ni non	18	4,46				
	Manquant	1	0,25				
Privé	0	0					

Tableau 13

*Résumé des variables sociodémographiques du questionnaire
Security Behavior Intentions Scale (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Application des conseils de la CNIL	Quelques-uns	164	40,59				Quelques-uns
	La plupart	132	32,67				
	Privé	50	12,38				
	Aucun	48	11,88				
	Tous	9	2,23				
	Manquant	1	0,25				
Connaissance du VPN	Oui	223	55,2				Oui
	Plutôt oui	74	18,32				
	Non	57	14,11				
	Plutôt non	38	9,41				
	Ni oui, ni non	11	2,72				
	Manquant	1	0,25				
	Privé	0	0				
Utilisation du VPN	Jamais	191	47,28				Jamais
	Parfois	93	23,02				
	Rarement	66	16,34				
	Toujours	30	7,43				
	Privé	23	5,69				
	Manquant	1	0,25				

Appendice H

Tableau 14. Résumé des variables sociodémographiques du
Test d'indépendance au contexte

Tableau 14*Résumé des variables sociodémographiques du Test d'indépendance au contexte*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Âge	Âge	149	91,41	22,24	4,48	21	21
	Manquant	12	7,36				
	Privé	2	1,23				
Sexe	Femme	112	68,71				Femme
	Homme	35	21,47				
	Manquant	12	7,36				
	Autre	2	1,23				
	Privé	2	1,23				
Statut marital	Célibataire	111	68,10				Célibataire
	Concubin	28	17,18				
	Manquant	12	7,36				
	Privé	10	6,13				
	Marié	2	1,23				
	Pacsé	0	0,00				
Nationalité	Française	123	75,46				Française
	Binationaux	15	9,20				
	Manquant	12	7,36				
	Étranger	8	4,91				
	Privé	5	3,07				
Niveau d'études	BAC	54	33,13				Licence
	Licence	62	38,04				
	Master	29	17,79				
	Manquant	12	7,36				
	Privé	3	1,84				
	BTS/DUT	2	1,23				
	Doctorat	1	0,61				
Domaine spécifique (<i>n</i> > 6)	Psychologie	25	15,34				Psychologie
	Droit	19	11,66				
	Biologie	15	9,20				
	Manquant	12	7,36				
	Santé	9	5,52				
	Privé	3	1,84				

Tableau 14*Résumé des variables sociodémographiques du Test d'indépendance au contexte (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Intéressement politique	Assez intéressé	56	34,36				Assez intéressé
	Peu intéressé	52	31,9				
	Très intéressé	23	14,11				
	Pas du tout intéressé	16	9,82				
	Manquant	12	7,36				
	Privé	4	2,45				
Orientation politique (un score de 0 représente un positionnement à « gauche » et 10 à « droite »)	Orientation	118	72,39	4,02	2,33	4	4
	Privé	24	14,72				
	Manquant	21	12,88				
Parti politique (<i>n</i> > 5)	Privé	72	44,17				Privé
	L'Europe Écologie Les Verts (EELV)	29	17,79				
	Manquant	12	7,36				
	Autre	11	6,75				
	La République En Marche	13	7,98				
	La France Insoumise	8	4,91				
CSP. de la mère	Parti Socialiste	7	4,29				Employés
	Employés	42	25,77				
	Professions intermédiaires	22	13,5				
	Privé	19	11,66				
	Cadres	18	11,04				
	Autres personnes sans activité professionnelle	16	9,82				
	Professions intellectuelles supérieures	15	9,2				
	Manquant	12	7,36				
	Agriculteurs	5	3,07				
	Commerçants	5	3,07				
	Ouvriers	5	3,07				
	Artisans	2	1,23				
	Chefs d'entreprise	2	1,23				
	Retraités	0	0				

Tableau 14*Résumé des variables sociodémographiques du Test d'indépendance au contexte (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode	
CSP. du père	Employés	33	20,25				Employés	
	Privé	21	12,88					
	Cadres	21	12,88					
	Ouvriers	21	12,88					
	Professions intellectuelles supérieures	12	7,36					
	Professions intermédiaires	12	7,36					
	Manquant	12	7,36					
	Chefs d'entreprise	11	6,75					
	Commerçants	6	3,68					
	Artisans	6	3,68					
	Agriculteurs	4	2,45					
	Autres personnes sans activité professionnelle	3	1,84					
	Finance annuelle	0-5000€	62	38,04				0-5000€
		5001-10000€	32	19,63				
Privé		31	19,02					
Manquant		12	7,36					
10001-15000€		10	6,13					
15001-20000€		10	6,13					
20001-25000€		2	1,23					
25001-30000€		2	1,23					
30001€ et plus	2	1,23						
Membre d'une association défendant la vie privée	Non	147	90,18				Non	
	Manquant	11	6,75					
	Oui	3	1,84					
	Privé	2	1,23					
Connaissance de la CNIL	Oui	60	36,81				Oui	
	Plutôt oui	38	23,31					
	Non	29	17,79					
	Plutôt non	15	9,2					
	Manquant	11	6,75					
	Ni oui, ni non	10	6,13					
Privé	0	0						

Tableau 14*Résumé des variables sociodémographiques du Test d'indépendance au contexte (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Application des conseils de la CNIL	Quelques-uns	59	36,2				Quelques-uns
	La plupart	56	34,36				
	Privé	21	12,88				
	Aucun	13	7,98				
	Manquant	11	6,75				
	Tous	3	1,84				
Connaissance du VPN	Oui	91	55,83				Oui
	Plutôt oui	31	19,02				
	Non	17	10,43				
	Manquant	11	6,75				
	Plutôt non	10	6,13				
	Ni oui, ni non	3	1,84				
	Privé	0	0				
Utilisation du VPN	Jamais	74	45,4				Jamais
	Parfois	35	21,47				
	Rarement	25	15,34				
	Manquant	11	6,75				
	Toujours	10	6,13				
	Privé	8	4,91				

Appendice I

Tableau 16. Résumé des variables sociodémographiques du
Questionnaire de caractérisation

Tableau 16*Résumé des variables sociodémographiques du Questionnaire de caractérisation*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Âge	Âge	149	86,63	22,24	4,48	21	21
	Manquant	21	12,21				
	Privé	2	1,16				
Sexe	Femme	112	65,12				Femme
	Homme	35	20,35				
	Manquant	21	12,21				
	Autre	2	1,16				
	Privé	2	1,16				
Statut marital	Célibataire	111	64,53				Célibataire
	Concubin	28	16,28				
	Manquant	21	12,21				
	Privé	10	5,81				
	Marié	2	1,16				
	Pacsé	0	0,00				
Nationalité	Française	123	71,51				Française
	Manquant	21	12,21				
	Binationaux	15	8,72				
	Étranger	8	4,65				
	Privé	5	2,91				
Niveau d'études	BAC	54	31,40				Licence
	Licence	62	36,05				
	Master	29	16,86				
	Manquant	21	12,21				
	Privé	3	1,74				
	BTS/DUT	2	1,16				
	Doctorat	1	0,58				
Domaine spécifique (<i>n</i> > 6)	Psychologie	25	14,53				Psychologie
	Manquant	21	12,21				
	Droit	19	11,05				
	Biologie	15	8,72				
	Santé	9	5,23				
	Privé	3	1,74				

Tableau 16

Résumé des variables sociodémographiques du Questionnaire de caractérisation (suite)

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenn e	Écart- type	Médiane	Mode
Intéressement politique	Peu intéressé	52	30,23				Assez intéressé
	Assez intéressé	56	32,56				
	Très intéressé	23	13,37				
	Manquant	21	12,21				
	Pas du tout intéressé	16	9,30				
	Privé	4	2,33				
Orientation politique (un score de 0 représente un positionnement à « gauche » et 10 à « droite »)	Orientation	127	73,84	4,02	2,33	4	4
	Privé	24	13,95				
	Manquant	21	12,21				
Parti politique (<i>n</i> > 5)	Privé	72	41,86				Privé
	L'Europe Écologie Les Verts (EELV)	29	16,86				
	La République En Marche	13	7,56				
	Manquant	12	6,98				
	Autre	11	6,40				
	La France	8	4,65				
	Insoumise						
Parti Socialiste	7	4,07					
CSP. de la mère	Employés	42	24,42				Employés
	Professions intermédiaires	22	12,79				
	Privé	19	11,05				
	Cadres	18	10,47				
	Autres personnes sans activité professionnelle	16	9,3				
	Professions intellectuelles supérieures	15	8,72				
	Manquant	12	6,98				
	Agriculteurs	5	2,91				
	Commerçants	5	2,91				
	Ouvriers	5	2,91				
	Artisans	2	1,16				
	Chefs d'entreprise	2	1,16				
	Retraités	0	0				

Tableau 16*Résumé des variables sociodémographiques du Questionnaire de caractérisation (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode		
CSP. du père	Employés	33	19,19				Employés		
	Cadres	21	12,21						
	Ouvriers	21	12,21						
	Privé	21	12,21						
	Professions intellectuelles supérieures	12	6,98						
	Professions intermédiaires	12	6,98						
	Manquant	12	6,98						
	Chefs d'entreprise	11	6,4						
	Commerçants	6	3,49						
	Artisans	6	3,49						
	Agriculteurs	4	2,33						
	Autres personnes sans activité professionnelle	3	1,74						
	Finance annuelle	0-5000€	62	36,05					0-5000€
		5001-10000€	32	18,60					
Privé		31	18,02						
Manquant		12	6,98						
10001-15000€		10	5,81						
15001-20000€		10	5,81						
20001-25000€		2	1,16						
25001-30000€		2	1,16						
30001€ et plus		2	1,16						
Membre d'une association défendant la vie privée	Non	147	85,47				Non		
	Manquant	20	11,63						
	Oui	3	1,74						
	Privé	2	1,16						

Tableau 16*Résumé des variables sociodémographiques du Questionnaire de caractérisation (suite)*

Variable	Modalité	<i>n</i>	Fréquence (%)	Moyenne	Écart-type	Médiane	Mode
Connaissance de la CNIL	Oui	60	34,88				Oui
	Plutôt oui	38	22,09				
	Non	29	16,86				
	Manquant	20	11,63				
	Plutôt non	15	8,72				
	Ni oui, ni non	10	5,81				
	Privé	0	0,00				
Application des conseils de la CNIL	Quelques-uns	59	34,30				Quelques-uns
	La plupart	56	32,56				
	Privé	21	12,21				
	Manquant	20	6,40				
	Aucun	13	7,56				
	Tous	3	1,74				
Connaissance du VPN	Oui	91	52,91				Oui
	Plutôt oui	31	18,02				
	Manquant	20	11,63				
	Non	17	9,88				
	Plutôt non	10	5,81				
	Ni oui, ni non	3	1,74				
	Privé	0	0,00				
Utilisation du VPN	Jamais	74	43,02				Jamais
	Parfois	35	20,35				
	Rarement	25	14,53				
	Manquant	20	11,63				
	Toujours	10	5,81				
	Privé	8	4,65				