



HAL
open science

L'exploration visuelle des pages Web : de la variabilité interindividuelle aux patterns d'exploration

Gautier Drusch

► **To cite this version:**

Gautier Drusch. L'exploration visuelle des pages Web : de la variabilité interindividuelle aux patterns d'exploration. Sociologie. Université de Lorraine, 2017. Français. NNT : 2017LORR0403 . tel-03534188

HAL Id: tel-03534188

<https://hal.univ-lorraine.fr/tel-03534188>

Submitted on 19 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Université de Lorraine
École Doctorale Stanislas



THÈSE

pour obtenir le grade de
Docteur en Ergonomie

L'EXPLORATION VISUELLE DES PAGES WEB : DE LA VARIABILITÉ INTER-INDIVIDUELLE AUX PATTERNS D'EXPLORATION

présentée par

Gautier DRUSCH

le 18 décembre 2017

dirigée par

J.M. Christian BASTIEN, Université de Lorraine

Membres du jury :

J.M. Christian BASTIEN	Directeur	Professeur	Université de Lorraine
Jean-Michel BOUCHEIX	Rapporteur	Professeur	Université Bourgogne Franche-Comté
Anne BOYER	Examineur	Directeur de recherche	Université de Lorraine
Christophe KOLSKI	Rapporteur	Professeur	Université de Valenciennes et du Hainaut-Cambrésis
Charles TIJUS	Examineur	Professeur	Université Paris 8

Remerciements

Tout d'abord, je remercie chaleureusement mon directeur de thèse, Christian Bastien qui m'a toujours soutenu depuis notre rencontre, pour ses précieux conseils, sa disponibilité et son humanité. Toutes ces années de travail à ses côtés ont été un vrai plaisir.

Je remercie les membres du laboratoire PErSEUs pour leur sympathie et leur dévouement pendant toutes mes années d'études.

Je remercie également Briec Conan Guez pour l'aide qu'il m'a apportée concernant la classification automatique ainsi que Stéphane Paris pour ses conseils concernant la procédure Mean-Shift.

Pour toutes ces années d'amitié et de collaborations fructueuses, je remercie Mélanie Becker dont le parcours s'est déroulé presque parallèlement au miens. Merci à l'ensemble des participants des expériences de cette thèse et aux élèves de Master qui y ont contribué.

Ma reconnaissance va enfin à ma famille, et tout particulièrement Caroline, pour leur patience et leur amour.

Table des matières

I Les comportements visuels sur le web : connaissances et méthodes	11
1 L'étude des comportements visuels sur le Web	12
1.1 Mouvements oculaires et systèmes d'oculométrie	12
1.1.1 Pourquoi l'oculométrie ?	12
1.1.1.1 Des justifications théoriques	12
1.1.1.2 Des justifications pratiques	13
1.1.2 Les mouvements oculaires	14
1.1.2.1 Le réflexe optocinétique	14
1.1.2.2 Les fixations	14
1.1.2.3 Les saccades	15
1.1.3 La technique du reflet cornéen	16
1.1.4 Limites de la technique du reflet cornéen	17
1.1.5 Représentation des fixations et des saccades	18
2 Connaissances sur le comportement visuel des internautes sur le Web	21
2.1 Les comportements des utilisateurs sur les moteurs de recherche .	21
2.1.1 Aspects globaux des comportements visuels	21
2.1.2 L'importance des deux premiers résultats de recherche . .	22
2.1.3 L'Effet du rang du résultat de recherche sur les comporte- ments oculaires	23
2.1.4 Les styles de visualisation	23
2.2 L'influence des caractéristiques des pages Web	24
2.3 L'influence du type de site Web	26
2.4 L'influence des tâches	27
2.4.1 Le type de tâche	27
2.4.2 La répétition des tâches	28
2.5 Les caractéristiques des utilisateurs	29
2.5.1 Rôle du genre	29
2.5.2 Rôle de l'expérience antérieure	30

2.5.3	L'influence de l'âge	31
2.5.4	L'influence de la culture	32
2.6	Recherches basées sur la théorie du scanpath	33
2.7	Conclusions et problématique	34
3	Aspects méthodologiques de l'étude des comportements visuels	36
3.1	La définition des aires d'intérêts (<i>AOIs</i>)	36
3.1.1	Selon une grille uniforme	37
3.1.2	Par fonction	38
3.1.3	Par leur position dans la page	38
3.1.4	Selon le <i>DOM</i>	39
3.1.5	Selon la structure des données oculométriques : la procédure Mean-Shift	39
3.2	L'analyse des données	40
3.2.1	Approche statique	40
3.2.1.1	L'extraction de connaissances dans l'approche statique	41
3.2.1.2	Conclusions et limites de l'approche statique	42
3.2.2	Les méthodes de visualisation	42
3.2.3	Méthodes dynamiques	45
3.2.3.1	La distance de Levenshtein et ses variantes	46
3.2.3.2	L'algorithme « <i>MultiMatch</i> »	47
3.3	Conclusions sur les méthodes	49
II	A la recherche d'une méthode d'analyse	50
4	Différentes méthodes d'analyse pour les patterns	51
4.1	Méthodologie expérimentale	52
4.1.1	Matériel et terrain	52
4.1.2	Procédure	52
4.1.3	Questionnaire de fin d'expérience	57
4.1.4	Echantillon	57
4.2	Méthodologie de détection des patterns comportementaux	58
4.2.1	Généralités sur la classification automatique	58
4.2.2	Le partitionnement autour des médoïdes (PAM)	59
4.2.3	La caractérisation des classes	60
4.3	<i>Heatmaps</i> des données	60
4.4	Classification basée sur les indicateurs oculométriques	63
4.4.1	Analyse des données	63
4.4.1.1	Matrice de distances basée sur des indicateurs oculométriques	63

4.4.1.2	Partitionnement autour des médoïdes (<i>PAM</i>) . . .	63
4.4.2	Résultats	63
4.4.3	Conclusions	63
4.5	Classification basée sur les positions des fixations	64
4.5.1	Analyse des données	64
4.5.1.1	Matrice de distances basée sur la position des fixations	64
4.5.1.2	Partitionnement autour des médoïdes (<i>PAM</i>) . . .	66
4.5.2	Résultats	66
4.5.3	Conclusions	66
4.6	Classification basée sur la distance de Hausdorff	67
4.6.1	Analyse des données	67
4.6.1.1	La distance de Hausdorff	67
4.6.1.2	Partitionnement autour des médoïdes (<i>PAM</i>) . . .	68
4.6.2	Résultats	68
4.6.2.1	Caractérisation des classes	70
4.6.3	Conclusions	71
4.7	Analyse par T-patterns basée sur la procédure Mean-Shift	72
4.7.1	Détection automatique des aires d'intérêts	72
4.7.2	L'analyse par T-patterns	72
4.7.2.1	Principe général	72
4.7.2.2	Principe de détection	73
4.7.2.3	Paramètres de détection dans le logiciel Theme TM	73
4.7.3	Analyse des données et résultats	74
4.7.3.1	Préparation des données	74
4.7.3.2	La procédure Mean-Shift	74
4.7.3.3	L'analyse par T-patterns	74
4.7.4	Conclusions	76
4.8	Conclusions et limites	77
5	Proposition et évaluation d'une nouvelle méthode d'analyse	79
5.1	Problématiques liées à la définition des <i>AOIs</i>	79
5.2	La détection automatique des <i>AOIs</i>	82
5.2.1	Description formelle de la procédure <i>Mean-Shift</i>	83
5.2.2	Méthodologie expérimentale	84
5.2.2.1	Matériel	85
5.2.2.2	Variables	85
5.2.2.3	Protocole	86
5.2.3	Résultats	87
5.2.3.1	Effet de l'ordre des éléments	87
5.2.3.2	Effet de l'animation	88
5.2.3.3	Effet de la taille des éléments	89

5.2.3.4	Effet de la distance entre les éléments	91
5.2.4	Conclusions : limites de la procédure <i>Mean-Shift</i>	91
5.2.4.1	Le choix de la bande passante (paramètre h)	92
5.2.4.2	Le qualité de la classification	93
5.2.4.3	La détection des <i>outliers</i>	94
5.2.5	Proposition d'amélioration du <i>Mean-Shift</i> pour les données oculométriques (<i>HMS-ET</i>)	94
5.2.5.1	Construction de la hiérarchie	94
5.2.5.2	Détection des <i>outliers</i>	95
5.2.5.3	Sélection de la meilleure structure	95
5.2.6	Evaluation du <i>HMS-ET</i>	97
5.2.6.1	Détection des outliers	97
5.2.6.2	Sélection de la bande passante	99
5.2.6.3	Conclusion	100
5.2.7	Conclusions sur la détection automatique des AOIs	100
5.3	Définition d'une méthode pour l'analyse des patterns	101
5.3.1	Limites des méthodes actuelles de comparaisons des scan- paths	101
5.3.1.1	Les méthodes de visualisation	101
5.3.1.2	La distance de Levenshtein	101
5.3.1.3	La distance de Levenshtein pondérée par les dis- tances inter- <i>AOIs</i>	102
5.3.1.4	Les techniques de type « <i>Common subsequences</i> »	102
5.3.1.5	Transitions entre <i>AOIs</i>	103
5.3.1.6	Le MultiMatch	103
5.3.1.7	Conclusions	103
5.3.2	Proposition d'une méthode pour comparer les scanpaths	103
5.3.2.1	Etape 1 : détection des <i>AOIs</i> via le <i>HMS-ET</i>	104
5.3.2.2	Etape 2 : Labellisation des fixations	104
5.3.2.3	Etape 3 : Simplification des scanpaths	105
5.3.2.4	Etape 4 : Comparaison des scanpaths selon la méthode <i>SE-ET</i> (String Edit for Eye-Tracking)	105
5.3.3	Evaluation de la performance des méthodes	107
5.3.3.1	Implémentation des méthodes	108
5.3.3.2	Méthodologie	109
5.3.3.3	Résultats	110
5.3.3.4	Conclusions	112
5.4	Conclusion générale	112

III	Existent-ils des patterns de comportements visuels sur le Web ?	114
6	Expérience 1 : identification de patterns et étude de leur stabilité à court terme	116
6.1	Identification des patterns	116
6.1.1	Conclusions	123
6.2	Stabilité des patterns	124
6.2.1	Conclusions	129
6.3	Conclusions	130
7	Expérience 2 : étude longitudinale	131
7.1	Matériel	131
7.2	Stimuli	131
7.3	Procédure	132
7.4	Echantillon	133
7.5	Résultats	133
7.5.1	Identification des patterns comportementaux	133
7.5.2	Stabilité des patterns dans le temps	134
7.5.3	Stabilité entre les sites Web dans le temps	136
7.6	Conclusions	136
8	Expérience 3 : influence des caractéristiques des participants sur les patterns	138
8.1	Matériel	138
8.2	Stimulus	138
8.3	Echantillon	139
8.4	Procédure	139
8.5	Résultats	140
8.5.1	Identification des patterns	140
8.5.2	Etude de la stabilité	142
8.6	Conclusions	142
9	Expérience 4 : influence du design sur les patterns	143
9.1	Stimuli	143
9.2	Matériel	145
9.3	Procédure	145
9.4	Participants	147
9.5	Résultats	147
9.5.1	Education Nationale	147
9.5.2	Service Public	148
9.6	Conclusions	149

10 Discussion générale	150
10.1 Synthèse de la démarche	150
10.2 Discussion sur les apports méthodologiques	151
10.2.1 La définition des AOIs	151
10.2.2 L'identification des patterns	152
10.2.3 La caractérisation des patterns	153
10.3 Discussion sur les patterns de comportements visuels	154
10.3.1 A propos des patterns identifiés	154
10.3.2 A propos de la stabilité des patterns	156

Introduction générale

Les recherches en oculométrie du Web se concentrent autour de deux thématiques complémentaires. D'une part, l'oculométrie est utilisée pour l'évaluation des sites Web. Il s'agit ici le plus souvent de trouver les meilleurs indicateurs de performance oculométriques, censés refléter les problèmes d'utilisabilité d'une interface. Il est ainsi possible de comparer ces indicateurs sur deux interfaces différentes afin de décider laquelle est la plus adaptée aux utilisateurs (voir par exemple [Goldberg, 1999]). D'autre part, l'oculométrie est utilisée afin d'étudier le comportement des internautes dans le but de proposer des modèles comportementaux de référence. Ces travaux sont utiles à la fois d'un point de vue fondamental et pratique puisque des recommandations pour la conception peuvent ainsi être formulées. A cet égard, les études en oculométrie nous ont fourni certaines connaissances. L'étude concernant le Web qui est probablement la plus citée est celle de Nielsen qui a montré que les internautes explorent souvent les pages Web selon un pattern en forme de «*F*». Ainsi, les internautes balaièrent d'abord horizontalement le haut de la page, pour ensuite effectuer un plus petit balayage, lui aussi horizontal, pour enfin balayer verticalement la partie gauche de la page. Sur la base de ces travaux, Nielsen fournit des recommandations pour rédiger des pages Web où il suggère de placer les informations importantes dans les deux premiers paragraphes, et d'utiliser au maximum les listes de points.

Nous pensons que ce pattern «moyen» ou «agrégé» ne correspond en fait à aucun internaute. Néanmoins, cette approche fait sens pour les concepteurs qui cherchent des modèles comportementaux de référence, permettant de s'abstraire au moins en partie de la variabilité comportementale des utilisateurs. Finalement, entre la posture consistant à schématiser l'ensemble des comportements oculaires des internautes et celle consistant à envisager ces comportements comme étant, par essence, individuels, il y aurait une position intermédiaire envisageant les comportements comme des typologies de patterns. Ainsi, nous postulons qu'il existe des patterns comportementaux qui se définissent comme des schémas comportementaux plus ou moins similaires pour un ensemble d'individus. Si ce type de pattern existe, il convient en premier lieu de définir une méthode pour les détecter puis les analyser. En second lieu, il s'agit de savoir s'ils sont stables

de manière intra- et inter-individuelle dans des conditions similaires, puis enfin d'en connaître les déterminants.

L'objectif de cette thèse est double : (1) développer une méthode d'analyse permettant de détecter des patterns comportementaux oculaires sur les pages Web à partir des données oculométriques et (2) découvrir si de tels patterns existent.

Cette thèse se compose de trois parties. La première partie fait état des connaissances issues de l'oculométrie appliquée au Web puis des méthodes d'analyse utilisées. La deuxième décrit notre recherche concernant les méthodes d'analyse des patterns comportementaux, puis celle que nous avons développée et retenue. Enfin, la troisième partie rapporte les résultats de cette analyse par patterns sur différentes expérimentations.

Première partie

Les comportements visuels sur le web : connaissances et méthodes

Chapitre 1

L'étude des comportements visuels sur le Web

1.1 Mouvements oculaires et systèmes d'oculométrie

1.1.1 Pourquoi l'oculométrie ?

La technique de l'oculométrie est spectaculaire dans le sens premier du terme. En effet, personne ne reste insensible à la vue d'une exploration visuelle en temps réel ou encore à la visualisation des parcours d'exploration visuelle de pages Web. Mais tout aussi impressionnante soit-elle, cette technique n'est pas simple à mettre en œuvre et les données qu'elle génère sont souvent difficiles à analyser. Ceci dit, l'oculométrie offre évidemment des possibilités intéressantes lorsqu'elle est utilisée de façon pertinente.

1.1.1.1 Des justifications théoriques

- Le recours à la technique oculométrique repose sur deux postulats forts :
- il y aurait un système visuel commun à tous les humains. En d'autres termes, tous les humains partageraient la même architecture neurale visuelle, et les processus opérant au sein de celle-ci seraient similaires ([Ware, 2003]) ;
 - l'attention serait directement liée à l'orientation du regard. Ce second postulat, justifiant directement le recours à l'oculométrie comme outil de mesure indirecte de l'activité cognitive humaine, correspond à l'hypothèse « *eye-mind* » développée par Just et Carpenter ([Just & Carpenter, 1980]). En ce sens, les données oculométriques peuvent nous fournir des indices à la fois qualitatifs et quantitatifs sur l'attention. Ainsi, la fixation d'un

objet par le regard témoignerait de l'attention qu'a le sujet pour cet objet et la durée de la fixation témoignerait de la quantité d'informations encodée ([Poole & Ball, 2005]).

Plus récemment, le modèle de Halverson et Hornof ([Halverson & Hornof, 2007]) vise à expliquer la variété des mouvements oculaires dans le cadre d'une théorie de la recherche visuelle dans les interactions Hommes-ordinateurs. A l'inverse, notamment, des modèles de Findlay et Walker ([Findlay & Walker, 1999]), de Treisman et son équipe ([Treisman & Gelade, 1980]; [Treisman & Gormican, 1988] pour la théorie originale, [Quinlan, 2003]) ou encore de Duncan ([Duncan, 1998]), ce modèle ne postule pas que les stratégies de recherche visuelle sont liées à des structures visuelles particulières ou à la saillance des éléments. Il y aurait ainsi trois caractéristiques de la recherche visuelle.

D'une part, les mouvements oculaires tendent à s'orienter vers des objets proches du champ visuel. Cela implique que les stratégies de recherche visuelle ne se font pas de manière hasardeuse, et indépendamment des caractéristiques des objets visuels. D'autre part, les objets fixés ne sont pas systématiquement identifiés, ce qui explique qu'une cible puisse être fixée plusieurs fois avant d'être identifiée comme telle. Et enfin, les mouvements oculaires ne s'initient qu'une fois que les objets fixés ont été identifiés. En d'autres termes, les durées de fixations ne sont pas contrôlées directement par les stratégies de recherche visuelle, mais plutôt par le temps requis pour percevoir l'objet fixé.

1.1.1.2 Des justifications pratiques

L'oculométrie est qualifiée de méthode « *online* » (recueil de données pendant l'interaction), en opposition aux méthodes « *offline* » du type verbalisations consécutives, entretiens, etc. ([Jamet et al., 2008]). L'intérêt des méthodes *online* réside dans l'objectivité de la mesure. Outre son caractère objectif, l'oculométrie permet d'atteindre une finesse d'analyse importante. En effet, il est aisé de savoir, par exemple, si un internaute a bien vu certains éléments d'une page Web, s'il n'a pas cliqué sur un lien parce qu'il ne l'a pas vu ou au contraire parce que l'intitulé ne l'incitait pas à le faire. L'utilisation de l'oculométrie permet donc de répondre à certaines questions de façon précise et objective. Évidemment, cette technique n'est pas miraculeuse et n'a pas pour but de remplacer les autres approches et méthodes utilisées pour connaître les comportements des Internautees ou pour évaluer les sites Web ([Bastien et al., 2010]).

Du côté de ses applications, l'oculométrie a été utilisée pour tenter de comprendre plus finement certaines catégories d'activité mentale ([Rudmann et al., 2003]) ou les processus cognitifs, notamment ceux impliqués dans la résolution de problèmes (cf. [Yoon & Narayanan, 2004] pour une revue). Cette technique a fait l'objet de nombreuses études sur les comportements de lecture (cf. [Rayner, 1998] pour une revue). Plus récemment, c'est sur

Internet que les recherches se sont développées. Ainsi, Fukuda et Bubb ([Fukuda & Bubb, 2003]), en combinant les deux types de méthodes sur le Web, concluent que les évaluations subjectives dépendent surtout du point de vue de l'utilisateur, alors que les méthodes *online* permettent de réaliser des études dont les résultats sont directement applicables car plus objectives.

L'oculométrie permet d'étudier de nombreux phénomènes comme l'exploration de menus ([Aaltonen et al., 1998, Byrne & Anderson, 1999, Fleetwood & Byrne, 2006]), l'interaction avec les bannières publicitaires ([Bayles, 2002, Burke et al., 2004, Burke et al., 2005]), la lecture des hypertextes ([Halverson & Hornof, 2004]), les marque-pages des navigateurs ([Poole et al., 2005]) ou encore l'évaluation de l'utilisabilité des logiciels ou sites Web ([Goldberg, 1999, Pretorius et al., 2005, Nakamichi et al., 2006, Ehmke & Wilson, 2007] pour une revue).

1.1.2 Les mouvements oculaires

1.1.2.1 Le réflexe optocinétique

Les yeux ont pour caractéristique d'être continuellement en mouvement. En effet, afin de maintenir constant le signal venant de l'environnement vers la rétine, les yeux effectuent des mouvements rapides et de faible amplitude nommés nystagmus physiologique ou réflexe optocinétique (à distinguer du nystagmus pathologique). Rayner ([Rayner, 1998]) rapporte deux autres types de légers mouvements oculaires : les dérives (drifts) et les microsaccades. Les dérives seraient de petites pertes de contrôle du système oculomoteur par le système nerveux : l'œil se met à dériver lentement hors du point d'attention, puis une rapide saccade (la microsaccade) le réoriente.

1.1.2.2 Les fixations

La littérature portant sur les travaux en oculométrie fait souvent état de données issues des fixations oculaires. Le terme « fixation » est un abus de langage mais, par convention, est considéré comme fixation un mouvement oculaire n'excédant pas une aire de 3ř (+/- $0,5\text{ř}$) d'angle visuel autour d'une cible pendant une durée minimale de 100 ms ([Goldberg, 1999]) et une vitesse comprise entre 15 et 100 degrés par seconde ([Jacob & Karn, 2003]). Il y aurait au moins trois processus intervenant en parallèle lors d'une fixation oculaire ([Goldberg, 1999]). Suite à une saccade, le stimulus visuel est encodé pendant 100 à 150 ms afin d'identifier l'objet de la fixation. Ensuite, le reste du champ visuel est traité par la vision périphérique afin de repérer l'objet de la prochaine fixation. Enfin, la saccade suivante est planifiée et préparée. La fixation (incluant les trois processus que nous venons de décrire) dure en moyenne 250 à 300 ms.

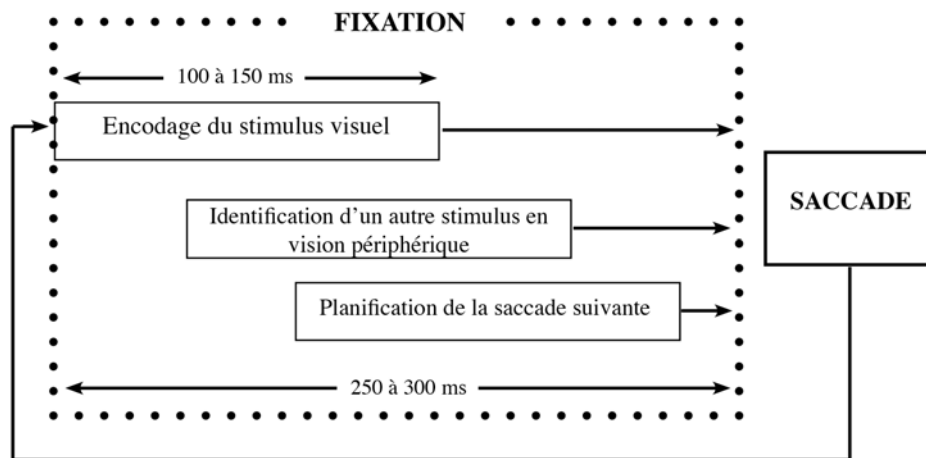


FIGURE 1.1 – Les trois processus intervenant lors d’une fixation oculaire (adapté de [Goldberg, 1999]).

1.1.2.3 Les saccades

Généralités

La fonction de la saccade est de permettre d’explorer le champ visuel. C’est le psychologue Russe Yarbus ([Yarbus, 1967]) qui a montré en premier que la perception visuelle d’une scène ne se fait pas par une exploration linéaire de ses éléments, mais plutôt par une succession de saccades et de fixations. Les saccades interviennent dans la lecture au sens large (lecture de scènes, de mots...) et dans les activités de recherche. La saccade est provoquée par un rapide mouvement oculomoteur permettant le passage d’un objet du champ visuel à un autre. La vitesse de cette saccade est de l’ordre de 500° d’angle visuel par seconde ([Rayner, 1998, Baccino, 2004]) et couvre de 1 à 40° d’angle visuel pour une durée de 30 à 120 ms ([Jacob, 1991]). La vitesse d’une saccade oculaire est fonction de la distance du point de fixation à un autre point cible. La vitesse de la saccade augmente très rapidement lors de son exécution pour atteindre un pic de vitesse ([Baccino, 2004]) qui va décroître progressivement jusqu’à atteindre le point de fixation suivant.

Deux éléments importants caractérisent la saccade :

- Lors d’une saccade, il semblerait que l’œil soit quasiment aveugle en raison de la trop grande rapidité de ce mouvement. Ce phénomène a été nommé par Matin ([Matin, 1974]) la suppression saccadique.
- Une fois que le mouvement saccadique est engagé, il serait impossible d’en modifier la trajectoire ([Baccino, 2004]).

Cela implique que les saccades sont des mouvements balistiques et stéréotypés ([Duchowski, 2007]). En effet, il a été observé que certains patterns de mouvements peuvent être répétés de manière stéréotypée. En outre, la trajectoire de la

saccade ne serait plus modifiable car elle aurait été préprogrammée avant son exécution. Cela expliquerait la relative précision des saccades : un calcul de la position de l'objet-cible serait formulé a priori.

Un élément supplémentaire vient étayer cette hypothèse. En effet, avant une saccade, il y a un temps de latence d'environ 200 ms pouvant correspondre au temps de traitement nécessaire à la préparation du mouvement saccadique.

Relation entre fixations et saccades

L'existence des saccades oculaires s'explique par leur relation aux fixations. Elles ne correspondent qu'au déplacement d'un point de fixation par la vision fovéale à un autre point de fixation. Ce sont donc les caractéristiques rétiniennes de l'œil qui nécessitent le recours aux saccades afin d'explorer le champ visuel.

La vision centrale correspond à trois structures de la rétine ([Duchowski, 2007]) :

- La fovéola (ou fovéa centrale) : c'est la zone qui contient la plus forte densité de cônes. Cela se traduit par une forte acuité visuelle en vision centrale mais l'angle visuel concerné n'est que de 1,3°. Plus l'angle visuel s'éloigne de cette zone, plus l'acuité visuelle sera basse.
- La fovéa : la densité de cônes y est aussi élevée, et l'angle visuel couvert par cette zone est de 5°, correspondant à 3% d'un écran de 21" à 60 cm de distance. Au delà de cet angle visuel, l'acuité diminue de moitié.
- La macula représente 16,7° d'angle visuel. L'acuité visuelle est la plus élevée à 2° d'angle visuel autour du point de fixation du regard et la vision parafovéale étend cette zone de forte acuité jusqu'à 5°. Jusqu'à 30° d'angle visuel autour du point de fixation, l'acuité est encore correcte, mais au delà, la résolution est si basse qu'on ne peut détecter que des mouvements d'ambiance.

Ainsi, les yeux parcourent sans cesse le champ visuel car c'est l'unique façon de voir distinctement chaque partie de ce champ.

1.1.3 La technique du reflet cornéen

Le système le plus utilisé aujourd'hui pour suivre les mouvements oculaires sur les écrans est la technique du reflet cornéen. Il nécessite de disposer d'une lampe infrarouge et d'une caméra infrarouge.

Aujourd'hui, les appareillages peuvent être directement inclus dans l'écran du micro-ordinateur ou déportés, et ne nécessitent plus d'être attachés sur la tête des participants.

Cette technique consiste à envoyer un faisceau de lumières infrarouges invisible à l'œil, filmé par une caméra. Ce faisceau se reflète sur la rétine rendant la pupille lumineuse (à l'instar des « yeux rouges » que l'on peut voir sur les photographies

de visages), permettant d'en délimiter les contours. Le reflet de ce faisceau est mouvant car il fixe constamment les orientations de la pupille. En parallèle, au moins un autre faisceau lumineux (le reflet cornéen), de plus petite taille, est projeté sur la cornée, en dehors de la pupille. Ce dernier reflet reste fixe, permettant un calcul vectoriel de la position relative du point de regard par rapport à la pupille via un algorithme.

Pour que l'utilisation de cette technique soit fiable, elle nécessite au préalable de chaque expérimentation qu'au moins une session de calibration soit effectuée. Celle-ci consiste à faire visualiser au sujet une série de points mouvants qu'il doit suivre du regard puis les fixer lorsqu'ils s'arrêtent. Les coordonnées de ces points sont définies à l'avance.

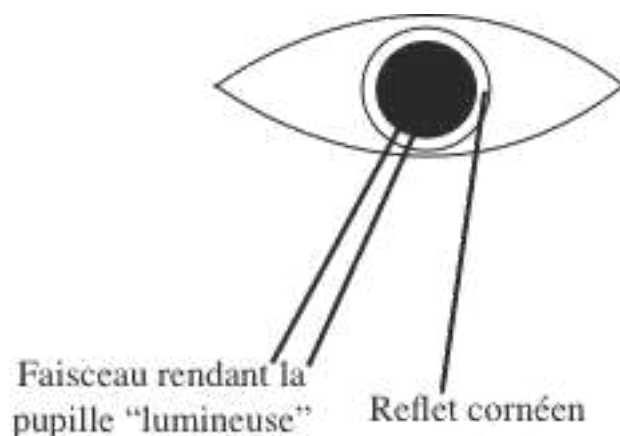


FIGURE 1.2 – Illustration de la technique du reflet cornéen.

Ce système capture ainsi à une certaine fréquence (typiquement entre 30 et 300 Hz) les positions du regard sur un écran (coordonnées X,Y et parfois Z). Ces données doivent faire l'objet d'un traitement supplémentaire pour les transformer, par exemple, en fixations (coordonnées X,Y associées à une durée). Différents algorithmes ont été proposés dans la littérature pour identifier les mouvements oculaires «offline» (pour une revue, voir [Salvucci & Goldberg, 2000]) ou «online», c'est à dire en temps réel ([Jambon & Luengo, 2012]).

1.1.4 Limites de la technique du reflet cornéen

Il y a quelques années déjà, Schnipke et Todd ([Schnipke & Todd, 2000]) rapportaient que seuls 37,5% des participants aux études utilisant l'oculométrie pouvaient être calibrés. Ce faible taux de succès n'était pourtant pas rapporté dans la littérature sur le sujet. Les auteurs en viennent à supposer que les chercheurs du domaine sélectionnent a priori les sujets ayant le plus de chance de réussir la phase de calibration. Il est évident qu'il s'agit là d'un biais de

sélection important. Chen (2001) pointait la nécessité de recalibrer sans cesse afin d'éviter un phénomène de dérives.

Jacob et Karn ([Jacob & Karn, 2003]) en concluent que 10 à 20% des sujets ne peuvent participer à ce type d'études. Par exemple, sur les 32 participants d'une étude, 7 ne pouvaient faire l'objet d'une calibration (environ 22%) avec un système Tobii 1750 ([Rodden et al., 2008]).

Bien que les progrès techniques soient significatifs, les problèmes de calibration subsistent et seraient dus à plusieurs facteurs comme une pupille ne reflétant pas assez la lumière, un iris trop clair donc impossible à distinguer de la pupille, une peau reflétant trop la lumière, une pupille trop large, des paupières trop fermées, voire des cils occultant la pupille. D'autres pertes de données sont à déplorer en raison d'incidents techniques ou de déplacements de la tête du participant en dehors du champ de la caméra infrarouge ([Schnipke & Todd, 2000, Nguyen et al., 2002]).

Ces différents problèmes ont amené Schnipke et Todd ([Schnipke & Todd, 2000]) à se poser des questions quant à la représentativité des échantillons, mais aussi quant au coût en termes de temps dus au remplacement des participants ne pouvant faire l'objet d'une calibration. Cela peut aussi conduire les chercheurs à étudier de faibles échantillons de participants.

1.1.5 Représentation des fixations et des saccades

Comme nous l'avons vu plus haut, à partir des données brutes issues de l'oculomètre, on calcule les fixations et les saccades oculaires. Chaque fixation est caractérisée par sa position sur le plan, mais aussi par sa durée alors que la saccade se caractérise par son amplitude et son angle. On appelle « *scanpath* » la succession de fixations et de saccades. La Figure 1.3 présente un exemple de ce que peuvent fournir les outils d'analyse des données oculométriques : les fixations sont représentées par des cercles dont la taille varie en fonction de leurs durées respectives et les numéros indiquent l'ordre des fixations. À partir des données oculométriques, il est possible de calculer un certain nombre d'indicateurs, parmi lesquels on trouve : (i) la durée moyenne des fixations par page Web (en ms), (ii) le taux de fixations par page Web, (iii) le taux de saccades par page Web, (iv) l'amplitude moyenne des saccades par page Web (en pixels), et (v) la longueur moyenne des scanpaths par page Web (en pixels).



FIGURE 1.3 – Un exemple de scanpath.

À partir de ces données, certains auteurs étudient, par exemple, l'effet de la structure des pages Web, de l'âge, du genre ou encore de l'expérience des internautes sur ces indices ; d'autres encore les utilisent pour évaluer l'utilisabilité des sites.

À ces mesures viennent s'ajouter des méthodes plus qualitatives de représentation des scanpaths. Les plus connues sont les fameuses « *heatmaps* » ou cartes de chaleurs dont un exemple sur une page Web est donné dans la Figure 1.4. Cela consiste à représenter l'intensité de données oculométriques d'un ou plusieurs individus (comme le nombre de fixations ou la durée des fixations) à l'aide d'un nuancier de couleurs selon un découpage matriciel du stimulus. Ainsi, dans notre exemple, la couleur rouge représente les zones contenant le plus de fixations et les zones vertes, celles qui en contiennent le moins (les zones transparentes ne contiennent aucune fixation).



FIGURE 1.4 – Heatmap du scanpath de la figure précédente.

Chapitre 2

Connaissances sur le comportement visuel des internauts sur le Web

2.1 Les comportements des utilisateurs sur les moteurs de recherche

L'utilisation des moteurs de recherche est à ce point courante que bon nombre d'utilisateurs en font leur page de démarrage par défaut. Pas étonnant, dans ce contexte, que l'utilisation des moteurs de recherche ait suscité autant d'études visant au final à guider les concepteurs quant à l'organisation des informations qui y sont affichées. Les paragraphes suivants présentent brièvement les études ayant abordé différents de ces moteurs de recherche sur l'utilisation de leurs résultats.

2.1.1 Aspects globaux des comportements visuels

Il est intéressant de noter dans un premier temps que dans 96% des cas, les internautes ne regardent pas au-delà de la première page des résultats de recherche ([Lorigo et al., 2008a]). De plus, au-delà du 7ème résultat de recherche, peu d'attention visuelle est allouée, et cela serait dû au simple fait que, généralement, il faut « scroller » (i.e., faire défiler la page avec l'ascenseur) pour les visualiser. En outre, dans 40% des cas, les internautes reformulent leurs requêtes quelle que soit la pertinence des résultats affichés, et après avoir visualisé seulement 3 résultats en moyenne. Ainsi, si les 3 premiers résultats qu'il a vus ne lui semblent pas pertinents, l'internaute a tendance à reformuler sa requête ([Lorigo et al., 2008a]).

De façon générale, les internautes ont tendance à ne regarder que les pages de résultats de recherche qui précèdent le clic. C'est-à-dire que s'ils sélectionnent par exemple le résultat no 4, ils auront éventuellement regardé les résultats 1, 2 et 3, mais plus rarement ceux qui sont en dessous du no 4 ([Granka et al., 2004]). Malgré cela, seuls 34% des scanpaths sont linéaires et 19% sont strictement linéaires. Cela signifie que plus de la moitié des séquences de visualisation contiennent des régressions ou des sauts de résultats. Les internautes ne regardent donc pas linéairement les résultats de recherche s'affichant sur une page.

Quel que soit le lien sur lequel les participants cliquent, ils visualisent généralement les 4 premiers résultats de recherche. Si le lien sélectionné est plus bas que le 4ème lien, les participants visualisent la plupart des liens qui le précèdent. Puis, après avoir visualisé environ 8 résultats de recherche sans cliquer, ils reformulent leur requête. Le pattern comportemental commun observé est le « *hub and spoke* », où l'internaute a tendance à cliquer sur un lien puis à revenir rapidement sur la page des résultats de recherche avec le bouton « précédent » de son navigateur. À la deuxième visite sur la page de résultats, ce sont toujours les premiers qui reçoivent le plus d'attention visuelle mais au fur et à mesure, la distribution des fixations oculaire évolue en faveur des résultats affichés plus bas. Au bout d'un certain temps, ce sont donc les résultats 3, 4 et 5 qui deviennent le centre de l'attention ([Cutrell & Guan, 2007a]).

2.1.2 L'importance des deux premiers résultats de recherche

Granka et ses collègues ([Granka et al., 2004]) ont montré que les résultats de recherche de Google de rang 1 et 2 sont les plus regardés par les internautes, au détriment des rangs inférieurs. Ces deux premiers résultats sont visualisés en proportion quasiment égales, et sont les plus fréquemment sélectionnés. Les internautes s'intéressent donc prioritairement à ceux-ci. Outre le fait qu'ils soient plus visualisés et sélectionnés, les deux premiers résultats servent aussi de point de référence. S'ils ne sont pas jugés immédiatement comme pertinents, l'internaute aura tendance à regarder les résultats suivants tout en réévaluant les deux premiers régulièrement ([Lorigo et al., 2006]).

L'oculométrie a permis de découvrir un fait surprenant : bien que les internautes regardent autant les résultats 1 et 2, ils cliquent dans 48% des cas sur le premier lien contre seulement 8% sur le second. De plus, il semblerait que ce soit réellement un effet de la position du résultat de recherche car même en inversant les résultats 1 et 2 artificiellement, les internautes continuent de cliquer sur le premier plus fréquemment ([Lorigo et al., 2008a]). Ainsi, le premier résultat est sélectionné 3 fois plus souvent que les autres, même s'il est en réalité le deuxième (inversion artificielle). En revanche, si l'ensemble des résultats est inversé, on remarque un plus faible taux de clic sur les liens 1 et 2 (qui reste

néanmoins plus élevé que pour les autres rangs). En somme, quelle que soit la condition, l'attention visuelle accordée aux deux premiers résultats de recherche reste constante et très élevée [Pan et al., 2007].

2.1.3 L'Effet du rang du résultat de recherche sur les comportements oculaires

En manipulant artificiellement le rang du résultat de recherche pertinent, Guan et Cutrell ([Guan & Cutrell, 2007]) ont montré que plus celui-ci est bas, plus les participants passent de temps à réaliser la tâche. Cela montre bien que ces derniers ne croient pas aveuglément en l'exactitude du premier résultat.

Pour une tâche de recherche d'information, lorsque le bon lien est en dessous du 2ème rang, le taux de clics exacts tombe à 20%. De plus, il s'avère que lorsque la cible est à une position basse (en dessous du 4ème rang), les participants cliquent fréquemment sur le premier lien (46%). L'explication de ce phénomène s'est faite en partie par le biais de l'oculométrie. Les auteurs ont observé que plus les résultats pertinents de recherche sont bas dans la page, plus leur probabilité d'être visualisés est basse. Néanmoins, cela n'explique pas pourquoi le taux de clics exacts est si faible. En effet, voir des liens non pertinents ne devrait pas inciter l'internaute à cliquer dessus. Finalement, l'explication réside dans la confiance qu'ont les internautes envers les moteurs de recherche. Ils semblent parir du principe que le moteur de recherche proposera forcément les résultats pertinents en premier ([Guan & Cutrell, 2007]).

Dans le cas où les résultats de recherche sont inversés artificiellement, les utilisateurs ont plus de difficultés à trouver le lien pertinent (temps plus longs, plus de fixations, plus de résultats inspectés, plus de régressions). Pourtant, après enquête, il s'avère que les participants ont tendance à attribuer ces difficultés à autre chose qu'au moteur de recherche (difficultés des tâches, mauvaises requêtes...). En résumé, quels que soient les facteurs pris en compte (type de tâche, position, pertinence), c'est la disposition des résultats de recherche qui explique les comportements d'attention et de sélection.

2.1.4 Les styles de visualisation

Outre les comportements moyens des utilisateurs sur les listes de résultats des moteurs de recherches, il y aurait des typologies de comportements. C'est ce que proposent Aula, Majaranta et Rähkä ([Aula et al., 2005]). Ces auteurs ont différencié les comportements des internautes en fonction de leur façon d'évaluer les résultats fournis par le moteur de recherche. Il y aurait ainsi les évaluateurs économes, et les évaluateurs exhaustifs. Ce qui différencie ces deux types d'évaluateurs, c'est le nombre moyen de résultats de la liste de recherche visualisés, puis la durée moyenne des fixations. Il y aurait ainsi 46% d'évaluateurs

« économes » et 54% d'« exhaustifs ». Dans 50% des cas, les évaluateurs « économes » ont visualisé au mieux la moitié des résultats de recherche visibles (6 à 7 étaient visibles sans « scroll »). Dans la plupart des cas, les évaluateurs « exhaustifs » visualisaient plus de la moitié des résultats visibles ou scrollaient pour voir tous les résultats avant de réaliser leur première action. En outre, le temps pour réaliser la première action était plus court, tout comme la durée moyenne des fixations oculaires, pour les évaluateurs « économes ».

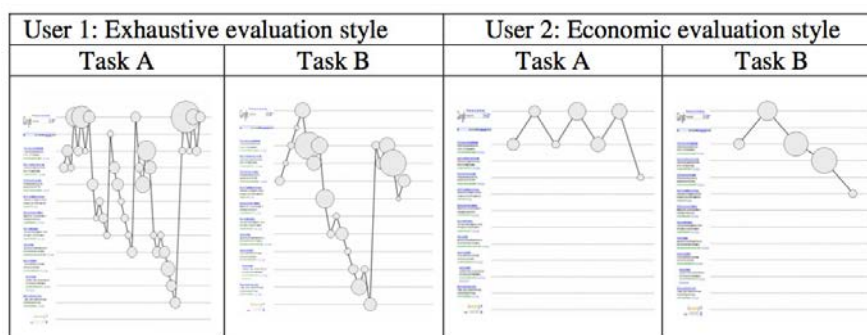


FIGURE 2.1 – Illustration des deux types d'évaluateurs décrits par Aula, Majaranta et Raihä ([Aula et al., 2005]).

En termes de performance, les « économes » sont plus rapides que les « exhaustifs », mais uniquement si les résultats de recherche sont pertinents. Les auteurs suggèrent que les « exhaustifs » manquent d'expertise, et sont donc plus dépendants des résumés des résultats de recherche pour prendre la décision de suivre un lien.

2.2 L'influence des caractéristiques des pages Web

Pan et son équipe ([Pan et al., 2004]) ont montré que les internautes utilisent des indices de visualisation issus de la première page d'un site Web pour visiter les autres pages de ce même site et que le comportement des utilisateurs change au fur et à mesure du temps passé à la navigation au sein d'un même site. Cela peut s'interpréter comme une adaptation au style de navigation induit par les sites. Les auteurs indiquent aussi que les variations des scanpaths sont influencées par le degré de complexité des pages web. Plus la page est complexe, plus les scanpaths varient, moins elle est complexe (e.g., Google), plus les scanpaths (inter-sujets) sont similaires.

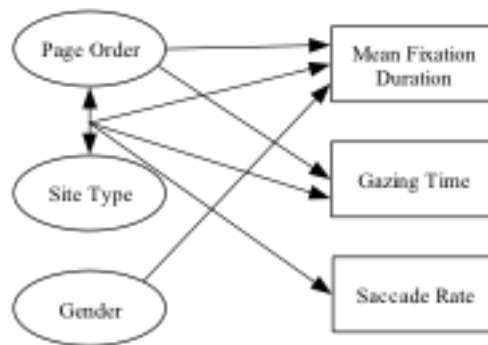


Figure 5. Overall Results

FIGURE 2.2 – Vue d’ensemble des résultats de Pan et col. ([Pan et al., 2004]).

Yesilada et ses collègues ([Yesilada et al., 2008]) proposent le modèle TAF (Travel Analysis Framework), pour rendre compte des comportements des utilisateurs d’internet. Selon ce modèle, l’utilisateur se déplace dans les pages web à l’instar d’un voyageur. Pour se repérer, le voyageur virtuel utilise des points de passage, des points d’orientation et des assistants à la navigation regroupant respectivement des éléments des interfaces web.



FIGURE 2.3 – Les repères à la navigation selon le modèle TAF ([Yesilada et al., 2008]).

Les résultats de leur étude montrent que 70% de leurs participants utilisent un point de référence (soit le logo, soit l’entête, soit le titre du site) lorsqu’ils

entrent dans un site Internet. Néanmoins, ces résultats ne s'appliquent pas à Google en raison de son interface dépouillée et familière pour leur échantillon d'utilisateurs. De plus, c'est le contenu des sites qui reçoit le plus de fixations, puis l'entête et la colonne de gauche, confirmant les résultats de Goldberg et ses collègues ([Goldberg et al., 2002]). Tzanidou et son équipe ([Tzanidou et al., 2005]) avaient aussi montré que la fixation initiale des participants se trouvait systématiquement soit en haut à gauche, soit sur la partie centrale supérieure des pages, indépendamment de la localisation du lien-cible. La colonne de droite et le bas de page enregistrent très peu de fixations. Le nombre et l'organisation des éléments de contour déterminent la façon dont les utilisateurs complètent leur navigation intrapage. Ce sont bien les éléments de contours, plutôt que la localisation des objets qui orientent l'attention. Par exemple, les colonnes de droite sont rarement fixées, mais si le contenu du site occupe aussi l'espace de droite, la partie droite de l'écran l'est aussi beaucoup. Autre élément intéressant de leur recherche, le rôle et la localisation des points de référence et la distribution de l'attention en fonction des éléments de contour étaient constants au fil des tâches. En d'autres termes, ce n'est pas seulement la tâche à réaliser qui guide l'utilisateur lors de sa navigation, mais aussi la structure des pages qu'il visite.

2.3 L'influence du type de site Web

Habuchi, Takeuchi et Kitajima ([Habuchi et al., 2007]) ont testé trois catégories de sites Web (publicitaires, d'actualités et de sociétés) en faisant varier la localisation de leurs zones de fonctions (navigation, contenu et publicité) à gauche, au milieu ou à droite des pages. Il semblerait que les sites Web d'actualités enregistrent moins de fixations que les deux autres types de sites mais que leur zone de contenu enregistre plus de fixations alors qu'il n'y a pas de différences pour le nombre de fixations interzones pour les autres catégories de sites Web. Les sites Web de sociétés et les sites publicitaires enregistrent moins de fixations dans les zones situées à droite des pages, ce qui n'est pas le cas des sites d'actualités.

Ces auteurs concluent que les utilisateurs adoptent une stratégie visuelle basée sur la fonction pour les sites Web dont le contenu est clair, alors qu'ils adoptent une stratégie basée sur la localisation pour les sites dont le contenu n'est pas explicite pour l'utilisateur.

2.4 L'influence des tâches

2.4.1 Le type de tâche

Il semble évident que les comportements d'exploration visuelle d'une page Web diffèrent selon que l'internaute cherchera une information précise ou qu'il explorera sans objectif particulier une même page Web. Le type de tâche a donc aussi fait l'objet de recherches.

Pour valider la théorie du scanpath sur le Web (cf. infra), la tâche proposée aux participants par Josephson et Holmes ([Josephson & Holmes, 2002]) était libre puisqu'elle consistait simplement à observer des pages Web fixes (il s'agit donc d'images complexes). Or, Goldberg et ses collègues ([Goldberg et al., 2002]) faisaient remarquer que le type de tâche influence grandement le comportement oculaire des internautes. Il s'est avéré dans leur expérience que les différences de scanpaths observées étaient dues au type de tâches réalisées, et non aux participants. Néanmoins, Pan et son équipe ([Pan et al., 2004]) n'ont pas observé de différences sur les durées de fixations entre le groupe soumis à des tâches et le groupe ne devant pas réaliser de tâche spécifique sur les sites proposés. Sur les résultats de listes de moteurs de recherche, Rele et Duchowski ([Rele & Duchowski, 2005]) n'ont pas non plus constaté de différences de moyennes des durées de fixations entre les tâches de navigation et de recherche d'information.

En s'intéressant à l'« *interruption overload* », Iqbal et Bailey ([Iqbal & Bailey, 2004]) ont cherché à développer une méthode visant à détecter le type de tâche que réalise un utilisateur. Pour ce faire, les auteurs ont ainsi enregistré différents indicateurs oculométriques pour les comparer entre les différentes tâches soumises à leurs participants. Leurs principaux résultats sont les suivants : (a) les patterns oculaires seraient influencés par la difficulté de la tâche ; (b) le temps passé à fixer une région d'intérêt serait un indicateur de sa complexité visuelle ; (c) les données oculométriques permettraient de décider où et comment présenter des notifications ; et enfin, (d) les utilisateurs ignoreraient, c'est-à-dire ne traiteraient pas, certaines régions d'intérêt importantes.

En se basant sur la taxonomie de Broder ([Broder, 2002]), Terai et ses collègues ([Terai et al., 2008]) ont testé l'influence du type de tâche sur la lecture des pages Web. Pour cela, ils ont distingué les tâches de recherche d'information des tâches transactionnelles. Les auteurs rapportent que le temps de lecture sur chaque page était plus court pour les tâches transactionnelles que pour les tâches de recherche d'information.

Les durées de fixation par zones d'intérêt sont aussi influencées par le type de tâche. C'est ce qu'ont montré Buscher, Cutrell et Morris ([Buscher et al., 2009]) en découpant leurs pages Web en zones de dimensions égales (à l'exception du bas de page). Bien qu'il y ait des différences intéressantes en fonction du type de tâche (butinage d'information et reconnaissance), c'est principalement lors de la

première seconde de visualisation, considérée comme révélatrice des attentes de localisation des utilisateurs, que ces différences sont les plus claires (cf. Figure 2.4).

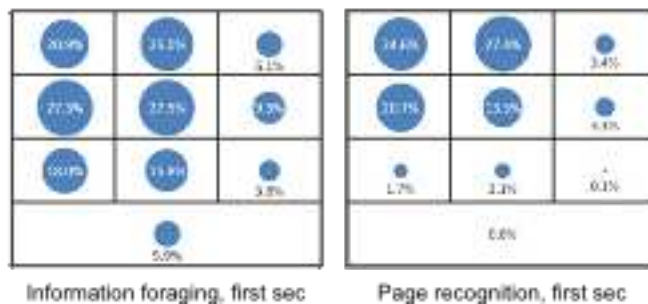


FIGURE 2.4 – Fréquences de visualisation des 10 aires d'intérêt sur toutes les pages durant la première seconde de visualisation en fonction du type de tâche (Buscher, Cutrell et Morris, 2009).

Outre l'influence du type de tâche, les auteurs ont aussi cherché à décrire les comportements oculaires des utilisateurs de façon globale. La partie droite des pages Web est bien moins fixée que les autres zones, à l'inverse des quatre zones en haut à gauche.

2.4.2 La répétition des tâches

Tzanidou et ses collègues ([Tzanidou et al., 2005]) ont utilisé le lien « à propos de nous » de pages d'accueil, comme stimulus. La tâche consistait à trouver ce lien sans que les utilisateurs sachent à l'avance où il se trouve dans les pages présentées. Chaque participant devait répéter dix fois la même tâche. La mesure oculométrique a porté sur le temps passé à fixer la cible, la localisation de la fixation, la fixation initiale (là où le point de fixation se trouve durant les 50 premières ms lors de l'apparition d'une page), le point d'entrée (les points de fixations effectués durant les 250 première ms) et le scanpath. Les résultats montrent que les utilisateurs s'adaptent à une localisation inattendue du lien « à propos de nous » (c'est-à-dire en bas des pages), allant jusqu'à changer d'avis sur leurs attentes de localisation après plusieurs expositions (des questionnaires ont aussi été soumis aux participants à différents moments de l'expérimentation). Néanmoins, c'est lorsque le lien « à propos de nous » apparaît en haut des pages qu'il est le plus rapidement trouvé. Le point d'entrée n'était pas constant entre les participants et entre les différentes pages, bien qu'aucun point d'entrée n'ait été enregistré dans les parties droites des pages.

La moitié des participants ont modifié leurs patterns de visualisation au fur et à mesure des sessions d'expositions aux pages web (et donc des changements de localisation des liens-cibles).

2.5 Les caractéristiques des utilisateurs

Les internautes diffèrent beaucoup entre eux : ils diffèrent en termes d'âge, d'expérience, de genre... Mais ces différences ont-elles un impact sur la façon avec laquelle ils explorent les sites Web ?

2.5.1 Rôle du genre

Pan et son équipe ([Pan et al., 2004]) faisaient remarquer que très peu de recherches antérieures à leur étude se sont attardées aux comportements oculaires des utilisateurs d'Internet. Leur objectif était d'identifier les déterminants des comportements oculaires sur Internet à partir de l'étude des fixations, du temps de fixation, du taux de saccades oculaires et, dans une moindre mesure, du scanpath selon la méthode de la distance d'édition. Leurs résultats indiquent que les moyennes des durées de fixation sur le Web sont significativement plus élevées pour les hommes que pour les femmes. Ces résultats sont confirmés par Lorigo et ses collègues ([Lorigo et al., 2006]) sur le moteur de recherche Google.

Dans une étude exploratoire, Hembrooke, Feusner et Gay ([Hembrooke et al., 2006]) ont agrégé les scanpaths des utilisateurs à partir de la méthode de la distance d'édition, en ne tenant compte que des similarités inter-utilisateurs. Deux types de tâches étaient proposés à deux groupes différents : une tâche libre de butinage et une tâche induite où des consignes étaient données. L'objectif de cette étude était d'évaluer l'utilité d'une comparaison de patterns de scanpaths moyens. Pour cela, les auteurs ont comparé les scanpaths individuels aux scanpaths moyens en fonction du type de tâche réalisé et du sexe des participants. Leurs résultats mettent en exergue l'utilité de la comparaison des scanpaths par la méthode de la distance d'édition. En effet, les scanpaths diffèrent selon les conditions expérimentales : le groupe « butinage » n'a pas effectué de saccades oculaires courtes, qui sont des indicateurs de lecture et de focalisation de l'attention, alors que le groupe « consignes » effectue de courtes saccades au centre de la page Web. Les auteurs ont aussi pu discriminer le genre des sujets par cette méthode.

Djamasbi et ses collègues ([Djamasbi et al., 2007]) se sont intéressé spécifiquement à l'utilisation des « *bricklets* ». Il s'agit de petites fenêtres disposées à différents endroits du site comportant des informations spécifiques destinées à rendre la navigation plus rapide et plus facile pour l'utilisateur. Ces fenêtres ont pour but d'apporter des informations importantes à l'utilisateur et d'attirer son attention. Il peut s'agir de liens vers des sites récemment visités à partir du site de référence, des commentaires utiles, des informations sur l'usage du site, etc. L'étude a été réalisée en laboratoire avec 17 hommes et 19 femmes âgés entre 20 et 60 ans. L'expérimentation a été réalisée avec un oculomètre. Les expérimentateurs ont demandé aux participants de répondre à différentes tâches

de recherche d'information sur un site financier. Les chercheurs testaient alors différents designs de « *bricklets* », avec fond noir ou clair, avec ou sans image. En fonction du placement de ces « *bricklets* », des aires d'intérêts avaient été définies au préalable. L'oculomètre permettait ainsi d'enregistrer le nombre de fixations (fixation supérieur à 300 millisecondes) dans chacune de ces aires. À la fin de l'expérience, les sujets devaient répondre à un questionnaire permettant d'évaluer l'attrait visuel des différents éléments et également d'identifier le nombre de bricklets vu pendant l'expérience.

Les résultats présentés, notamment, sous forme de carte de chaleur, ont montré que les hommes jugent plus attrayantes les bricklets comportant des images. Les femmes ont des fixations plus nombreuses sur les bricklets à fond clair que sur celles ayant un fond foncé. Par ailleurs, les femmes ont tendance à se souvenir beaucoup plus facilement des bricklets avec un fond foncé.

Buscher, Cutrell et Morris ([Buscher et al., 2009]) ont cherché à décrire plus finement les différences de comportements oculaires entre les hommes et les femmes. Pour cela, ils ont découpé des pages Web en 10 zones d'intérêt, dont 9 de taille égale (la dixième correspond au bas de page) et ont soumis les participants à deux types de tâches : une tâche de reconnaissance et une tâche de recherche d'information.

La tâche de reconnaissance a généré des durées de fixation plus longues pour les femmes que pour les hommes. Les différentes aires d'intérêt ont enregistré des durées de fixation qui varient selon le genre des participants lors de la tâche de recherche d'information. Les hommes passent plus de temps à regarder la zone en haut à droite et le bas de page que les femmes, mais moins de temps à regarder la zone centrale du haut de la page.

2.5.2 Rôle de l'expérience antérieure

Dans une recherche sur la répétition des tâches, Tzanidou et son équipe ([Tzanidou et al., 2005]) ont observé que la moitié des participants ont modifié leurs patterns de visualisation au fur et à mesure des sessions d'expositions aux pages web (et donc des changements de localisation des liens-cibles). Il s'est avéré que ces utilisateurs étaient les plus expérimentés sur l'usage d'Internet.

Habuchi, Takeuchi et Kitajima ([Habuchi et al., 2006]) ont cherché à voir si le niveau d'expérience sur le Web influençait les comportements oculaires des sujets. Pour cela, ils ont utilisé le même paradigme que Noton et Stark ([Noton & Stark, 1971a]) qui consiste à présenter une première fois les stimuli (ici, 15 pages Web) pendant 20 secondes, puis une deuxième fois, mais mélangé avec de nouveaux stimuli (ici, 30 pages Web) pendant 5 secondes. Les auteurs ont ainsi montré, sur un échantillon très restreint, que les utilisateurs expérimentés avaient des comportements oculaires plus consistants que les autres pendant les deux phases de l'expérience (observation et reconnaissance). Cependant, ces

mêmes auteurs n'ont pas trouvé d'effet significatif du niveau d'expérience dans une expérience ultérieure ([Habuchi et al., 2007])

2.5.3 L'influence de l'âge

Dans ce type de problématique, les chercheurs comparent généralement un groupe de personnes jugées comme sénior avec un autre groupe plus jeune. Actuellement, les publications portent surtout sur les différences entre les deux groupes en termes de comportements oculaires lorsque certains éléments des pages Web varient. Le but est d'affiner les recommandations ergonomiques existantes pour améliorer le confort de navigation des séniors.

A partir de trois sites Web aux structures et caractéristiques différentes, Fukuda et Bubb ([Fukuda & Bubb, 2003]) ont comparé les comportements oculaires d'un échantillon de séniors avec un échantillon de personnes plus jeunes. La longueur des mouvements oculaires est deux fois plus importante pour les séniors que pour les plus jeunes lorsque les pages Web contiennent beaucoup de petits éléments. De plus, les durées de fixations sont plus longues pour les séniors, surtout sur les petits éléments des pages Web (strictement inférieurs à 11 pixels).

Grahame, Laberge et Scialfa ([Grahame et al., 2004]) ont inclus l'oculométrie dans leur procédure expérimentale visant à déterminer l'influence de la structure des pages Web sur la recherche visuelle chez les séniors. La mesure oculométrique retenue par les auteurs était le nombre de fixations et le nombre de mouvements oculaires. De façon globale, les personnes âgées font six fois plus de fixations que les plus jeunes, et effectuent plus de fixations lorsque la taille des polices de caractère est petite. En outre, les personnes âgées effectuent plus de mouvements oculaires que les jeunes lorsque les pages Web sont encombrées, et contiennent beaucoup de liens. L'absence de lien-cible engendre aussi une hausse du nombre de fixations plus importante pour les séniors. Enfin, les personnes âgées ont une tendance prononcée à fixer le centre et le bas des pages Web pour trouver un lien-cible.

Outre les séniors, des différences en termes de comportements oculaires ont été notés entre les personnes âgées de plus ou moins 30 ans ([Buscher et al., 2009]). Les participants de leur étude ayant plus de 30 ans enregistrent des temps de regard plus élevés que ceux ayant moins de 30 ans. Les plus jeunes fixaient aussi plus longuement la région centrale des pages Web et moins longuement le centre gauche.

Djamasbi et ses collègues ([Djamasbi et al., 2010]) ont quant à eux cherché à mieux comprendre ce qui influence les jugements de la « génération Y » (jeunes adultes âgés de 18 à 32 ans) lorsqu'ils évaluent l'attractivité d'un site Web (sans toutefois inclure dans leur étude un groupe contrôle). Il semblerait que ceux-ci ont tendance à fixer les visages sur les images des sites Web en tout premier,

mais uniquement si ces visages représentent une personne célèbre (cf. Fig. 8.8). La première fixation sur une aire d'intérêt a lieu sur le centre ou le haut de la page (visibles sans scroll), et la plupart du temps sur un logo, un visage ou un élément de navigation. Les quatre premières fixations concernent le logo ou le menu du haut. Enfin, s'il y a une large image sur la page Web, elle sera fixée en premier au détriment des éléments textuels alors que s'il n'y a pas de large image, c'est le menu de navigation qui sera fixé. Pour résumer, cette étude suggère que les internautes de la « génération Y » préfèrent les pages Web où sont présentes de larges images, des images de personnes célèbres, et des textes courts.



FIGURE 2.5 – Heatmaps des pages Web jugées attractives par des internautes de la « génération Y » [DJA 10].

2.5.4 L'influence de la culture

Les styles cognitifs semblent différer selon la culture envisagée. Les différences seraient plus perceptibles entre les cultures occidentales et asiatiques. Ainsi, Masuda et Nisbett ont montré que les asiatiques ont un style « holistique », c'est à dire que les éléments de l'environnement d'une scène visuelle sont importants pour eux alors que les occidentaux ont un style plus « analytique », centré sur les objets. Nisbett et Norenzayan (2002) ont exposé plus en détail ces deux styles cognitifs. Partant de ces travaux, Dong et Lee ([Dong & Lee, 2008]) ont proposé une expérience visant à explorer les patterns de visualisation des pages Web de Chinois, Coréens et Américains. Leurs résultats montrent que les patterns de visualisation reflètent les styles cognitifs décrits par l'équipe de Nisbett. Plus précisément, l'échantillon américain semble employer une stratégie de visualisation analytique (les aires d'intérêt sont explorées une à une), avec un pattern de visualisation en « 5 » alors que les asiatiques emploient une stratégie plutôt holistique avec un pattern de visualisation en « 0 ».

2.6 Recherches basées sur la théorie du scanpath

Le but de ce type d'étude est de valider un modèle des mouvements oculaires permettant d'expliquer les comportements des utilisateurs sur Internet. Au départ, ces travaux s'appliquaient sur des images fixes et relativement simples. Aujourd'hui, certains chercheurs se demandent si ce modèle est valable sur des images complexes telles que les sites Internet, mais aussi dans quelle mesure ce modèle permet de repérer les stratégies oculaires d'exploration des scènes visuelles.

Noton et Stark ([Noton & Stark, 1971a]) ont cherché à étudier les processus sériels de la perception visuelle humaine. Plus spécifiquement, selon eux, si les processus sont sériels, alors on doit être capables d'en étudier l'ordre. Les auteurs se sont donc intéressés à l'ordre et la localisation des fixations en omettant volontairement d'enregistrer le nombre de fixations nécessaire à la visualisation d'images (ces images étaient présentées en condition de faible visibilité). Les résultats de leur expérience ont montré que les sujets avaient tendance à adopter un pattern de visualisation fixe pour une même image, ce qu'ils ont nommé « scanpath ».

Ils en ont déduits que le scanpath ne peut pas être réduit à un simple mouvement oculaire permettant d'aligner des zones de la scène visuelle sur la rétine, mais que le scanpath est une composante de la mémoire permettant la reconnaissance des objets. En effet, après la présentation d'un stimulus visuel inconnu à un sujet, ce dernier encode des informations sensorielles et mnésiques (encodage du scanpath). Lors d'une seconde présentation du stimulus (phase de reconnaissance), le sujet fait correspondre sa représentation interne du stimulus avec le stimulus présenté. En d'autres termes, le sujet va reproduire la séquence de mouvements oculaires mémorisée en vérifiant que les caractéristiques de la scène visuelle encodée sont bien présentes. Une fois que le scanpath est reconstitué, le sujet reconnaît l'objet. Ainsi, le scanpath peut être assimilé à une forme de connaissance procédurale. Cette conclusion est étayée par deux observations lors de leur expérience. D'une part, les scanpaths étaient inconsistants entre les sujets, ce qui tend à montrer que ces derniers ne sont pas uniquement déterminés par les caractéristiques de la scène visuelle. D'autre part, le scanpath d'un sujet donné différait selon les images présentées, indiquant que le scanpath n'est pas un phénomène stéréotypé, mais plutôt un pattern de visualisation consistant pour une image particulière.

Le scanpath est ainsi envisagé comme un processus de type « top-down ». Toutefois, cette approche du scanpath a été affinée, voire contredite ultérieurement par Groner, Walder et Groner ([Groner et al., 1984]). Ces auteurs distinguent deux types de scanpaths : les scanpaths locaux, subissant une influence de type « bottom-up » où les saccades sont guidées par les éléments proximaux repérés par la vision périphérique et les scanpaths globaux, influencés par un processus de type « top-down », intervenant pour explorer globalement une scène visuelle.

C'est ainsi que pour ces auteurs, lorsqu'une recherche visuelle est dirigée par une tâche spécifique, les scanpaths observés seront globaux. Le scanpath décrit par Noton et Stark ([Noton & Stark, 1971a]) correspond aux scanpaths locaux.

Josephson et Holmes ([Josephson & Holmes, 2002]) ont proposé de tester la théorie de Noton et Stark sur le Web. Pour cela, ils ont demandé à des sujets de visualiser trois pages Web différentes (issues respectivement d'un portail, d'un site publicitaire et d'un site d'actualités) à 48 heures d'intervalle. Afin de comparer les scanpaths intra- et inter-sujets, ils ont utilisé la méthode de la distance d'édition qui consiste à transformer les scanpaths en un alphabet permettant de construire un indice de dissimilarité (qui tenait compte aussi des distances spatiales). Leurs résultats indiquent qu'il y a plus de différences entre les scanpaths inter-sujets qu'entre les scanpaths intra-sujets pour les pages Web ne respectant pas ou peu la convention de lecture « droite-gauche, haut-bas » (le portail et la page publicitaire). A contrario, les scanpaths inter-sujets se ressemblent pour la page issue du site d'actualités. Il semblerait donc que les sujets aient leurs propres habitudes de visualisation mais que d'autres variables interviendraient dans la définition du scanpath comme la mémoire ou les caractéristiques des pages Web. Néanmoins, il convient de signaler que les résultats de Josephson et Holmes s'appuient sur des statistiques descriptives ne permettant pas de décider avec certitude de la similarité ou non de deux scanpaths.

2.7 Conclusions et problématique

A de rares exceptions près, la plupart des études que nous avons citées tentent de schématiser globalement les comportements oculaires des internautes. En effet, les auteurs cherchent à savoir par exemple comment se distribuent les fixations sur les résultats d'un moteur de recherche ou encore si les seniors ont des comportements différents des plus jeunes. Pourtant, il est possible qu'il y ait des sous-structures occultées par cette description globale. En d'autres termes, il se peut que des catégories ou sous-catégories de comportements existent dans cette globalité.

Cette approche pourrait nous permettre d'affiner nos connaissances car il est possible qu'indépendamment des caractéristiques des individus, des « styles » ou « patterns » de visualisation existent. Le potentiel d'une telle découverte aurait des implications pratiques dans la mesure où le comportement visuel de l'internaute, étant donné son contexte et ses caractéristiques, fait aujourd'hui l'objet d'une abstraction issue de données agrégées. Ainsi, les recommandations actuelles concernant la conception des pages Web, ou encore la formation à la recherche d'informations pourraient être améliorées par une granularité plus fine de cet internaute abstrait. Néanmoins, les connaissances issues de ce type de travaux sont rares, et doivent faire l'objet de recherches d'un point de vue méthodologique,

comme nous le verrons dans le chapitre suivant.

La problématique de cette thèse est donc double :

1. Comment identifier des patterns comportementaux oculaires sur les pages Web ? En d'autres termes, quelle méthodologie adopter pour répondre à cette question ?
2. Existent-ils des patterns comportementaux de visualisation des pages Web ?

Chapitre 3

Aspects méthodologiques de l'étude des comportements visuels

Malgré la nécessité de comprendre les comportements visuels des internautes, la question des méthodes d'analyse reste prépondérante. En effet, comment assurer que les modèles comportementaux proposés se basent sur des méthodes d'analyse fiables ? En d'autres termes, a-t-on les bons outils pour analyser les données oculométriques ?

Les fabricants d'oculomètres proposent leurs propres solutions. Il est ainsi possible de définir des aires d'intérêts correspondant à des parties des pages Web faisant l'objet d'hypothèses de recherche. Ensuite, un certain nombre d'indicateurs oculométriques sont enregistrés au sein de ces zones (ex. nombre de fixations, durée des fixations, durée avant la première fixation,...). Les outils existants proposent aussi de visualiser les scanpaths de différentes manières. Comme on peut s'en douter, ces outils « standards » sont très limités, notamment dans la comparaison des scanpaths.

Dans les sections suivantes, nous allons présenter les méthodes d'analyse les plus couramment utilisées dans la littérature afin d'identifier celles qui nous seraient utiles pour répondre à notre problématique.

3.1 La définition des aires d'intérêts (*AOIs*)

Plusieurs méthodes existent pour définir les *AOIs* : selon une grille uniforme, par leur fonction, par leur position dans la page, ou encore selon le *DOM* (Document Object Model, c'est à dire la description des objets HTML dans le cas des pages Web). Généralement, les auteurs définissent les *AOIs* en fonction de

leurs besoins, cette pratique ne faisant pas l'objet d'un consensus.



FIGURE 3.1 – Exemple de zones d'intérêts définies a priori sur une page Web ([Drusch et al., 2012]).

3.1.1 Selon une grille uniforme

Il est possible de découper une page Web selon une grille uniforme. Cette méthode a l'avantage de ne pas nécessiter d'avoir une connaissance a priori du stimulus présenté aux participants. Cependant, il est évident que cette méthode est limitée car aucun sens ne peut être attribué aux zones ainsi créées. A notre connaissance, très peu d'études ont utilisé cette approche.

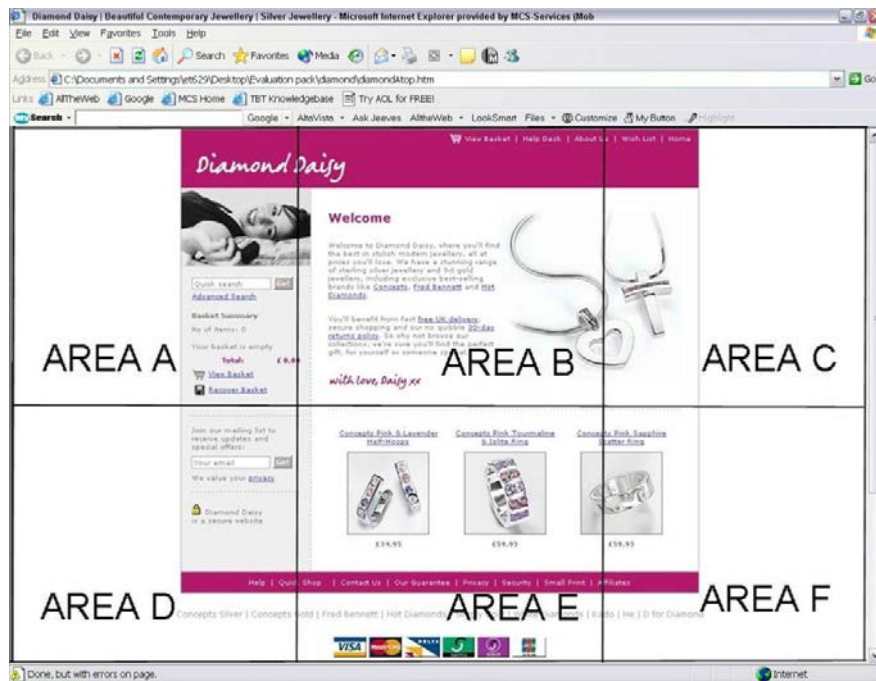


FIGURE 3.2 – Exemple de zones d'intérêts définies selon une grille uniforme ([Tzanidou et al., 2005]).

3.1.2 Par fonction

Yesilada, Jay, Stevens et Harper ([Yesilada et al., 2008]) ont proposé un découpage des pages Web en différentes sections dans le but de comprendre de quelle façon les utilisateurs utilisent les éléments visuels des pages Web pour réaliser leurs tâches. Pour cela, ils se sont basés sur le *Travel Analysis Framework (TAF)* qui considère la navigation sur le Web comme un voyage. Dans le *TAF*, trois grandes catégories d'éléments existent : les *Way Points* qui favorisent la prise de décision (ex. menus, logos, liens), les *Orientation Points* permettent de s'orienter et les *Travel Assistants* utilisés pour se ré-orienter (e.g., le moteur de recherche). Ces éléments ont été identifiés comme aires d'intérêts sur plusieurs pages Web afin d'enregistrer la part d'attention visuelle qui leur est allouée selon différentes tâches.

3.1.3 Par leur position dans la page

Cette méthode de définition des *AOIs* se rencontre le plus souvent dans l'étude des moteurs de recherche. En effet, la question qui se pose fréquemment est celle de l'ordre de visualisation des résultats des moteurs de recherche. Ainsi, Granka, Joakims et Gay ([Granka et al., 2004]) se demandent si le rang des résultats de recherche de Google a une influence sur leur exploration visuelle. Ils ont donc

découpé chaque liste en aires d'intérêts puis les ont identifi   selon leur position et ont enfin enregistré divers indicateurs oculom  triques par zone. Il est possible d'affiner les analyses en   tablissant des sous-aires d'int  r  ts correspondant aux diff  rents   l  ments de la liste (le titre, le r  sum   du site et l'URL du site) comme l'ont fait Cutrell et Guan ([Cutrell & Guan, 2007b]).

3.1.4 Selon le *DOM*

Le *DOM* (*Document Object Model*) d'une page Web correspond aux diff  rentes balises caract  risant chaque   l  ment (visible ou non pour l'utilisateur) de la page Web. Buscher, Cuttrel et Morris ([Buscher et al., 2009]) ont cherch      cr  er un mod  le computationnel afin de pr  dire la saillance des   l  ments du *DOM*. Pour cela, ils ont consid  r   chaque   l  ment comme une aire d'int  r  t puis les ont group  s selon leurs caract  ristiques. Ils ont aussi choisi de retirer de l'analyse les   l  ments larges, comme la balise « *body* », englobant l'ensemble de la page Web. Une s  rie d'indicateurs oculom  triques ont ensuite   t   enregistr  s sur ces aires d'int  r  ts.

3.1.5 Selon la structure des donn  es oculom  triques : la proc  dure Mean-Shift

L'id  e ici est de faire abstraction du stimulus pr  sent   aux participants en d  tectant les *AOIs* directement    partir des donn  es oculom  triques (g  n  ralement les fixations).

Santella et DeCarlo ([Santella & DeCarlo, 2004]) ont propos   une m  thode de classification, bas  e sur l'algorithme *mean-shift*, afin de d  finir *a posteriori*, et de fa  on automatique, les aires d'int  r  t. Il sont partis de l'id  e selon laquelle un tel algorithme devait :

- produire des r  sultats consistants, c'est    dire ne d  pendant pas d'une phase d'initialisation al  atoire (   l'instar du *k-means clustering*)
- ne pas n  cessiter de conna  tre    l'avance le nombre de classes/*AOIs* (classification «non param  trique»)
-   tre robuste, c'est    dire que les donn  es extr  mes ne doivent pas affecter l'ensemble de la classification

Les techniques de classification habituellement utilis  es, telles que l'*Expectation-Maximization* ou le *k-means clustering* n  cessitent, dans leurs impl  mentations de base, de conna  tre par avance le nombre de classes/*AOIs*. Ils ont ainsi introduit le mean-shift clustering pour d  tecter automatiquement,    partir des donn  es de fixations, les *AOIs* visualis  es par un ou plusieurs participants.

La proc  dure mean-shift d  place it  rativement les points de fixation vers des zones de plus forte densit      l'aide d'une fonction *kernel* ayant une taille de bande passante *h* afin d'y appliquer une m  thode de classification. Le principal

avantage est qu'il n'est pas nécessaire de définir le nombre d'aires d'intérêt a priori. Duchowski et ses collègues ([Duchowski et al., 2010]) ont appliqué cette méthode dans leur logiciel de visualisation et de comparaison des scanpaths.

Cette méthode fournit de bons résultats comme le montre la Figure 3.3. Cependant elle nécessite de définir correctement la fenêtre h .

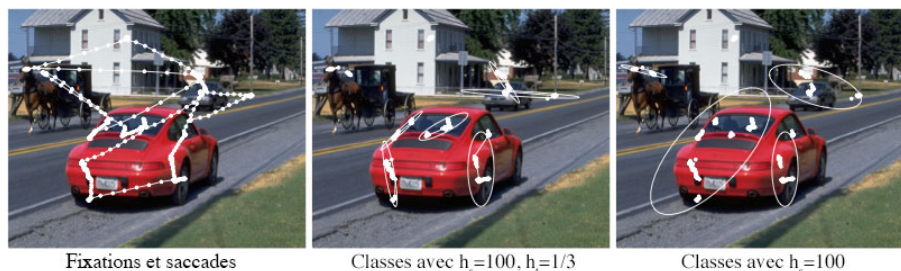


FIGURE 3.3 – Résultats de la procédure mean-shift pour un participant (adapté de [Santella & DeCarlo, 2004]).

3.2 L'analyse des données

La définition des aires d'intérêt n'est que la première étape avant l'analyse des scanpaths. Il existe deux approches différentes, correspondant à deux manières d'envisager les données oculométriques, que nous allons présenter : l'approche statique, qui permet d'extraire des informations quantitatives, et l'approche dynamique, destinée initialement à comprendre le cheminement visuel des utilisateurs.

3.2.1 Approche statique

L'analyse de données dans l'approche statique se contente de représenter « statiquement » les données, principalement dans le but de comparer quantitativement l'effet de diverses variables sur des données oculométriques. Certaines études se centrent sur des indicateurs globaux comme le temps de fixation par page Web, ou encore le nombre de fixations total. La plupart définissent des *AOIs* et enregistrent des mesures « statiques » au sein de celles-ci. Dans les métriques utilisées, certaines sont liées aux fixations (par exemple, le nombre de fixations, la durée des fixations, la durée séparant l'ouverture de la page de la première fixation d'une zone spécifique), d'autres aux saccades (par exemple, le nombre de saccades, l'amplitude des saccades) ou encore liées au regard (par exemple, le temps de regard sur une zone).

Dans cette méthode, les auteurs utilisent plusieurs mesures pour comparer les stratégies utilisateurs. Le plus souvent, il s'agira de comparer certaines données relevées entre les *AOI*, parmi lesquelles :

- La durée totale des fixations par *AOI*
- La moyenne des durées de fixation par *AOI*
- Le nombre total de fixations par *AOI*
- Le taux de fixations par *AOI*
- L'ordre de fixation des *AOI*
- La probabilité de visualisation de chaque *AOI*
- Localisation du premier regard (lors des 50 premières millisecondes)
- Localisation du point d'entrée (lors des 250 premières millisecondes)

C'est ainsi qu'il est possible pour certains auteurs de connaître quelles zones des pages Web servent de point de référence à l'orientation ou encore celles qui sont entièrement occultées par les utilisateurs.

Buscher, Cutrell et Morris ([Buscher et al., 2009]) ont proposé de prendre en compte de nouvelles mesures :

L'impact médian de la fixation par *AOI*

Cet impact est calculé en repérant tous les éléments du *DOM* qui sont à l'intersection d'une fixation, définie par un cercle de 2r de diamètre. Puis, un impact de distance est calculé, correspondant à l'aire de la fixation qui est couverte par l'élément du *DOM* (qui est une *AOI*). Si l'élément recouvre entièrement le cercle de fixation, cet impact sera de 1. Enfin, cet impact est pondéré par la durée de la fixation. L'impact médian de chaque *AOI* peut ainsi être calculé, et il correspondra, selon les auteurs, au taux d'attention visuelle qui lui est alloué.

La fréquence de visualisation

Il s'agit du taux d'utilisateurs qui regardent une *AOI* dans une page sur le nombre d'utilisateurs ayant vu cette page.

Le temps médian de la première fixation par *AOI*

La question ici est de savoir combien de temps il faut aux utilisateurs pour visualiser chaque *AOI* à partir du moment où ils ont ouvert la page Web.

3.2.1.1 L'extraction de connaissances dans l'approche statique

L'étude de Yesilada et son équipe ([Yesilada et al., 2008]) illustre bien le principe général des méthodes statiques dans une approche « *top-down* ». En effet, le postulat de départ est que certains éléments des sites Web sont susceptibles de jouer un rôle dans les comportements oculaires des internautes. Partant de cette idée, les chercheurs vont segmenter en aires d'intérêts ces divers éléments pour enregistrer divers indicateurs oculométriques et ainsi en connaître leur distribution. Les éléments du TAF ont été identifiés comme aires d'intérêts sur plusieurs pages Web afin d'enregistrer la part d'attention visuelle qui leur est

allouée selon différentes tâches. Leur étude montre notamment que le logo, le *header* et les titres sont fixés par 70% des participants à l'entrée sur la page. Mais le logo n'est fixé directement que dans 30% des cas, faisant du *header* et du titre les points de référence lors de l'entrée sur un site Web. En outre, le centre de la page est plus fréquemment regardé que le bas de page (pied de page).

Avec ce type méthodes, de nombreux auteurs ont étudié l'effet de différentes variables sur la distribution de mesures oculométriques statiques sur une page Web ou entre aires d'intérêt. Ainsi, on peut relever des études concernant l'effet du type de site Web (Habuchi, Takeuchi et Kitajima, 2007 ; Pan et al., 2004), le type de tâche (Tzanidou et al., 2005 ; Cutrell et Guan, 2007 ; Buscher, Cutrell et Morris, 2009) ou encore des variables liées à l'utilisateur comme l'âge (Fukuda et Bubb, 2003 ; Grahame, Laberge et Scialfa, 2004 ; Djamasbi et al., 2010), le genre (Pan et al., 2004 ; Lorigo et al., 2006 ; Buscher, Cutrell et Morris, 2009) ou la culture (Dong et Lee, 2008).

3.2.1.2 Conclusions et limites de l'approche statique

La comparaison des scanpaths est très limitée par l'usage de méthodes statiques. Les indicateurs retenus seront forcément très globaux et ne prennent pas en compte l'aspect dynamique d'un scanpath, qui correspond à une succession temporelle de fixations et de saccades oculaires. C'est ainsi que des mesures tenant compte de cet aspect sont apparues dans la littérature.

3.2.2 Les méthodes de visualisation

Ces méthodes vont permettre aux chercheurs de visualiser les données oculométriques sous différents angles, imperceptibles dans leur format initial, afin qu'ils en identifient des structures à partir de la perception qu'ils en ont. De nombreuses techniques ont été développées ces dernières années ([Blascheck et al., 2014]), nous présentons ici les études qui visent à identifier des patterns de visualisation.

Nielsen ([Nielsen, 2006]) a agrégé les heatmaps de plus de deux-cent participants qui ont visualisé différents sites Web. Il a observé un pattern comportemental qu'il qualifie d'« universel », le « *F-shape* » (Figure 3.4). Ainsi, les utilisateurs visualiseraient n'importe quelle page Web en balayant le haut de la page dans un premier mouvement horizontal, suivi d'un deuxième moins large et plus bas dans la page pour finir l'exploration dans un mouvement vertical descendant et longeant le côté gauche.



FIGURE 3.4 – Le « F-shape » ([Nielsen, 2006]).

Aula, Majaranta et Raiha ([Aula et al., 2005]) voulaient analyser les « styles de visualisation » des internautes sur les moteurs de recherche. Ils ont donc développé une méthode de visualisation statique présentant l'ordre dans lequel chaque aire d'intérêt est visualisée. Pour comparer les scanpaths, ils ont imprimé les visualisations qu'ils ont ensuite inspecté séparément puis ensemble. Après accord entre les auteurs, ils ont identifié deux patterns comportementaux : les « économes » et les « exhaustifs ».

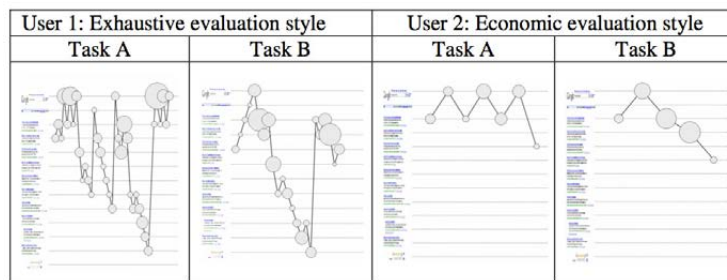


FIGURE 3.5 – Les styles de visualisation identifiés par Aula, Majaranta et Raiha ([Aula et al., 2005]).

Pellacini, Lorigo et Gay ([Pellacini et al., 2006]) puis Lorigo et ses collègues ([Lorigo et al., 2008b]), dans le contexte de la visualisation des listes de résultats de moteurs de recherche, ont développé une vue circulaire des scanpaths. Les aires d'intérêts sont représentées de façon circulaire par des lettres (de A à J dans la Figure 3.6a). Autour de ces lettres figurent des cercles dont la taille dépend du nombre de visites qu'a reçue chaque aire d'intérêt. Puis, en périphérie de l'ensemble, les transitions entre les zones sont représentées aussi de façon circulaire.

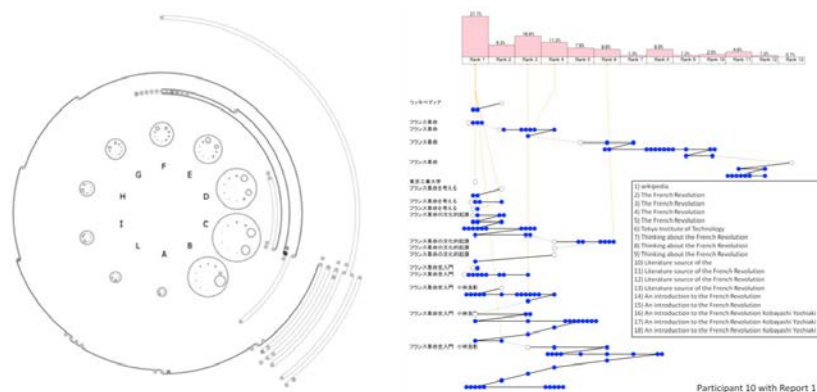


FIGURE 3.6 – Visualisation des scanpaths sur des résultats de moteurs de recherche : (a) sous forme circulaire (Pellacini, Lorigo et Gay, 2006) et (b) sous forme d’histogramme (Egusa et al, 2008).

Egusa et son équipe ([Egusa et al., 2008]) ont tenté de représenter les comportements visuels des utilisateurs lorsqu’ils interagissent avec un moteur de recherche (Figure 3.6b). Ici les rectangles présents dans la partie supérieure correspondent, dans l’ordre, aux rangs des résultats de recherche. Au dessus de ceux-ci, des histogrammes quantifient la proportion de fixations pour chaque rang.

En dessous de ces éléments, les requêtes sont listées et les fixations sont représentées par des cercles (blancs pour la première fixation et bleus pour les autres). Les lignes en pointillés entre les fixations signifient que l’utilisateur a quitté la page Web du moteur de recherche. Chaque ligne jaune partant des rangs correspond à un clic.

Dong et Lee ([Dong & Lee, 2008]) ont cherché à comparer la perception des pages Web dans une perspective interculturelle. Pour cela, ils ont conçu des graphiques en radars censés refléter les patterns de visualisation d’Américains, de Chinois et de Coréens. Deux types de graphiques, basés sur deux critères différents, ont ainsi été conçus : un premier correspondant au type de pattern de visualisation (exploration séquentielle, circulaire, centrée sur les éléments de navigation, sur le titre...) et un second à la forme des patterns (en « 0 », en « 5 », en « N », en « Z », en « X »). Ces deux critères ont été analysés visuellement à partir de l’impression des scanpaths. Au final, deux graphiques par groupe de participants et synthétisant les deux critères ont permis de procéder à une analyse comparative.

Tsang et Swindells ([Tsang et al., 2010]) ont proposé de «tagger» chaque AOI (c’est à dire lui donner un sens) puis d’observer le cheminement de groupes de scanpaths identifiés comme étant similaires ou supposés comme tel. Leur logiciel, eSeeTrack, propose une visualisation des scanpaths sous forme d’arbre où des groupes peuvent être distingués. Dans la figure 3.7 l’arbre des scanpaths

commence par l'AOI *vital_screen* puis se subdivise en différentes branches permettant d'appréhender de manière simplifiée le chemin visuel de groupes d'individus. Ici, les AOIs (noeuds de l'arbre) en couleur sont celles visualisées par des personnes définies comme expertes. Ainsi, il est possible d'étudier les différences entre individus experts et non experts lors d'une tâche afin d'en extraire des connaissances et des recommandations.

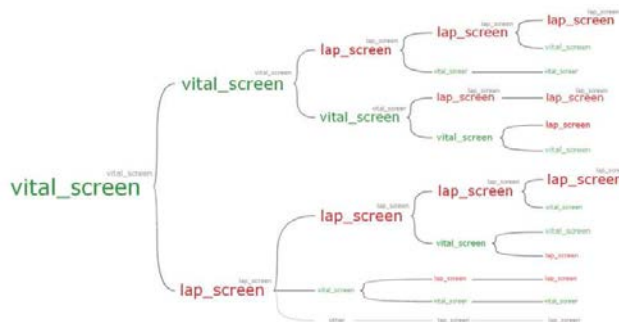


FIGURE 3.7 – Visualisation sous forme d'arbres de fixations sur des AOIs taggées lors d'une tâche de simulation chirurgicale ([Tsang et al., 2010]).

3.2.3 Méthodes dynamiques

Ce type de méthode est actuellement en émergence. Plusieurs auteurs ont constaté qu'il y a un certain besoin de méthodes objectives, et ne faisant pas intervenir directement la perception des chercheurs, permettant de comparer les scanpaths.

Les approches dynamiques sont essentiellement basées sur des méthodes d'alignement et l'usage d'AOIs. Il s'agit de comparer les scanpaths deux à deux après les avoir transformés en chaînes de caractères (chaque caractère correspondant à une AOI). Il en existe deux types, la première est dite « locale », où l'objectif est de trouver des sous-patterns de visualisation communs entre deux scanpaths ou plus ([Hembrooke et al., 2006, West et al., 2006, Kübler et al., 2014]). Différentes méthodes ont été proposées, mais elles se basent généralement sur l'identification de sous-séquences communes entre les scanpaths.

La seconde, plus fréquemment utilisée, est dite « globale » et permet de comparer des scanpaths dans leur globalité. Dans cette catégorie d'alignement, la méthode de la distance d'édition (ou distance de Levenshtein) a été introduite dans ce domaine par Brandt et Stark ([Brandt & Stark, 1997]). Celle-ci permet de comparer les scanpaths deux à deux pour calculer une distance qui prend en compte les coûts d'insertion, de remplacement et de suppression nécessaires à la transformation d'une chaîne en une autre.

3.2.3.1 La distance de Levenshtein et ses variantes

La première étape consiste à définir des aires d'intérêt et leur attribuer un identifiant unique. Puis, chaque point de fixation de chaque scanpath est associé à une aire, en fonction de sa position. Ensuite, chaque scanpath est représenté par une chaîne de caractère représentant l'ordre des fixations de chaque aire d'intérêt. Enfin, il suffit de calculer un indice de similarité S_s pour chaque paire scanpaths en prenant en compte les coûts d'insertion, de substitution et de suppression liés à la transformation d'une chaîne en une autre. S_s varie de 0 (similarité parfaite) à 1 (dissimilarité parfaite).

Par exemple, pour les scanpaths $S_1 = ABCDE$ et $S_2 = ABBCE$, la distance d'édition sera de 2 et l'indice de similarité $S_s(S_1, S_2) = 0.4$.

Utilisé de cette façon ([Brandt & Stark, 1997]), la distance d'édition ne nous renseigne pas suffisamment sur la différence qu'il y a entre deux scanpaths. C'est pour cette raison que des améliorations ont été proposées dans la littérature.

Prise en compte des zones communes entre scanpaths

Privitera et Stark ([Privitera & Stark, 2002]), puis Duchowski et son équipe ([Duchowski et al., 2010]) ont associé à l'indice de similarité S_s , un autre indice S_p qui reflète le nombre de zones communes entre deux scanpaths (0 pour aucune zone en commun, 1 pour toutes les zones en commun). Dans l'exemple précédent, quatre zones sont communes (sur cinq) entre les deux scanpaths, ce qui nous donne un indice $S_p(S_1, S_2) = 0.8$.

Pour mieux analyser les résultats, les scanpaths peuvent être catégorisés selon leur nature et le stimulus (Figure 3.8). R pour les scanpaths Répétitifs, où le même sujet visualise le même stimulus à différents moments; L pour Local, des sujets différents visualisent le même stimulus; I pour Idiosyncratique, le même sujet visualise des stimulus différents; G pour Global, où différents sujets visualisent différents stimuli.

Same Subject	Different Subjects		Same Subject	Different Subjects
Repetitive 0.64	Local 0.54	Same Picture	Repetitive 0.42	Local 0.28
Idiosyncratic 0.34	Global 0.28	Different Pictures	Idiosyncratic 0.21	Global 0.16
S_p			S_s	
	Random 0.21			Random 0.04

FIGURE 3.8 – Exemple de catégorisation des scanpaths ([Privitera & Stark, 2002]).

Prise en compte de la distance inter-zones

Les premiers auteurs à avoir pris en compte la distance inter-zones sont Joseph-

son et Holmes ([Josephson & Holmes, 2002]). Pour eux, assigner des coûts de substitution uniformes ne permet pas de représenter la variabilité des scanpaths. En effet, des coûts uniformes attribuent la même valeur à deux zones adjacentes qu'à deux zones éloignées. Dans leur article, ils ont donc décidé que le coût de substitution d'un caractère par un autre serait inversement proportionnel à leur proximité physique. Ainsi, pour calculer un coût de substitution, il suffit de compter le nombre de zones séparant deux points de fixation puis de diviser ce nombre par le nombre maximum de zones pouvant séparer deux points de fixation. Par exemple, pour deux zones adjacentes visualisées A et B, où le nombre maximal de zones séparant deux points de fixation est de 5, le coût de substitution serait de 0.2.

A	B	C
D	E	F
G	H	I

FIGURE 3.9 – Exemple d'aires d'intérêt ([Takeuchi & Habuchi, 2007]).

Takeuchi et Habuchi ([Takeuchi & Habuchi, 2007]) ont amélioré cette méthode en précisant sa limite. En effet, le calcul du coût de substitution de Josephson et Holmes repose sur le concept d'adjacence et non de distance physique. Par exemple, le coût de substitution de la zone A par la zone B serait le même (0.33) que celui de la zone A par la zone E en prenant en compte l'adjacence (Figure 3.9). Or, la distance physique des centres de ces paires de zones diffère, comme pourrait le montrer le calcul des distances en pixels.

Pour Takeuchi et Habuchi, la fonction du coût de substitution doit se baser sur la distance entre le centre des zones (qui est fonction du nombre de zones et la taille de l'écran). De cette façon, le coût de substitution est réellement proportionnel à la distance inter-zones.

3.2.3.2 L'algorithme «*MultiMatch*»

L'objectif initial de cet algorithme est de permettre la comparaison des scanpaths d'un point de vue géométrique. En d'autres termes, il est censé répondre à la question : «ces deux scanpaths sont-ils similaires?».

Le *MultiMatch* est une mesure de similarité multidimensionnelle inter-scanpaths, basée sur les vecteurs ([Jarodzka et al., 2010, Foulsham et al., 2012]). Dans cette méthode, les scanpaths sont vus comme une série de vecteurs correspondants à des saccades «idéales» de déplaçant entre des fixations munies d'une durée. Cette représentation permettrait de conserver à la fois la forme des

scanpaths ainsi que l'ordre, la position, et la durée des fixations. L'algorithme procède en trois phases :

1. Simplification des scanpaths :

- Les saccades en dessous d'un seuil d'amplitude défini (typiquement 4r) sont combinées avec la ou les suivantes.
- Les saccades qui ont un angle de moins de 45r par rapport à la précédente sont combinées.

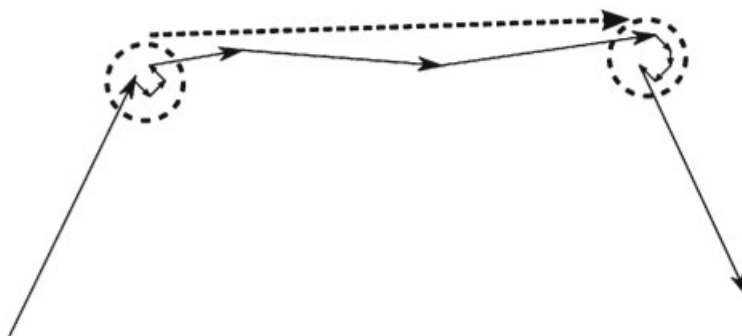


FIGURE 3.10 – Etape de simplification des scanpaths dans l'algorithme Multi-Match. Les cercles en pointillés représentent le critère d'amplitude et la flèche en pointillés le critère angulaire ([Dewhurst et al., 2012]).

2. Alignement des scanpaths :

- Chaque vecteur d'un scanpath est comparé avec les vecteurs d'un autre scanpath, ce qui génère une matrice de comparaisons.
- L'alignement séquentiel le plus proche entre les scanpaths est détecté à l'aide de l'algorithme du plus court chemin de Dijkstra.

3. Comparaison des scanpaths :

Sur la base des scanpaths alignés, chaque paire de vecteurs fait l'objet de comparaisons. Différentes métriques sont proposées :

- Différence vectorielle
- Différence d'amplitude
- Différence angulaire
- Distance euclidienne entre les fixations
- Différence en durée entre les fixations

Cette méthode a l'avantage de ne pas nécessiter la définition d'*AOIs a priori*. Cependant, le contexte du stimulus n'est pas pris en compte, ce que permettent justement les *AOIs*, ce qui la rend moins intéressante pour l'identification de patterns d'exploration visuelle. De plus, l'étape de simplification proposée conduit à la suppression de fixations (par la combinaison de saccades), surtout pour le critère angulaire, qui pourraient pourtant être importantes.

3.3 Conclusions sur les méthodes

Analyser des données oculométriques est une tâche complexe, d'autant plus que le choix de la méthode peut conditionner les résultats des études. En effet, par exemple, la modification des *AOIs* peut engendrer des résultats potentiellement différents entre groupes d'individus étudiés. C'est pour cette raison que certains chercheurs tentent de trouver le moyen de s'en affranchir. L'autre question est de savoir comment analyser la dynamique des scanpaths. L'application des méthodes dynamiques pour mieux comprendre les comportements des internautes demeure rare, mais le point principal porte, selon nous, sur leur fiabilité et leur capacité à modéliser de façon exhaustive les scanpaths. Nous aborderons ces limites dans le chapitre 5.

Jusqu'à présent, peu d'études ont cherché à identifier des typologies de patterns comportementaux, et la plupart d'entre elles se basent sur des méthodes de visualisation. Bien que cette approche soit séduisante, il manque des critères clairs de validation des patterns identifiés et une méthode pour les détecter de façon automatique (de manière «*bottom-up*») sur de large jeux de données. Dans le prochain chapitre, nous présentons les travaux que nous avons mené pour obtenir des éléments de réponses à ces questions.

Deuxième partie

A la recherche d'une méthode
d'analyse

Chapitre 4

Différentes méthodes d'analyse pour les patterns

Afin d'identifier des typologies de patterns comportementaux visuels, nous avons mené une expérimentation où des internautes devaient réaliser différentes tâches de recherche d'information sur différents sites Web. Nous leur avons aussi demandé de répéter certaines tâches immédiatement afin d'étudier la stabilité de ces patterns. Cette étude visait à obtenir un échantillon d'internautes aussi homogène que possible afin de savoir si, indépendamment de toute variable socio-démographique, de tels patterns comportementaux existent. Dans ce chapitre, nous rapportons les différentes méthodes d'analyse que nous avons appliquées pour atteindre notre objectif sur les pages d'accueil de deux des sites étudiés. De façon générale, et à l'exception de la dernière (*T-patterns*), elles comportent trois étapes :

1. Appliquer une méthode de comparaison des scanpaths (distance inter-scanpaths) ;
2. Générer une matrice de distances inter-scanpaths ;
3. Appliquer une méthode de classification sur la matrice de l'étape (2) pour identifier des classes de scanpaths bien distinctes ;
4. Si l'étape (3) a donné des résultats, caractériser les classes identifiées.

Nous avons choisi des méthodes de comparaison des données oculométriques qui envisagent des aspects différents des scanpaths : les indicateurs oculométriques, la position des fixations et les *T-patterns*. De cette façon, nous avons été en mesure de prendre en compte les durées, les positions et l'ordre des fixations selon la méthode appliquée.

Nous verrons finalement que certaines méthodes nous ont permises d'atteindre, au moins partiellement, notre objectif, mais que des doutes subsistent quant à leur pertinence.

4.1 Méthodologie expérimentale

4.1.1 Matériel et terrain

Nous avons utilisé le système d'oculométrie déporté Tobii T120, basé sur la technique du reflet cornéen et capturant les images de la réflexion à 60Hz. C'est le logiciel Tobii Studio (v.1.3.14) qui enregistrait les mouvements oculaires, avec un filtre définissant une fixation comme ayant une durée minimale de 100ms dans un rayon de 35 pixels.

Il s'agissait d'une procédure expérimentale de laboratoire, réalisée dans la salle d'observation du laboratoire Pergolab. Les participants étaient installés face au système Tobii. Ce dernier était posé sur une table stabilisée et ajustée en hauteur. En outre, la distance entre l'écran et la chaise était constamment d'environ 70cm.

La navigation sur les sites Web était rendue possible par la mise à disposition d'une souris. Le clavier était contrôlé par l'expérimentateur pour lancer les différentes phases de l'expérimentation mais aussi pour minimiser une éventuelle perte de données oculométriques engendrée par les frappes clavier.

4.1.2 Procédure

Lors de l'accueil des participants, nous leur précisions qu'ils auraient à répondre à des questions en navigant sur un site Internet particulier sans avoir l'usage du clavier. De plus, nous leur expliquions à l'avance qu'ils subiraient cinq phases de calibration et ce en quoi elles consistaient. En dehors des deux premières tâches d'exploration, les participants n'avaient pas de limites temporelles pour atteindre l'objectif fixé et pouvaient abandonner à tout moment.

Tous les participants réalisaient ensuite une tâche d'entraînement (exploration libre) sur deux sites Web puis 2 tâches (T) répétées (R) sur deux sites Web (SW) dans le même ordre selon ce plan expérimental :

$$S_{113} < SW_2 * T_2 * R_2 >$$

1. « Explorez librement pendant 15 secondes la page Web qui va apparaître sans cliquer sur aucun lien ». Page Web : www.priceminister.com (page d'accueil enregistrée).



FIGURE 4.1 – Page d'accueil du site Priceminister

2. « Explorez librement pendant 15 secondes la page Web qui va apparaître sans cliquer sur aucun lien ». Page Web : www.natureetdecouverte.com (page d'accueil enregistrée).



FIGURE 4.2 – Page d'accueil du site Nature et Découverte

Les tâches 3 à 6 étaient réalisées deux fois de suite.

3. « A l'aide du site Internet qui va s'afficher, répondez à la question suivante : quels sont les trois cas permettant de dissoudre un PACS? Quand vous aurez trouvé les réponses, citez-les à haute voix. ». Page Web de départ : www.service-public.fr (page d'accueil).

4. « A l'aide du site Internet qui va s'afficher, répondez à la question suivante : quels sont les deux premières conditions générales de recrutement pour accéder à la fonction publique (c'est à dire les conditions de recrutement pour être fonctionnaires)? Quand vous aurez trouvé les réponses, citez-les à haute voix. ». Page Web de départ : www.service-public.fr (page d'accueil).



FIGURE 4.3 – Page d'accueil du site Service Public

5. « A l'aide du site Internet qui va s'afficher, répondez à la question suivante : quelle est l'adresse du théâtre municipal du 15e arrondissement parisien ? Quand vous aurez trouvé la réponse, citez-la à haute voix. ». Page Web de départ : www.carnavalet.paris.fr (page d'accueil).

6. « A l'aide du site Internet qui va s'afficher, répondez à la question suivante : quel est le plein tarif d'une visite-conférence au musée Carnavalet ? Quand vous aurez trouvé la réponse, citez-la à haute voix. ». Page Web de départ : www.carnavalet.paris.fr (page d'accueil).

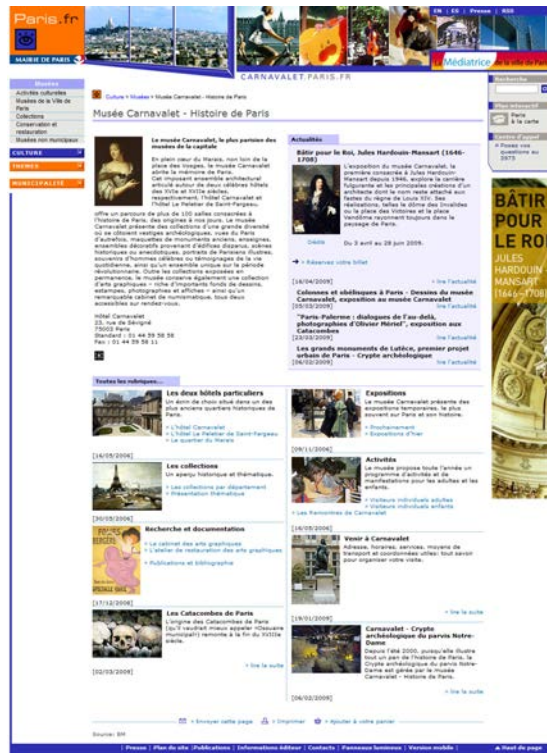


FIGURE 4.4 – Page d'accueil du site du Musée Carnavalet

Ces questions apparaissaient à l'écran, puis étaient suivies de l'apparition d'une page Web blanche où figurait un lien permettant d'accéder au site Web stimulus dans le but de réduire les temps de latence d'affichage et de positionner le regard du participant.

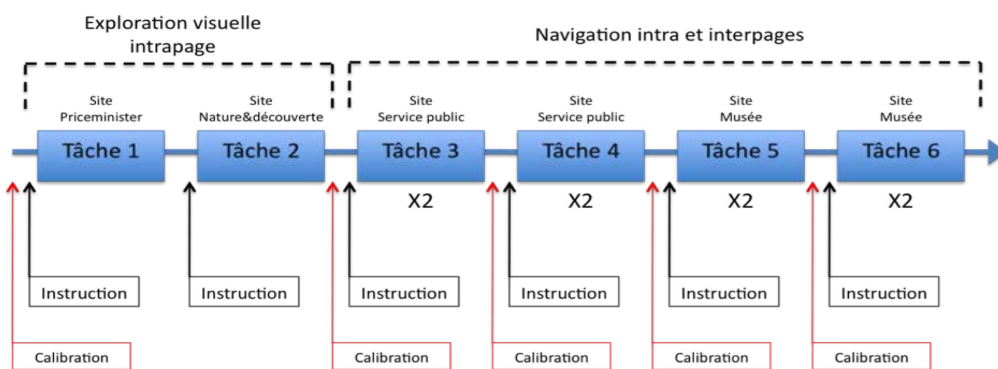


FIGURE 4.5 – Schéma récapitulatif de la procédure expérimentale de l'expérimentation.

Les différentes tâches n'ont pas fait l'objet d'un contrebalancement car, étant dans l'optique d'une identification de patterns, notre objectif était que tous les

participants soient placés dans des conditions strictement similaires. Ne testant pas l'influence de la structure des pages Web ou des tâches, l'effet d'ordre ne constitue pas un biais expérimental dans notre cas.

Chaque binôme de tâche était précédé d'une phase de calibration et chaque tâche de l'effacement de l'historique de navigation d'Internet Explorer (v. 7).

Une fois que les sujets avaient lu une consigne, nous leur demandions de nous dire ce qu'ils devaient faire. Suite aux différents pré-tests, il nous est paru nécessaire d'insister oralement et systématiquement sur certains mots-clefs afin que les sujets aient parfaitement assimilé l'objet de leur recherche.

Une tâche prenait fin :

- Si le sujet trouvait la bonne réponse.
- Si le sujet donnait une mauvaise réponse
- Si le sujet décidait d'abandonner.

De la sorte, les passations étaient entièrement standardisées.

Enfin, un questionnaire de fin d'expérience était rempli par les participants.

4.1.3 Questionnaire de fin d'expérience

Ce questionnaire nous a permis de contrôler a posteriori l'homogénéité de notre échantillon selon différents critères, mais aussi de vérifier qu'il n'y ait pas d'autres facteurs expliquant mieux les classes supposées que les données comportementales.

Ce questionnaire était découpé en plusieurs parties :

- Questions concernant l'étude : Nous vérifions que les participants n'étaient pas déjà habitués aux expérimentations sur Internet, et qu'ils n'étaient pas conscient de l'objectif de la nôtre. Nous évaluons le degré de connaissance antérieure des sites Web parcourus durant les passations et le degré de difficulté perçu des différentes tâches.
- Fiche signalétique : Recueil de données démographiques, évaluation des problèmes ophtalmologiques, ancienneté et fréquence de l'usage d'Internet, sites Web visités le plus souvent et navigateur Web utilisé.
- Connaissance d'Internet : Il s'agit de sept items où les sujets doivent se positionner en termes d'expertise par rapport à des affirmations. Cette partie du questionnaire a été inspirée et traduite de Lim et Lee ([Lim & Lee, 2000]).

4.1.4 Echantillon

L'échantillon total s'élève à 113 participants. A ce nombre, nous avons soustrait 16 participants à cause de problèmes de calibration, puis 9 autres en raison d'un temps de latence d'affichage du stimulus trop long. En effet, pour ces

participants, la page web n'affichait que la bannière d'en-tête lors des premiers temps. L'échantillon faisant l'objet d'une analyse est donc de 88 participants.

Le tableau ci-dessous présente les caractéristiques de notre échantillon :

Variabiles	Données sur les variables
Âge moyen (écart-type)	23,64 ans (3,72)
Proportion d'hommes et de femmes	47,7% d'hommes / 52,3% de femmes
Port de lunettes	50,5% de l'échantillon porte des lunettes ou des lentilles
Ancienneté de l'utilisation d'Internet	Plus de 3 ans pour 97,7% de l'échantillon
Fréquence d'utilisation d'Internet	Au moins 1 heure par semaine pour 88,6% de l'échantillon.
Fréquence d'utilisation du site www.servicepublic.fr	Très rarement ou rarement pour 99% de l'échantillon

TABLE 4.1 – Caractéristiques de l'échantillon.

4.2 Méthodologie de détection des patterns comportementaux

Afin de détecter des patterns comportementaux dans les données de fixations fournies par l'oculomètre de notre expérimentation, nous avons choisi de nous baser sur une méthode de classification automatique non supervisée.

4.2.1 Généralités sur la classification automatique

La classification automatique sur les sujets a pour but de classer des individus dans des groupes les plus homogènes possible. De cette façon, l'homogénéité des classes sera maximale alors que l'homogénéité entre les classes sera minimale. En d'autres termes, en prenant en compte un ensemble de variables donné, les sujets appartenant à une classe seront relativement semblables et les sujets n'appartenant pas aux mêmes classes seront peu semblables. La classification automatique est en fait une technique nous permettant de simplifier la réalité. En effet, nous sommes capables de résumer une information en deux voire trois dimensions, mais au delà, il nous est impossible d'appréhender la complexité de la réalité. Or, sur un jeu de données, il y a autant de dimensions qu'il y a de variables. La classification automatique s'exécute sur des variables ou directement sur une matrice de dissimilarité. Les variables servant à classer les sujets sont dites actives. Il est alors possible d'ajouter des variables passives qui n'interviennent pas dans la classification, mais permettent de caractériser plus finement les classes.

4.2.2 Le partitionnement autour des médoïdes (PAM)

Le PAM ([Kaufman & Rousseeuw, 1990]) est une méthode pour classifier des données en minimisant l'erreur quadratique, c'est à dire la distance entre les points d'une classe et le point désigné pour être le centre de cette classe (le médoïde). Ces médoïdes sont les meilleurs représentants de leur classe (il sont les objets pour lesquels la dissimilarité moyenne aux autres objets de leur classe est la plus faible).

Cette méthode est ici appliquée sur les matrices de dissimilarités obtenues à l'aide des méthodes décrites dans les sections suivantes.

Le PAM procède en deux phases :

- «*Build*» : une classification initiale est obtenue en sélectionnant itérativement les objets représentatifs (médoïdes) jusqu'à ce que k objets aient été trouvés.
- «*Swap*» : la sélection de tous les objets représentatifs et donc de la classification est affinée. Cela se fait en calculant les effets de l'échange d'un médoïde par un non-médoïde.

Finalement, il est possible de savoir quel nombre de médoïdes produit la meilleure classification en calculant les coefficients de silhouette. Ces coefficients sont calculés pour deux classes de cette façon :

Pour deux classes A et C :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Où :

$\mathbf{a}(i)$ = dissimilarité moyenne entre le point i et tous les autres point dans A ;

$\mathbf{b}(i) = \min d(i, C)$ ($C \neq A$) ;

$\mathbf{d}(i, C)$ = dissimilarité moyenne entre le point i et tous les autres points dans C .

Le coefficient de silhouette s'étend de -1 (mauvaise classe) à 1 (bonne classe). Plus précisément, les résultats s'interprètent de cette façon ([Kaufman & Rousseeuw, 1990]) :

- Entre 0.71 et 1 : une structure forte a été identifiée
- Entre 0.51 et 0.70 : une structure raisonnable a été identifiée
- Entre 0.26 et 0.50 : la structure est faible et pourrait être artificielle.
- Moins de 0.25 : aucune structure n'a été détectée.

En calculant le $s(i)$ moyen pour tous les éléments d'une même classe, il est possible de connaître la robustesse globale d'une classification. Ainsi, en réitérant une classification en faisant varier le nombre de classes, il suffit simplement de sélectionner la meilleure configuration qui produit le coefficient de silhouette moyen le plus élevé.

4.2.3 La caractérisation des classes

Afin de caractériser les classes, nous utilisons 9 indicateurs supplémentaires, classiquement utilisés dans la littérature, et qui peuvent faire l'objet de statistiques :

- Nombre de fixations.
- Temporelles : durée totale des scanpaths (en ms), durée moyenne des fixations des scanpaths (en ms).
- Métrique (distance) : longueur des scanpaths (en pixels), amplitude moyenne des saccades (en pixels).
- Topographiques :
 - Le barycentre X des scanpaths pondéré par la durée des points de fixation (en pixels). Par exemple, un barycentre X élevé signifie que l'utilisateur fixe plus et plus longtemps la partie droite du stimulus.
 - Le barycentre Y des scanpaths pondéré par la durée des points de fixation (en pixels). Par exemple, un barycentre Y élevé signifie que l'utilisateur fixe plus et plus longtemps la partie basse de la page (ici, le point (0,0) se situe en haut à gauche de la page).
 - La dispersion des scanpaths, pondérée par la durée des points de fixation. Un indice de dispersion élevé fait référence à un scanpath dispersé.
 - L'aire de l'enveloppe convexe (en pixels au carré). Une grande aire de l'enveloppe convexe fait référence à un scanpath comportant des points de fixation de grande amplitude.

4.3 *Heatmaps* des données

Nous avons généré la *heatmap* des positions des fixations enregistrées sur les pages Web présentées aux participants et faisant l'objet d'une analyse dans ce chapitre.

De façon globale pour la page d'accueil du service public (1243 x 911 pixels), nous constatons que la plupart des éléments de la page Web ont été visualisés par les 113 participants.



FIGURE 4.6 – Heatmap des données de fixations enregistrées sur la page d'accueil du service public lors de l'expérimentation.

Pour la page d'accueil du musée Carnavalet (1243 x 1407 pixels), les participants ont visualisé moins de zones, et se sont essentiellement concentrés sur le haut de la page (menus et titre de la page).



FIGURE 4.7 – Heatmap des données de fixations enregistrées sur la page d'accueil du musée Carnavalet lors de l'expérimentation.

On constate que les *heatmaps* de ces deux pages Web sont très différentes, cela est probablement dû à leurs différences de structures et de taille, mais aussi aux tâches que devaient réaliser les participants. En outre, nous n'avons pas identifié de pattern global en forme de «*F*».

De façon évidente, il n'est pas possible de savoir si certaines zones ont été plus ou moins visualisées par certaines catégories d'individus sans formulation d'hypothèses.

4.4 Classification basée sur les indicateurs oculométriques

Nous avons constaté que dans la littérature, de nombreuses études comparent des groupes d'individus sur la base de différents indicateurs oculométriques. Nous avons souhaité essayer la démarche inverse : trouver des groupes d'individus à partir de ces indicateurs. Cette section rapporte les résultats de cette démarche.

4.4.1 Analyse des données

4.4.1.1 Matrice de distances basée sur des indicateurs oculométriques

Dans un premier temps, pour chaque individu, nous avons calculé ses 9 indicateurs oculométriques décrits *supra*. Ensuite, l'ensemble de ces indicateurs (sous forme de tableau *individu X indicateurs*) ont été centré-réduits afin d'éviter la sur- ou la sous-pondération d'un indicateur (fonction *scale* sous *R*). Enfin, ce tableau a servi à générer une matrice de distances euclidiennes inter-individus (fonction *dist* sous *R*).

4.4.1.2 Partitionnement autour des médoïdes (*PAM*)

La matrice des distances entre toutes les paires de scanpaths a servi d'entrée au *PAM* afin d'identifier des patterns. Finalement, en utilisant le coefficient de silhouette, le nombre de classes a été défini.

4.4.2 Résultats

L'application de la procédure que nous avons décrite a permis d'identifier deux classes pour chacune des deux tâches analysées. Pour la tâche 3, le coefficient de silhouette moyen s'élève à 0.17, et pour la tâche 5, le coefficient de silhouette moyen s'élève à 0.50. Il semblerait qu'aucune structure n'ait pu être identifiée dans les données à l'aide de cette méthode.

4.4.3 Conclusions

L'objectif de cette méthode était d'utiliser les indicateurs oculométriques fréquemment utilisés dans la littérature pour comparer les groupes d'individus comme base d'une classification visant à détecter des patterns comportementaux. Il en ressort que les classes détectées ne sont pas suffisamment robustes, et donc que les indicateurs oculométriques ne sont probablement pas suffisamment discriminants. De plus, ces indicateurs ne prennent pas en compte de manière fine les positions des fixations oculaires, ce qui nous semble important dans la comparaison des scanpaths. Néanmoins, ces indicateurs restent utiles et

pourraient servir à caractériser des classes identifiées par une autre méthode. Dans les méthodes que nous allons décrire dans les parties suivantes, nous avons décidé d'utiliser les indicateurs pour expliquer les classes plutôt que comme base de la classification.

4.5 Classification basée sur les positions des fixations

Dans cette section, nous présentons une méthode de comparaison des scanpaths basée sur la localisation des fixations, indépendamment de leur ordre. En effet, il nous a paru intéressant de tester ce type d'approche avant de prendre en compte l'ordre des fixations, pour limiter la variabilité des scanpaths.

4.5.1 Analyse des données

4.5.1.1 Matrice de distances basée sur la position des fixations

Afin de nous affranchir de toute notion d'aires d'intérêts, nous avons choisi de produire une matrice de dissimilarité qui ne tient pas compte de l'ordre des points des scanpaths, mais de la distance entre les points, relative à leur position. Cette dernière a été produite en plusieurs étapes (illustrées par deux scanpaths fictifs) :

Pour un couple de scanpaths :

1. Pour chaque point (p) du premier scanpath, on a cherché la distance minimale à tous les points (p') du second scanpath. De cette façon, le point le plus proche de chaque point du premier scanpath est repéré.

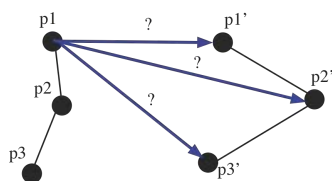


FIGURE 4.8 – Etape 1

2. La distance (D) la plus faible entre chaque paire de points est calculée



FIGURE 4.9 – Etape 2.

3. L'ensemble des distances entre les points du couple de scanpaths est agrégé par la moyenne des distances de ces points.

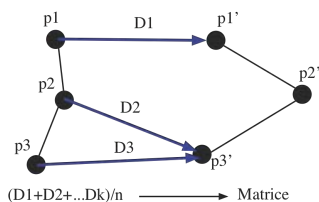


FIGURE 4.10 – Etape 3.

4. Cette agrégation a lieu deux fois, la matrice n'est donc pas symétrique.

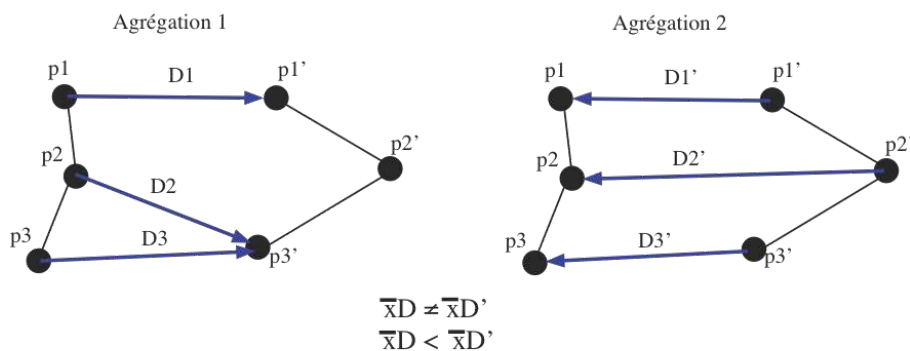


FIGURE 4.11 – Etape 4.

5. Afin d'obtenir une matrice symétrique, cette dernière est symétrisée par la distance maximale entre les moyennes des distances (dans notre exemple, ce serait la distance moyenne D' , donc l'agrégation 2, qui serait prise en compte).

6. L'ensemble des étapes précédemment décrites sont répétées jusqu'à ce que tous les couples de scanpaths aient été comparés.

7. On obtient ainsi une matrice de dissimilarités indiquant chaque couple de scanpaths.

De manière formelle, étant donné une paire de scanpaths SP_1 et SP_2 ayant chacun une liste de fixations $F = (f_1, f_2, \dots, f_n)$, la distance des positions est donnée par :

$$D_{sp} = \max \{d_{pos}(SP_1, SP_2), d_{pos}(SP_2, SP_1)\}; \quad (4.1)$$

Où :

$$d_{pos}(SP_1, SP_2) = \frac{1}{n} \sum_{f_i \in SP_1, f_j \in SP_2} \min(\|f_i, f_j\|^2)$$

Avec cette procédure appliquée à l'ensemble des données, nous obtenons une matrice des distances entre toutes les paires de scanpaths (Algorithme 4.1). Cette matrice est ensuite utilisée pour calculer les classes avec la méthode PAM.

Algorithme 4.1 Obtention d'une matrice de distance des positions des fixations entre toutes les paires de scanpaths.

Etant donné SP , une liste de scanpaths contenant des fixations.

Définir matrice[][] comme la matrice contenant les comparaisons de scanpaths.

pour chaque couple de scanpaths SP_i et SP_j **faire**

matrice[i][j] = distance des positions entre les fixations de SP_i et SP_j selon l'équation 4.1.

fin

4.5.1.2 Partitionnement autour des médoïdes (PAM)

Nous avons dans un premier temps établis une matrice des distances selon la procédure décrite dans le paragraphe précédent entre toutes les paires de scanpaths, puis nous avons analysé cette matrice avec la méthode PAM afin d'identifier des patterns. Finalement, en utilisant le coefficient de silhouette, le nombre de classes a été défini.

4.5.2 Résultats

L'application de la procédure que nous avons décrite a permis d'identifier deux classes pour chacune des deux tâches analysées. Pour la tâche 3, le coefficient de silhouette moyen s'élève à 0.25, et pour la tâche 5, le coefficient de silhouette moyen s'élève à 0.39. Il semblerait qu'aucune structure n'ait pu être identifiée dans les données à l'aide de cette méthode.

4.5.3 Conclusions

Là encore, deux classes ont été identifiées mais la structure de celles-ci ne semble pas stable. Il est possible que cette absence de stabilité soit due à une trop forte finesse dans la comparaison des scanpaths. En effet, la comparaison des scanpaths se base sur la distance euclidienne entre chaque point de fixation, sans lissage des données. Les scanpaths sont donc identifiés comme tous différents.

4.6 Classification basée sur la distance de Hausdorff ¹

La comparaison des positions des fixations n'a pas donné de résultats. Comme nous l'avons vu dans la section précédente, cela est peut-être dû à un trop faible «lissage» des données. Dans cette section, nous introduisons la distance de Hausdorff, une méthode de comparaison de formes, appliquée aux données oculométriques pour répondre à cette problématique.

4.6.1 Analyse des données

4.6.1.1 La distance de Hausdorff

La distance de Hausdorff ([Hausdorff, 1962]) est utilisée pour calculer la similarité entre des ensembles de points. Si nous considérons la projection des coordonnées de points de fixation sur un plan comme une figure géométrique, alors la distance de Hausdorff peut être utilisée pour calculer la distance entre les données de fixations de deux scanpaths. La distance d'Hausdorff se calcule de cette façon :

Etant donné deux ensembles non-vides de points $A = (a_1, \dots, a_n)$ et $B = (b_1, \dots, b_n)$ dans un espace métrique E , avec une distance d , la distance de Hausdorff H est donnée par :

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (4.2)$$

Où :

$$h(A, B) = \max(\min_{a \in A} d(a, b))$$

Ainsi, la distance de Hausdorff entre deux ensembles de points est la distance maximale entre un point d'un sous-ensemble et le point le plus proche d'un autre sous-ensemble (Figure 4.12).

1. Cette étude a fait l'objet d'une publication ([Drusch et al., 2011])

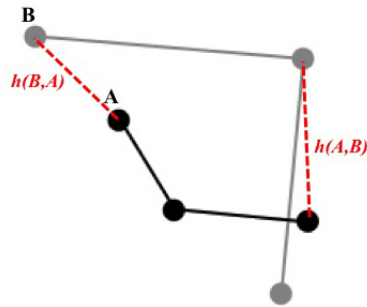


FIGURE 4.12 – Illustration de la distance de Hausdorff entre deux scanpaths fictifs A et B (ici, $H(A,B) = h(A,B)$).

Dans notre cas, nous devons normaliser la distance de Hausdorff par la diagonale de la page Web. Ainsi, la distance est une métrique de dissimilarité qui varie de 0 à 1. Avec cette procédure appliquée à l’ensemble des données, nous obtenons une matrice des distances entre toutes les paires de scanpaths (Algorithme 4.2). Cette matrice est ensuite utilisée pour calculer les classes avec l’approche par médoïdes que nous avons détaillé *supra*.

Algorithme 4.2 Obtention d’une matrice de distance de Hausdorff entre toutes les paires de scanpaths.

Étant donné SP , une liste de scanpaths contenant des fixations.

Définir $diagStimulus$ comme la diagonale du stimulus présenté aux participants.

Définir $matrice[][]$ comme la matrice contenant les comparaisons de scanpaths.

pour chaque couple de scanpaths SP_i et SP_j **faire**

$matrice[i][j] =$ distance de Hausdorff entre les fixations de SP_i et SP_j selon l’équation 4.2 normalisée par $diagStimulus$.

fin

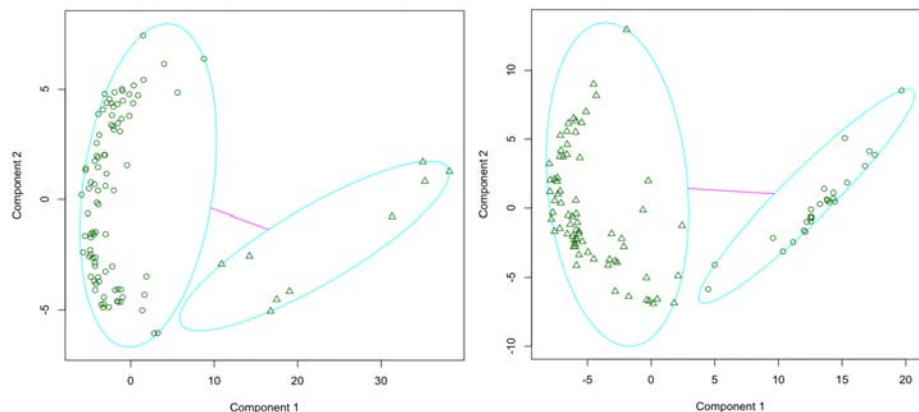
4.6.1.2 Partitionnement autour des médoïdes (PAM)

Nous avons dans un premier temps établis une matrice des distances de Hausdorff entre toutes les paires de scanpaths, puis nous avons analysé cette matrice avec la méthode PAM afin d’identifier des patterns. Finalement, en utilisant le coefficient de silhouette, le nombre de classes a été défini.

4.6.2 Résultats

L’application de la procédure que nous avons décrite a permis d’identifier deux classes pour chacune des deux tâches analysées. Dans la tâche 3, une des classes est composée de 91 participants et la seconde de 9 participants. Le coefficient de silhouette global est de 0.71, ce qui indique qu’une structure stable a été trouvée ([Kaufman & Rousseeuw, 1990]). Dans la tâche 5, les deux classes sont composées respectivement de 27 et 72 participants. Le coefficient de silhouette

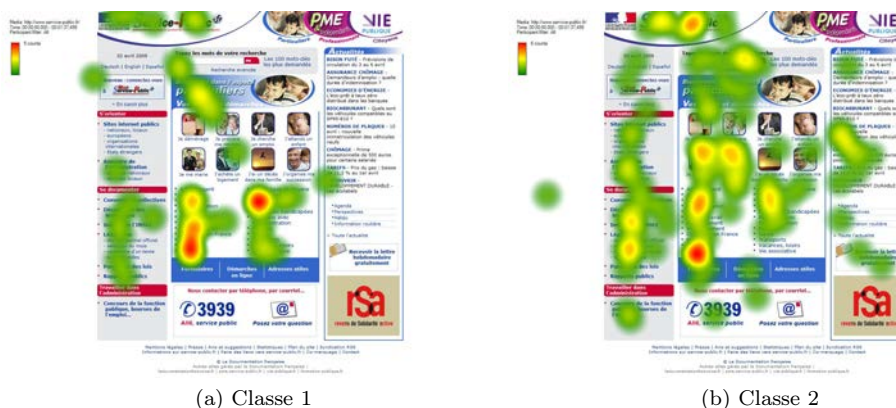
moyen est de 0.64, ce qui indique qu'une structure raisonnable a été détectée. Nous pouvons voir la séparation nette des classes dans la Figure 4.13.



(a) Tâche 3 (86,07% de la variabilité des points).(b) Tâche 5 (88,41% de la variabilité des points).

FIGURE 4.13 – Analyse multidimensionnelle des deux dimensions pour les tâches 3 (a) et 5 (b), avec les résultats de la classification superposés.

L'identification des médoïdes de chaque classe (les scanpaths représentatifs pour chaque classe) nous a permis de générer les *heatmaps* correspondantes (Figures 4.14 et 4.15). Comme nous pouvons le voir dans ces figures, les *heatmaps* de la tâche 3 diffèrent entre les deux classes. Ainsi, les participants de la classe 2 semblent faire une exploration plus exhaustive de la page Web. Ils visualisent plus de zones que les participants de la classe 1. De plus, les zones ayant le plus fort taux de fixations (en rouge) diffèrent d'une classe à une autre.



(a) Classe 1

(b) Classe 2

FIGURE 4.14 – Heatmaps des scanpaths représentatifs des deux classes de la tâche 3.

Pour la tâche 5, nous observons la même tendance : les deux classes de

scanpaths ont des distributions différentes des fixations. Nous constatons aussi que la classe 1 explore le bas de la page alors que les participants de la classe 2 l'ont complètement ignoré.

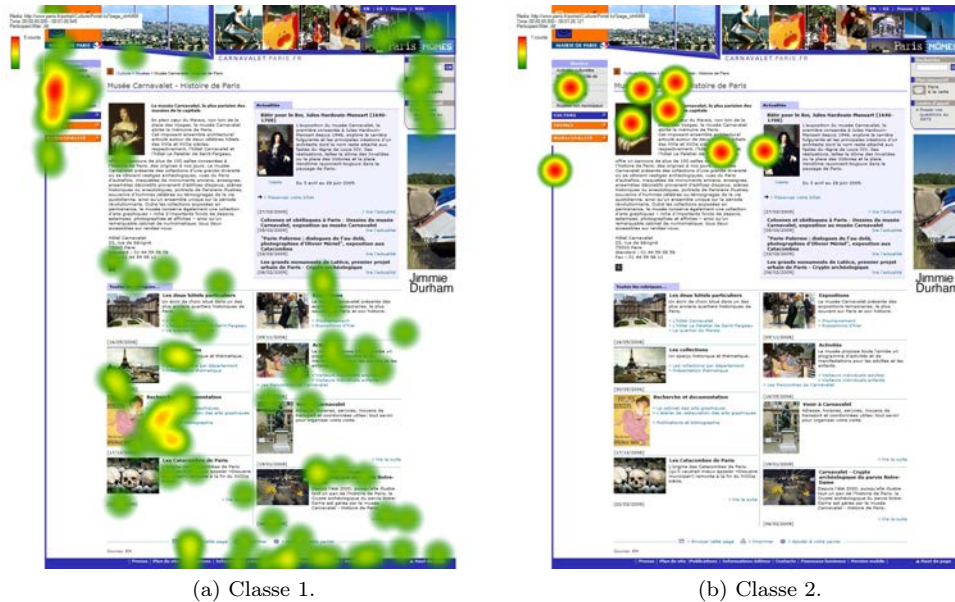


FIGURE 4.15 – Heatmaps des scanpaths représentatifs des deux classes de la tâche 5.

4.6.2.1 Caractérisation des classes

Dans la tâche 3, il apparaît que les participants de la classe 2 diffèrent significativement de ceux de la classe 1 sur deux indicateurs supplémentaires uniquement : la dispersion des scanpaths et l'aire de l'enveloppe convexe (Table 4.2).

<i>*p < .001</i>	Classe 1 (n=91)		Classe 2 (n=9)	
	Moyenne	Écart-type	Moyenne	Écart-type
Durée totale	33291.1	22401.1	54034.4	46537.5
Durée des fixations	421.8	76.4	414.7	58.0
Nombre de fixations	80.6	52.4	129.4	104.8
Longueur des scanpaths	9808.7	5868.4	16953.5	12296.6
Barycentre X	518.4	51.3	504.9	33.0
Barycentre Y	423.9	45.7	433.9	53.1
Dispersion*	237.1	34.6	268.5	20.5
Aire de l'enveloppe convexe*	32785.0	87496.5	54843	109495

TABLE 4.2 – Comparaisons statistiques entre les classes de la tâche 3 sur différents indicateurs.

Dans la tâche 5, les scanpaths de la classe 1 diffèrent de ceux de la classe 2 sur chaque métrique. Les scanpaths de la classe 1 sont plus longs en durée et en pixels. Leurs durées de fixations sont plus courtes. Ils ont aussi un plus grand nombre de fixations. Les mouvements oculaires des participants de la classe 1 sont plus orientés en bas à droite de la page que ceux de la classe 2. Les scanpaths de la classe 1 sont plus dispersés et couvrent une large portion de la page Web (Table 4.3).

<i>*p < .001</i>	Classe 1 (n=27)		Classe 2 (n=72)	
	Moyenne	Ecart-type	Moyenne	Ecart-type
Durée totale*	26589.6	12966.6	9191.0	5069.6
Durée des fixations*	347.6	37.5	468.4	128.6
Nombre de fixations*	76.7	37.1	21.2	14.7
Longueur des scanpaths*	13163.1	5954.1	3092.3	2852.8
Barycentre X*	469.9	78.4	285.7	113.5
Barycentre Y*	484.8	104.0	246.6	57.0
Dispersion*	454.6	53.2	214.3	101.5
Aire de l'enveloppe convexe*	854472.5	24780	163214	19948.3

TABLE 4.3 – Comparaisons statistiques entre les classes de la tâche 5 sur différents indicateurs.

4.6.3 Conclusions

Nous avons proposé d'utiliser la distance de Hausdorff pour comparer les paires de scanpaths. Cette comparaison nous a permis de générer une matrice de dissimilarité qui a été utilisée en entrée d'une méthode de classification basée sur le partitionnement autour des médoïdes (PAM). Cette approche a produit des résultats intéressants. Elle démontre que les utilisateurs n'ont pas nécessairement un pattern en forme de «*F*» lorsqu'ils cherchent une information sur une page Web. De plus, cette approche nous a permis d'identifier des groupes d'utilisateurs ayant des patterns de visualisation «similaires». Les analyses statistiques conduites sur différents indicateurs oculométriques indiquent de quelle façon ces groupes diffèrent. Il a aussi été possible de caractériser plus finement ces classes de scanpaths.

D'un point de vue méthodologique, la distance de Hausdorff semble être un outil d'analyse utile pour comparer les scanpaths. De plus, cette méthode ne requiert pas la définition d'aires d'intérêts *a priori*, et donc de formuler des hypothèses sur la distribution des fixations oculaires sur différents éléments d'une page Web. Néanmoins, cette méthode pourrait potentiellement être utilisée dans une approche «*top-down*» où les expérimentateurs spécifieraient par avance des groupes d'utilisateurs puis testeraient statistiquement les résultats du test avec une méthode de permutation proposée par Feusner et Lukoff ([Feusner & Lukoff, 2008]).

La sensibilité de la distance de Hausdorff aux données extrêmes peut être un avantage puisque cela permet de ne pas trop agréger les données. Lors de la conduite d'un test utilisateurs, il est important de détecter les cas extrêmes plutôt que de lisser complètement les données.

Néanmoins, cette approche a certaines limites. Par exemple, il n'est pas possible de prendre en compte l'ordre des fixations dans un scanpath. Cette méthode se base plutôt sur leur forme, alors que les scanpath ne se réduisent pas à un simple ensemble. En prenant en compte cet aspect de l'exploration visuelle des pages Web, plus de classes auraient probablement été détectées.

4.7 Analyse par T-patterns basée sur la procédure Mean-Shift²

Dans cette étude, nous avons combiné l'approche de Santella et DeCarlo (décrite dans la section 5.2.1 page 83), qui permet de définir des zones d'intérêt à partir des données de fixations, avec la méthode de Magnusson ([Magnusson, 2000]), basée sur l'analyse *T-patterns* (*time-patterns*) afin de détecter des patterns comportementaux visuels.

4.7.1 Détection automatique des aires d'intérêts

L'unique paramètre à définir pour identifier des aires d'intérêts avec la procédure mean-shift est la taille de la bande passante (aussi nommée fenêtre) h . Nous avons choisi de suivre des recommandations de Duchowski et ses collègues ([Duchowski et al., 2010]) en définissant ce paramètre afin qu'il corresponde à l'étendue de la vision fovéale (environ 5° d'angle visuel). A l'instar de leur étude, notre implémentation a été $h_s = 50$ pixels (à une résolution de 1280x1024, l'étendue spatiale est de 1,5° d'angle visuel).

4.7.2 L'analyse par T-patterns

4.7.2.1 Principe général

L'objectif d'une analyse par T-pattern est de découvrir la structure cachée d'un comportement. Elle permet de mettre en évidence la structure temporelle des comportements. Celle-ci a tendance à être récurrente. Dans la figure 4.16, nous pouvons voir un exemple fictif de T-pattern. Ici, chaque lettre correspond à un type d'événement. A l'intérieur de ces événements il y a un T-pattern A,B,C et D, masqué par d'autres événements (W et K).

². Cette étude a fait l'objet d'une publication ([Drusch & Bastien, 2012a]).

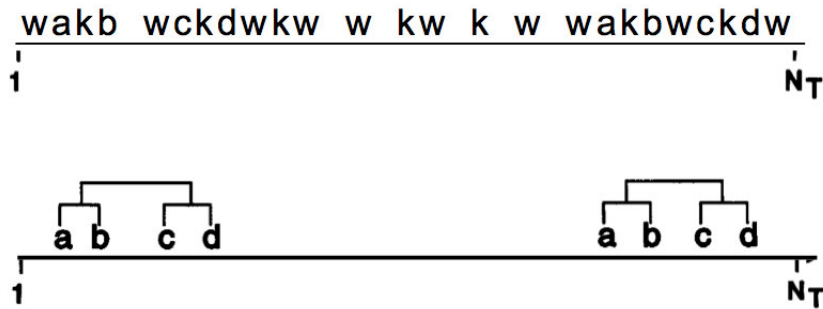


FIGURE 4.16 – Exemple de T-pattern ([Magnusson, 2000]).

4.7.2.2 Principe de détection

Un T-pattern est une combinaison d'événements qui, dans une série, sont répétés plusieurs fois dans le même ordre. La distance temporelle entre les composants d'un T-pattern est relativement invariante. Pour cela, une hypothèse nulle est formulée selon laquelle chaque composant du T-pattern est indépendant et distribué aléatoirement dans le temps. Selon Magnusson : «if A is an earlier and B a later component of the same recurring T- pattern, then, after an occurrence of A at t , there is an interval $[t + d1, t + d2]$ ($d2 - d1 > 0$) that tends to contain at least one occurrence of B more often than would be expected by chance.». La relation entre A et B est appelée «intervalle critique». Dans l'algorithme, les T-patterns les plus simples sont détectés en premier (dans notre exemple de la figure 4.16, (AB) et (CD)) et sont ensuite combinés dans des T-patterns plus larges ((AB)(CD)). Ces T-patterns plus larges peuvent éventuellement aussi être combinés. Finalement, les T-patterns partiels ((AB) et (CD)) sont éliminés du résultat final s'ils font partie de T-patterns plus larges ((AB)(CD)). Ainsi, le nombre et la composition des T-patterns dépend entièrement des données.

4.7.2.3 Paramètres de détection dans le logiciel ThemeTM

ThemeTM, un logiciel développé par Magnusson ([Magnusson, 1996, Magnusson, 2000]), a été utilisé pour réaliser les analyses. Pour chercher des T-patterns, certains paramètres doivent d'abord être ajustés. L'objectif est de trouver la meilleure configuration possible des paramètres afin d'obtenir un ensemble significatif de T-patterns. Les principaux paramètres disponibles dans le logiciel sont (manuel de Theme 5) :

- *Minimum occurrences* : le nombre minimal de fois qu'un T-pattern doit apparaître pour être détecté. En général, plus cette valeur est faible, plus les T-patterns détectés seront larges.
- *Significance level* : la probabilité maximale acceptable pour qu'un intervalle critique soit détecté par hasard uniquement. En général, plus cette valeur est faible, moins un T-pattern tend à être détecté par hasard.

- *Fast-Free limit* : le niveau hiérarchique en dessous duquel les T-patterns libres sont acceptés. Il est possible de définir l'intervalle critique comme «*fast*», où un type d'événement d'un T-pattern A est suivi par B relativement rapidement.
- *Minimum sample* : le pourcentage minimal d'échantillons dans lequel un T-pattern doit apparaître pour être détecté. Ce paramètre peut être utile quand les données contiennent plus d'un échantillon.

Dans cette étude, nous avons commencé l'analyse en définissant le «*Minimum occurrence*» à 4, et le «*Significance level*» à 0.005. La «*Fast-Free limit*» a été réglée sur l'algorithme «*Free*» et le «*Minimum sample*» à 0.

4.7.3 Analyse des données et résultats

Dans cette étude, nous avons analysé uniquement les données de la page d'accueil de la tâche 3.

4.7.3.1 Préparation des données

Nous avons identifié l'ensemble des zones d'intérêt via la procédure mean-shift (spatiale) à partir de l'ensemble des coordonnées de fixations de tous les participants sur la page Web. Ensuite, nous avons transformé chaque scanpath en une série d'aires d'intérêts, munies chacune d'un identifiant unique (ex. une lettre). Ainsi, chaque scanpath peut être considéré comme un ensemble d'enregistrements comportementaux contenant les occurrences des événements, qui sont les aires d'intérêt visualisées. Nous avons aussi simplifié les scanpaths en supprimant les répétitions. Finalement, nous avons exporté chaque scanpath dans un format compatible avec le logiciel Theme.

4.7.3.2 La procédure Mean-Shift

Trente-trois aires d'intérêts ont été identifiées à l'aide de la procédure Mean-Shift. Seulement 11 aires étaient sur la page Web. Les autres aires étaient généralement fixées une seule fois, par un seul participant. Cela est probablement dû à du bruit dans les données. Ainsi, certaines fixations ont été enregistrées en dehors de la page Web (par exemple, sur le navigateur).

4.7.3.3 L'analyse par T-patterns

Un total de 8497 occurrences d'événements a été analysé. Chaque aire d'intérêt est ici considérée comme un type d'événement d'un comportement et chaque occurrence d'une aire d'intérêt est une occurrence d'un événement. En paramétrant Theme avec un nombre minimal d'occurrences («*Minimum occurrences*») de 4, 13100 patterns ont été détectés dans les données. Ces T-patterns ont un nombre moyen d'occurrences de 6,22 (= 7,96) et une longueur moyenne de 7,2 occurrences

(= 2,1). Le T-pattern le plus long inclus est une séquence de 21 occurrences et le plus court n'en contient que 2. Le T-pattern le plus long n'apparaît que 4 fois dans tous les scanpaths (100 scanpaths au total).

Etant donné qu'il s'agit d'une analyse «*bottom-up*», certains paramètres doivent être définis empiriquement ([Magnusson, 2000]). Afin de trouver le bon compromis entre le nombre de T-patterns et leur longueur, en d'autres mots, pour obtenir un nombre réduit de T-patterns relativement longs, les paramètres doivent être ajustés en fonction des résultats.

Patterns	((1 3) 9) 4)	((1 2) 9) 4)	((3 2) (9 5))
Occurrences	45	41	40
Durée (ms)	363676	279040	224215
% durée	11	8	7
Nombre de participants	45	40	38
Patterns	((3 2) (6 1))	((6 1) (7 6))	((1 3) (2 1))
Occurrences	45	51	41
Durée (ms)	229068	215872	168601
% durée	7	6	5
Nombre de participants	40	41	37

TABLE 4.4 – Caractéristiques des T-patterns les plus longs (les numéros sont les identifiants des aires d'intérêt).

Sur la base des résultats précédents, nous n'avons modifié qu'un seul paramètre : le nombre minimal d'occurrences («*Minimum occurrences*») pour qu'un T-pattern soit détecté. Après quelques itérations, nous avons trouvé un nombre minimal d'occurrences satisfaisant ($N_{min} = 40$). Ce nombre minimal d'occurrences nous a permis d'obtenir un nombre réduit de patterns, mais avec une longueur suffisante.

Ainsi, nous avons obtenu 78 T-patterns. Le nombre moyen d'occurrence par T-pattern est de 83,32 (= 51,07). La longueur moyenne des T-patterns est de 2,74 (= 0,59). Le T-pattern le plus long contient 4 occurrences. Leurs caractéristiques sont présentés dans la table 4.4, et 4 d'entre eux sont représentés dans la figure 4.17.



FIGURE 4.17 – Illustration de 4 des six plus longs T-patterns identifiés (les numéros sont les identifiants des aires d'intérêt).

4.7.4 Conclusions

L'objectif de cette étude était d'appliquer une approche «bottom-up» à l'analyse et la comparaison des scanpaths sur le Web. Cette approche se base sur deux méthodes spécifiques : la procédure mean-shift pour identifier automatiquement les aires d'intérêt visualisées par les participants, et l'analyse T-pattern pour détecter des patterns. La combinaison de ces deux techniques s'est révélée pertinente puisque des patterns ont effectivement été relevés, sans avoir eu à définir *a priori* d'aires d'intérêt. De plus, la technique T-pattern permet l'identification de patterns visuels cachés. Comme l'indiquent Mast et Burmester ([Mast & Burmester, 2011]), les T-patterns révèlent des patterns inattendus sur de larges échantillons de données.

Néanmoins, l'analyse des données fut difficile en raison des ajustements nécessaires des paramètres de détection. En fait, il n'existe pas de mécanisme de validation pour s'assurer qu'un paramétrage permet de mettre en évidence des

patterns «signifiants». Une légère modification des paramètres peut produire un nombre très différent de patterns et il n’y a pas de solution pour décider quel paramètre retenir.

Un autre problème que nous avons relevé est lié à l’identification des participants concernés par les patterns. Comme indiqué dans la table 4.4, la somme du nombre de participants concernés par les différents patterns excède le nombre total de participants dans cette étude. Cela signifie qu’un participant peut avoir différents patterns. Ainsi, il devient difficile d’identifier des classes de participants en se basant sur leurs comportements oculaires.

Cette approche conviendrait finalement mieux à un objectif d’identification de sous-séquences comportementales indépendamment de l’identification de groupes de participants (comme ont pu le faire, mais avec une méthode différente, [Hembroke et al., 2006]).

4.8 Conclusions et limites

Nous avons observé que chaque méthode donne des résultats différents, que ce soit en termes de nombre de patterns identifiés, de stabilité des patterns, de données prises en comptes ou encore de types (et donc de finalité) de patterns (Table 4.5). Cela pose un problème majeur car nous ne disposons pas de critères pour déterminer laquelle donne des résultats fiables.

Les indicateurs oculométriques (I-O), la position des fixations (P-F) et la distance de Hausdorff (D-H) ne prennent pas vraiment en compte l’ordre des fixations. La position des fixations est prise en compte de façon différente selon la méthode, et la durée des fixations n’est envisagée que par les indicateurs oculométriques. La finalité n’est pas la même pour l’analyse par T-patterns car les participants peuvent se trouver dans plusieurs patterns identifiés. En ce sens, il s’agit plutôt de patterns «représentatifs».

Il nous apparaît que les AOIs ont une forte utilité pour les analyses car elles permettent de connaître les zones des pages Web visitées selon les différents patterns. En ce sens, une méthode reposant sur la procédure Mean-Shift nous paraît intéressante. Néanmoins, le Mean-Shift appliqué à l’identification d’AOIs comporte des limites importantes qui n’ont pour l’heure pas été mises en exergue dans la littérature oculométrique (ce point sera abordé en détails dans le Chapitre 5). En effet, nous ne sommes pas en mesure de savoir si les AOIs identifiées représentent de manière fiables les données oculométriques, alors que nous n’avons aucune connaissance a priori de leur distribution.

Dans le prochain chapitre, nous tenterons de répondre à l’ensemble de ces limites en (1) testant puis améliorant la procédure Mean-Shift pour la détection d’AOIs et en (2) proposant et validant une méthode de comparaison des scanpaths qui prenne en compte à la fois leurs aspects temporels et topographiques afin

d'identifier des typologies de patterns.

Méthode	Nombre de patterns	Stabilité moyenne	Ordre des fixations	Position des fixations	Durée des fixations	Types de patterns
I-O	2	0,33	Non	Non	Oui	Typologies
P-F	2	0,32	Non	Oui	Non	Typologies
D-H	2	0,67	Non	Oui	Non	Typologies
MS-TP	78	NA	Oui	Oui	Non	Représentatifs

Légende : *I-O* (Indicateurs oculométriques), *P-F* (Position des Fixations), *D-H* (Distance de Hausdorff), *MS-TP* (Mean-Shift puis T-patterns).

TABLE 4.5 – Comparaison des méthodes d'analyse des patterns oculométriques.

Chapitre 5

Proposition et évaluation d'une nouvelle méthode d'analyse¹

Comme nous avons pu le voir dans le chapitre précédent, chaque méthode génère des typologies de patterns différents sur le même jeu de données (en termes de nombre de patterns, de types de patterns ou encore de stabilité des classes obtenues). Ces résultats ne nous ont pas permis de choisir de manière fiable une méthode ou une autre. Il s'avère en fait que ces méthodes n'ont pas été comparées de manière formelle dans la littérature et encore moins validées dans l'objectif d'identifier des typologies de patterns comportementaux.

Dans ce chapitre, nous décrivons la méthode d'analyse que nous avons développé pour l'identification de patterns comportementaux visuels sur le Web. Nous avons choisi de conserver la notion d'*AOI* car elle a à la fois une utilité pratique et théorique (les scanpaths ne sont pas indépendants des stimuli observés). Ainsi, dans une première partie, nous faisons état des limites actuelles concernant la définition de ces *AOIs* pour ensuite proposer une nouvelle approche. Enfin, nous proposons puis validons une nouvelle méthode pour identifier des typologies de patterns comportementaux.

5.1 Problématiques liées à la définition des *AOIs*

Différents aspects méthodologiques concernant la définition de ces zones d'intérêts méritent d'être soulevés. En effet, bien que différentes approches existent, le problème de la définition de ces zones demeure.

1. Pour réaliser les différents traitements de données oculométriques de ce chapitre, nous avons développé une application Java dont les captures d'écran sont visibles en Annexe E.

Qui définit les zones d'intérêt ?

A l'heure actuelle, aucune étude n'a porté sur la question de savoir si un expérimentateur définit les zones d'intérêt de la même façon qu'un participant ou même qu'un autre expérimentateur. En effet, la perception de la structuration du contenu d'une page peut différer d'une personne à une autre. De la même manière, il est possible que le sens attribué à une zone particulière par le concepteur ou l'expérimentateur ne soit pas le même que celui d'un participant. Ces questions sont pour le moment sans réponse.

Forme et niveau de granularité

A notre connaissance, peu d'études justifient précisément le choix de leurs zones d'intérêt en termes de forme et de niveau de granularité. Par exemple, si un expérimentateur décide que le logo d'une page Web est une zone d'intérêt afin de recueillir la proportion des fixations en son sein, quelle doit être la forme de cette zone? Si les limites de cette zone sont changées, les patterns de visualisation peuvent changer, tout comme les conclusions qui peuvent être émises sur l'exploration visuelle des différents éléments de la page. De la même manière, que faire des fixations qui sont à la limite de cette zone?

Les figures 5.1 à 5.3 illustrent, à partir d'un exemple réel, les effets que peuvent avoir le choix de la forme et de la taille d'une zone d'intérêt pour le logo d'une page Web. Chacune de ces figures représentent une capture d'écran d'une partie d'une page Web, centrée sur le logo, ainsi qu'une représentation des fixations oculaires. Ces données sont issues d'une expérience portant sur 47 participants (cf. Section ?? page ??).

Dans la figure 5.1, nous avons défini une zone d'intérêt comme étant le plus petit rectangle possible couvrant le logo. Nous avons enregistré en moyenne 2,55 fixations concernant 26 participants pour une durée de fixation moyenne de 563ms.

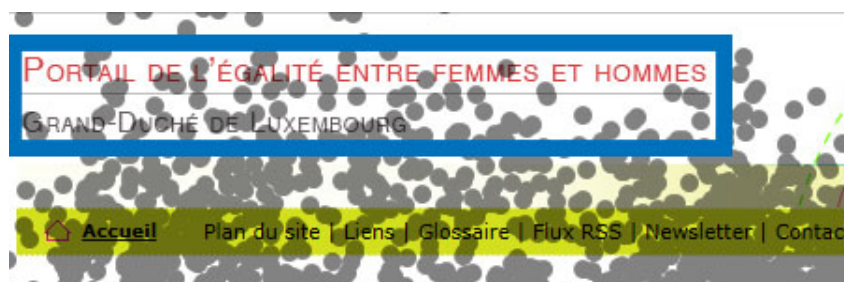


FIGURE 5.1 – Capture partielle d'une page Web avec une représentation des fixations et une zone d'intérêt (logo) définie comme le plus petit rectangle possible.

Dans la figure 5.2, nous avons défini la même zone d'intérêt, mais avec un rectangle plus large. Nous avons enregistré en moyenne 7,21 fixations concernant

34 participants pour une durée de fixation moyenne de 1618 ms.



FIGURE 5.2 – Capture partielle d'une page Web avec une représentation des fixations et une zone d'intérêt (logo) définie comme un rectangle plus large.

Dans la figure 5.3, nous avons défini la même zone d'intérêt, mais comme étant le plus petit polygone possible couvrant le logo. Nous avons enregistré en moyenne 1,4 fixations concernant 23 participants pour une durée de fixation moyenne de 260 ms.



FIGURE 5.3 – Capture partielle d'une page Web avec une représentation des fixations et une zone d'intérêt (logo) définie comme le plus petit polygone possible.

Le tableau suivant résume les résultats de ce comparatif. Il apparaît clairement que la mesure des indicateurs oculométriques varie en fonction du choix de la forme ou de la taille d'une zone d'intérêt.

Des tests de Student pour échantillons appariés ont montré que les différences concernant le nombre moyen de fixations dans la zone entre l'aire d'intérêt «petit rectangle» et «gros rectangle» ($t(46) = 4.801$, $p < 0.01$, 95% CI [2.71, 6.61], $d = 4.66$), entre l'aire d'intérêt «petit rectangle» et «polygone» ($t(46) = 2.595$, $p = 0.013$, 95% CI [0.26, 2.04], $d = 1.5$), et entre l'aire d'intérêt «gros rectangle» et «polygone» ($t(46) = 4.55$, $p < 0.01$, 95% CI [3.24, 8.38], $d = 5.81$) étaient statistiquement significatives.

Des tests de Student pour échantillons appariés ont montré que les différences concernant la durée moyenne des fixations dans la zone entre l'aire d'intérêt «petit rectangle» et «gros rectangle» ($t(46) = 4.754$, $p < 0.01$, 95% CI [0.61,

1.5], $d=1.05$), entre l'aire d'intérêt «petit rectangle» et «polygone» ($t(46) = 4.545$, $p=0.031$, 95% CI [0.03, 0.58], $d=0.3$), et entre l'aire d'intérêt «gros rectangle» et «polygone» ($t(46) = 4.239$, $p<0.01$, 95% CI [0.71, 2], $d=1.36$) étaient statistiquement significatives.

Type de zone	% participants ayant fixé la zone	Nombre moyen de fixations dans la zone	Durée moyenne des fixations (millisecondes)
Petit rectangle	55	2,55	563
Gros rectangle	72	7,21	1618
Polygone	49	1,4	260

TABLE 5.1 – Effet de différentes formes et tailles de zones d'intérêts sur la mesure de quelques indicateurs oculométriques.

Cette étude visait à comparer les différences de comportements oculaires sur le Web entre les hommes et les femmes. Nous avons donc aussi étudié l'effet du type d'aire d'intérêt sur cette comparaison (Table 5.2). Il s'avère que pour la durée des fixations, le gros rectangle ($t(45) = 2.933$, $p<.01$, 95% CI [0.67, 3.68], $d=2.32$) et, dans une moindre mesure, le petit rectangle ($t(45) = 2.096$, $p=.05$, 95% CI [-0.003, 1.88], $d=0.94$) conduiraient à conclure qu'il y a des différences statistiquement significatives entre les hommes et les femmes alors que le polygone ne mènerait pas à cette conclusion ($t(45) = 1.577$, $p=.131$, 95% CI [-0.11, 0.8], $d=0.34$).

Type de zone	Différence moyenne du nombre de fixations dans la zone	Différence moyenne de la durée des fixations dans la zone (secondes)
Petit rectangle	0.94*	3.93*
Gros rectangle	2.32**	9.62**
Polygone	0.34	1.97*

* $p .05$; ** $p .01$

TABLE 5.2 – Comparaison des moyennes du nombre de fixations et de la durée des fixations entre hommes et femmes selon différents types d'aires d'intérêts.

Nous pouvons conclure que le choix du type d'aire d'intérêt n'est pas anodin puisqu'il peut potentiellement faire varier les conclusions des études.

5.2 La détection automatique des *AOIs*

La procédure Mean-Shift est décrite comme étant la plus robuste pour détecter automatiquement des *AOIs*, à partir des données ([Santella & DeCarlo, 2004]). Cette approche est particulièrement intéressante pour l'objectif de la présente thèse, mais engendre divers questionnements nous empêchant de l'appliquer en

l'état. En effet, comme nous le verrons, cette procédure n'est en fait pas tout à fait «automatique», et n'a pas été réellement évaluée pour la détection *d'AOIs*. Ainsi, plusieurs questions actuellement sans réponses se posent lorsqu'on la met en oeuvre :

- Comment définir le paramètre de bande passante ?
- Le paramètre de bande passante utilisé dans la littérature est-il «universel» ?
- Les *AOIs* détectées reflètent-elles correctement les données ?
- N'y a-t-il pas du «bruit» dans les données, engendrant donc des *AOIs* qui n'existent pas vraiment ?

Dans cette section, nous présentons une expérimentation visant à tester les limites de la procédure *Mean-Shift* telle qu'elle est actuellement proposée dans la littérature. Les limites de cette approche seront exposées, puis nous en proposerons une version améliorée.

5.2.1 Description formelle de la procédure *Mean-Shift*

Etant donné un ensemble de N points :

$$\{x_j \mid j \in 1 \dots N\},$$

chaque point ayant un des labels K :

$$\{c_j \in 1 \dots K \mid j \in 1 \dots N\}.$$

Le processus consiste à bouger («*shift*») de façon itérative un point x_i vers une nouvelle localisation $s(x_i)$ correspondant à la moyenne («*mean*») pondérée des points adjacents en se basant sur la fonction *kernel* k :

$$s(x) = \frac{\sum_j k(x - x_j)x_j}{\sum_j k(x - x_j)} \quad (5.1)$$

Le *kernel* k peut être uniforme (chaque point est pondéré de la même façon) ou encore gaussien (la pondération d'un point dépend de sa distance spatiale et/ou temporelle au centre x). Le *kernel* k contient un paramètre h qui correspond à la taille de la fenêtre définissant les points qui vont servir à calculer la moyenne pondérée s . Ce paramètre h est important car il va indirectement être responsable du nombre et de la taille des classes obtenues.

Santella et DeCarlo proposent d'utiliser un *kernel* gaussien. Les fixations se caractérisent par une position et une durée, ainsi le point $x_i = (x_i, y_i, t_i)$.

Il est possible de prendre en compte à la fois les données spatiales et temporelles :

$$k_{spatiotemporel}([x_i, y_i, t_i]) = \exp\left(-\frac{x_i^2 + y_i^2}{h_s^2} - \frac{t_i^2}{h_t^2}\right),$$

ou uniquement les données spatiales :

$$k_{spatial}([x_i, y_i]) = \exp\left(-\frac{x_i^2 + y_i^2}{h_s^2}\right).$$

La procédure mean-shift est la suivante :

Algorithme 5.1 procédure mean-shift

Initialisation

$$\begin{aligned} & \text{pour } j \in 1 \dots N \\ & \quad y_j^0 = x_j \end{aligned}$$

Itération n

$$\begin{aligned} & \text{répéter jusqu'à convergence} \\ & \quad \text{pour } j \in 1 \dots N \\ & \quad \quad y_j^n = s(y_j^{n-1}) \end{aligned}$$

Au départ, chaque point est initialisé comme un centre de classe. Ensuite, chacun de ces centres est bougé selon l'équation 5.1 de manière itérative vers une nouvelle localisation de plus forte densité jusqu'à ce que cette localisation reste la même entre deux itérations (convergence, voir Figure 5.4). Enfin, il est nécessaire de fusionner les centres proches sur la base d'un certain seuil (correspondant généralement à h) car l'initialisation se fait sur tous les points.

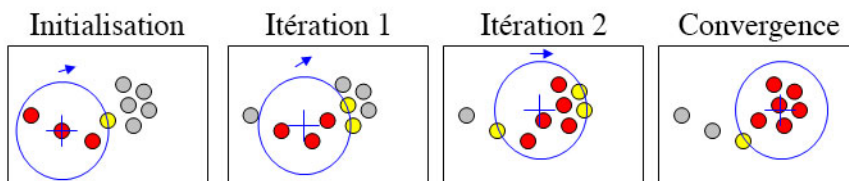


FIGURE 5.4 – Illustration de la création d'une classe (en deux itérations) par la procédure mean-shift sur des données spatiales (en bleu, la fenêtre h) (adapté de [Philippeau, 2008]).

5.2.2 Méthodologie expérimentale

L'objectif de cette expérimentation était de tester les limites éventuelles de la détections d'AOIs avec la procédure Mean-Shift, puis d'avoir une base de comparaison pour proposer une amélioration.

Pour cela, nous avons défini un ensemble d'images et d'animations se composant de points numérotés que les participants devaient suivre visuellement dans l'ordre alors qu'ils étaient enregistrés par un oculomètre (Tobii T120). La taille et l'ordre des points, la distance inter-points ainsi que la vitesse des animations variaient d'une condition à une autre.

5.2.2.1 Matériel

Une série d'images (de dimension 1280x800px) a été créée pour les besoins de l'expérimentation. Elles étaient toutes composées de 8 points noirs comportant un numéro blanc à l'intérieur. La taille de ce numéro est strictement proportionnelle au point qui la contient (= taille du point/2) et la police de caractères est le Times Roman. Les points sont toujours espacés de la même distance relative, quelle que soit leur taille. L'espacement de base est de 100px en X et $(1.5 * \text{espacementX} + \text{taille})$ px en Y.

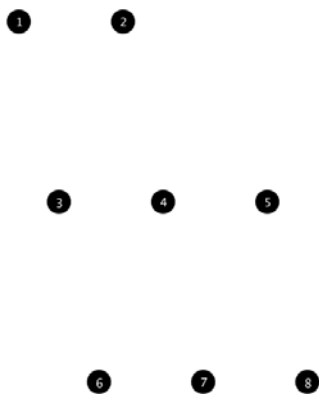


FIGURE 5.5 – Exemple de matériel

5.2.2.2 Variables

Taille des points (T, 2 modalités) :

- 30px (inférieur au champ de vision fovéale)
- 50px (équivalent au champ de vision fovéale)

Ordre (O, 2 modalités) :

- Origine (numéro 1 en haut à gauche)
- Inverse (numéro 1 en bas à droite)

Distance entre les points (D, 2 modalités) :

- Fixe (tous les points sont à une même distance les uns des autres, cf. supra)
- Variable (certains couples de points n'ont pas le même espacement que les autres en X, soit $0 * \text{taille}$, $0.5 * \text{taille}$, $1 * \text{taille}$).

Type d'animation (TA, 3 modalités) :

- Aucune = Image (tous les points sont affichés d'un coup sur une image). Dans ce cas, le stimulus disparaît sur décision du participant.
- Animation lente (les points apparaissent successivement avec un délai de 1 seconde, le point $n+1$ fait disparaître le point n). Le stimulus est affiché pendant 8 secondes.
- Animation rapide (les points apparaissent successivement avec un délai de 500 millisecondes, le point $n+1$ fait disparaître le point n). Le stimulus est affiché pendant 4 secondes.

5.2.2.3 Protocole

Toutes les modalités des variables sont combinées dans un plan à mesures répétées. Le plan expérimental est le suivant :

$$S_{30} < T_2 * O_2 * D_2 * TA_3 >$$

Ainsi, il y a 24 conditions expérimentales. Pour éviter un effet d'ordre, les conditions étaient présentées aléatoirement sans remise pour chaque participant.

Chaque participant a subi 2 procédures de calibration ainsi qu'une petite série d'essais. Chaque image-stimulus a été précédé d'une image comportant une croix noire sur fond blanc qui a servi à positionner le regard des participants (figure 5.6).

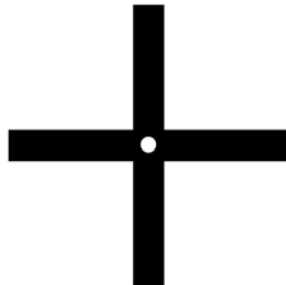


FIGURE 5.6 – image servant à positionner le regard du participant

Pré-consigne :

« Une image comportant une croix noire sur fond blanc va apparaître. Fixez le point blanc au centre de cette croix jusqu'à ce qu'elle disparaisse. Ensuite... »

Consigne pour les images :

« ... Une image contenant un ensemble de points numérotés de 1 à 8 va s'afficher. Regardez chaque point dans l'ordre puis appuyez sur la touche espace du clavier. »

Consigne pour les animations :

« ... Une animation dans laquelle vont apparaître successivement 8 points va s'afficher. Regardez chaque point. »

Entraînement (5 points) :

- Image fixe
- Animation lente
- Animation rapide

5.2.3 Résultats

5.2.3.1 Effet de l'ordre des éléments

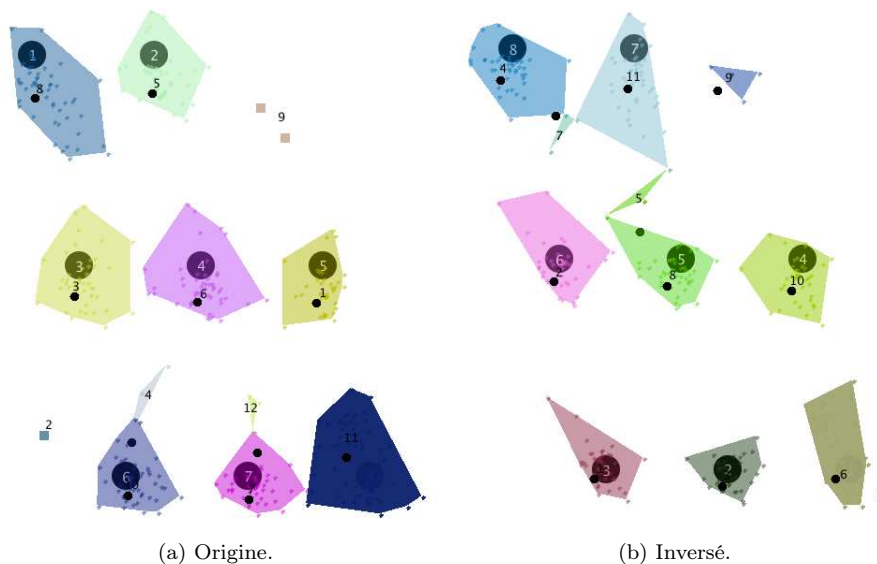


FIGURE 5.7 – Effet de l'ordre des éléments du stimulus.

L'ordre des éléments ne semble pas affecter les résultats de la procédure *Mean-Shift*. En effet, sur les deux résultats présentés, les *AOIs* identifiées sont relativement similaires. Cependant, on observe quelques *AOIs* de très petites tailles et éloignée des éléments du stimulus, qui sont probablement des *outliers*.

5.2.3.2 Effet de l'animation



FIGURE 5.8 – Effet de du type d'animation.

L'animation ne semble pas affecter les résultats de la procédure *Mean-Shift*. En effet, sur les trois résultats présentés, les *AOIs* identifiées sont relativement similaires. Cependant, on observe quelques *AOIs* de très petites tailles, qui sont probablement, là aussi, des *outliers*. On constate aussi que certaines *AOIs* sont trop étendues, englobant des points de fixations éloignés de leur centre.

5.2.3.3 Effet de la taille des éléments

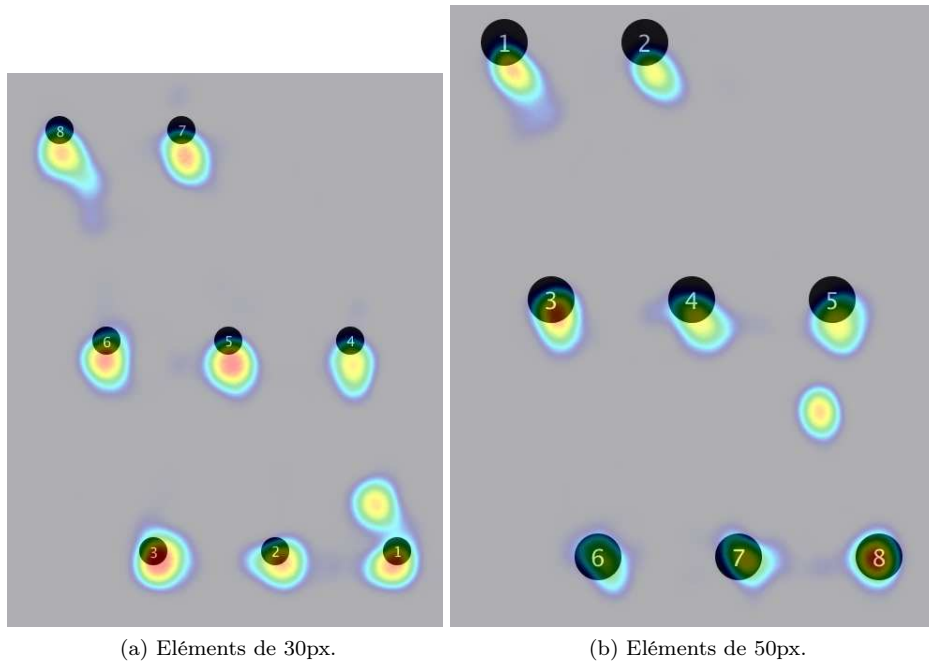


FIGURE 5.9 – Heatmap des données.

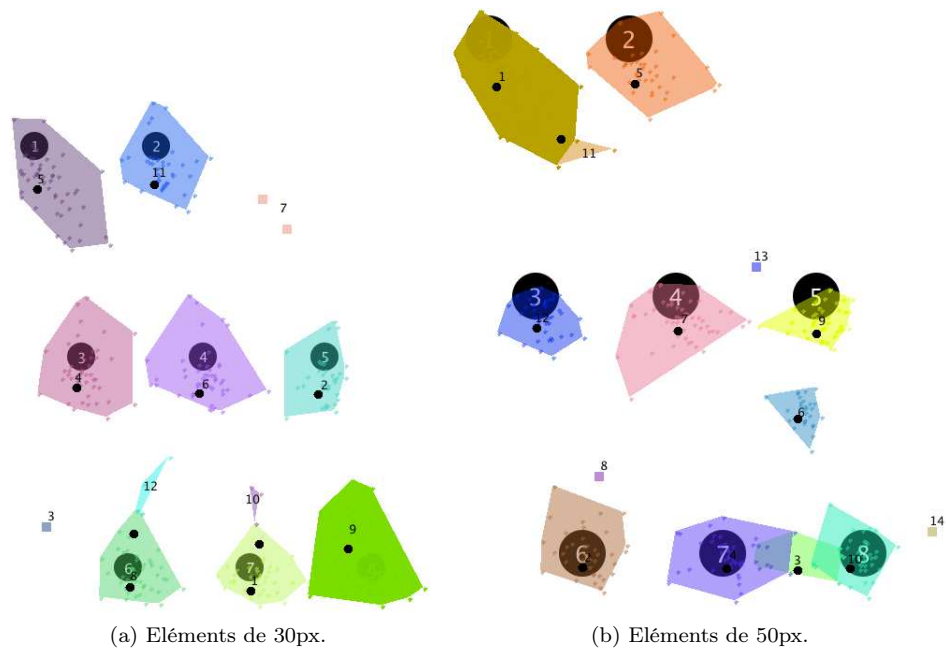


FIGURE 5.10 – Effet de la taille des éléments du stimulus.

Il semblerait qu'une *AOI* n'ait pas été identifiée pour les éléments de 30 pixels. En effet, on peut voir sur la heatmap qu'il y aurait en fait une *AOI* supplémentaire en bas à droite (cette *AOI* est en fait un artefact expérimental constitué d'une fixation intermédiaire suite à l'affichage du stimulus). Là aussi, certaines *AOIs* semblent trop étendues.

5.2.3.4 Effet de la distance entre les éléments

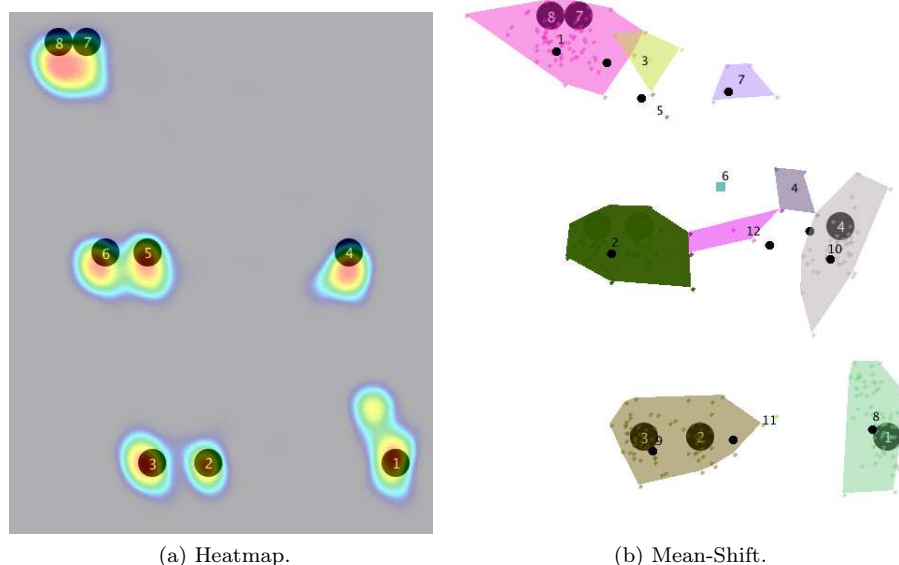


FIGURE 5.11 – Effet de la distance entre éléments du stimulus.

On observe que la réduction des distances inter-éléments provoque la fusion des *AOIs*. Ainsi, les éléments 2 et 3, 5 et 6 puis 7 et 8 sont groupés au sein d'une même *AOI*.

5.2.4 Conclusions : limites de la procédure *Mean-Shift*

Avec la distance variable, on se rend compte que la taille de la bande passante (et donc des *AOIs*), dépend surtout du stimulus, et plus précisément de l'espace-ment entre ses éléments. Contrairement à ce qui est couramment admis, ce n'est donc pas la taille de la vision fovéale qui doit fixer la taille de la bande passante, mais bien les caractéristiques du stimulus. Ce résultat complique considérablement la sélection de la bande passante car les éléments d'un stimulus peuvent varier en taille et en distances. De plus, les caractéristiques d'un stimulus ne sont pas toujours connus par avance, il convient donc de définir une méthode qui identifie la taille de la bande passante directement à partir des données oculométriques.

Selons nous, pour être réellement utile, la détection automatique des *AOIs* par la procédure mean-shift doit à la fois être réellement non-paramétrique, permettre la détection des *outliers* («fausses» *AOIs*) et fournir un indicateur global sur sa qualité. En l'état actuel des connaissances, ces trois exigences constituent les limites de la procédure mean-shift.

5.2.4.1 Le choix de la bande passante (paramètre h)

Le paramètre h est en fait la taille de la fenêtre ou bande passante permettant d'estimer la densité autour de chaque point. Ce paramètre nécessite d'être choisi par avance, et de ce choix vont dépendre le nombre et la forme des *AOIs*. Il n'y a pas de consensus à ce niveau, notamment pour les données oculométriques. A notre connaissance, seuls Duchowski et son équipe ([Duchowski et al., 2010]) ont proposé de fixer h au regard de l'étendue de la vision fovéale. Bien que théoriquement séduisante, leur proposition n'a pas été testée de façon systématique, alors que le choix de la bande passante est d'une importance critique.

Ainsi, lorsque ce paramètre est trop petit ($h = 10$ dans la figure 5.12), l'estimation de la densité sera trop fine avec pour résultat un grand nombre de petites classes. A contrario, lorsque ce paramètre est trop grand ($h = 50$ dans la figure 5.12), l'estimation de la densité ne sera pas assez fine avec pour résultat un petit nombre de grandes classes.

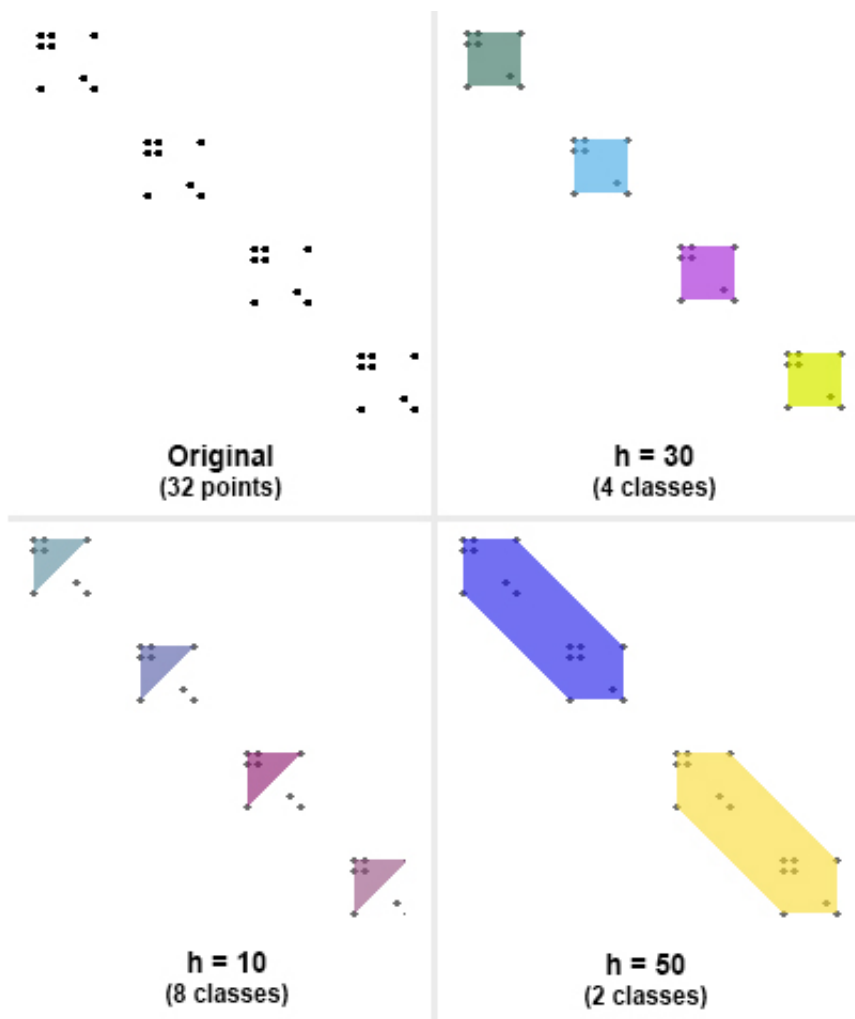


FIGURE 5.12 – Comparaison de l’effet de différentes valeurs du paramètre h de la procédure mean-shift sur des données fictives ($h=30$ représente la valeur «idéale» pour ce jeu de données).

5.2.4.2 Le qualité de la classification

La qualité de la classification est en fait liée au choix du paramètre h . En effet, comment juger de la meilleure configuration d’AOIs en fonction du choix de la taille de ce paramètre ? Dans la figure 5.12, nous savons par avance qu’il y a 4 AOIs et donc que $h = 30$ a probablement généré la meilleure structure de classes possible. En revanche, si nous n’avions pas d’idée sur la taille ou le nombre des AOIs, il nous faudrait un critère fiable permettant de savoir que $h = 30$ génère la meilleure configuration possible.

5.2.4.3 La détection des *outliers*

Les données oculométriques ne sont pas toutes utiles pour une analyse. En effet, il peut y avoir des fixations intermédiaires ou à la limite du stimulus. Généralement, on considère que ces *outliers* n'auront pas tendance à se regrouper, et seront donc dans de très petites classes. Ainsi, il suffit de supprimer les classes contenant moins de n fixations. Cependant, cette approche est limitée car il peut y avoir des *outliers* contenant plus de n fixations et n dépend donc du nombre de fixations total à analyser (n serait petit pour peu de fixations et plus grand pour un grand nombre de fixations). Ainsi, n devient un paramètre supplémentaire de la procédure *Mean-Shift*, et, là encore, il n'existe pas de critère faisant consensus pour fixer sa valeur (à notre connaissance, ce paramètre n'a pas fait l'objet de discussions dans la littérature).

5.2.5 Proposition d'amélioration du *Mean-Shift* pour les données oculométriques (*HMS-ET*)

Nous avons choisi de modifier la méthode initialement proposée par Vatturi et Wong ([Vatturi & Wong, 2009]) dans le cadre de la détection de catégories dans le domaine de la surveillance. Celle-ci nous a parue intéressante car, selon ses auteurs, elle permet de détecter les *outliers*. Il s'agit d'une version hiérarchique de la procédure mean-shift (*Hierarchical Mean-Shift*, *HMS*) qui produit une hiérarchie de classes en «floutant» les données de façon répétée. Ce «floutage» est accompli en exécutant itérativement la procédure mean-shift en incrémentant la taille de la bande passante, et en utilisant les centres de l'itération précédente pour l'itération suivante. Le *HMS* produit une hiérarchie de classes similaire à un dendrogramme, mais basé sur le Mean-Shift.

5.2.5.1 Construction de la hiérarchie

Dans un premier temps, la bande passante initiale est sélectionnée en trouvant la plus faible distance (supérieure à zéro) h_{min} entre deux points des données. Le niveau le plus bas de la hiérarchie est composé des points individuels. L'itération initiale du *HMS* applique le *Mean-Shift* avec la bande passante h_{min} sur ces points. Chaque point est assigné à sa propre classe de taille 1 ou est fusionné à une classe plus grande. Dans les itérations suivantes, les centres de ces classes deviennent les points du niveau suivant de la hiérarchie et la bande passante est incrémentée en multipliant la valeur de la bande passante par k (valeur incrémentale de la bande passante). La procédure *Mean-Shift* avec la bande passante incrémentée est appliquée sur ces points (qui sont les centres du niveau inférieur de la hiérarchie) afin de produire un nouvel ensemble de classes. Ce processus continue jusqu'à ce qu'il ne reste qu'une seule classe. Ainsi, les données sont «floutées» de manière répétée par le centre des classes des points.

Algorithme 5.2 Mean-Shift hiérarchique

Étant donné S , l'ensemble des données et h la taille de la bande passante ;
Définir $h = h_{min}$ la distance minimale supérieure à zéro entre deux points dans S ;
Définir $Herarchy[0].ClusterCenters = S$;
 $l = 0$;
répéter
 $newClusterCenters = MeanShift(Herarchy[l].ClusterCenters, h)$;
 Définir $Herarchy[l+1].ClusterCenters = newClusterCenters$;
 $h = h + 1.1$;
 $l = l + 1$;
jusqu'à $size(Herarchy[l].ClusterCenters) == 1$;

5.2.5.2 Détection des *outliers*

Le critère d'*outlierness* est basé sur le concept de durée de vie d'une classe. La durée de vie se définit comme l'étendue des échelles logarithmiques dans laquelle une classe survit. De façon similaire, Vatturi et Wong définissent la durée de vie dans le *HMS* comme l'étendue des bandes passantes logarithmiques dans laquelle une classe survit (la différence logarithmique entre la bande passante au moment où la classe est formée et le moment où elle est fusionnée avec une autre classe). Ainsi, les classes qui ont de longues durées de vie et peu de points ont de hautes valeurs d'*outlierness*.

L'*outlierness* se calcule de cette façon, pour une classe C_i :

$$Outlierness_i = \frac{\text{durée de vie de } C_i}{\text{nombre de points dans } C_i} \quad (5.2)$$

Dans la phase initiale du *HMS*, les classes contiennent uniquement les données d'origine. A chaque nouveau niveau hiérarchique, il est possible que certaines classes ne fusionnent pas avec d'autres, et persistent dans plusieurs niveaux hiérarchiques, ce qui augmente leur durée de vie et donc leur probabilité d'être détectées comme *outliers*.

5.2.5.3 Sélection de la meilleure structure

Pour détecter la meilleure structure (ensemble de classes) de classification dans la hiérarchie, nous avons choisi de nous référer à un indicateur de stabilité (les indicateurs habituels n'étant pas adaptés au *Mean-Shift Clustering*, voir [Kovács et al., 2005]). Le critère le plus simple aurait été de retenir la structure stabilisée au plus bas niveau hiérarchique. Autrement dit, lorsque deux niveaux hiérarchiques différents, après incrémentation de bande passante, aboutissent à la même configuration de classes. Néanmoins, il est possible que ce critère ne soit jamais atteint, et donc qu'aucune structure ne puisse être détectée. Nous proposons un indicateur se rapprochant de ce concept, mais qui permet toujours

de détecter une structure.

Cet indicateur évalue le coût de déplacement (e.g. la distance euclidienne) des centres des classes entre deux niveaux hiérarchiques. Ce coût est le plus bas (0) lorsque deux niveaux hiérarchiques produisent strictement la même classification (stabilité), en d'autres termes, lorsque l'incrément de la bande passante produit les mêmes résultats. Ainsi, pour chaque niveau hiérarchique dans le HMS, il est possible de calculer son indicateur de stabilité.

Etant donné une hiérarchie ayant l niveaux $H = (H_1, H_2 \dots H_l)$ contenant chacun un ensemble de n classes $C = (C_1, C_2 \dots C_n)$ ayant chacune un centre v , l'indicateur de stabilité pour un niveau hiérarchique j est donné par :

$$stab(H_j) = \sum_{C_i \in H_j, C_k \in H_{j+1}} \min(\|v_i - v_k\|^2) \quad (5.3)$$

La meilleure configuration sera sélectionnée de la façon suivante :

$$\min[0] \{stab(H)\} \quad (5.4)$$

Il faut noter que le critère de stabilité peut être atteint à plusieurs niveaux de la hiérarchie, c'est à dire qu'il est possible qu'à plusieurs endroits, l'incrément de la bande passante produise le même résultat.

De cette façon, nous pouvons détecter le niveau hiérarchique le plus stable au plus bas niveau et ne conserver que le résultat du Mean-Shift de ce niveau, associé à sa propre bande passante. L'avantage de cette implémentation est qu'elle permet d'utiliser n'importe quel critère de stabilité.

Algorithme 5.3 Sélection de la meilleure configuration.

Construire la hiérarchie des classes en utilisant le Mean-Shift hiérarchique décrit dans l'algorithme 5.2.

Définir ht comme étant sa hauteur.

pour $i = ht$ **jusqu'à** 0 **faire**

 Pour chaque ensemble de classes C à la hauteur i calculer son niveau de stabilité par rapport à $i+1$ selon l'équation 5.3;

 Pour chaque classe C_j à la hauteur i calculer sa valeur d'*outlierness* selon l'équation 5.2;

fin

Trouver le premier niveau hiérarchique ayant le meilleur niveau de stabilité (équation 5.4);

Supprimer les classes de ce niveau ayant une valeur d'*outlierness* 1.

Il est aussi intéressant de supprimer les classes du niveau sélectionné détectées comme *outliers*. Chaque classe de ce niveau hiérarchique se voit attribuer un niveau d'*outlierness*, mais peut devenir *outlier* dans un niveau supérieur, ce qui

aurait pour effet la non-détection d'un *outlier*. Ainsi, nous retenons le plus haut niveau d'*outlierness* que peut atteindre une classe dans toute la hiérarchie pour détecter les *outliers*.

5.2.6 Evaluation du *HMS-ET*

Dans cette section, nous évaluons le *HMS-ET* sur les données de l'expérimentation. Afin de juger de son apport, nous comparons sa capacité à détecter les outliers et à choisir correctement la bande passante par rapport à la procédure «classique» où la taille de la bande passante est fixée à 1.5 degrés d'angle visuel (50 pixels dans notre cas).

5.2.6.1 Détection des outliers

Malgré le caractère «guidé» des tâches de notre expérimentation, nous avons observé que les participants émettent fréquemment des fixations intermédiaires avant de fixer leur cible. Nous pouvons voir un exemple de ce phénomène dans la figure 5.13.

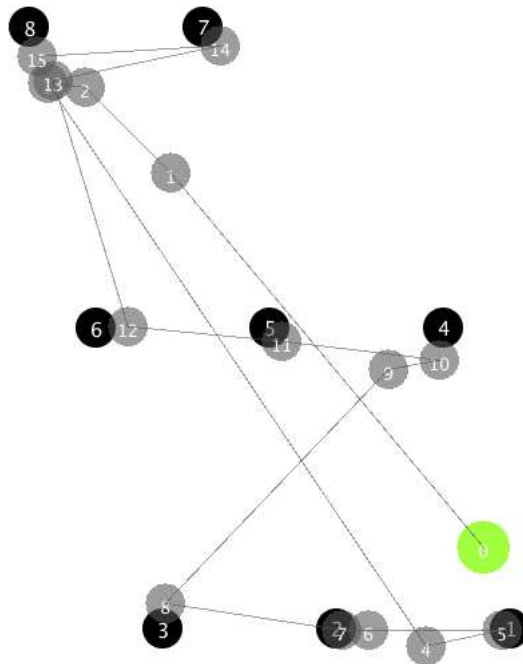


FIGURE 5.13 – Exemple de scanpath avec fixations intermédiaires.

Ces fixations rendent particulièrement difficile la détection des *AOIs* et conduisent la procédure *Mean-Shift* à sur-évaluer leur taille, à les différencier de manière peu précise, et à proposer de (parfois) nombreuses *AOIs* qui ne

sont pas toujours utiles pour l'analyse des données (*outliers*). On peut voir une illustration de ce constat dans la figure 5.14.

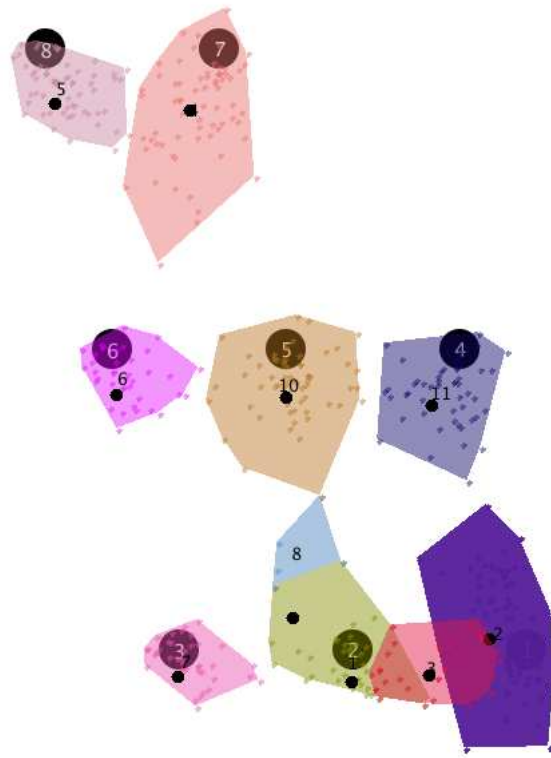
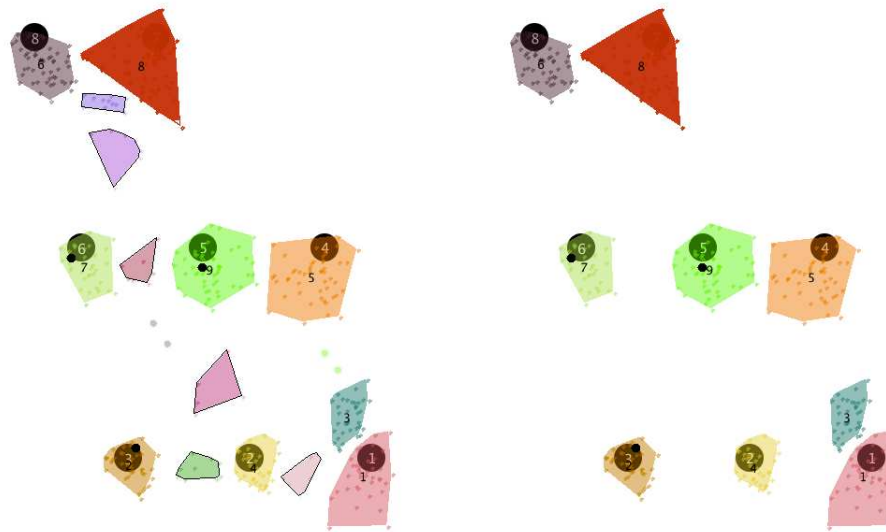


FIGURE 5.14 – Résultat du Mean-Shift avec $h=50$.

En analysant ces mêmes données avec le *HMS-ET*, il a été possible de détecter les *outliers* (Figure 5.15a) et de sélectionner des *AOIs* bien distinctes (Figure 5.15b).



(a) HMS-ET avec *outliers* (AOIs avec bordures noires).

(b) HMS-ET final.

FIGURE 5.15 – Résultat du *HMS-ET* avec et sans l’affichage des *outliers*.

5.2.6.2 Sélection de la bande passante

Nous avons vu que l’espacement entre les éléments d’un stimulus influe sur les résultats du *Mean-Shift*. Plus précisément, les éléments proches ont tendances à être fusionnés dans une même *AOI*. Nous avons donc testé le *HMS-ET* dans ce type de configuration. Il apparait clairement qu’il identifie correctement les *AOIs* (Figure 5.16b). Le *HMS-ET* semble s’adapter localement à la densité des points de fixation.

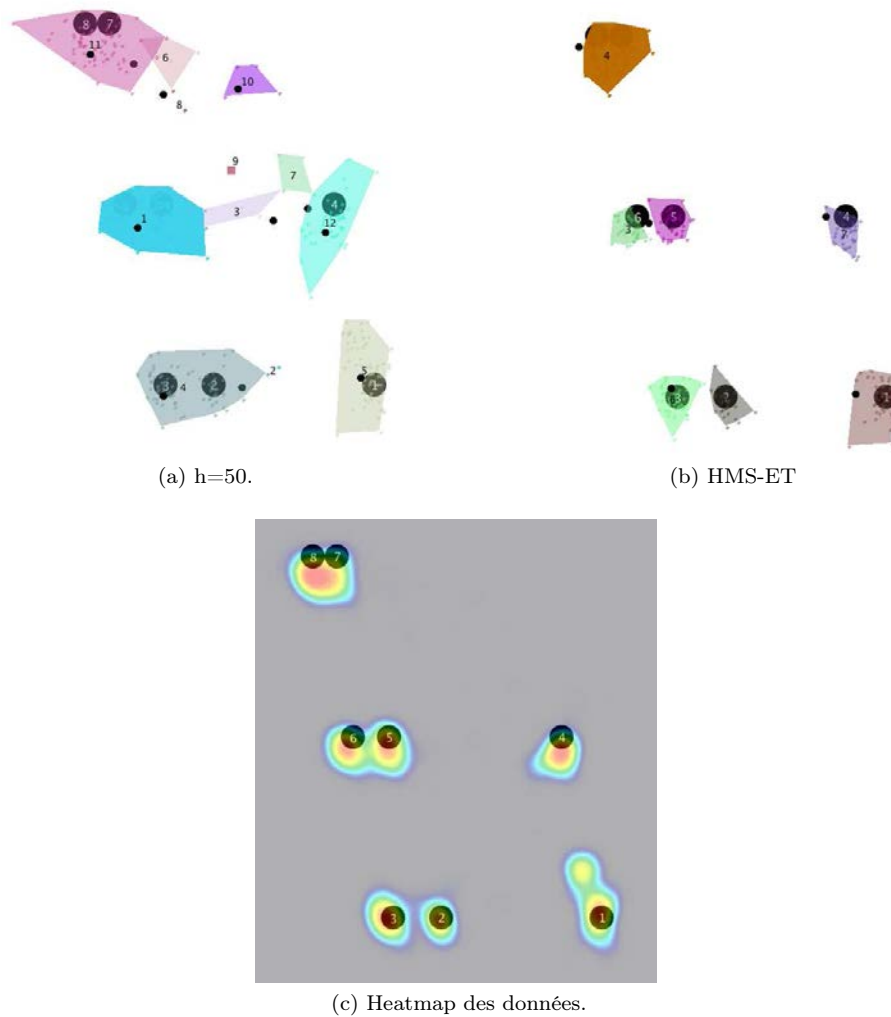


FIGURE 5.16 – Comparaison entre le Mean-Shift avec bande passante fixe ($h=50$) et notre proposition.

5.2.6.3 Conclusion

Les résultats de notre comparatif montrent que le *HMS-ET* semble supérieur à la procédure *Mean-Shift* «classique». Il répond parfaitement à notre problématique de détection non-paramétrique des *AOIs*, tout en identifiant (et supprimant) le «bruit» dans les données oculométriques.

5.2.7 Conclusions sur la détection automatique des *AOIs*

Nous sommes partis du constat que la procédure *Mean-Shift* proposée dans la littérature comportait des limites la rendant inapplicable en l'état. D'une part, le paramètre de bande passante ne fait pas l'objet d'un consensus pour

les données oculométriques, et d'autre part, la question des *outliers* n'a jamais été traitée. Nous avons proposé une amélioration hiérarchique du *Mean-Shift* (le *HMS-ET*) permettant de sélectionner la structure la plus stable en identifiant puis supprimant les *outliers*. L'évaluation du *HMS-ET* nous a permis de confirmer l'apport en fiabilité de cette méthode sur des données réelles. Concernant le choix de la structure, nous avons constaté visuellement sur nos données que le critère de stabilité que nous proposons semble efficace. Cependant, nous ne savons pas si ce critère est adapté à tous les types de jeux de données, ce point mériterait plus de recherches.

5.3 Définition d'une méthode pour l'analyse des patterns

Pour analyser les patterns oculométriques, nous devons définir une métrique permettant de comparer les scanpaths entre eux. Comme nous avons pu le voir dans le chapitre précédent, une telle métrique peut potentiellement sur-simplifier les scanpaths et conduire à une perte de finesse d'analyse. Dans cette section, nous verrons quelles sont les limites des méthodes actuelles et comment nous proposons de les résoudre.

5.3.1 Limites des méthodes actuelles de comparaisons des scanpaths

5.3.1.1 Les méthodes de visualisation

Ces méthodes sont particulièrement intéressantes pour comparer qualitativement deux scanpaths visuellement. Cependant, elles s'avèrent inadaptées à de gros échantillons de données, les rendant difficilement utilisables pour l'identification de patterns comportementaux.

5.3.1.2 La distance de Levenshtein

La principale limite de la distance de Levenshtein dans sa version originale est qu'elle n'est pas adaptée à la comparaison de scanpaths car les aspects temporels ne sont pas pris en compte. Par exemple, si deux individus visualisent les mêmes *AOIs* mais dans des ordres différents, la distance de Levenshtein va les sur-différencier. Dans la figure 5.17, les deux scanpaths fictifs ont visité les *AOIs* dans des ordres strictement différents. Dans ce cas, la distance de Levenshtein sera de 1 ($\frac{4}{4}$ = remplacement de chaque caractère), la conclusion serait que ces deux scanpaths sont entièrement différents, de la même façon que s'ils n'avaient aucune *AOI* en commun.

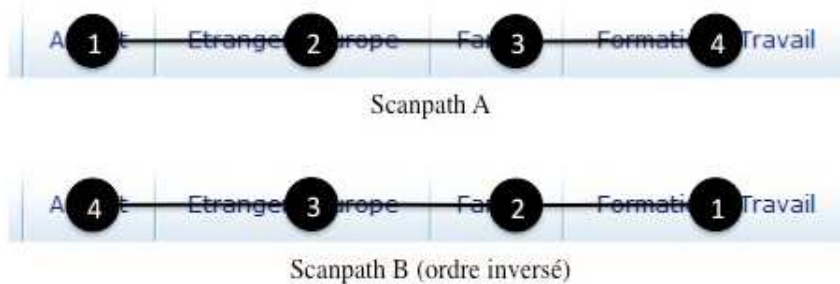


FIGURE 5.17 – Deux scanpaths fictifs visualisant les mêmes *AOIs* dans un ordre strictement différent.

Une autre limite est que la durée des fixations n'est pas non plus prise en compte, ainsi, si deux scanpaths visualisent les mêmes *AOIs* dans le même ordre, mais avec des durées de fixations différentes, la distance de Levenshtein sera de 0 (scanpaths strictement similaires).

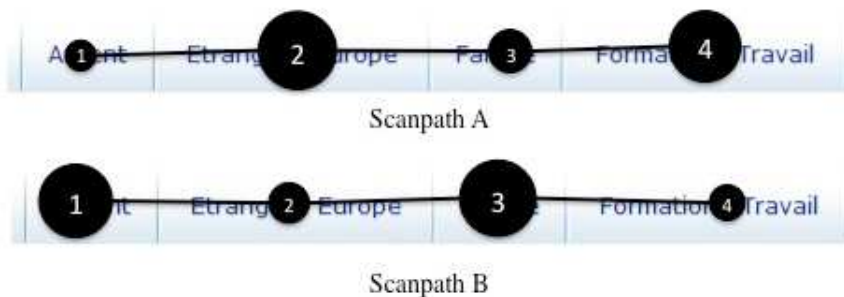


FIGURE 5.18 – Deux scanpaths fictifs visualisant les mêmes *AOIs* dans le même ordre mais avec des durées de fixations différentes (la taille des cercles représente la durée des fixations).

5.3.1.3 La distance de Levenshtein pondérée par les distances inter-*AOIs*

L'amélioration proposée par Takeuchi [Takeuchi & Habuchi, 2007] pondère la substitution d'une *AOI* par une autre par la distance qui les sépare. Pourtant, rien n'indique que deux *AOIs* proches se ressemblent plus que deux *AOIs* distantes. En effet, si un menu est proche d'un logo, mais éloigné du contenu, cela ne signifie par pour autant que le menu ressemble plus au logo qu'au contenu.

5.3.1.4 Les techniques de type «*Common subsequences*»

Ces méthodes sont intéressantes pour identifier un pattern «représentatif», mais ont moins d'intérêt pour comparer les scanpaths comme nous avons pu le voir dans les chapitres précédents.

5.3.1.5 Transitions entre *AOIs*

Les matrices de transitions n'ont qu'un faible intérêt dans notre cas, car en se focalisant sur des paires d'*AOIs*, elle prennent en compte uniquement l'aspect temporel à $n+1$ ([Josephson & Holmes, 2002]).

5.3.1.6 Le MultiMatch

La principale limite du MultiMatch porte sur la phase de simplification des scanpaths. celle-ci consiste à fusionner deux saccades lorsque celles-ci prennent la même direction. Cela a pour effet de supprimer des fixations qui pourraient être importantes pour l'analyse des données. Par exemple, il arrive fréquemment que l'oeil balaye horizontalement les items d'un menu, il ne resterait donc plus que la fixation de départ (premier item) et la fixation d'arrivée (dernier item), sans prise en compte des fixations intermédiaires (items au milieu du menu).



FIGURE 5.19 – Exemple de simplification d'un scanpath fictif via le MultiMatch.

5.3.1.7 Conclusions

Nous pensons que les méthodes que nous avons citées ne conviendront pas à notre objectif qui est de conserver les caractéristiques des scanpaths (notamment l'ordre et la durée des fixations). Ce constat nous a conduit à développer une nouvelle méthode, dérivée de la distance de Levenshtein, que nous allons décrire et évaluer dans la section suivante.

5.3.2 Proposition d'une méthode pour comparer les scanpaths

Nous proposons de définir une méthode en 4 étapes pour comparer les scanpaths. Celle-ci se base sur des *AOIs* définies par le *HMS-ET* puis simplifie les scanpaths pour les comparer selon une version modifiée de la distance de Levenshtein qui prend en compte de manière plus fidèle l'ordre de visualisation des *AOIs* ainsi que la durée des temps de regard (fixations groupées au sein des *AOIs*).

5.3.2.1 Etape 1 : détection des AOIs via le HMS-ET

La première étape consiste à détecter automatiquement les AOIs à l'aide du HMS-ET sur l'ensemble des données oculométriques de tous les participants, puis d'attribuer un identifiant unique à chacune d'entre elles.

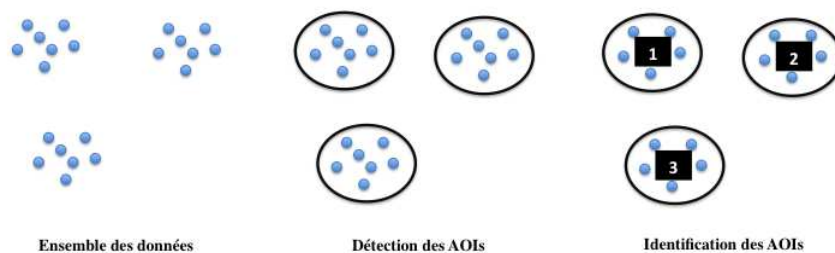


FIGURE 5.20 – Détection et identification des AOIs.

5.3.2.2 Etape 2 : Labellisation des fixations

Dans la deuxième étape, chaque fixation de chaque scanpath reçoit un label correspondant à l'identifiant de l'AOI à laquelle elle appartient (Algorithme 5.4). Si cette fixation n'a eu lieu dans aucune AOI, elle se voit :

- Attribuer le label de l'AOI la plus proche si le centre de celle-ci n'est pas éloigné de plus de 2 fois la taille de la bande passante finale.
- Supprimer dans le cas contraire (il s'agit d'un *outlier*). Il faut noter que cette étape n'est pas obligatoire et dépend des besoins d'analyses.

Algorithme 5.4 Labellisation des fixations d'un scanpath.

Etant donné F la liste des fixations d'un scanpath.

Etant donné $AOIs$ la liste des AOIs (définie par le HMS-ET) de l'ensemble des fixations de tous les participants dont chacune est munie d'un identifiant unique.

Etant donné h la taille de la bande passante globale du HMS-ET.

pour chaque fixation f_i dans F **faire**

si f_i appartient à une AOI d' $AOIs$

Labelliser f_i avec l'identifiant de l'AOI.

sinon

Définir $AOI_{nearest}$ comme l'AOI $\in AOIs$ la plus proche de f_i .

Définir $dist_{nearest}$ comme la distance entre f_i et le centre d' $AOI_{nearest}$.

si $dist_{nearest} \leq (2h)$

Labelliser f_i avec l'identifiant d' $AOI_{nearest}$.

sinon

Supprimer f_i .

fin

De cette façon, chaque scanpath est transformé en une chaîne de caractères aisément comparable.

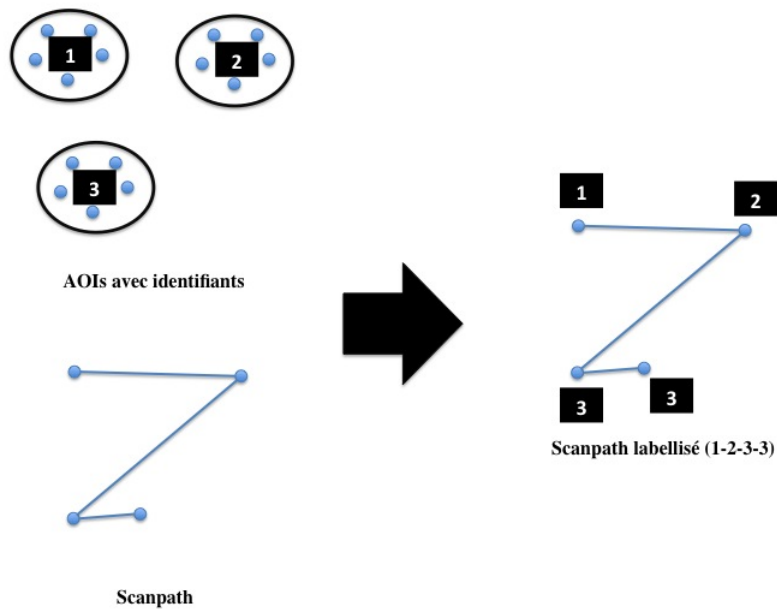


FIGURE 5.21 – Labellisation d'un scanpath.

5.3.2.3 Etape 3 : Simplification des scanpaths

Cette étape consiste à fusionner les fixations consécutives appartenant aux mêmes AOIs. Cette procédure permet de faciliter l'alignement et la comparaison des scanpaths ([Brandt & Stark, 1997]).

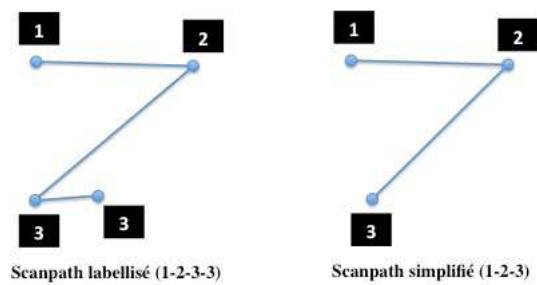


FIGURE 5.22 – Simplification d'un scanpath.

5.3.2.4 Etape 4 : Comparaison des scanpaths selon la méthode *SE-ET* (String Edit for Eye-Tracking)²

La version du *Mean-Shift* que nous utilisons ne conserve pas les aspects temporels des scanpaths. Nous avons fait le choix d'introduire cette notion lors de la comparaison des scanpaths. Ainsi, nous proposons une version modifiée de

2. Cette méthode a fait l'objet d'une publication ([Drusch & Bastien, 2012b]).

la distance de Levenshtein (Damerau–Levenshtein) qui intègre une fonction de coût adaptée aux données oculométriques correspondant à la transposition d’un caractère. Nous proposons de prendre en compte à la fois la distance temporelle entre deux fixations de deux scanpaths sur une même *AOI* et la différence entre les durées de ces deux fixations. Ainsi, la fonction de coût permet de considérer que deux scanpaths seront identiques uniquement s’ils ont visualisé strictement les mêmes *AOIs* dans le même ordre et avec les mêmes durées de fixations.

Pour une fixation labellisée par son *AOI* correspondante dans le scanpath 1 de taille n à la position i ayant une durée de fixation de dur_i et dans le scanpath 2 de taille m à la position j ayant une durée de fixation de dur_j :

Dans un premier temps, nous trouvons la plus faible distance temporelle entre la fixation sur une *AOI* au moment i sur le scanpath 1 et la fixation sur cette même *AOI* sur le scanpath 2 :

$$d_t = \min(|i - j|)$$

Ensuite, nous calculons la distance entre les durées de ces deux fixations de façon à ne pas trop pénaliser une forte différence (afin d’éviter de n’avoir que des couples de scanpaths très dissimilaires, les durées de fixations étant très variables) :

$$d_f = 1 - \exp\left(-\left(\frac{|dur_i - dur_j|}{\max(dur_i, dur_j)}\right)^2\right)$$

Ainsi, le coût de substitution est en fait un coût de transposition prenant en compte à la fois les positions et durées relatives des fixations :

$$cost = \left\{ \left\{ \begin{array}{l} \frac{d_t}{\max(n,m)-df} \quad d_t > 0 \\ \frac{d_f}{\max(n,m)} \quad sinon \end{array} \right\} \quad \begin{array}{l} label \in sp2 \\ sinon \end{array} \right\}$$

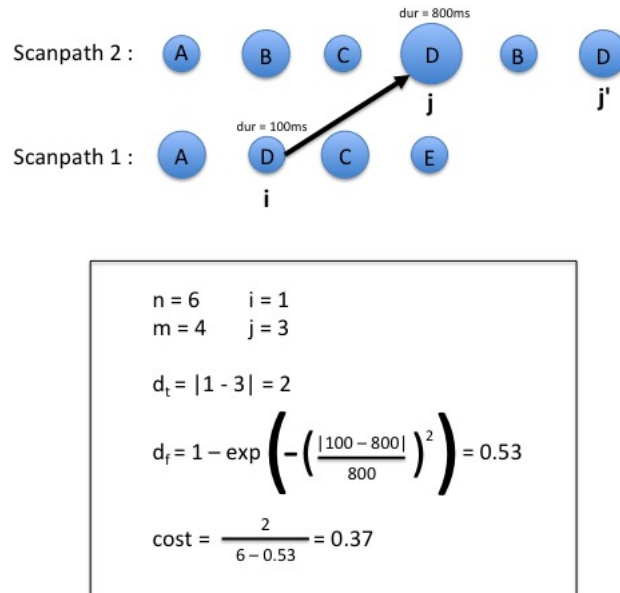


FIGURE 5.23 – Illustration de la fonction de coût du *SE-ET* pour le label d’AOI «D» à la position $i = 1$ dans le scanpath 1.

Notre approche présente ainsi plusieurs avantages :

- Deux *AOIs* à la même position n’auront un coût de 0 que si les durées de fixations sont similaires
- Les durées de fixations pondèrent légèrement le coût de transposition, mais la fonction peut être ajustée si les objectifs de recherche portent essentiellement sur cette métrique
- L’adjacence entre les *AOIs* n’est envisagée que d’un point de vue temporel
- Aucun paramètre n’est nécessaire, la fonction de coût étant indépendante du stimulus présenté aux individus
- Des scanpaths similaires et de même longueur mais ayant visités des *AOIs* dans un ordre différent ne seront pas identifiés comme étant strictement différents

5.3.3 Evaluation de la performance des méthodes

Dans cette section, nous évaluons la performance de différentes méthodes de comparaison des scanpaths pour l’identification de patterns comportementaux au *SE-ET*. Pour cela, nous avons combiné les données de certaines sessions de l’expérimentation visant à valider la procédure *Mean-Shift* (cf. section 5.2.2 page 84). L’objectif est de tester la précision de différents algorithmes pour

détecter, à partir des données (réelles), des patterns préalablement connus (patterns «induits»). Plus spécifiquement, nous testons leur capacité à distinguer correctement des groupes de scanpaths potentiellement différents selon trois conditions :

- Durées variables : les données combinées d’une session de visualisation où tous les éléments apparaissaient en même temps (image) et d’une autre où la vitesse d’apparition des éléments était contrôlée (animation). Cette condition devrait faire apparaître deux patterns de visualisation semblables en termes de positions et d’*AOIs* visitées, mais différents en termes de durées de fixations.
- Positions variables : les données combinées d’une session de visualisation où la distance entre les éléments était fixe et d’une autre où cette distance était variable. Cette condition devrait faire apparaître deux patterns de visualisation semblables en termes de durées de fixations, mais différents en termes de positions et d’*AOIs* visitées.
- Durée et positions variables : les données combinées d’une session de visualisation où la vitesse d’apparition des éléments était rapide (animation rapide) selon une distance fixe avec une session où la vitesse d’apparition des éléments était lente (animation lente) selon une distance variable. Cette condition devrait faire apparaître deux patterns de visualisation différents en termes de durées de fixations et en termes de positions et d’*AOIs* visitées.

Chaque condition devrait donc engendrer deux classes principales si l’on applique sur les données une méthode de comparaison des scanpaths puis une méthode de classification. Afin de juger de la performance des méthodes de comparaison, il suffit de compter combien de scanpaths ont été correctement classifiés pour chacune d’entre elles.

5.3.3.1 Implémentation des méthodes

Nous avons choisi de comparer le *SE-ET* à 4 méthodes différentes. Deux d’entre elles se basent sur les *AOIs* (*AOIs* communes et distance de Levenshtein) et deux autres s’en affranchissent (MultiMatch et position des fixations) Cette section décrit de quelle façon nous les avons implémentées.

***AOIs* communes**

Cette méthode consiste simplement à compter combien d’*AOIs* visitées sont communes pour deux scanpaths. Pour en faire un score de dissimilarité, pour deux scanpaths *SP1* et *SP2* :

$$1 - \frac{\text{nombre d' } AOIs \text{ communes entre } SP1 \text{ et } SP2}{\text{nombre d' } AOIs \text{ uniques dans } SP1 \text{ et } SP2}$$

Distance de Levenshtein

Il s'agit de l'implémentation de base de la distance de Levenshtein où les coûts d'insertion et de suppression sont fixes.

MultiMatch

Notre implémentation du MultiMatch diffère de celle d'origine ([Jarodzka et al., 2010, Dewhurst et al., 2012]) sur deux points :

- L'étape de simplification ne se base pas sur l'angle vectoriel entre deux fixations, mais sur la procédure de simplification que nous avons décrite dans la section 5.3.2.3. De cette façon, nous pallions à la limite principale de cette méthode qui est la suppression de fixations.
- Afin de générer un score de dissimilarité unique, la comparaison des scanpaths simplifiés se fait uniquement sur la différence angulaire et d'amplitude des saccades.

Position des fixations

Il s'agit de la méthode décrite dans la section 4.5.1.1 où les scanpaths sont comparés uniquement selon la position de leurs fixations, indépendamment des *AOIs* identifiées.

5.3.3.2 Méthodologie

Données

Afin de comparer les performances des méthodes, nous avons dans un premier temps combiné les données de sessions expérimentales suffisamment différentes pour générer des catégories de scanpaths distincts. Chaque méthode a ainsi reçue en entrée l'ensemble des scanpaths de 2 sessions distinctes, chaque scanpath étant identifié de façon unique par le couple *identifiant sujet + identifiant session*.

Comparaison des scanpaths

Chaque méthode a permis de créer une matrice de dissimilarité pour chaque couple de scanpaths.

Classification des scanpaths

Chaque matrice a permis de créer une classification ascendante hiérarchique (agglomération par la méthode de Ward) où nous avons identifié les deux classes principales puis compté le nombre de scanpaths incorrectement classifiés («erreurs»).

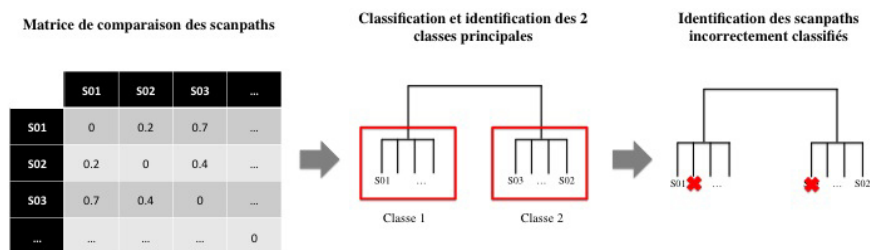


FIGURE 5.24 – Etapes de la méthodologie d’évaluation de la performance des méthodes.

5.3.3.3 Résultats

Les résultats de notre comparatif montrent que de façon générale, le *SE-ET* obtient de meilleures performances (Figure 5.25). En effet, le taux d’erreurs de classification est systématiquement plus bas que ce soit pour la condition durées (19,3% contre 35,5 à 43,5%), la condition positions (1,6% contre 3,2 à 8%) ou la condition positions et durées (0% contre 3,2 à 4,8%).

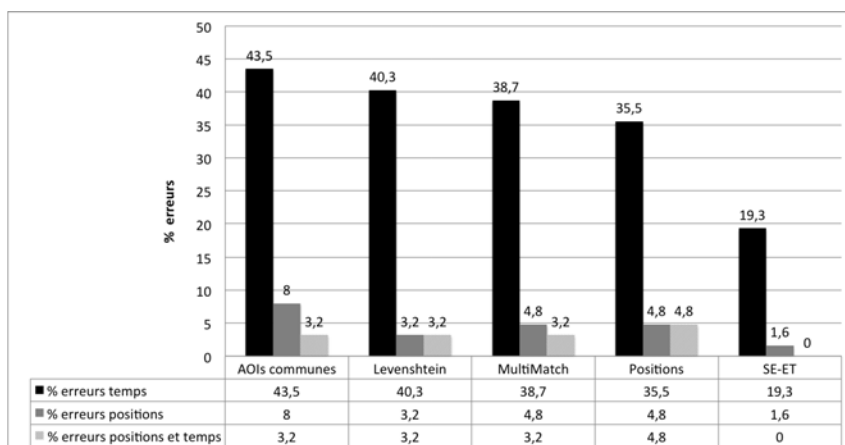


FIGURE 5.25 – Comparaison des méthodes.

Afin d’évaluer la robustesse des classes identifiées avec le *SE-ET*, nous avons calculé, pour chaque classifications ascendantes hiérarchiques, leur niveau de significativité via un rééchantillonnage multidimensionnel par bootstrap avec 10 000 répliquions. Cette méthode est particulièrement efficace pour juger de l’existence réelle des classes identifiées ([Shimodaira, 2002, Shimodaira, 2004, Suzuki & Shimodaira, 2004]). Pour réaliser ces analyses, nous avons utilisé la fonction *pvclust* du package du même nom sous *R*.

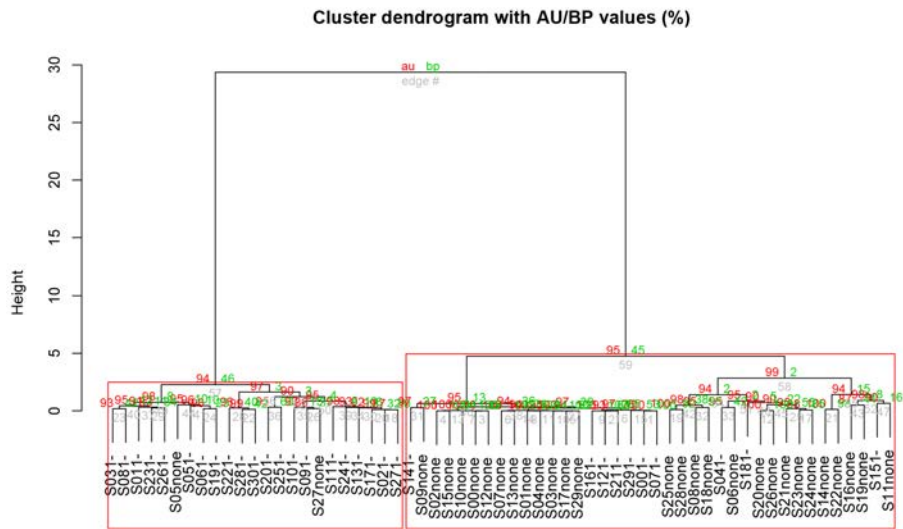


FIGURE 5.26 – Dendrogramme de la classification durées.

Concernant la condition durées, on constate qu’il existe deux classes englobant l’ensemble des données.

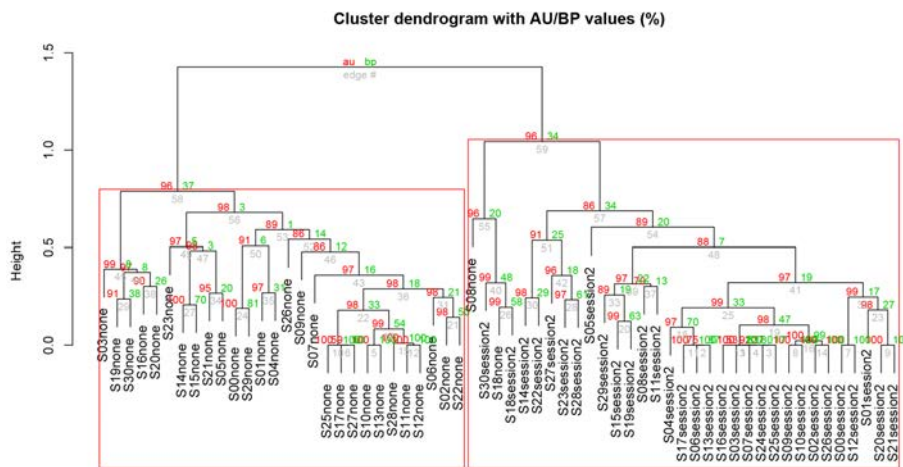


FIGURE 5.27 – Dendrogramme de la classification position.

Concernant la condition positions, on constate aussi deux classes englobant l’ensemble des données.

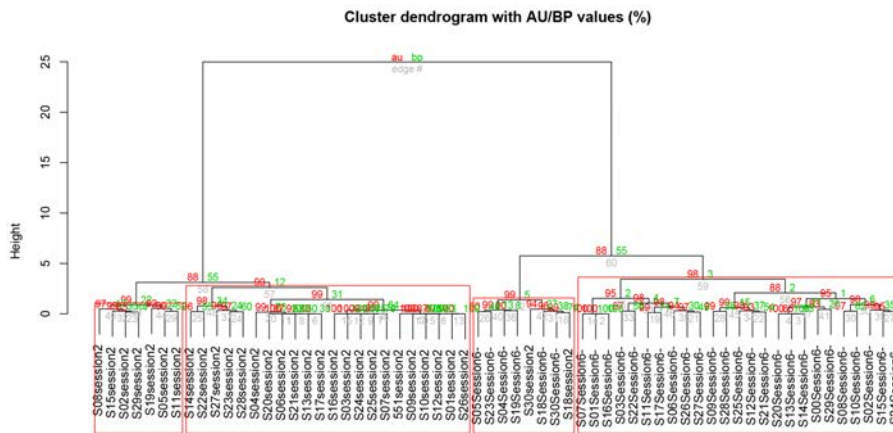


FIGURE 5.28 – Dendrogramme de la classification position et durée.

La condition positions et durées a engendré deux classes principales distinguant les deux catégories de scanpaths et regroupant la majorité des données, puis deux plus petites classes.

5.3.3.4 Conclusions

Nous avons proposé une nouvelle méthode d'analyse des scanpaths dans le but d'identifier des patterns de visualisation. Cette méthode se base sur une version hiérarchique et adaptée aux données oculométriques de la procédure *Mean-Shift* (le *HMS-ET*) combinée à une version améliorée de la distance de Levenshtein (le *SE-ET*) prenant en compte à la fois la distance temporelle et les durées de fixations entre les scanpaths en un seul indicateur. Nous avons évalué cette proposition par rapport différentes méthodes utilisées dans la littérature pour la comparaison des scanpaths. Il s'avère que le *SE-ET* obtient de meilleures performances globales pour l'identification de patterns comportementaux visuels et répond aux limites des méthodes actuelles.

5.4 Conclusion générale

Ce chapitre visait à déterminer une méthode fiable pour l'identification de patterns comportementaux visuels à partir des données oculométriques («*data-driven*»). Nous avons finalement opté pour la combinaison de deux approches complémentaires : l'identification automatique d'*AOIs* et la comparaison de scanpaths.

L'identification automatique d'*AOIs* nous a parut être d'une grande importance pratique car les patterns de visualisation ont potentiellement plus d'utilité lorsqu'on est capable de dire quelles zones d'un stimulus sont plus ou moins

visitées par telle ou telle catégorie de scanpaths. De plus, le caractère automatique de cette approche permet d'éviter d'avoir à formuler des hypothèses sur la distribution des données oculométriques, mais aussi de limiter les biais inhérents à la définition a priori des *AOIs*. Cependant, nous avons démontré que l'approche de base qui consiste à fixer arbitrairement le paramètre de bande passante de la procédure *Mean-Shift* pour identifier les *AOIs* comporte des limites importantes remettant en cause sa fiabilité. Pour cette raison, nous avons proposé une modification de la procédure *Mean-Shift* adaptée aux données oculométriques (le *HMS-ET*) que nous avons évalué. Il s'avère que le *HMS-ET* est d'une plus grande fiabilité que la procédure *Mean-Shift* «classique», et permet en outre une identification fiable des *outliers* dans les données, facilitant notamment la comparaison des scanpaths.

La comparaison des scanpaths est une problématique aujourd'hui bien documentée dans la littérature. De nombreuses approches ont été proposées, mais aucune ne fait l'objet d'un consensus, et cela est d'autant plus vrai que rares sont les études comparatives de ces méthodes sur des données réelles. De façon générale, nous avons constaté que la plupart de ces approches se basent plus ou moins sur la distance de Levenshtein, en considérant les scanpaths comme des chaînes de caractères. Bien que des améliorations aient été proposées, aucune ne permettait de prendre en compte correctement les aspects temporels des scanpaths. Pour ces raisons, nous avons développé notre propre version de cet algorithme (le *SE-ET*), qui prend en compte la distance temporelle et la durée des fixations de deux scanpaths à comparer.

Ainsi, la combinaison de ces deux approches nous permet à la fois de conserver la notion d'*AOIs*, et de comparer les scanpaths en prenant en compte au mieux leurs aspects temporels. Enfin, nous avons évalué avec succès cette méthode globale d'identification des patterns visuels, qui identifie correctement des patterns connus à l'avance selon différentes conditions expérimentales.

Troisième partie

Existent-ils des patterns de
comportements visuels sur le
Web ?

Dans la partie précédente, nous avons développé et validé une méthode d’analyse des patterns de comportements visuels. Maintenant, il s’agit de savoir si, sur le Web, de tels patterns existent puis s’ils sont stables. Pour répondre à cette question, nous avons mené une série d’expérimentations qui font varier différents facteurs. Dans une première expérimentation, nous avons proposé une série de tâches que des participants devaient répéter immédiatement. Dans la deuxième expérimentation, les participants devaient répéter ces tâches mais dans un intervalle de temps plus long (2 jours). Pour ces deux expérimentations, nous avons sélectionné des participants très homogènes (mêmes tranches d’âge et habitudes du Web). La troisième expérimentation comporte des participants moins homogènes (diversités d’âges et d’habitudes du Web). Enfin, la dernière expérimentation avait pour but de faire varier le design des pages Web présentées.

Au travers de ces expérimentations, nous avons cherché à étudier à la fois l’existence des patterns mais aussi leur stabilité ainsi que les conditions dans lesquelles ils pourraient potentiellement varier.

Pour chacune de ces expérimentations, l’analyse a été réalisée selon 5 étapes :

1. Extraction des données oculométriques enregistrées sur les pages d’accueil de chaque tâche
2. Identification des *AOIs* à l’aide du *HMS-ET*
3. Création d’une matrice de similarité basée sur la comparaison des scanpaths deux à deux par le *SE-ET*
4. Application d’une classification ascendante hiérarchique pour identifier des patterns de comportements visuels
5. Caractérisation des patterns par divers indicateurs oculométriques.

Pour identifier la partition «optimale» de la classification, nous avons choisi d’utiliser la méthode *Dynamic Hybrid Cut* ([Langfelder et al., 2008]). Cette méthode présente l’avantage de sélectionner automatiquement le nombre de classes, d’identifier des sous-classes éventuelles et de détecter les *outliers*. Il s’agit en fait d’un hybride de la classification hiérarchique et du *PAM*.

Afin de caractériser les classes (patterns), nous avons choisi de représenter différents indicateurs à l’aide de l’application Web *Bertifier* ([Perin et al., 2014]). Cette application se base sur l’analyse matricielle de Jaques Bertin en encodant les valeurs d’un tableau visuellement (dans notre cas, des cercles de taille variable) puis en groupant les colonnes et lignes similaires. Outre les indicateurs oculométriques décrits dans le chapitre 3, nous avons ajouté :

- Le «back» qui représente un retour vers une *AOI* visualisée précédemment
- La «revisite» qui correspond à la proportion de visualisation des mêmes *AOIs*
- Le nombre d’*AOIs* distinctes visualisées
- Le nombre de fixations par *AOI*

Chapitre 6

Expérience 1 : identification de patterns et étude de leur stabilité à court terme

Cette expérimentation qui incluait 113 participants, visait d'une part à servir de base à l'identification de patterns de comportements visuels, et d'autre part à étudier la stabilité des patterns à court terme. En effet, les participants devaient répéter (R) immédiatement chaque tâche (session) du protocole. Cette expérimentation est décrite dans le chapitre 4.

6.1 Identification des patterns

Dans un premier temps, nous avons analysé les données de la première session de l'expérimentation (page d'accueil du site Service Public) afin de voir si la méthode proposée dans le chapitre 4 fournit des résultats.

19 *AOIs* ont été identifiées à l'aide du *HMS-ET*.



FIGURE 6.1 – AOIs identifiées à l'aide du HMS-ET

3 classes de scanpaths ont été identifiées sur la base du SE-ET, et sont visibles dans la figure 6.2.

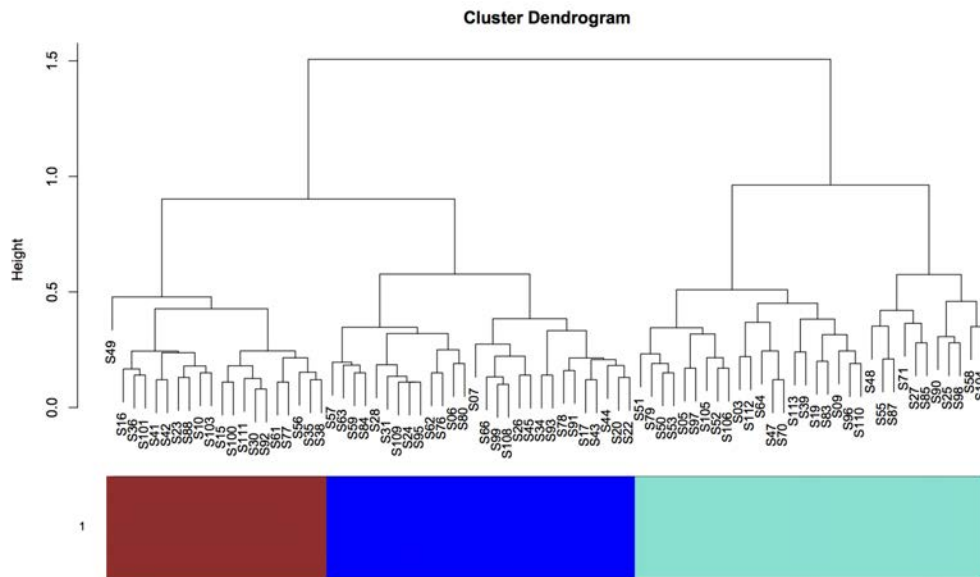


FIGURE 6.2 – Dendrogramme de la classification.

La caractérisation des scanpaths nous permet de voir que les trois classes sont bien distinctes. Afin de voir si un pattern est plus «performant» qu'un autre, nous avons ajouté un indicateur : le taux de sélection du bon lien sur la page d'accueil.

	G1 (n=20)	G2 (n=28)	G3 (n=32)
Barycentre Y	541,05	393,14	398,72
Nombre de fixations par AOI	5,72	3,70	2,73
Durée totale	55869,90	30769,68	18577,47
Nombre de fixations	133,10	76,36	42,56
Longueur des scanpaths	15868,80	9587,10	5560,16
Aire de l'enveloppe convexe	421432,88	363239,04	274547,50
Barycentre X	541,05	530,04	505,75
Nombre d'AOIs distinctes	23,10	20,64	15,47
Dispersion	246,67	247,33	228,13
Nombre de backs	73,85	44,14	23,84
Taux de revisite	68	55	39
Back pondéré	0,56	0,58	0,57
Amplitude des saccades	120,06	126,42	133,39
Durée des fixations	416,37	405,80	444,50
Taux de sélection du bon lien	44,44	33,33	53,33

TABLE 6.1 – Moyenne des indicateurs entre les groupes sur différents indicateurs.

INDICATEURS	1	2	3
BARY	●	•	●
NBFXPARAOI	●	●	•
DURTOTALE	●	●	•
NBFXIATION	●	●	•
LONGSP	●	●	
HULL	●	●	
BARX	●	●	•
AOISDISTINCTES	●	●	•
DISPERSION	●	●	•
BACK	●	●	●
REVISITE	●	●	●
BACKPOND	●	●	●
AMPLSACMOY	•	●	●
DURFIXMOY	●	•	●
TXBONLIEN	●	•	●

FIGURE 6.3 – Bertifier des patterns.

Afin de vérifier si ces groupes existent aussi dans une autre tâche pour ce site, nous avons procédé à la même analyse sur la deuxième tâche que les participants devaient réaliser. Cette fois, 29 *AOIs* ont été détectées (Figure 6.4).

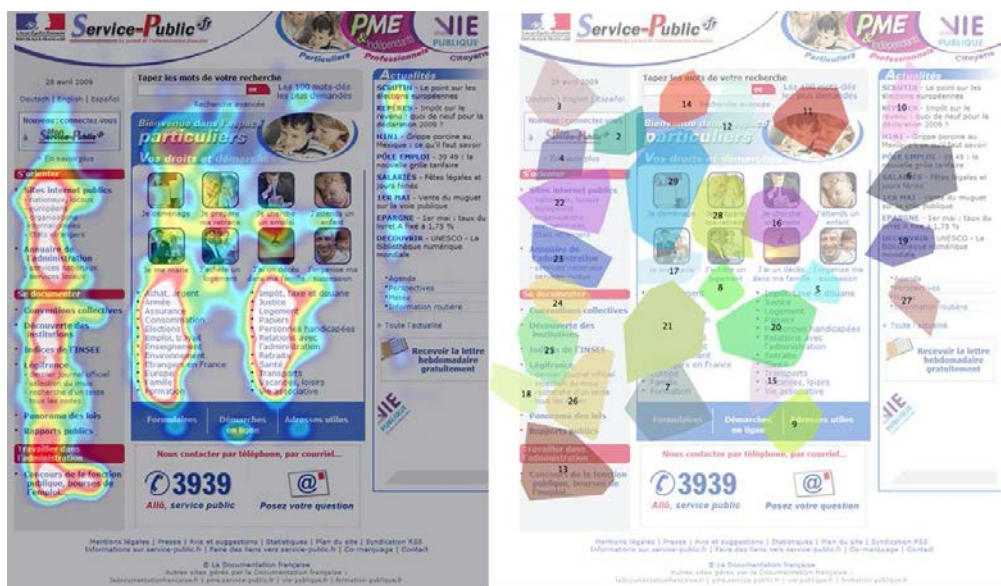


FIGURE 6.4 – Heatmap (à gauche) et AOIs (à droite) identifiés.

A nouveau, trois groupes de scanpaths ont été identifiés, le dendrogramme est très similaire à celui décrit précédemment.

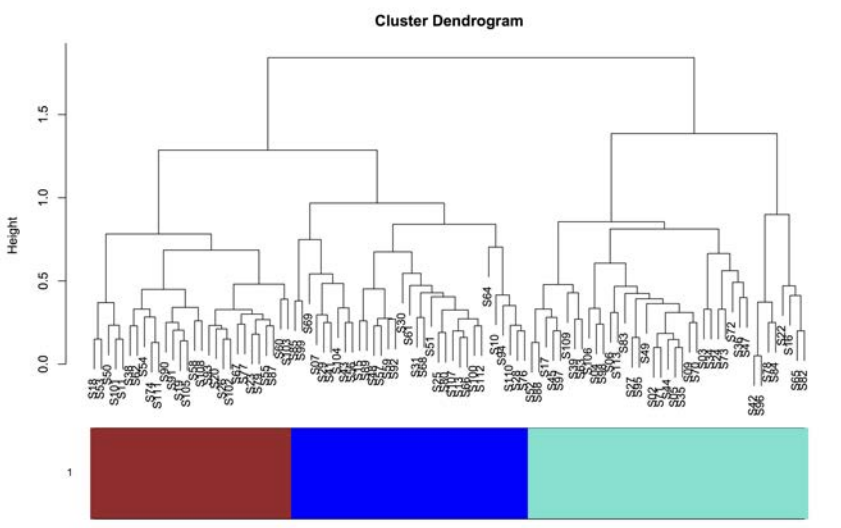


FIGURE 6.5 – Heatmap (à gauche) et AOIs (à droite) identifiés.

La représentation des indicateurs des trois groupes à l'aide de l'application Bertifier conduit là encore à des résultats très similaires à ceux décrits précédemment (les données chiffrées figurent en Annexe A).

INDICATEURS	1	2	3
DISPERSION	●	●	·
AMPLSACMOY	●	●	·
BACKPOND	●	●	●
AOISDISTINCTES	●	●	·
HULL	●	●	
LONGSP	●	●	
DURTTOTALE	●	●	·
NBFXIATION	●	●	·
BACK	●	●	·
REVISITE	●	●	●
NBFXPARAOI	●	●	●
BARX	●	·	●
DURFIXMOY	·	●	●
BARY	·	●	●
TXBONLIEN	●	●	·

FIGURE 6.6 – Bertifier des patterns.

Afin de mieux illustrer les trois classes, nous avons isolé les scanpaths de chaque groupe afin de réaliser leurs heatmaps respectives ainsi qu'un *HMS-ET*.

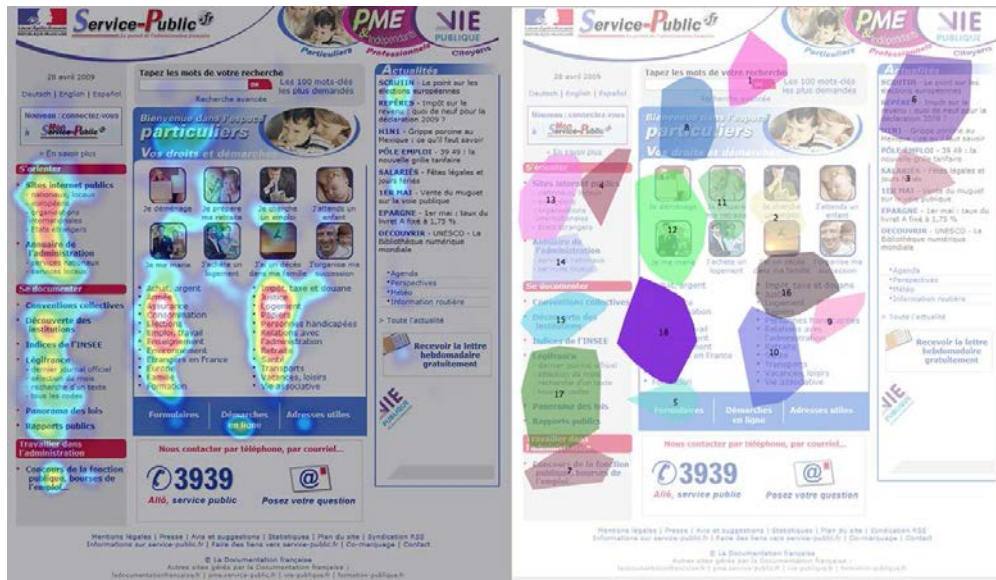


FIGURE 6.7 – Heatmap (à gauche) et AOIs (à droite) identifiées dans le groupe 1.

Nous constatons que les scanpaths du groupe 1 explorent la page Web de façon relativement exhaustive. Ils se caractérisent par un grand nombre de fixations dispersées, des saccades de grande amplitude ainsi qu'une tendance à

revisiter fréquemment les mêmes AOIs. La durée des fixations de ces scanpaths est en revanche relativement courte.



FIGURE 6.8 – Heatmap (à gauche) et AOIs (à droite) identifiés dans le groupe 2.

Les scanpaths du groupe 2 couvrent la page Web de manière moins exhaustive et semblent, pour cette tâche, se centrer principalement sur le côté gauche. Ils ont moins tendance à revisiter les mêmes AOIs et leurs durées de fixations sont un peu plus longues. L'amplitude de leurs saccades sont quasiment identiques à celles du groupe 1 mais ils contiennent peu de fixations.



FIGURE 6.9 – Heatmap (à gauche) et AOIs (à droite) identifiés dans le groupe 3.

Les scanpaths du groupe 3 se caractérisent par une très faible dispersion, peu de fixations et des saccades de faible amplitude. Leurs durées de fixations sont en revanche plus élevées que les autres groupes. Lors de cette tâche, les scanpaths se sont concentrés principalement sur le milieu de la page Web.

6.1.1 Conclusions

Il apparaît que les trois groupes de scanpaths sont bien distincts et que les données globales ne représentent finalement pas de façon fidèle la variabilité des comportements visuels des participants. En effet, les individus du groupe 1 visualisent les pages Web de façon exhaustive et disparate, en revenant fréquemment sur des éléments déjà vus pour enfin choisir leur cible alors que ceux du groupe 2 ont plutôt tendance à visualiser rapidement certains éléments sans trop y revenir. Les individus du groupe 3, quant à eux, se focalisent exclusivement sur quelques éléments proches les uns des autres et choisissent très rapidement leur cible. En outre, il semblerait que la performance, ici mesurée par le taux de sélection du bon lien sur la page d'accueil, ne soit pas associée à une classe particulière.

Il serait maintenant intéressant de savoir si ces groupes sont liés aux caractéristiques des individus comme leur niveau d'expérience d'internet ou encore des données socio-démographiques. Cependant, avant de procéder à ce type d'analyse, il est plus pertinent de savoir si ces patterns sont stables dans le temps. En d'autres termes, est-ce que les individus ont tendance à rester dans le même groupe lorsqu'on leur fait réaliser d'autres tâches et sur différents sites Web ? Dans la prochaine section, nous répétons les mêmes analyses sur l'ensemble

de cette expérimentation dont le protocole comportait un ensemble de tâches répétées sur différents sites Web de structures différentes.

6.2 Stabilité des patterns

La question est maintenant de savoir si ces groupes sont stables dans le temps. Nous avons donc appliqué la procédure décrite précédemment à l'ensemble des sessions de l'expérimentation.

Nous avons obtenu le même nombre de classes pour chaque session. Nous constatons aussi que la représentation Bertifier aboutit à des similitudes importantes entre toutes les sessions. Les trois groupes identifiés sont quasiment identiques entre les 8 sessions de l'expérimentation (les dendrogrammes et les données chiffrées figurent en Annexe A).

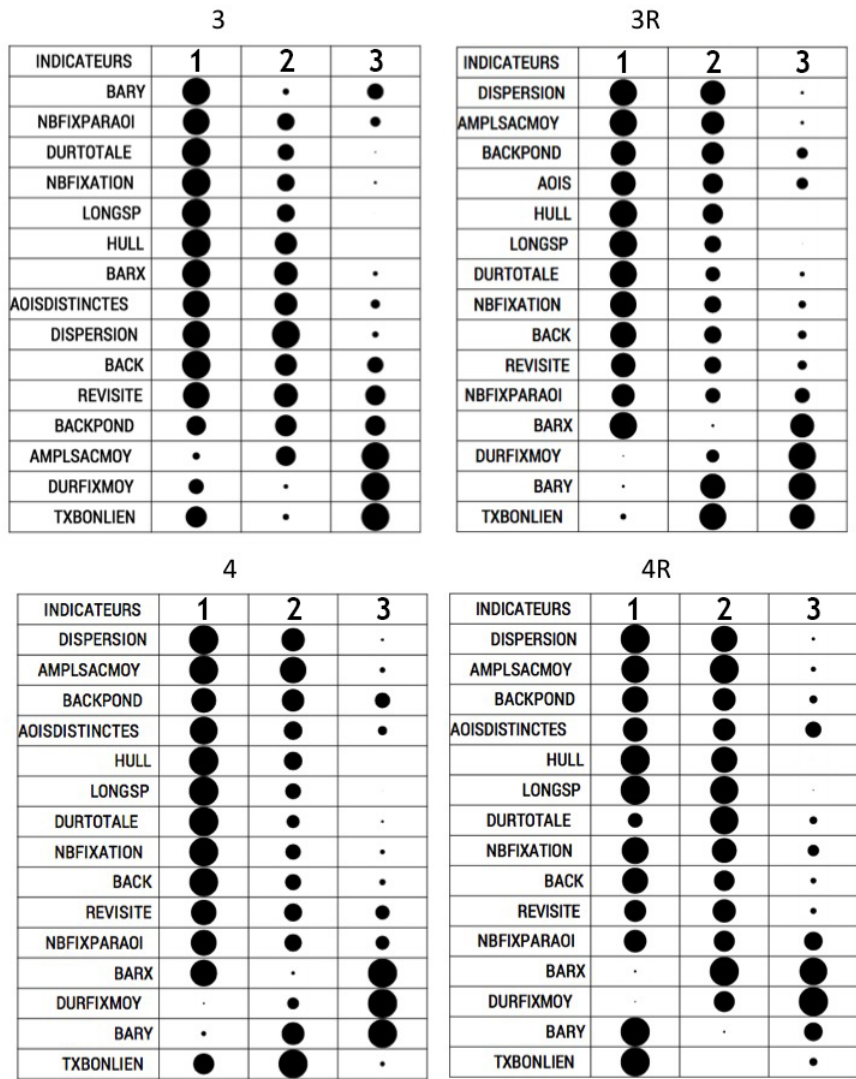


FIGURE 6.10 – Bertifier des sessions 3, 3R, 4 et 4R.

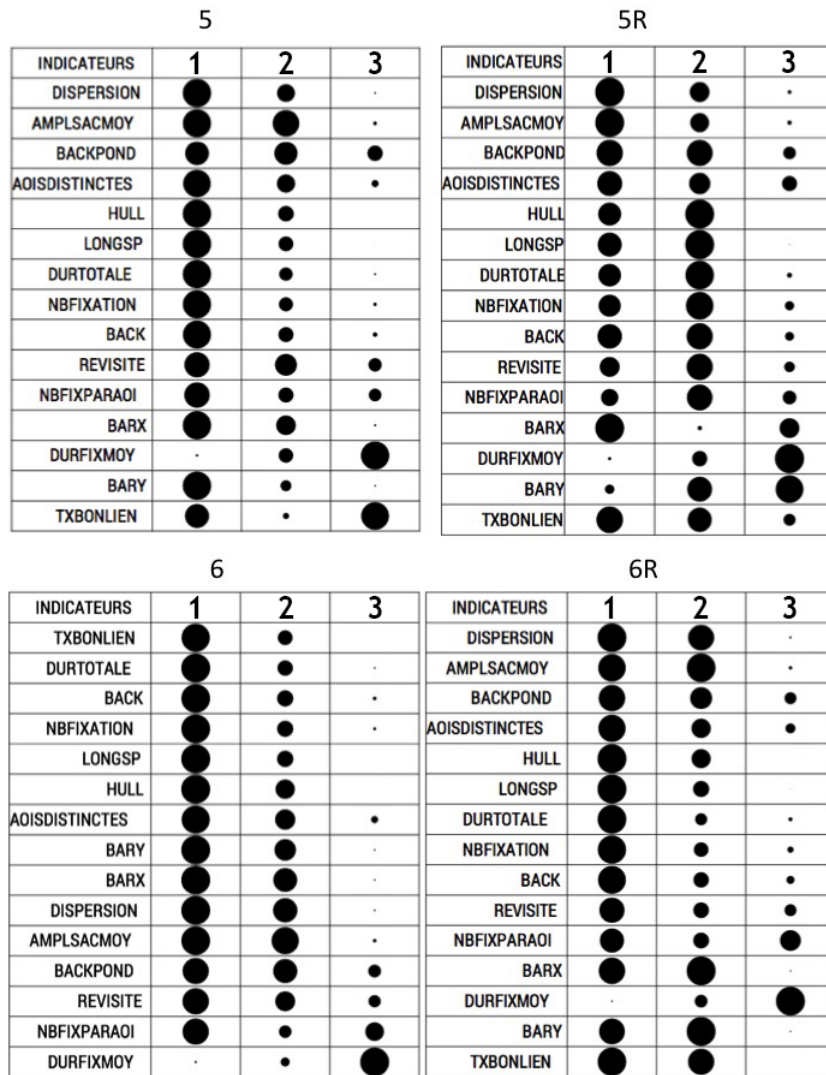


FIGURE 6.11 – Bertifier des sessions 5, 5R, 6 et 6R.

Bien que les groupes soient identiques, nous ne savons pas si ceux-ci contiennent plus ou moins les mêmes individus à chaque fois. En d'autres termes, est-ce qu'un individu qui est dans le groupe 1 dans une session restera dans le même groupe lors de la session suivante ?

Afin de le savoir, nous avons calculé le taux de stabilité entre les sessions, c'est à dire la propension des individus à rester dans le même groupe. Celle-ci se calcule de la manière suivante pour deux Sessions S_i au temps i et S_{i+n} :

$$\frac{\text{nombre de participants dans le même groupe entre } S_i \text{ et } S_{i+n}}{\text{nombre de participants total}} \times 100$$

Comme nous pouvons le voir dans la figure 6.12, la stabilité (propension à rester dans le même groupe) dépend des sites et des tâches.

Sur le site du Service Public :

- la répétition de la première tâche (3 vers 3R) engendre une distribution aléatoire des participants entre les groupes (33%).
- le passage de la première tâche répétée vers la quatrième tâche (3R vers 4) ainsi que la répétition de la quatrième (4 vers 4R) engendre une stabilité légèrement supérieure au hasard (39%).

La transition entre les deux sites (4R vers 5) engendre une distribution aléatoire des participants entre les groupes (34%).

Sur le site du Musée Carnavalet :

- la répétition de la première tâche (5 vers 5R) engendre une stabilité supérieure au hasard (41%).
- le changement de tâche (5R vers 6) semble faire légèrement changer les participants de groupes (29%)
- la répétition de la dernière tâche (6 vers 6R) engendre la plus forte stabilité des groupes (54%).

De façon globale, la stabilité des groupes est de :

- 36% sur le site du Service Public
- 47% sur le site du Musée Carnavalet

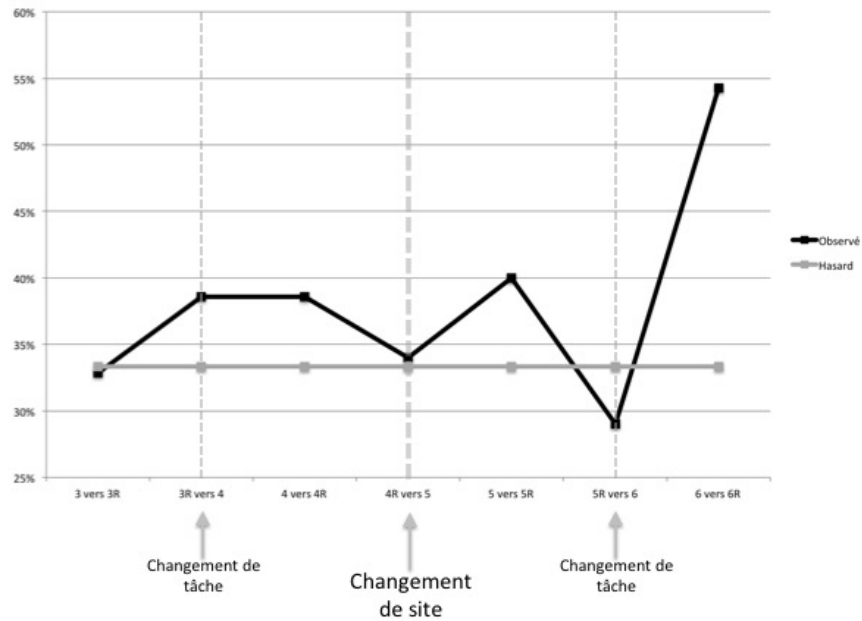


FIGURE 6.12 – Proportion de scanpaths restant dans le même groupe selon les sessions.

En dehors de leur répétition, nous avons voulu savoir si le changement de tâche et de site était à l'origine d'une instabilité des groupes.

Ainsi, nous constatons que :

- le passage de la première à la deuxième tâche sur le site du Service Public (3 vers 4) engendre une stabilité de 37%
- le changement de site (4 vers 5) engendre une stabilité de 36%
- le changement de tâche sur le site du Musée Carnavalet engendre une stabilité de 47%

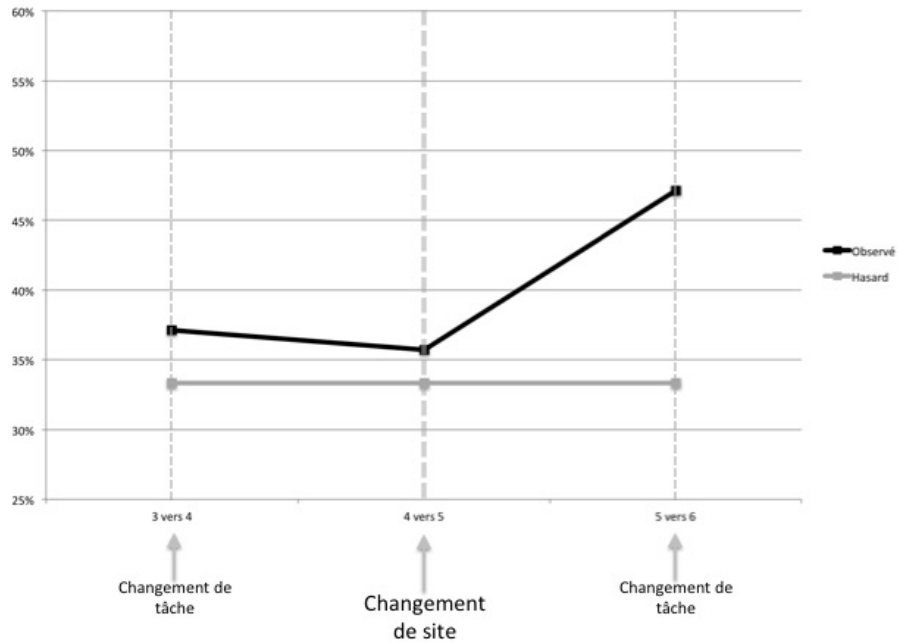


FIGURE 6.13 – Proportion de scanpaths restant dans le même groupe selon les sessions hors répétitions.

6.2.1 Conclusions

Nos résultats laissent apparaître qu’il n’y a pas vraiment de tendance concernant la stabilité des patterns en fonction de la répétition d’une tâche, du changement de tâche ou de site. Les patterns sont plus stables lors du changement et de la répétition des tâche sur le site du Musée Carnavalet que sur le site du Service Public. En outre, les patterns sont globalement plus stables sur le site du Musée Carnavalet.

Finalement, nous constatons que globalement, les patterns identifiés ne sont pas très stables. En effet, la stabilité maximale est de 54% lors d’une répétition de tâche, ce qui signifie tout de même que 46% des participants ont changé de groupe. Il apparaît donc que les individus ont tendance à adopter un comportement différent selon les tâches et les sites, mais selon 3 patterns distincts.

6.3 Conclusions

La méthode d'identification des patterns de comportements visuels que nous avons décrite dans le Chapitre 4 s'est révélée fructueuse. Nous avons pu identifier trois patterns bien distincts qui révèlent la structure sous-jacente des données agrégées globalement. Cependant, deux points nous ont étonnés dans ces résultats. En premier lieu, les mêmes patterns semblent émerger sur différents sites et différentes tâches. On aurait pu penser que les patterns se modifieraient sensiblement, et que des sous-groupes apparaîtraient. En second lieu, ces patterns sont effectivement stables, mais pas pour un individu donné. Cela signifie qu'un individu ne va pas toujours adopter le même pattern, alors qu'il explore le même site ou réalise la même tâche.

On aurait pu penser, par exemple, qu'il existe des individus qui ont tendance à explorer exhaustivement les pages Web de façon stable, au moins lorsqu'un site Web leur est présenté pour la première fois. En effet, la théorie du scanpath stipule que ce dernier est «le chemin habituellement préféré» pour chaque individu, alors que nos résultats tendent à montrer que, sur le Web, il n'y aurait pas de préférence, mais plutôt l'adoption d'un pattern ou d'un autre en fonction du contexte du site, de la tâche et de sa répétition.

Il est cependant possible que ces résultats soient dûs à notre protocole expérimental. En effet, les participants devaient répéter les mêmes tâches consécutivement, dans un court laps de temps. Il est possible qu'un phénomène d'apprentissage à court terme nous empêche d'observer la stabilité des patterns. Dans le prochain chapitre, nous présentons une étude longitudinale où les participants devaient répéter des tâches mais dans un délai de plusieurs jours.

Chapitre 7

Expérience 2 : étude longitudinale¹

Nous avons vu dans le précédent chapitre que les patterns ne semble pas être stables pour un individu donné, et que cela était peut-être dû à notre protocole expérimental. Nous avons donc mené une étude où les participants devaient réaliser trois fois des tâches sur deux sites Web sur une semaine.

7.1 Matériel

Nous avons utilisé le système d’oculométrie déporté Tobii T120 (réglé sur une fréquence d’échantillonnage de 60 Hz) avec une résolution d’écran de 1280x1024 pixels. Le logiciel Tobii Studio (v.2.3.2.0) a été utilisé pour enregistrer puis exporter les données oculométriques. La durée minimale d’une fixation a été fixée à 100 ms dans un rayon de 35 pixels. Le dispositif expérimental était placé sur un bureau. Durant les passations, les participants s’asseyaient face à l’écran et avaient uniquement le droit d’utiliser la souris, le clavier était sous contrôle de l’expérimentateur pour éviter de perdre des données lors des frappes sur le clavier. La distance entre le participant et l’écran était de 60 cm.

7.2 Stimuli

Pour réaliser cette étude nous avons choisi deux sites web distincts, un site web culturel et un site web gouvernemental. Les participants ont parcouru l’ensemble des sites mais seules les pages d’accueil ont été analysées.

Page d’accueil du site Web culturel (musée Carnavalet) :

1. Cette étude a fait l’objet d’une publication ([Drusch et al., 2012])



FIGURE 7.1 – Page d'accueil du musée Carnavalet.

Page d'accueil du site Web gouvernemental (France Diplomatie) :



FIGURE 7.2 – Page d'accueil de France Diplomatie.

7.3 Procédure

L'expérience fut menée en laboratoire. Nous informions les participants qu'ils réaliseraient quatre tâches de recherche d'information sur le web et que leurs performances n'étaient pas mesurées. Aucune limite de temps ne leur était imposée.

Les tâches à réaliser, sur deux sites Web différents, étaient les suivantes :

1. « Qui fut le premier donateur du Musée Carnavalet ? ». Page Web de départ : carnavalet.paris.fr
2. « En quelle année fut créé le CNED (Centre National d'Enseignement à Distance) ? ». Page Web de départ : diplomatie.gouv.fr

Nous avons demandé aux participants de venir 3 fois sans les informer du contenu des rendez-vous : les lundi (t1), mercredi (t2) et vendredi (t3) de la même semaine. A chaque session, ils effectuaient les tâches dans le même ordre. Avant chaque tâche nous effectuions une calibration. La consigne apparaissait à l'écran, nous demandions aux participants de la lire à voix haute et nous nous assurons de sa bonne compréhension. Les tâches débutèrent à partir de la page d'accueil des sites. Une tâche prenait fin lorsque :

- le participant donnait la bonne réponse ;
- le participant donnait une mauvaise réponse ;
- le participant décidait d'abandonner.

A la fin de la troisième passation (t3) nous faisons passer un questionnaire aux participants afin de contrôler l'homogénéité de notre échantillon et de vérifier s'il n'y avait pas de biais. Le questionnaire était découpé en trois parties :

La fiche signalétique

Nous avons recueilli les données démographiques ainsi que des informations relatives aux problèmes ophtalmologiques.

L'utilisation d'Internet

Nous demandions aux participants d'évaluer leur niveau en informatique et nous recueillions des informations liées à leur usage du web (nombres d'années d'utilisation, temps passé par jour et leur(s) principale(s) activité(s) sur le Web). Cette rubrique comprenait également deux questions contrôle : une première pour savoir s'ils étaient habitués aux expérimentations sur le web et une seconde nous permettant de vérifier s'ils connaissaient l'objectif de notre étude. La connaissance antérieure des sites Web parcourus et le degré de difficulté perçu fut également évalué.

Le niveau d'expertise avec Internet

Cette rubrique reprenait 7 items du questionnaire de Lim et Lee ([Lim & Lee, 2000]). Les sujets devaient répondre à des affirmations en se positionnant en termes d'expertise.

7.4 Echantillon

Trente-et-un étudiants en sciences humaines ont été recrutés pour cette étude. Les données de 4 sujets ont été exclus car la calibration fut impossible pour l'un d'entre eux et les 3 autres ne sont pas venus à toutes les sessions du protocole.

Age moyen (écart type) : 21.46 (2.24).

Genre : 65.4% de femmes et 34.6% d'hommes.

Niveau en informatique : 69.2% déclarent avoir un « bon niveau ».

Nombre d'années d'utilisation d'internet : + de 3 ans pour l'ensemble de l'échantillon.

Temps passé par jour sur Internet : 92.3% déclarent passer au moins 1H / jour sur internet.

7.5 Résultats

7.5.1 Identification des patterns comportementaux

A l'instar de l'expérience précédente, nous avons identifié 3 classes pour les deux sites Web (les dendrogrammes et les données chiffrées figurent en Annexe B). Ces classes semblent assez similaires sur les 3 temps et entre les sites Web :

- La première classe se caractérise par une exploration exhaustive des pages Web avec une tendance prononcée à revenir sur les mêmes *AOIs* plusieurs fois.
- La deuxième classe comporte des scanpaths comportant moins de fixations mais avec des saccades de plus forte amplitude.
- La troisième classe contient les scanpaths les moins étendus sur les pages Web, avec de faibles amplitudes de saccades et de longues durées de fixations.

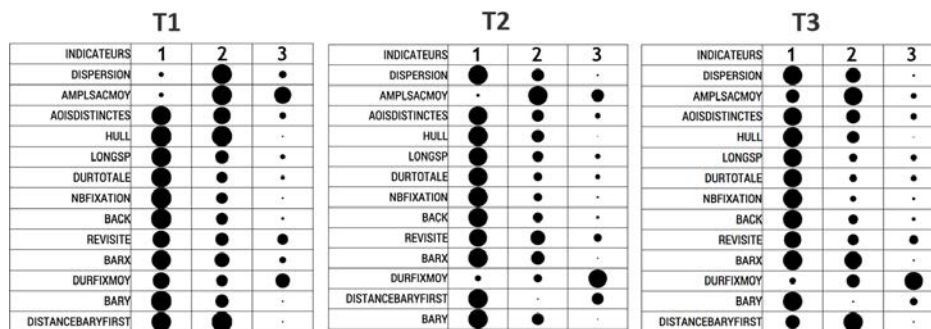


FIGURE 7.3 – Comparaison des patterns sur les 3 temps sur le site Web du Musée Carnavalet.

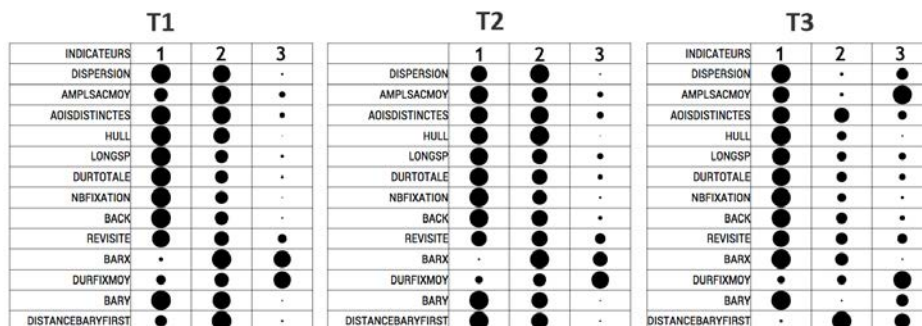


FIGURE 7.4 – Comparaison des patterns sur les 3 temps sur le site Web de France Diplomatie.

7.5.2 Stabilité des patterns dans le temps

Pour la tâche réalisée sur le site du musée Carnavalet, il semblerait que plus on s'éloigne dans le temps, plus les participants ont tendance à changer de classes. En effet, entre les T1 et T2, la stabilité des patterns est au-dessus du hasard (36%) alors qu'entre T2 et T3 (27%) puis T1 et T3 (24%), la proportion de scanpaths restant dans le même groupe est plus faible que le hasard.

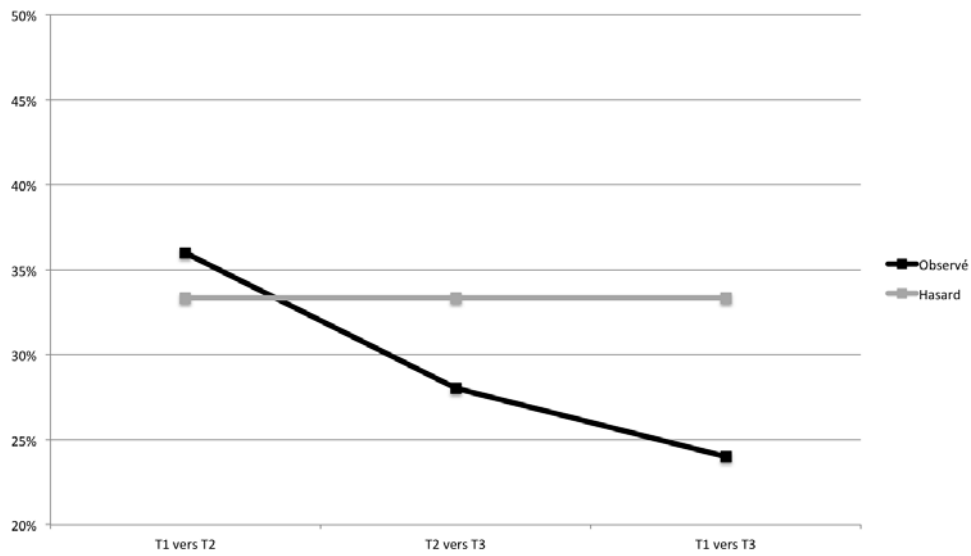


FIGURE 7.5 – Proportion de scanpaths restant dans le même groupe dans le temps.

Pour la tâche réalisée sur le site de France Diplomatie, nous observons un phénomène inverse. Cette fois-ci, les classes semblent devenir plus stables au fur et à mesure du temps : entre T1 et T2 le taux est de 24%, entre T2 et T3 le taux est de 27% et entre T1 et T3 il est de 32%. En revanche, à aucun moment la stabilité ne dépasse le seuil du hasard.

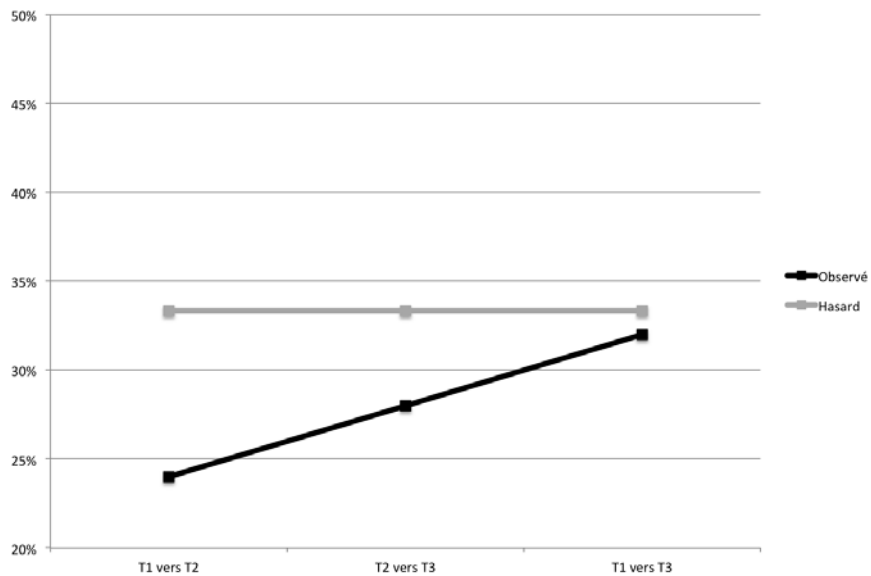


FIGURE 7.6 – Proportion de scanpaths restant dans le même groupe dans le temps.

7.5.3 Stabilité entre les sites Web dans le temps

Nous avons aussi cherché à vérifier si les patterns présentaient une certaine stabilité lors du changement de sites Web et donc de tâche pour chaque temps.

- A T1, la stabilité est proche du hasard (32%).
- A T2, les scanpaths plutôt tendance à rester dans le même groupe (40%).
- A T3, les scanpaths ont tendance à changer de groupe (28%)

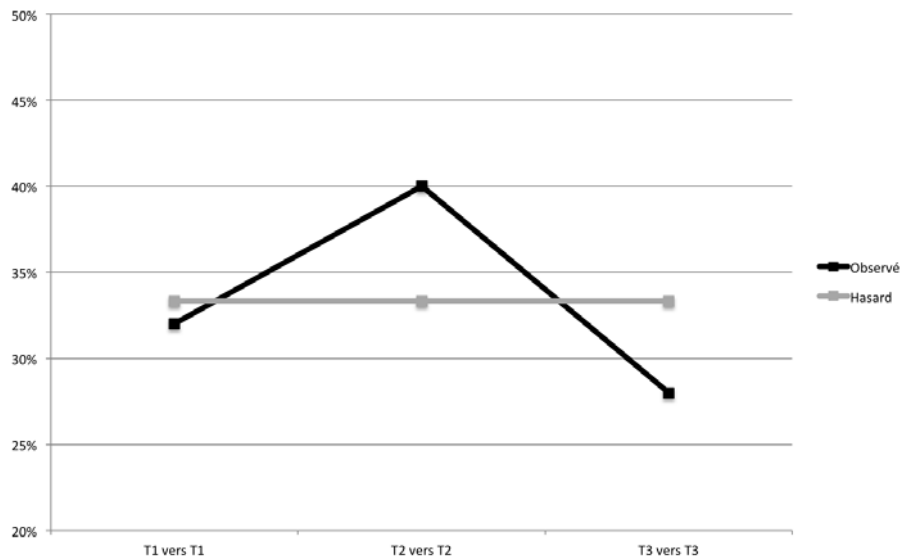


FIGURE 7.7 – Proportion de scanpaths restant dans le même groupe entre les sites Web dans le temps.

7.6 Conclusions

Cette expérimentation avait un double objectif. D'une part, elle nous a permis de vérifier si les patterns identifiés dans l'expérimentation précédente (Chapitre 6) pouvaient être reproduits dans des conditions différentes, et d'autre part de répondre à la question de la stabilité des patterns sur un temps plus long.

Les résultats de l'identification des patterns montrent que ceux-ci sont très similaires entre les deux expériences. En effet, nous avons trouvé à nouveau 3 classes qui ont les mêmes caractéristiques quelque soit le site Web et le temps envisagé. En outre, les participants n'étaient pas les mêmes entre les deux études, et pourtant, nous observons toujours ces 3 mêmes patterns.

Concernant la stabilité des patterns, nous avons observé une tendance au changement de patterns au fur et à mesure du temps, ce qui n'était globalement pas le cas dans l'expérience précédente. A court terme (c'est à dire lorsqu'un participant répète la même tâche immédiatement), les patterns semblent avoir

une certaine tendance à la stabilité (entre 36 et 47%), alors qu'à plus long terme, c'est l'inverse qui se produit (entre 24 et 36%). Dans tous les cas nous n'avons pas observé de tendance marquée à la stabilité ou à l'instabilité.

Pour ces deux expériences, nous avons sélectionné à dessein des participants ayant des profils homogènes : des étudiants d'environ une vingtaine d'années et ayant une bonne expérience du Web. Peut-être que cela explique la similitude des résultats que nous avons obtenus lors de l'identification des patterns. Dans le chapitre suivant, nous tentons de répondre à cette question à l'aide d'une expérimentation où les participants ont des profils variés, et n'ont pas été sélectionnés dans le milieu universitaire.

Chapitre 8

Expérience 3 : influence des caractéristiques des participants sur les patterns

Cette étude a été réalisée pour le compte du Ministère de l'Égalité des Chances Luxembourgeois. Les participants sélectionnés avaient des profils variés et ne faisaient pas partie du milieu universitaire. Ainsi, l'objectif de cette expérimentation était de voir si on peut retrouver les patterns précédemment identifiés (Chapitres 6 et 7) sur des participants moins homogènes.

8.1 Matériel

Nous avons utilisé le système d'oculométrie déporté Tobii T120 (réglé sur une fréquence d'échantillonnage de 60 Hz) avec une résolution d'écran de 1280x1024 pixels. Le logiciel Tobii Studio (v.3.0) a été utilisé pour enregistrer puis exporter les données oculométriques. La durée minimale d'une fixation a été fixée à 100 ms dans un rayon de 35 pixels. Le dispositif expérimental était placé sur un bureau. Durant les passations, les participants s'asseyaient face à l'écran et avaient uniquement le droit d'utiliser la souris, le clavier était sous contrôle de l'expérimentateur pour éviter de perdre des données lors des frappes sur le clavier. La distance entre le participant et l'écran était de 60 cm.

8.2 Stimulus

Pour réaliser cette étude nous avons choisi un site Web du gouvernement Luxembourgeois . Les participants ont parcouru l'ensemble du site mais seule la page d'accueil a été analysée.

Page d'accueil du site Web (Ministère de l'Égalité des Chances) :



FIGURE 8.1 – Page d'accueil du site du Ministère de l'Égalité des Chances Luxembourgeois.

8.3 Echantillon

47 participants ont été recrutés dans le cadre d'une étude pour le Ministère de l'Égalité des Chances Luxembourgeois. 7 participants n'ont pas fait l'objet d'une analyse en raison de problèmes de calibration de l'oculomètre.

Age : les participants ont entre 21 et 60 ans. 22% ont entre 21 et 30 ans, 43% entre 31 et 40 ans, 20% entre 41 et 50 ans et 15% entre 51 et 60 ans.

Genre : 23 femmes (57,5%) et 17 hommes (42,5%).

Niveau d'étude : 25% ont un niveau inférieur au Bac, 10% ont le niveau Bac, 12% ont un niveau Bac+1 à Bac+2, 23% Bac+3 à Bac+4, 30% ont au moins un Bac+5.

Usage d'internet : 100% déclarent utiliser Internet depuis plus de 6 ans et tous l'utilisent dans le cadre de leur travail.

Compétences liées à Internet : 10% jugent leurs compétences passables, 25% assez bonnes, 27% bonnes, 25% très bonnes et 13% excellentes.

8.4 Procédure

L'expérience fut menée dans une salle sans fenêtres dédiée à l'expérimentation dans les locaux du Ministère. Nous informions les participants qu'ils réaliseraient des tâches de recherche d'information sur le Web et que leurs performances n'étaient pas mesurées. Aucune limite de temps ne leur était imposée.

Les tâches à réaliser, dans un ordre aléatoire sans remise, étaient les suivantes (depuis l'accueil du site www.mega.public.lu) :

1. « On fait souvent la différence entre sexe et genre. A partir de la page d'accueil du portail MEGA qui va s'afficher, trouvez la page qui explique ces deux termes.»
2. « À partir de la page d'accueil du portail MEGA qui va s'afficher, trouvez la page contenant les mesures qui peuvent être mises en place dans les communes pour l'égalité des chances.»
3. « À partir de la page d'accueil du portail MEGA qui va s'afficher, trouvez la page du programme pour créatrices d'entreprises.»
4. « À partir de la page d'accueil du portail MEGA qui va s'afficher, trouvez les documents pour une demande de patronage ou de subventions auprès du Ministère de l'Égalité des chances.»
5. « À partir de la page d'accueil du portail MEGA qui va s'afficher, trouvez le lien permettant d'accéder aux statistiques sur l'écart salarial entre hommes et femmes au Luxembourg»

La consigne apparaissait à l'écran, puis nous demandions aux participants de la lire à voix haute et nous nous assurons de sa bonne compréhension avant de commencer. Une tâche prenait fin lorsque :

- le participant donnait la bonne réponse ;
- le participant donnait une mauvaise réponse ;
- le participant décidait d'abandonner.

8.5 Résultats

8.5.1 Identification des patterns

Quatre patterns ont été identifiés pour toutes les tâches (sessions). Les dendrogrammes et les données chiffrées figurent en Annexe C.

- La première classe se caractérise par des scanpaths longs et étendus sur les pages Web et de courtes durées de fixations.
- La deuxième classe comporte des scanpaths moins exhaustifs mais ayant une amplitude des saccades élevée.
- La troisième classe contient des scanpaths courts, avec des saccades de faible amplitude et des durées de fixations courtes.
- La quatrième classe se caractérise par une faible couverture de la page Web, mais avec des durées de fixations élevées.

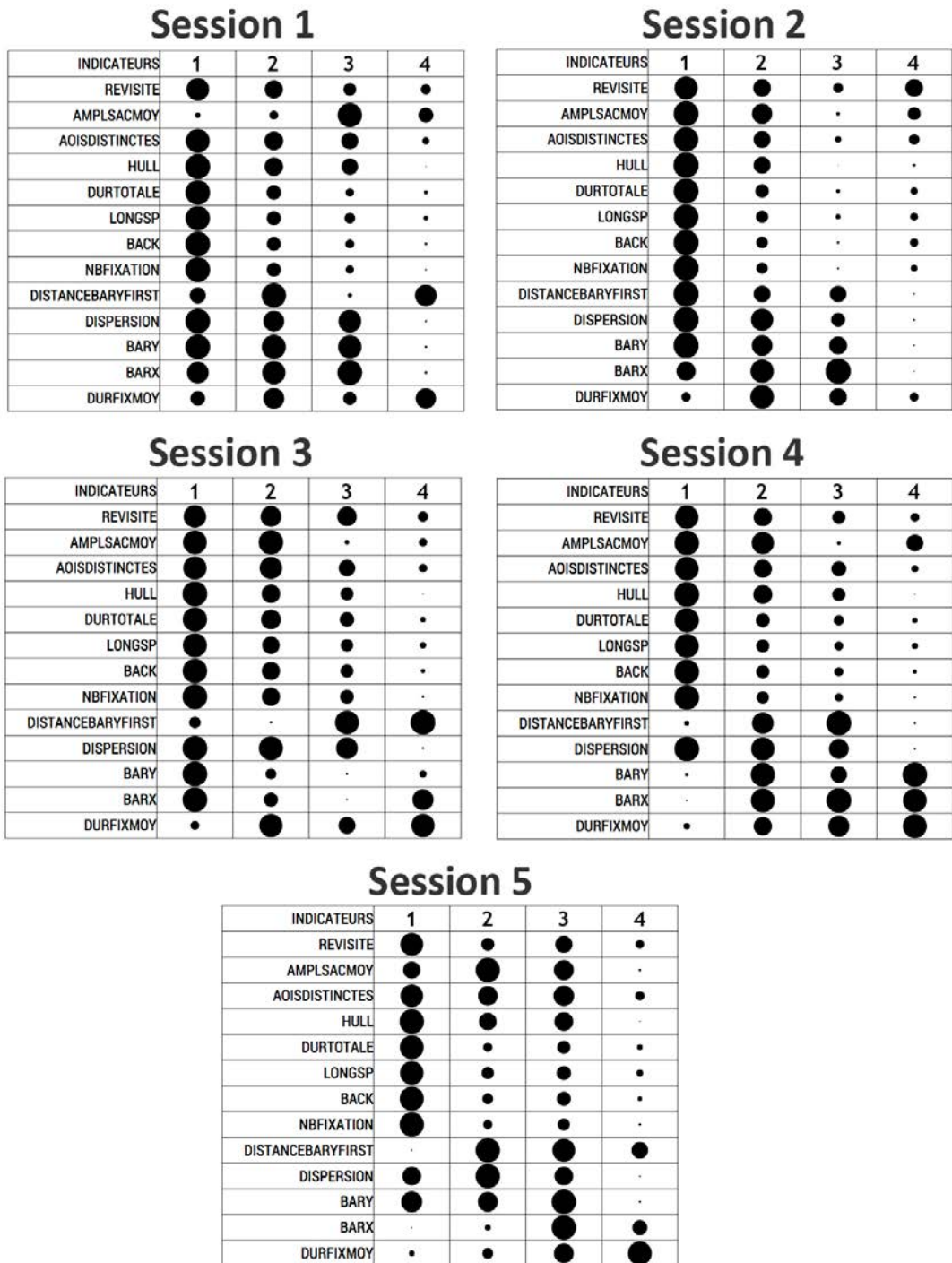


FIGURE 8.2 – Résultats de toutes les sessions de l'expérimentation.

8.5.2 Etude de la stabilité

Les participants ayant réalisé les tâches dans un ordre aléatoire, nous ne présentons pas l'évolution de la stabilité. La figure 8.3 présente la stabilité entre chaque couples de sessions. De façon générale, la stabilité varie de 22,5% à 40%. La stabilité moyenne est de 30% (légèrement au-dessus du hasard).

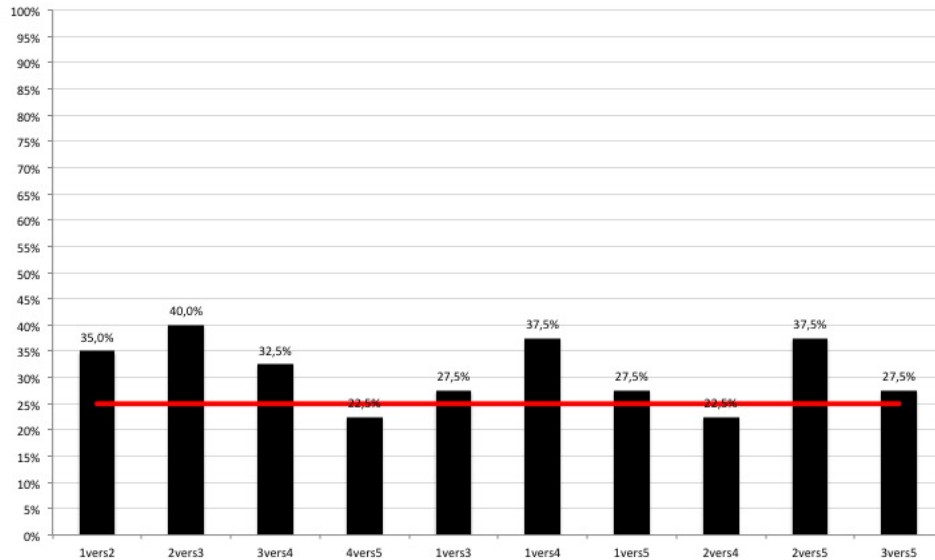


FIGURE 8.3 – Stabilité des patterns entre les sessions de l'expérimentation (la ligne rouge représente le hasard).

8.6 Conclusions

Dans les expérimentations précédentes, nous avons identifié systématiquement 3 patterns de comportements visuels similaires sur des échantillons de participants homogènes et issus du milieu universitaire. Dans cette expérience, où les participants avaient des profils plus variés, 4 patterns sont observés. Il semblerait que le profil des participants ait une certaine influence sur le nombre de patterns. Cependant, ces résultats sont à prendre avec précaution car d'autres facteurs ont pu influencer ces résultats (site Web, difficulté des tâches...).

Concernant la stabilité, à nouveau, elle s'est révélée faible (30% en moyenne). Bien que les profils aient été plus variés dans cette étude, les patterns ne sont pas stables, ce qui laisse envisager que les participants adoptent un style de visualisation ou un autre en fonction du contexte et non de leurs caractéristiques intrinsèques.

Chapitre 9

Expérience 4 : influence du design sur les patterns

9.1 Stimuli

Pour réaliser cette étude nous avons choisi deux sites web gouvernementaux qui étaient modifiés dynamiquement. Pour cela, nous avons créé un serveur *proxy* qui modifiait à la volée la structure des pages selon la condition expérimentale. Ainsi, les deux sites étaient présentés en deux versions aux participants.

Pour les deux sites, nous avons inversé :

- Le logo (initialement à gauche) et le moteur de recherche (initialement à droite)
- Les colonnes de gauche et de droite



FIGURE 9.1 – Eléments du design ayant été modifiés.

Page d'accueil du de l'Education Nationale :



FIGURE 9.2 – Page d'accueil du site de l'Education Nationale (originale à gauche et modifiée à droite)

Page d'accueil du site du Service Public :



FIGURE 9.3 – Page d’accueil du site du Service public en deux version (originale à gauche, modifiée à droite)

9.2 Matériel

Nous avons utilisé le système d’oculométrie déporté Tobii T120 (réglé sur une fréquence d’échantillonnage de 60 Hz) avec une résolution d’écran de 1280x1024 pixels. Le logiciel Tobii Studio (v.2.3.2.0) a été utilisé pour enregistrer puis exporter les données oculométriques. La durée minimale d’une fixation a été fixée à 100 ms dans un rayon de 35 pixels. Le dispositif expérimental était placé sur un bureau. Durant les passations, les participants s’asseyaient face à l’écran et avaient uniquement le droit d’utiliser la souris, le clavier était sous contrôle de l’expérimentateur pour éviter de perdre des données lors des frappes sur le clavier. La distance entre le participant et l’écran était de 60 cm.

9.3 Procédure

L’expérience fut menée en laboratoire. Nous informions les participants qu’ils réaliseraient quatre tâches de recherche d’information sur le web et que leurs performances n’étaient pas mesurées. Aucune limite de temps ne leur était imposée.

L’expérimentation comprenait 4 parties, et chaque participant devait en passer 2 en fonction du groupe dans lequel il se trouvait. Ainsi un sujet, se retrouvait face à 2 versions originales de deux sites distincts ainsi que face à ces 2 mêmes versions mais modifiées. Nous contrebalançons les versions d’un même site puis l’ordre des sites (Tableau 9.1). Toutefois, nous ne contrebalançons

pas l'interaction tâche/version. Chaque tâche était précédée d'une phase de calibration et de l'effacement de l'historique de navigation d'Internet Explorer (v. 7).

Les consignes étaient les suivantes :

1. « A l'aide du site Internet qui va s'afficher, répondez à la question suivante : Trouvez la date du début des vacances scolaires en Lorraine. Quand vous aurez trouvé la réponse, citez-la à haute voix » (page de départ : www.education.gouv.fr). Cette tâche était réalisée deux fois de suite sur les deux versions du site.
2. « A l'aide du site Internet qui va s'afficher, répondez à la question suivante : Quels sont les trois cas permettant de dissoudre un PACS? Quand vous aurez trouvé la réponse, citez-la à haute voix » (page de départ : www.service-public.fr). Cette tâche était réalisée deux fois de suite sur les deux versions du site.

Les participants cherchaient ensuite les réponses sur les différents sites et la tâche prenait fin lorsque le participant citait la bonne réponse à voix haute, donnait une mauvaise réponse à voix haute ou alors lorsqu'il décidait d'abandonner.

Groupe	Essai 1	Essai 2	Essai 3	Essai 4
	Site 1		Site 2	
1	V1	V2	V1	V2
2	V2	V1	V2	V1
	Site 2		Site 1	
3	V1	V2	V1	V2
4	V2	V1	V2	V1

TABLE 9.1 – Récapitulatif de la procédure expérimentale.

A la fin des passations, nous faisons passer un questionnaire aux participants afin de contrôler l'homogénéité de notre échantillon et de vérifier s'il n'y avait pas de biais. Le questionnaire était découpé en trois parties :

La fiche signalétique

Nous avons recueilli les données démographiques ainsi que des informations relatives aux problèmes ophtalmologiques.

L'utilisation d'Internet

Nous demandions aux participants d'évaluer leur niveau en informatique et nous recueillions des informations liées à leur usage du web (nombres d'années d'utilisation, temps passé par jour et leur(s) principale(s) activité(s) sur le Web). Cette rubrique comprenait également deux questions contrôle : une première pour savoir s'ils étaient habitués aux expérimentations sur le web et une seconde nous permettant de vérifier s'ils connaissaient l'objectif de notre étude. La connaissance antérieure des sites Web parcourus et le degré de difficulté perçu fut également évalué.

Le niveau d'expertise avec Internet

Cette rubrique reprenait 7 items du questionnaire de Lim et Lee ([Lim & Lee, 2000]). Les sujets devaient répondre à des affirmations en se positionnant en termes d'expertise.

9.4 Participants

78 participants ont réalisé l'expérience.

Age moyen (écart type) : 21.60 (2.18).

Genre : 48,7% de femmes et 51,3% d'hommes.

Nombre d'années d'utilisation d'internet : + de 3 ans pour l'ensemble de l'échantillon.

Temps passé par jour sur Internet : 98.7% déclarent passer au moins 1H / jour sur internet.

9.5 Résultats

Les dendrogrammes et les données chiffrées de cette expérience figurent en Annexe D.

9.5.1 Education Nationale

Dans la première version (V1) du site Web, 3 classes ont été identifiées.

- La première classe se caractérise par une exploration exhaustive de la page Web, avec de fréquentes revisites d'AOIs avec de nombreuses fixations et une durée totale du scanpath élevée. L'amplitude des saccades est plus faible que dans les autres classes.
- La deuxième classe est intermédiaire. Les scanpaths qui la compose sont moins exhaustifs mais l'amplitude des saccades est élevée.
- La troisième classe comporte les scanpaths les moins exhaustifs et longs, mais avec des durées de fixations plus élevées que dans les autres classes.

Dans la deuxième version (V2) du site Web, 3 classes ont été à nouveau identifiées.

- La première classe est similaire à celle de la V1, à l'exception de la distance entre le barycentre et la première fixation. Les scanpaths de cette classe ont aussi eu moins tendance à visiter la droite de la page.
- La deuxième classe est similaire à celle de la V1, mais elle est plus dispersée et les scanpaths ont eu plus tendance à descendre dans la page.
- La troisième classe ressemble aussi beaucoup à celle de la V1, avec une distance entre le barycentre et la première fixation plus élevée et une plus grande tendance à aller sur la droite de la page.

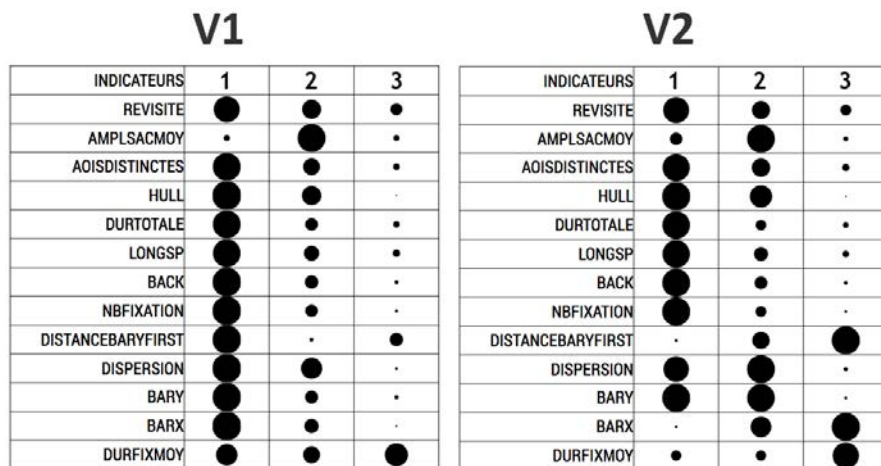


FIGURE 9.4 – Comparaison des patterns sur les 2 versions du design du site de l’Education Nationale.

Le taux de stabilité des patterns entre la V1 et la V2 est de 30% (légèrement en-dessous du hasard).

9.5.2 Service Public

Dans la première version du site Web, 3 classes ont été identifiées :

- La première classe se caractérise par une exploration exhaustive de la page Web, avec des saccades de grande amplitude, des scanpaths longs mais des durées de fixations courtes avec une tendance forte à revisiter les mêmes *AOIs*.
- La deuxième classe est intermédiaire entre les classes 1 et 3. Elle se caractérise par une plus grande distance entre le barycentre et la première fixation.
- La troisième classe contient les scanpaths les moins exhaustifs, avec le moins de fixations, mais avec une durée des fixations plus élevée.

Dans la deuxième version du site Web, 2 classes ont été identifiées :

- La première classe comporte les scanpaths les plus exhaustifs, et de plus longue durée avec une amplitude des saccades plus courte.
- La seconde classe se caractérise par des scanpaths courts, avec peu de fixations, mais avec une amplitude des saccades plus élevée et des durées de fixations plus longues.

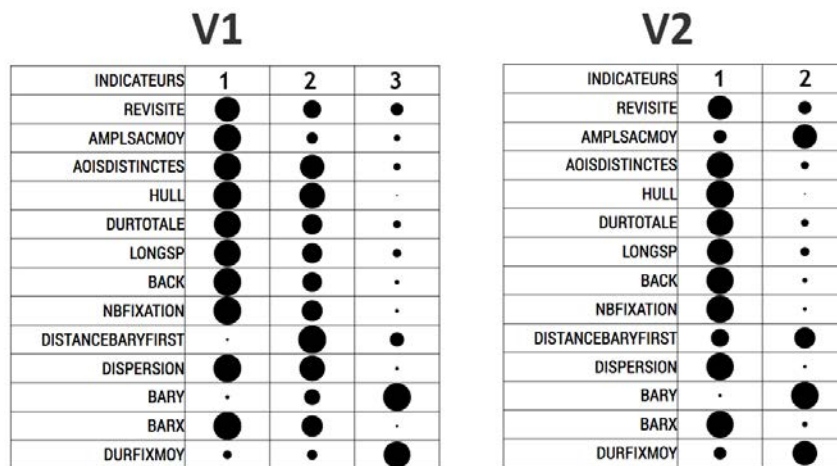


FIGURE 9.5 – Comparaison des patterns sur les 2 versions du design du site du Service Public.

Le taux de stabilité n'a pas pu être calculé puisque le nombre de classes diffère selon les versions du site.

9.6 Conclusions

Pour le site de l'Education Nationale, la modification de la page Web semble avoir un effet limité. En effet, 3 classes ont été identifiées sur les deux versions, et celles-ci ont peu de différences sur les indicateurs oculométriques. En revanche, sur le site du Service Public, la deuxième version du site Web a engendré la perte de la classe 2, qui est habituellement une classe intermédiaire entre les classes 1 et 3.

Chapitre 10

Discussion générale

10.1 Synthèse de la démarche

Lors de cette thèse, nous avons identifié des travaux qui mettaient en exergue des patterns de comportements visuels sur le Web. L'étude la plus populaire, qui présente le «F-Shape» comme un pattern «universel» sur le Web, en se basant sur l'agrégation de l'ensemble des données de fixations des participants sous forme de heatmaps, laisse supposer *in fine* que les individus explorent visuellement plus ou moins tous de la même façon les pages Web, et ce quelque soit leur contexte de navigation (tâche, page Web...) ou leurs caractéristiques propres (démographiques, expérience du Web...).

Nous le voyons, ce type d'approche a des implications épistémologiques. En effet, cela pré-suppose notamment que des styles d'exploration visuelle ne peuvent être détectés et donc que les internautes sont homogènes entre eux.

D'un point de vue méthodologique, nous avons constaté que les méthodes d'analyse des données oculométriques existantes n'étaient pas conçues pour détecter des typologies de patterns, ce qui peut expliquer pourquoi un pattern unique est identifié. C'est pour cette raison que l'apport de la présente thèse est principalement méthodologique.

D'un point de vue pratique, il est commode de n'avoir qu'un seul pattern de visualisation à prendre en compte. Néanmoins, si justement il y a des styles ou typologies de patterns différents selon la tâche, la page Web ou encore les caractéristiques des individus, ce pattern unique risque de ne correspondre à personne. De plus, l'agencement des pages Web qui en découle risque de ne pas être adapté aux utilisateurs pour lesquels ces dernières sont destinées.

Finalement, nous avons voulu savoir comment faire pour identifier des typologies de patterns comportementaux visuels, et tenter d'en identifier selon diverses conditions expérimentales. Dans une première partie, nous avons appliqué et combiné des méthodes existantes afin de voir si des patterns pouvaient être

identifiés. Les résultats obtenus ont montrés des limites importantes qui nous ont servi ensuite de base pour proposer une méthode plus adaptée. Enfin, nous avons mené une série d'expérimentations destinées d'une part à appliquer la méthode que nous avons proposée sur le Web, et d'autre part à tester dans quelles conditions des patterns peuvent être détectés ainsi que leur stabilité.

Dans un premier temps, nous discuterons des apports méthodologiques de cette thèse pour enfin se centrer sur la question des patterns de comportements visuels.

10.2 Discussion sur les apports méthodologiques

10.2.1 La définition des AOIs

Nous avons montré que la forme et la granularité des *AOIs* ont une influence directe sur les résultats d'une analyse et peuvent même conduire à des conclusions opposées. Pour répondre à ce problème, nous avons choisi une approche *bottom-up* qui consiste à détecter les *AOIs* directement à partir des données de fixation.

Néanmoins, la procédure *Mean-Shift* utilisée pour définir *a priori* les *AOIs* n'avait pas fait l'objet d'une validation systématique. Les rares travaux l'ayant utilisé proposaient de fixer la taille de la fenêtre du *kernel* en relation avec l'étendue de la vision fovéale. Nous avons montré par une expérimentation que cette taille théorique n'est en fait pas adaptée et qu'elle est plutôt relative à la structure du stimuli présenté aux participants. Finalement, le paramètre théorique semble occulter la vision parafovéale. Par exemple, si un élément unique d'une page Web est plus grand que l'étendue du champ de la vision fovéale, l'oeil ne percevra pas clairement ses contours, et sera obligé de réaliser plusieurs fixations pour en avoir une image nette. Cela ne signifie pas pour autant que cet élément est composé de plusieurs *AOIs*.

Une autre limite de la procédure *Mean-Shift*, qui n'a pourtant pas été, à notre connaissance, discutée dans la littérature, est la présence de «bruit» dans les données. Il arrive fréquemment que des fixations visuelles soient émises en dehors de tout élément d'une page Web ou encore qu'il y ait des fixations intermédiaires entre plusieurs éléments. Ces fixations vont pourtant être identifiées comme *AOIs* alors qu'elles n'en sont pas. Nous avons constaté, durant nos expérimentations, que certains participants utilisent le pointeur de la souris pour scroller, et qu'ils ont tendance à fixer la barre de scroll de façon intermittente lorsqu'ils réalisent cette action. Ce comportement génère différents effets de bord : une ou plusieurs «*AOI(s)*» sur la barre de scroll et des fixations intermédiaires entre la page Web et cette barre de scroll. Combien d'*AOIs* supplémentaires sont détectées durant la procédure *Mean-Shift* du fait de ce type de comportement ?

Pour ces raisons, nous avons développé une version de la procédure *Mean-Shift* (*HMS-ET*) qui a le double avantage de définir à partir des données la

taille de la fenêtre du *kernel* et d'identifier les *outliers*. Nous avons montré sa supériorité dans diverses situations expérimentales.

Maintenant se pose la question de son usage dans d'autres types de recherche que la présente thèse. De nombreux travaux utilisent les *AOIs a priori* (c'est à dire définies par le chercheur) car des hypothèses sont émises, par exemple, sur l'usage d'un élément d'une page Web en fonction de différentes variables. Dans ce type de cas, le *HMS-ET* pourrait être difficile à utiliser, car il est difficile d'associer un sens à une *AOI* définie à partir des données. Une solution intermédiaire pourrait consister à utiliser les *AOIs* identifiées automatiquement pour ensuite dessiner les *AOIs* utiles aux hypothèses formulées. Des travaux additionnels sont nécessaires pour évaluer la pertinence de cette approche.

10.2.2 L'identification des patterns

La question de la méthode d'analyse des données oculométrique pour identifier des typologies de patterns est loin d'être triviale. En effet, il suffit de choisir une méthode de comparaison des scanpaths, puis de les classifier, et des classes sont la plupart du temps obtenues. En effet, dans la Partie II, nous avons appliqué un ensemble de méthodes qui ont toutes produites des résultats. Pourtant, nous l'avons vu, ces résultats différaient selon les méthodes et nous n'avons pas pu savoir laquelle choisir, étant donné le manque de critères fiables de validation. Cela montre, s'il en était besoin, qu'il est nécessaire de valider une démarche méthodologique d'analyse de données avant de conclure sur des résultats.

Cela pose aussi la question des données prises en compte en entrée de la classification, et plus généralement les données prises en compte pour toute analyse oculométrique. Dans certains cas, il est opportun de se limiter, par exemple, à la durée des fixations, si une hypothèse porte sur ce point précis. En revanche, il y a bien des cas où des différences statistiquement significatives peuvent être trouvées, simplement en sélectionnant l'une ou l'autre dimension des scanpaths. Ainsi, quelles données et indicateurs prendre en compte pour comparer des populations d'individus? Cela est valable aussi pour l'identification de patterns, le *F-Shape*, par exemple, ne concerne que le nombre de fixations : quelle forme serait apparue si les auteurs avaient choisi la durée des fixations au lieu de leur fréquence? Nous avons résolu en partie ce problème en validant notre méthode (*SE-ET*) sur une expérimentation qui consistait à provoquer des patterns de visualisation. Ensuite, nous avons appliqué strictement la même méthode sur plusieurs expérimentations, ce qui nous a permis de reproduire les résultats. Mais comment être certains que les classes obtenues ne sont pas de simples artefacts, malgré les précautions méthodologiques, et en l'absence de théorie sur le sujet? Il reste encore à expliquer pour quelles raisons ces classes existent.

10.2.3 La caractérisation des patterns

La dimension temporelle, ainsi que la dynamique de l'exploration visuelle ont été prises en compte dans nos analyses et donc dans les classes qui ont été identifiées dans les diverses expérimentations. Cependant, ces aspects n'ont pas été pris en compte directement pour la caractérisation des classes et leur visualisation, puisque nous avons choisi de présenter une visualisation d'indicateurs oculométriques via l'application *Bertifier*. En réalité, nous avons fait plusieurs tentatives d'illustrations et d'analyses, notamment de séquences, mais la variabilité de ces dernières est tellement grande que de nouveaux travaux dédiés spécifiquement à ce sujet devraient être menés. La figure 10.1, illustre ce phénomène : en ordonnée, ce sont les identifiants des *AOIs* identifiées dans la première session de l'expérimentation du Chapitre 6 ; en abscisse, le temps (en ms). Chaque point correspond à la visualisation d'une *AOI* par un participant et chacune des 3 couleurs illustre une classe identifiée. On peut voir que la classe en rouge concentre les scanpaths les plus exhaustifs, longs (en durée), avec une tendance à la «revisite» (ce que l'on sait à l'aide des indicateurs), alors que la classe en bleu concentre les scanpaths les plus courts. On peut aussi voir que toutes les *AOIs* ont été visitées par toutes les classes, mais à des moments différents, selon une dynamique différente. Ce type de visualisation peut s'avérer utile mais ne renseigne finalement pas plus que les données des indicateurs.

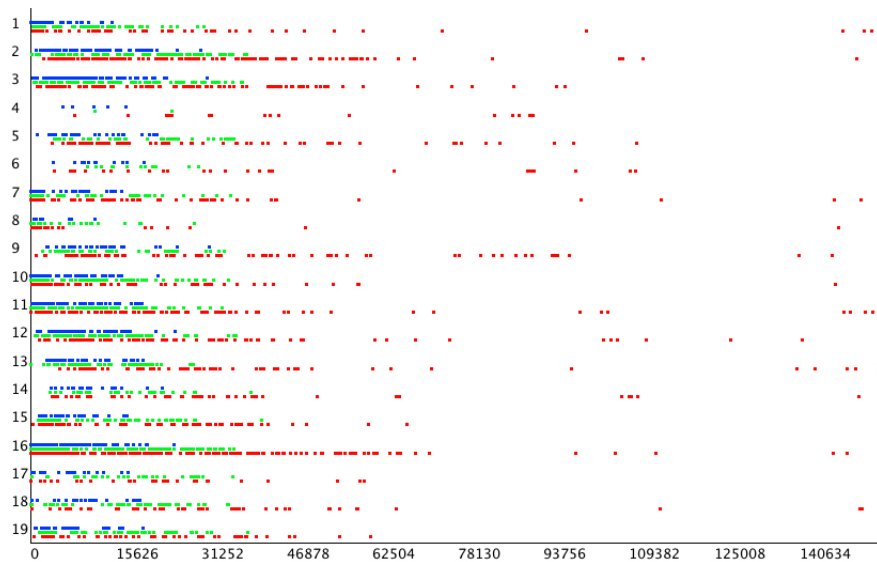


FIGURE 10.1 – Illustration des séquences de visualisation des *AOIs* pour 3 classes de scanpaths.

10.3 Discussion sur les patterns de comportements visuels

10.3.1 A propos des patterns identifiés

Afin d'identifier des patterns de comportements visuels et d'étudier leur stabilité, nous avons mené quatre expérimentations.

Dans la première, les participants devaient réaliser deux tâches sur deux sites Web en les répétant immédiatement. Nous avons identifié 3 patterns distincts pour chacune des tâches, répétées ou non. Nous avons remarqué que les patterns sont à peu près les mêmes quelque soit la tâche ou le site envisagé, et même après la répétition d'une tâche. Nous aurions pu penser que la répétition d'une tâche engendrerait une réduction du nombre de patterns, puisque cette répétition a tendance à faire varier à la baisse les indicateurs oculométriques (moins de fixations, moins d'*AOIs*...) mais ce ne fut pas le cas.

Dans la deuxième expérimentation, la procédure expérimentale était similaire, mais les participants répétaient les tâches à deux jours d'intervalle. Là aussi, nous avons identifié trois patterns de structure proche quel que soit le site ou la répétition de la tâche. Les participants n'étaient pas les mêmes et pourtant les patterns ressemblent beaucoup à ceux de la première expérimentation.

La troisième expérimentation, à la différence des deux précédentes, n'était pas réalisée sur des participants issus du milieu universitaire. Nous avions souhaité reproduire nos analyses sur une population moins homogène. Cette fois-ci, quatre patterns ont été identifiés sur chacune des cinq tâches proposées.

La dernière expérimentation visait à faire varier l'agencement des pages Web présentées aux participants (à nouveau issus du milieu universitaire). Nous avons modifié l'organisation des pages Web afin que les participants réalisent la même tâche pour chaque site Web, mais avec des variations de structure dans un ordre contrebalancé. Globalement, nous avons retrouvé les trois patterns des deux premières expérimentations.

Sauf exception, nous avons identifié sur l'ensemble des tâches et sites des expérimentations au moins 3 patterns distincts.

Le premier pattern se caractérise par une exploration exhaustive des pages Web. Le nombre de fixations, la durée totale des scanpaths, le nombre d'*AOIs* vues sont élevés, la dispersion et le taux revisite des *AOIs* sont aussi élevés. En revanche, la durée des fixations est basse. Il semblerait que ce pattern comporte des scanpaths qui explorent longuement et exhaustivement les pages Web, mais sans s'attarder sur un élément en particulier.

Le deuxième pattern semble être une version atténuée du premier. Relativement au premier, il comporte moins de fixations, des scanpaths moins longs et une plus faible tendance à la revisite des *AOIs* mais avec une durée des fixations souvent légèrement supérieure. Les scanpaths de ce pattern explorent moins

longuement les pages Web mais peuvent s'attarder sur certains éléments en particulier.

Le troisième pattern est l'inverse des deux autres. Les scanpaths sont courts (en durée et en nombre de fixations), peu dispersés, et les AOIs sont peu revisitée. La durée des fixations est la plus élevée des trois patterns. Il semblerait que ce pattern comporte des scanpaths qui explorent minutieusement un faible nombre d'éléments des pages Web.

Nous avons vu lors de la première expérimentation qu'il n'y a pas de pattern plus performant que les autres. On pourrait penser qu'une exploration exhaustive conduite plus facilement à une bonne performance, ou au contraire que ce type d'exploration reflète une difficulté dans le choix de la cible. On pourrait aussi penser que le troisième pattern est celui emprunté par les individus les plus habitués à explorer les pages Web, et donc les plus performants, mais ce n'est pas le cas. Il semblerait que l'un ou l'autre des patterns soit plus performant selon les circonstances.

Ce constat ne nous permet pas d'émettre des recommandations pour l'apprentissage à la perception des pages Web, ni pour une conception des pages Web qui favoriserait l'adoption d'un pattern plutôt qu'un autre.

Nos résultats ne vont pas dans le sens d'un pattern unique et universel, comme le *F-Shape*, tant en termes d'analyse de forme (cf. chapitre 4.6) que d'analyse portant sur la dynamique des scanpaths. D'une part, nous avons retrouvé la plupart du temps au minimum trois patterns distincts sur chacune des expérimentations, mais en plus, un pattern supplémentaire est apparu lorsque nous avons invité des participants aux profils plus hétérogènes. Nous pensons que de nouvelles recherches sur ce point seraient utiles. En effet, quels caractéristiques des individus sont susceptibles de faire varier les patterns? Néanmoins, il est possible que ce ne soient finalement pas les caractéristiques des individus qui aient provoqué l'apparition d'un pattern supplémentaire, mais peut-être les caractéristiques des tâches (étaient-elles plus difficiles que celles des autres expérimentations?), ou encore des pages Web.

Concernant les pages Web, notre étude sur la modification de leurs caractéristiques sont mitigés. Pour le premier site de l'expérimentation, les patterns n'ont pas changé fondamentalement, alors que pour le deuxième site, le deuxième pattern semble avoir fusionné avec un autre. Il est possible finalement que les patterns «exhaustifs» et «minutieux» soient dominants (comme le suggèrent [Aula et al., 2005]), et que selon le contexte (caractéristiques des individus, difficulté de la tâche, structure des pages Web...), d'autres patterns émergent. Là encore, nous n'avons pas la réponse et de nouveaux travaux seraient nécessaires.

10.3.2 A propos de la stabilité des patterns

Dans chacune de nos expérimentations, nous avons évalué la stabilité des patterns, c'est à dire la propension des individus à adopter le même pattern entre différentes sessions de navigation. Selon les expérimentations, elle était plus ou moins élevée, mais globalement, elle reste assez basse (Figure 10.2).

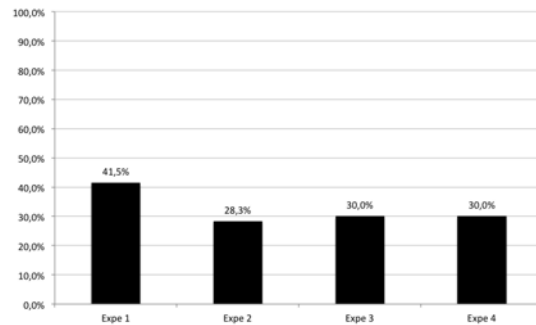


FIGURE 10.2 – Stabilité moyenne des patterns selon les expérimentations de la thèse.

La stabilité des patterns n'avait jamais été étudiée. Dans leur étude sur les résultats de moteurs de recherche, [Aula et al., 2005] rapportent qu'il existe deux styles de visualisation : les «économiques», qui seraient plus efficaces dans une tâche de recherche d'information alors que les «exhaustifs» seraient moins performants. Les auteurs concluent que les «exhaustifs» manquent probablement d'expertise et sont dépendants des résumés des résultats de recherche pour prendre une décision. Ces auteurs font donc le lien entre styles de visualisation et caractéristiques intrinsèques des individus. Pourtant, leur étude n'a pas évalué la stabilité des patterns. Nos résultats montrent que les patterns ne sont en fait pas stables (ou très modérément), un même individu pouvant emprunter des patterns différents entre plusieurs tâches comme pour une même tâche, et ce, quelle que soit la temporalité envisagée. Cette relative absence de stabilité ne plaide pas en faveur d'une explication par des variables intrinsèques aux individus, mais plutôt d'un contexte particulier qui va provoquer l'adoption d'un pattern ou d'un autre. Pour expliquer l'existence des patterns, il conviendrait maintenant de réaliser des expériences qui font varier de façon systématique le contexte des participants.

La faible stabilité des patterns que nous avons observée ne va pas dans le sens de la théorie du Scanpath ([Noton & Stark, 1971b]) en ce qui concerne l'exploration de pages Web. En effet, cette théorie stipule que les individus ont tendance à emprunter le même chemin visuel lorsqu'ils sont soumis de manière répétée à un même stimulus. Nous n'avons pas observé de forte tendance à la stabilité dans la répétition des tâches dans nos expérimentations. Nos résultats vont plutôt dans le sens de Josephson et Holmes ([Josephson & Holmes, 2002]),

qui avaient observé que la similarité des scanpaths inter-participants pouvait être plus élevée que la similarité des scanpaths intra-participant chez les 8 sujets de leur étude.

Bibliographie

- [Aaltonen et al., 1998] Aaltonen, A., Hyrskykari, A., & Rähä, K.-J. (1998). 101 Spots, or How Do Users Read Menus? *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '98*.
- [Aula et al., 2005] Aula, A., Majaranta, P., & Riihinen, K. (2005). Eye-Tracking Reveals the Personal Styles for Search Result Evaluation. *Human-Computer Interaction-INTERACT 2005*, (pp. 1058–1061).
- [Baccino, 2004] Baccino, T. (2004). *La lecture électronique*. Sciences et technologies de la connaissance. Presses universitaires de Grenoble.
- [Bastien et al., 2010] Bastien, J. M. C., Drusch, G., & Dinet, J. (2010). Connaitre les comportements des internautes : méthodes et outils. In L. Calderan, B. Hidoine, & J. Millet (Eds.), *L'utilisateur numérique*. ADBS.
- [Bayles, 2002] Bayles, M. E. (2002). Designing Online Banner Advertisements : Should We Animate? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '02* (pp. 363–366). New York, NY, USA : ACM.
- [Blascheck et al., 2014] Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2014). State-of-the-Art of Visualization for Eye Tracking Data. *Eurographics Conference on Visualization (EuroVis)*.
- [Brandt & Stark, 1997] Brandt, S. & Stark, L. (1997). Spontaneous eye-movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, 9(1), 27–38.
- [Broder, 2002] Broder, A. (2002). A taxonomy of web search. *ACM Sigir Forum*, 36(2), 3–10.
- [Burke et al., 2004] Burke, M., Gorman, N., Nilsen, E., & Hornof, A. (2004). Banner ads hinder visual search and are forgotten. In *Extended abstracts of the 2004 conference on Human factors and computing systems - CHI '04* (pp. 1139). New York, New York, USA : ACM Press.
- [Burke et al., 2005] Burke, M., Hornof, A., Nilsen, E., & Gorman, N. (2005). High-cost banner blindness : Ads increase perceived workload, hinder visual

- search, and are forgotten. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(4), 445.
- [Buscher et al., 2009] Buscher, G., Cutrell, E., & Morris, M. (2009). What do you see when you're surfing?: using eye tracking to predict salient regions of web pages. In *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 21–30). : ACM.
- [Byrne & Anderson, 1999] Byrne, M. & Anderson, J. (1999). Eye tracking the visual search of click-down menus. *Proceedings of the . . .*, (pp. 402–409).
- [Cutrell & Guan, 2007a] Cutrell, E. & Guan, Z. (2007a). An eye-tracking study of information usage in Web search : Variations in target position and contextual snippet length. In *Proc. CHI 2007*.
- [Cutrell & Guan, 2007b] Cutrell, E. & Guan, Z. (2007b). What are you looking for?: an eye-tracking study of information usage in web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 416). : ACM.
- [Dewhurst et al., 2012] Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it : scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior research methods*, 44(4), 1079–100.
- [Djamasbi et al., 2010] Djamasbi, S., Siegel, M., & Tullis, T. (2010). Generation Y, web design, and eye tracking. *International Journal of Human-Computer Studies*, 68(5), 307–323.
- [Djamasbi et al., 2007] Djamasbi, S., Tullis, T., Hsu, J., Mazuera, E., Osberg, K., & Bosch, J. (2007). Gender preferences in web design : usability testing through eye tracking. In *Proceedings of the Thirteenth Americas Conference on Information Systems (AMCIS)* (pp. 1–8).
- [Dong & Lee, 2008] Dong, Y. & Lee, K.-P. (2008). A Cross-Cultural Comparative Study of Users Perceptions of a Webpage : With a Focus on the Cognitive Styles of Chinese , Koreans and Americans Eye Tracking in Usability Testing. *International Journal of Design*, 2(2), 19–30.
- [Drusch et al., 2011] Drusch, G., Bastien, J., & Dinet, J. (2011). From gaze plots to eye fixation patterns using a clustering method based on Hausdorff distances. In *23rd French Speaking Conference on IHM*, number Ea 4432 (pp. 24–27). Sophia Antipolis, France : ACM.
- [Drusch & Bastien, 2012a] Drusch, G. & Bastien, J. M. C. (2012a). Analyzing visual scanpaths on the Web using the mean shift procedure and T-pattern detection. In *Proceedings of the 2012 Conference on Ergonomie et Interaction homme-machine - Ergo'IHM '12* (pp. 181–184). New York, New York, USA : ACM Press.

- [Drusch & Bastien, 2012b] Drusch, G. & Bastien, J. M. C. (2012b). Analyzing Web pages visual scanpaths : between and within tasks variability. *Work (Reading, Mass.)*, 41 Suppl 1, 1559–1566.
- [Drusch et al., 2012] Drusch, G., Bastien, J. M. C., & Boisadan, A. (2012). Évolution des indicateurs oculométriques lors de la répétition de tâches de recherche d'information sur le web : résultats préliminaires. In *ErgoIHM 2012* Biarritz, France.
- [Duchowski, 2007] Duchowski, A. (2007). *Eye tracking methodology : Theory and practice*. Springer-Verlag New York Inc.
- [Duchowski et al., 2010] Duchowski, A. T., Driver, J., Jolaoso, S., Tan, W., Ramey, B. N., & Robbins, A. (2010). Scanpath comparison revisited. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10*, volume 1 (pp. 219). New York, New York, USA : ACM Press.
- [Duncan, 1998] Duncan, J. (1998). Converging levels of analysis in the cognitive neuroscience of visual attention. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 353(1373), 1307–17.
- [Egusa et al., 2008] Egusa, Y., Takaku, M., Terai, H., Saito, H., Kando, N., & Miwa, M. (2008). Visualization of User Eye Movements for Search Result Pages. In *Proceedings of EVIA 2008 (NTCIR-7 Pre-Meeting Workshop)* (pp. 42–46).
- [Ehmke & Wilson, 2007] Ehmke, C. & Wilson, S. (2007). Identifying web usability problems from eye-tracking data. In *Proceedings of the 21st British CHI Group Annual Conference on HCI 2007 : People and Computers XXI : HCI... but not as we know it-Volume 1* (pp. 119–128). : British Computer Society.
- [Feusner & Lukoff, 2008] Feusner, M. & Lukoff, B. (2008). Testing for statistically significant differences between groups of scan patterns. *Proceedings of the 2008 symposium on Eye tracking research & applications - ETRA '08*, (pp.43).
- [Findlay & Walker, 1999] Findlay, J. M. & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *The Behavioral and brain sciences*, 22(4), 661–74; discussion 674–721.
- [Fleetwood & Byrne, 2006] Fleetwood, M. & Byrne, M. (2006). Modeling the Visual Search of Displays : A Revised ACT-R Model of Icon Search Based on Eye-Tracking Data. *Human-Computer Interaction*, 21(2), 153–197.
- [Foulsham et al., 2012] Foulsham, T., Dewhurst, R., & Nyström, M. (2012). Comparing scanpaths during scene encoding and recognition : A multi-dimensional approach. 5(4), 1–14.

- [Fukuda & Bubb, 2003] Fukuda, R. & Bubb, H. (2003). Eye tracking study on Web-use : Comparison between younger and elderly users in case of search task with electronic timetable service. *PsychNology Journal*, 1(3), 202–228.
- [Goldberg, 1999] Goldberg, J. (1999). Computer interface evaluation using eye movements : methods and constructs. *International Journal of Industrial Ergonomics*, 24(6), 631–645.
- [Goldberg et al., 2002] Goldberg, J., Stimson, M., Lewenstein, M., Scott, N., & Wichansky, A. (2002). Eye tracking in web search tasks : design implications. In *Proceedings of the 2002 symposium on Eye tracking research & applications*, number 650 (pp. 51–58). : ACM.
- [Grahame et al., 2004] Grahame, M., Laberge, J., & Scialfa, C. (2004). Age differences in search of Web page : the effects of links size, link number and clutter. *Human factors*, 46(3), 385–398.
- [Granka et al., 2004] Granka, L. a., Joachims, T., & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. In *Proceedings of the 27th annual international conference on Research and development in information retrieval - SIGIR '04* (pp. 478). New York, New York, USA : ACM Press.
- [Groner et al., 1984] Groner, R., Walder, F., & Groner, M. (1984). Looking at Faces : Local and Global Aspects of Scanpaths. In A. G. Gale & F. Johnson (Eds.), *Theoretical and Applied Aspects of Eye Movement Research Selected/Edited Proceedings of The Second European Conference on Eye Movements*, volume 22 of *Advances in Psychology* (pp. 523–533). North-Holland.
- [Guan & Cutrell, 2007] Guan, Z. & Cutrell, E. (2007). An eye tracking study of the effect of target rank on web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 420). : ACM.
- [Habuchi et al., 2006] Habuchi, Y., Takeuchi, H., & Kitajima, M. (2006). The influence of web browsing experience on web-viewing behavior. In *Proceedings of the 2006 symposium on Eye tracking research & applications - ETRA '06*, number March (pp.47). New York, New York, USA : ACM Press.
- [Habuchi et al., 2007] Habuchi, Y., Takeuchi, H., & Kitajima, M. (2007). Where Do Web Visitors Look and at What ? An Eye-Tracking Study. In *HCI International 2007 Posters* (pp. 1084–1088).
- [Halverson & Hornof, 2004] Halverson, T. & Hornof, A. J. (2004). Link colors guide a search. In *Extended abstracts of the 2004 conference on Human factors and computing systems - CHI '04* (pp. 1367). New York, New York, USA : ACM Press.
- [Halverson & Hornof, 2007] Halverson, T. & Hornof, A. J. (2007). A minimal model for predicting visual search in human-computer interaction. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, (pp. 431).

- [Hausdorff, 1962] Hausdorff, F. (1962). *Set theory*. New York, NY, USA.
- [Hembrooke et al., 2006] Hembrooke, H., Feusner, M., & Gay, G. (2006). Averaging scan patterns and what they can tell us. In *Proceedings of the 2006 symposium on Eye tracking research & applications - ETRA '06*, volume 1 (pp.41). New York, New York, USA : ACM Press.
- [Iqbal & Bailey, 2004] Iqbal, S. & Bailey, B. (2004). Using Eye Gaze Patterns to Identify User Tasks. In *The Grace Hopper Celebration of Women in Computing*, volume 3 (pp. 5–10). : Citeseer.
- [Jacob & Karn, 2003] Jacob, R. & Karn, K. (2003). Eye tracking in human-computer interaction and usability research : Ready to deliver the promises. In H. Hyona, J., Radach, R., & Deubel (Ed.), *The minds eye : cognitive and applied aspects of eye movement research*, number 1905 (pp. 573–605). Amsterdam : Elsevier.
- [Jacob, 1991] Jacob, R. J. K. (1991). The Use of Eye Movements in Interaction Techniques : What You Look At is What You Get. *ACM Transactions on Information Systems*, 9(3), 152–169.
- [Jambon & Luengo, 2012] Jambon, F. & Luengo, V. (2012). Analyse oculométrique « on-line » avec zones d'intérêt dynamiques : application aux environnements d'apprentissage sur simulateur. In *Proceedings of the 2012 Conference on Ergonomie Et Interaction Homme-machine, Ergo'IHM '12* (pp. 1 :1—1 :8). New York, NY, USA : ACM.
- [Jamet et al., 2008] Jamet, E., Lemercier, C., & Février, F. (2008). Les methodes offline. In A. Chevalier & A. Tricot (Eds.), *Ergonomie Cognitive des documents électroniques*. PUF, Paris.
- [Jarodzka et al., 2010] Jarodzka, H., Holmqvist, K., & Nyström, M. (2010). A vector-based, multidimensional scanpath similarity measure. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, volume 1 (pp. 211–218). : ACM.
- [Josephson & Holmes, 2002] Josephson, S. & Holmes, M. E. (2002). Attention to repeated images on the World-Wide Web : another look at scanpath theory. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc*, 34(4), 539–48.
- [Just & Carpenter, 1980] Just, M. & Carpenter, P. (1980). A theory of reading : from eye fixations to comprehension. *Psychological review*, 87(4), 329–354.
- [Kaufman & Rousseeuw, 1990] Kaufman, L. & Rousseeuw, P. (1990). *Finding groups in data an introduction to cluster analysis*.
- [Kovács et al., 2005] Kovács, F., Legány, C., & Babos, A. (2005). Cluster validity measurement techniques. *6th international symposium of Hungarian . . .*

- [Kübler et al., 2014] Kübler, T. C., Kasneci, E., & Rosenstiel, W. (2014). SubMatch : scanpath similarity in dynamic scenes based on subsequence frequencies. *ETRA '14 Proceedings of the Symposium on Eye Tracking Research and Applications*, (pp. Pages 319—322).
- [Langfelder et al., 2008] Langfelder, P., Zhang, B., & Horvath, S. (2008). Defining clusters from a hierarchical cluster tree : The Dynamic Tree Cut package for R. *Bioinformatics*, 24, 719–720.
- [Lim & Lee, 2000] Lim, K. F. & Lee, J. (2000). IT Skills of university undergraduate students enrolled in a first year unit. *Australian Journal of Educational Technology*, 16(3), 215 :238.
- [Lorigo et al., 2008a] Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F., & Pan, B. (2008a). Eye tracking and online search : Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*, 59(7), 1041–1052.
- [Lorigo et al., 2008b] Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F., & Pan, B. (2008b). Eye tracking and online search : Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*, 59(7), 1041–1052.
- [Lorigo et al., 2006] Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., & Gay, G. (2006). The influence of task and gender on search and evaluation behavior using Google. *Information Processing & Management*, 42(4), 1123–1131.
- [Magnusson, 1996] Magnusson, M. S. (1996). Hidden real-time patterns in intra- and inter-individual behavior : description and detection. *European Journal of Psychological Assessment*, 12(2).
- [Magnusson, 2000] Magnusson, M. S. (2000). Discovering hidden time patterns in behavior : T-patterns and their detection. *Behavior Research Methods, Instruments, & Computers*, 32(1), 93–110.
- [Mast & Burmester, 2011] Mast, M. & Burmester, M. (2011). Exposing repetitive scanning in eye movement sequences with t-pattern detection. In *Proceedings IADIS International conference . . .* (pp. 137–145).
- [Matin, 1974] Matin, E. (1974). Saccadic suppression : A review and an analysis. *Psychological bulletin*, 81(12), 899–917.
- [Nakamichi et al., 2006] Nakamichi, N., Shima, K., Sakai, M., & Matsumoto, K.-i. (2006). Detecting low usability web pages using quantitative data of users' behavior. *Proceeding of the 28th international conference on Software engineering - ICSE '06*, (pp. 569).
- [Nguyen et al., 2002] Nguyen, K., Wagner, C., Koons, D., & Flickner, M. (2002). Differences in the Infrared Bright Pupil Response of Human Eyes. In *Pro-*

- ceedings of the 2002 Symposium on Eye Tracking Research & Applications*, ETRA '02 (pp. 133–138). New York, NY, USA : ACM.
- [Nielsen, 2006] Nielsen, J. (2006). F-Shaped Pattern For Reading Web Content.
- [Noton & Stark, 1971a] Noton, D. & Stark, L. (1971a). Scanpaths in eye movements during pattern perception. *Science*, 171.
- [Noton & Stark, 1971b] Noton, D. & Stark, L. (1971b). Scanpaths in Eye Movements during Pattern Perception. *Science*, 171(3968), 308–311.
- [Pan et al., 2004] Pan, B., Hembrooke, H., Gay, G., Granka, L., Feusner, M., & Newman, J. (2004). The determinants of web page viewing behavior : an eye-tracking study. In *Proceedings of the 2004 symposium on Eye tracking research & applications*, volume 1 (pp. 154). : ACM.
- [Pan et al., 2007] Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust : Users’ decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, 12(3), 801–823.
- [Pellacini et al., 2006] Pellacini, F., Lorigo, L., & Gay, G. (2006). *Visualizing Paths in Context*. Technical report, Citeseer, Dartmouth.
- [Perin et al., 2014] Perin, C., Dragicevic, P., & Fekete, J.-d. (2014). Revisiting Bertin Matrices : New Interactions for Crafting Tabular Visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 20, 2082–2091.
- [Philippeau, 2008] Philippeau, X. (2008). Filtre Mean Shift.
- [Poole & Ball, 2005] Poole, A. & Ball, L. (2005). Eye tracking in human-computer interaction and usability research : current status and future prospects. *Encyclopedia of human computer interaction*, (pp. 211–219).
- [Poole et al., 2005] Poole, A., Ball, L., & Phillips, P. (2005). In search of salience : A response-time and eye-movement analysis of bookmark recognition. In *Proceedings of HCI 2004* (pp. 363–378).
- [Pretorius et al., 2005] Pretorius, M., Calitz, A., & van Greunen, D. (2005). The added value of eye tracking in the usability evaluation of a network management tool. In *Proceedings of SAICSIT 2005* (pp. 1–10). White River, South Africa : South African Institute for Computer Scientists and Information Technologists.
- [Privitera & Stark, 2002] Privitera, C. & Stark, L. (2002). Algorithms for defining visual regions-of-interest : Comparison with eye fixations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(9), 970–982.
- [Quinlan, 2003] Quinlan, P. T. (2003). Visual feature integration theory : Past, present, and future. *Psychological bulletin*, 129(5), 643–673.
- [Rayner, 1998] Rayner, K. (1998). Eye movements in reading and information processing : 20 years of research. *Psychological bulletin*, 124(3), 372–422.

- [Rele & Duchowski, 2005] Rele, R. & Duchowski, A. (2005). Using eye tracking to evaluate alternative search results interfaces. In *Human Factors and Ergonomics Society Annual Meeting Proceedings*, volume 49 (pp. 1459–1463). : Human Factors and Ergonomics Society.
- [Rodden et al., 2008] Rodden, K., Fu, X., Aula, A., & Spiro, I. (2008). Eye-mouse coordination patterns on web search results pages. In *Proceeding of the twenty-sixth annual CHI conference extended abstracts on Human factors in computing systems - CHI '08* (pp. 2997). New York, New York, USA : ACM Press.
- [Rudmann et al., 2003] Rudmann, D. S., McConkie, G. W., & Zheng, X. S. (2003). Eyetracking in cognitive state detection for HCI. *Proceedings of the 5th international conference on Multimodal interfaces - ICMI '03*, (pp. 159).
- [Salvucci & Goldberg, 2000] Salvucci, D. & Goldberg, J. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications* (pp. 71–78). New York, New York, USA : ACM.
- [Santella & DeCarlo, 2004] Santella, A. & DeCarlo, D. (2004). Robust clustering of eye movement recordings for quantification of visual interest. In *Proceedings of the Eye tracking research & applications symposium on Eye tracking research & applications - ETRA'2004* (pp. 27–34). New York, New York, USA : ACM Press.
- [Schnipke & Todd, 2000] Schnipke, S. K. & Todd, M. W. (2000). Trials and tribulations of using an eye-tracking system. *CHI '00 extended abstracts on Human factors in computer systems - CHI '00*, (pp. 273).
- [Shimodaira, 2002] Shimodaira, H. (2002). An Approximately Unbiased Test of Phylogenetic Tree Selection. *Systematic Biology*, 51(3), 492–508.
- [Shimodaira, 2004] Shimodaira, H. (2004). Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *Annals of Statistics*, 32(6), 2616–2641.
- [Suzuki & Shimodaira, 2004] Suzuki, R. & Shimodaira, H. (2004). An application of multiscale bootstrap resampling to hierarchical clustering of microarray data : How accurate are these clusters. *The Fifteenth International Conference on . . .*, (pp. 0–1).
- [Takeuchi & Habuchi, 2007] Takeuchi, H. & Habuchi, Y. (2007). A quantitative method for analyzing scan path data obtained by eye tracker. In *IEEE Symposium on Computational Intelligence and Data Mining, 2007. CIDM 2007*, number Cidm (pp. 283–286). : Ieee.
- [Terai et al., 2008] Terai, H., Saito, H., Egusa, Y., Takaku, M., Miwa, M., & Kando, N. (2008). Differences between informational and transactional tasks

- in information seeking on the web. In *Proceedings of the second international symposium on Information interaction in context - IiX '08* (pp. 152). New York, New York, USA : ACM Press.
- [Treisman & Gelade, 1980] Treisman, A. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12, 97–136.
- [Treisman & Gormican, 1988] Treisman, a. & Gormican, S. (1988). Feature analysis in early vision : evidence from search asymmetries. *Psychological review*, 95(1), 15–48.
- [Tsang et al., 2010] Tsang, H. Y., Tory, M., & Swindells, C. (2010). eSeeTrack - Visualizing sequential fixation patterns. *IEEE Transactions on Visualization and Computer Graphics*, 16, 953–962.
- [Tzanidou et al., 2005] Tzanidou, E., Petre, M., Minocha, S., & Grayson, A. (2005). Combining eye tracking and conventional techniques for indications of user-adaptability. . . . *Interaction-INTERACT 2005*, (pp. 753 – 766).
- [Vatturi & Wong, 2009] Vatturi, P. & Wong, W.-K. (2009). Category detection using hierarchical mean shift. *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '09*, (pp. 847).
- [Ware, 2003] Ware, C. (2003). Design as applied perception. In John Carroll (Ed.), *HCI Models, Theories, and Frameworks*. (pp. 10–26). Morgan Kaufman.
- [West et al., 2006] West, J., Haake, A., Rozanski, E., & Karn, K. (2006). eye-Patterns : Software for Identifying Patterns and Similarities Across Fixation Sequences. In *Proceedings of the 2006 symposium on Eye tracking research & applications* (pp. 149–154). : ACM.
- [Yarbus, 1967] Yarbus (1967). *Eye movements and vision*. New York, plenum pre edition.
- [Yesilada et al., 2008] Yesilada, Y., Jay, C., Stevens, R., & Harper, S. (2008). Validating the use and role of visual elements of web pages in navigation with an eye-tracking study. In *Proceeding of the 17th international conference on World Wide Web* (pp. 11–20). New York, New York, USA : ACM.
- [Yoon & Narayanan, 2004] Yoon, D. & Narayanan, N. H. (2004). Predictors of Success in Diagrammatic Problem Solving. In. In A. Blackwell, K. Marriott, & A. Shimojima (Eds.), *Diagrammatic Representation and Inference* (pp. 301–315). Springer.

Liste des publications pendant la thèse

Revue internationale avec comité de lecture

Becker, M., Bastien, J. M. C., & Drusch, G. (2015). Patterns comportementaux sur la page d'accueil et performances de navigation sur les sites Web. *Psychologie Française*, 60, 159-171.

Drusch, G. & Bastien, J.M.C. (2012). Analyzing Web pages visual scanpaths : between and within subjects variability. *Work : a journal of Prevention, Assessment and Rehabilitation*, 41, 1559-1566.

Communications à des conférences internationales avec actes

Drusch, G., Bastien, J. M. C., & Paris, S. (2014). Analysing eye-tracking data : from scanpaths and heatmaps to the dynamic visualisation of areas of interest. In T. Ahram, W. Karwowski & T. Marek (Eds.), *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics (AHFE 2014, Krakow, Poland, 19-23 July 2014)*.

Becker, M., Bastien, J. M. C., & Drusch, G. (2013). Navigation sur le web : de nouveaux indicateurs pour l'identification de patterns comportementaux. 7ème Conférence ÉPIQUE 2013 (Bruxelles, 10-12 juillet), Bruxelles.

Drusch, G., & Bastien, J. M. C. (2012). Analyzing visual scanpaths on the Web using the mean shift procedure and T-pattern detection : a bottom-up approach. In *Ergo'IHM '12. Proceedings of the 2012 Conference on Ergonomie et Interaction homme-machine (Biarritz, 17-19 Octobre)* (pp. 181-184). New

York, NY : ACM.

Drusch, G., Bastien, J. M. C., & Dinet, J. (2011). From gaze plots to eye fixation patterns using a clustering method based on Hausdorff distances. In 23ème Conférence Francophone sur l'Interaction Homme-Machine (Nice-Sophia Antipolis, 24-27 Octobre). New York : ACM.

Bastien, J. M. C., & Drusch, G. (2011). Exploring Web pages... Are all the people really exhibiting an F-shaped pattern? In ERGODESIGN USIHC - 11º Congresso Internacional de Ergonomia e Usabilidade de Interfaces Humano-Computador (Manaus, Brasil, 15-19 May).

Communications à des conférences nationales avec actes

Drusch, G., Bastien, J.M.C & Boisadan, A. (2012). Évolution des indicateurs oculométriques lors de la répétition de tâches de recherche d'information sur le web : résultats préliminaires. ErgoIHM'2012, Biarritz, France.

Drusch, G., Bastien, J.M.C, Becker, M., Rodriguez, P. (2012). Evalyzer : un outil d'évaluation à distance de l'expérience utilisateur. ErgoIHM'2012, Biarritz, France.

Dinet, J., Kitajima, M., Bastien, J. M. C., & Drusch, G. (2011). How young pupils read search engines results pages? Paper presented at the 21st Annual Meeting of the Society for Text and Discourse, Poitiers, France.

Drusch, G., & Bastien, C. (2011). Analyser les scanpaths : quelles méthodes pour quels résultats? In C. Bastien, J. Cegarra, A. Chevalier, J. Dinet, N. Gregori & V. Grosjean (Eds.), *Epique'2011. Sixième Colloque de Psychologie Ergonomique* (pp. 435-448). Nancy : Presses Universitaires de Nancy.

Chapitres d'ouvrages

Drusch, G., Bastien, C., & Dinet, J (2011). L'exploration visuelle de pages web. In J. Dinet & C. Bastien (Eds.), *L'ergonomie des objets et des environnements physiques et numériques* (pp. 219-244). Paris : Lavoisier.

Bastien, J. M. C., Drusch, G., & Dinet, J. (2010). Connaître les comportements des internautes : méthodes et outils. In L. Calderan, B. Hidoine & J.

Millet (Eds.), L'usager numérique. Séminaire INRIA (27 septembre – 1er octobre 2010 – Anglet) (pp. 39-62). Paris : ADBS

Annexes

Annexe A

Indicateurs et dendrogrammes de l'expérience 1

A.1 Session 3

Indicateurs	1	2	3
NbFixation	133,1	76,3571429	42,5625
DurTotale	55869,9	30769,6786	18577,4688
DurFixMoy	416,374	405,8	444,499688
LongSp	15868,8035	9587,10036	5560,16
AmplSacMoy	120,061	126,418571	133,391875
BarX	541,05	530,035714	505,75
BarY	411,35	393,142857	398,71875
Dispersion	246,665	247,328214	228,131563
Hull	421432,875	363239,036	274547,5
Back	73,85	44,1428571	23,84375
Revisite	0,678	0,54928571	0,38875
BackPond	0,56234598	0,58435048	0,56942523
AoisDistincte	23,1	20,6428571	15,46875
NbFlxParAoi	5,71817135	3,69874918	2,73055822
TxBonLien	44,4444444	33,3333333	53,3333333

FIGURE A.1 – Indicateurs de la session 3 pour les 3 groupes

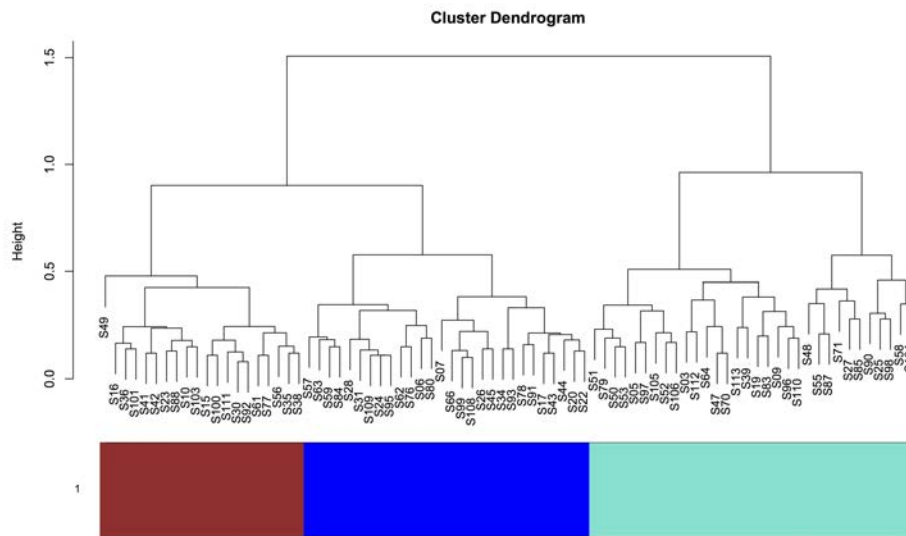


FIGURE A.2 – Dendrogramme de la Session 3

A.2 Session 3R

Indicateurs	1	2	3
NbFixation	16,4411765	9,23809524	5,15384615
DurTotale	6584,79412	4325,38095	3469,88462
DurFixMoy	407,133529	481,439524	740,097308
LongSp	2311,36176	1156,96286	457,751923
AmplSacMoy	139,892941	123,499524	84,1019231
BarX	527,529412	443,666667	508,423077
BarY	374,794118	444,52381	457,461538
Dispersion	169,970294	157,458095	100,778077
Hull	99617,3676	60973,0476	13916,75
Back	9,17647059	4,0952381	1
Revisite	0,38882353	0,18571429	0,05576923
BackPond	0,56953073	0,46075695	0,15992063
Aois	6,52941176	4,95238095	2,92307692
NbFixParAoi	2,49652406	1,83922902	1,84157509
TxBonLien	69,6969697	90	87,5

FIGURE A.3 – Indicateurs de la session 3R pour les 3 groupes

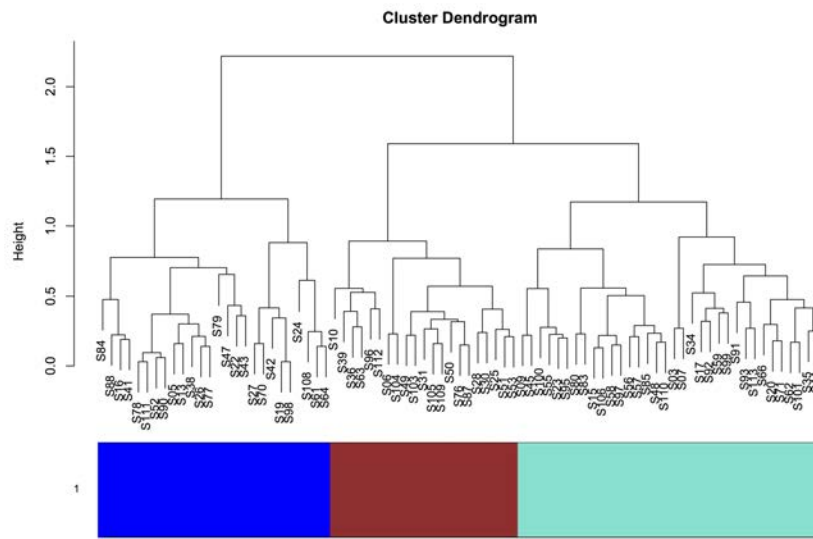


FIGURE A.4 – Dendrogramme de la Session 3R

A.3 Session 4

Indicateurs	1	2	3
NbFixation	50,1785714	23,4545455	13,6923077
DurTotale	21673,2143	11143,4242	8593,28205
DurFixMoy	427,811429	473,242121	715,384359
LongSp	6034,55786	2709,57727	1328,79615
AmplSacMoy	123,8275	119,658485	98,2151282
BarX	512,892857	460,242424	523
BarY	411,892857	432,090909	445,025641
Dispersion	231,069643	194,595152	131,3
Hull	274924,25	141197,606	53476,0897
Back	27,25	12,030303	6
Revisite	0,4175	0,26575758	0,20820513
BackPond	0,57031545	0,52468368	0,41762348
AoisDistincte	13,6071429	8,54545455	5,51282051
NbFixParAoi	3,71725973	2,7797593	2,49303012
TxBonLien	25	43,75	5,55555556

FIGURE A.5 – Indicateurs de la session 4 pour les 3 groupes

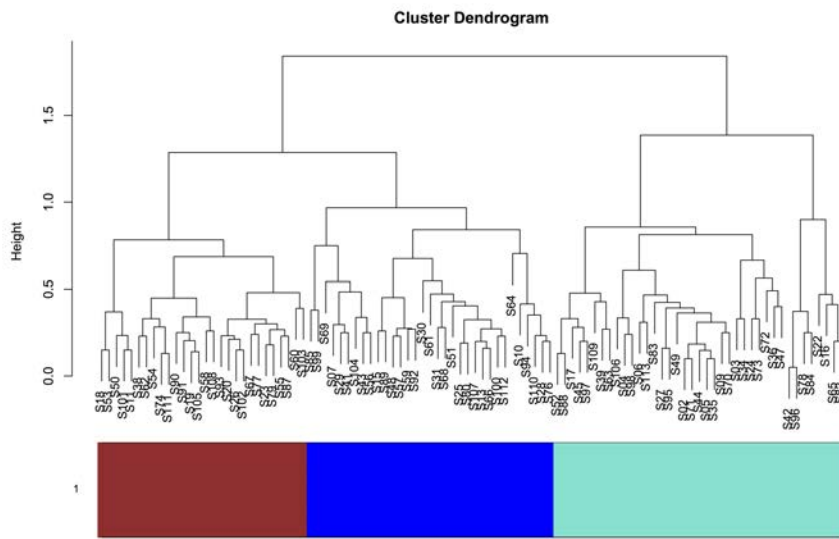


FIGURE A.6 – Dendrogramme de la Session 4

A.4 Session 4R

Indicateurs	1	2	3
NbFixation	8,08333333	7,25	3,62962963
DurTotale	3401,94444	4360,5	3145,25926
DurFixMoy	459,61	707,948929	947,977407
LongSp	873,606667	826,293929	275,508148
AmplSacMoy	98,2522222	100,998214	69,8366667
BarX	355,638889	500,5	486,037037
BarY	610,055556	443,071429	510,296296
Dispersion	137,879167	123,831786	63,9477778
Hull	45950,9028	37621,3929	5220,24074
Back	3,72222222	2,32142857	0,18518519
Revisite	0,15833333	0,17642857	0,01222222
BackPond	0,39752568	0,30382981	0,0377425
AoisDistincte	3,11111111	2,75	1,81481481
NbFixParAoi	2,59768519	2,49404762	2,35802469
TxBonLien	96,969697	0	7,40740741

FIGURE A.7 – Indicateurs de la session 4R pour les 3 groupes

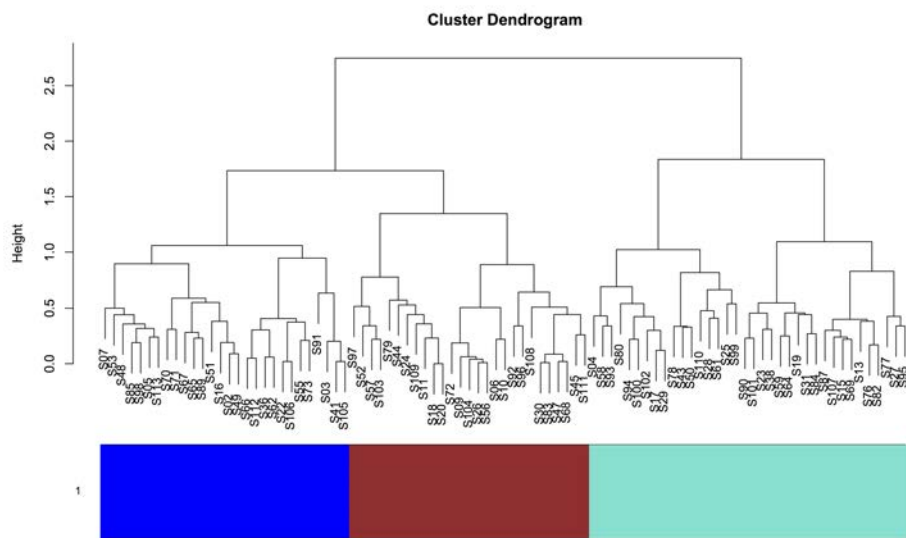


FIGURE A.8 – Dendrogramme de la Session 4R

A.5 Session 5

Indicateurs	1	2	3
NbFixation	82,68	32,15	16,2777778
DurTotale	28771,48	12347,6	7630,85185
DurFixMoy	349,9408	387,351	493,850741
LongSp	14353,3756	5413,1355	2042,66574
AmplSacMoy	177,712	171,6955	121,873704
BarX	513,2	413,95	321,481481
BarY	566,72	287,45	239,944444
Dispersion	471,204	310,311	203,434074
Hull	904799,5	345806,275	96611,2963
Back	42,84	16,3	6,38888889
Revisite	0,4848	0,397	0,2312963
BackPond	0,52635783	0,51507259	0,39136134
AoisDistincte	20,6	12,05	6,25925926
NbFixParAoi	4,09322611	2,81341126	2,59977219
TxBonLien	47,826087	33,33333333	53,0612245

FIGURE A.9 – Indicateurs de la session 5 pour les 3 groupes

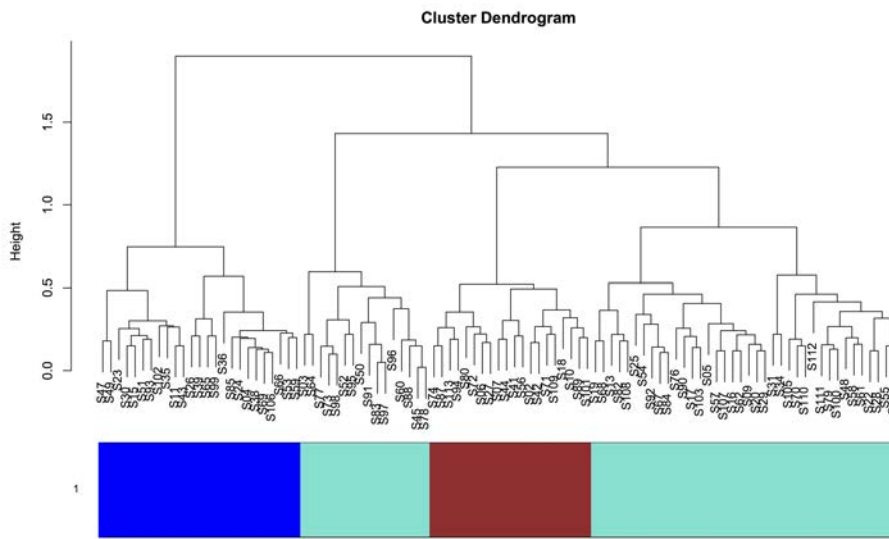


FIGURE A.10 – Dendrogramme de la Session 5

A.6 Session 5R

Indicateurs	1	2	3
NbFixation	12,5757576	9,22727273	4,4047619
DurTotale	5487,06061	4215,13636	2255,85714
DurFixMoy	489,926364	467,810909	547,889762
LongSp	1541,76636	1158,79591	362,414048
AmplSacMoy	93,5084849	118,198182	75,285
BarX	226,666667	268	245,5
BarY	275,272727	269,181818	276,857143
Dispersion	107,910303	139,877727	80,1438095
Hull	81590,9546	55392,75	12470,5
Back	4,45454546	3,81818182	0,5
Revisite	0,44121212	0,24954545	0,07071429
BackPond	0,43812171	0,44340035	0,11693122
AoisDistincte	3,42424242	4,13636364	2,42857143
NbFixParAoi	3,18903319	2,21515152	1,91071429
TxBonLien	78,125	78,9473684	75

FIGURE A.11 – Indicateurs de la session 5R pour les 3 groupes

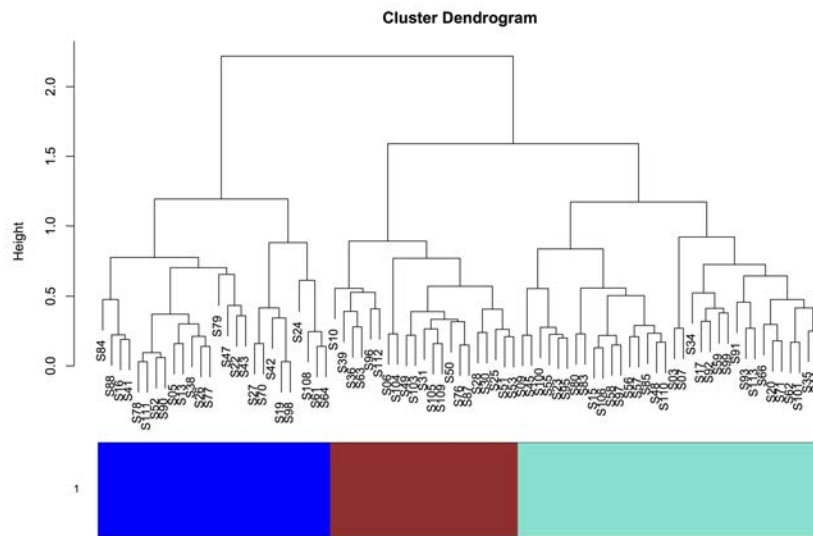


FIGURE A.12 – Dendrogramme de la Session 5R

A.7 Session 6

Indicateurs	1	2	3
NbFixation	100,413793	38,2	10,3695652
DurTotale	34268,9655	13533,84	5291,69565
DurFixMoy	344,414828	361,8096	531,92413
LongSp	15027,2517	5416,872	921,052174
AmplSacMoy	149,803448	142,7096	86,4771739
BarX	509,965517	413,16	212,326087
BarY	527,137931	397,28	238,456522
Dispersion	364,931035	285,6136	99,0347826
Hull	658413,707	308009,02	30170,7826
Back	63,1034483	21,84	3,39130435
Revisite	0,59448276	0,4032	0,2473913
BackPond	0,63233709	0,56714006	0,31219444
AoisDistincte	17,6206897	9,8	2,47826087
NbFixParAoi	5,7387185	4,08083667	4,67644928
TxBonLien	19,0476191	5,26315789	0

FIGURE A.13 – Indicateurs de la session 6 pour les 3 groupes

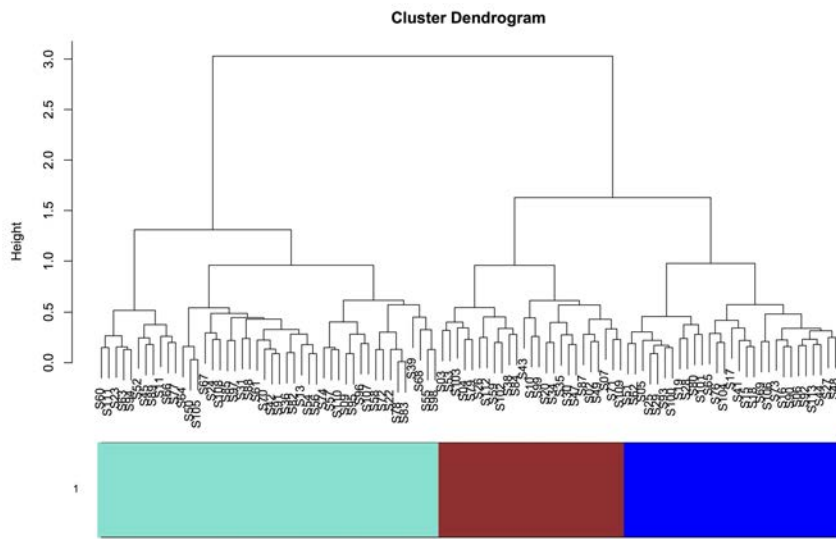


FIGURE A.14 – Dendrogramme de la Session 6

A.8 Session 6R

Indicateurs	1	2	3
NbFixation	25,4130435	10	5,33333333
DurTotale	8140,95652	3807,76923	2969,88889
DurFixMoy	331,56	394,331538	654,142963
LongSp	3505,16957	1446,24538	499,842593
AmplSacMoy	142,746304	148,288077	81,8266667
BarX	610,326087	686,730769	230,074074
BarY	600,26087	684,346154	272,555556
Dispersion	280,699565	243,962692	90,0162963
Hull	234425,75	112409,827	18944,2963
Back	12,5	3,88461538	1,03703704
Revisite	0,30217391	0,11923077	0,06888889
BackPond	0,543711	0,38942633	0,15083774
AoisDistincte	9,02173913	5,26923077	2,48148148
NbFixParAoi	2,60806641	1,961279	2,30864198
TxBonLien	37,5	31,25	0

FIGURE A.15 – Indicateurs de la session 6R pour les 3 groupes

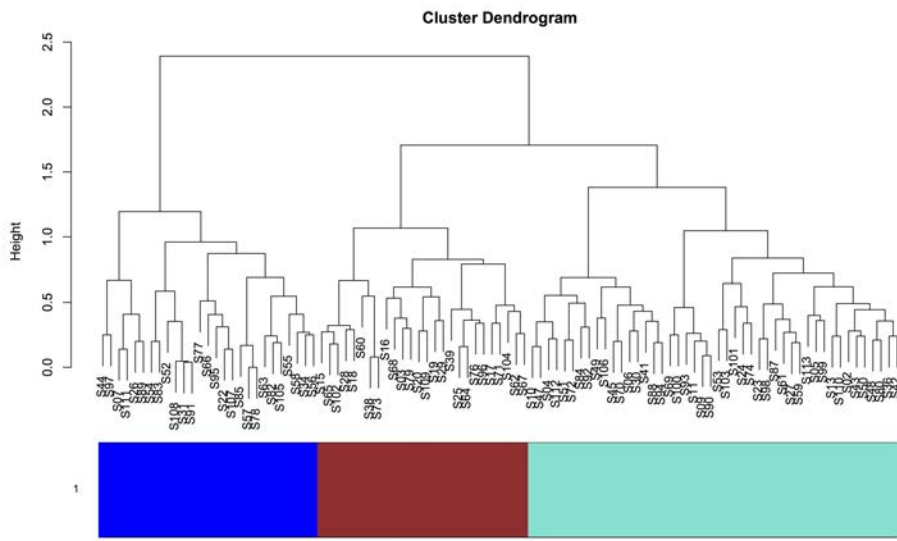


FIGURE A.16 – Dendrogramme de la Session 6R

Annexe B

Indicateurs et dendrogrammes de l'expérience 2 (étude longitudinale)

B.1 Musée Carnavalet

INDICATEUR:	1	2	3
NbFixation	152,214286	87,375	55,875
DurTotale	35198,6429	18939,375	12635,625
DurFixMoy	224,123571	213,0575	218,45375
LongSp	23001,8336	15013,0813	9396,91125
AmplSacMoy	162,493571	177,81	173,6
BarX	583,785714	580,625	576,875
BarY	575,5	507	459
Dispersion	358,547143	372,56375	359,79625
DistanceBary	586,65785	598,976866	433,2136
Hull	720302,25	718061,688	583751,875
AoisDistincte	23,8571429	22	16
Back	80,0714286	47,875	31,5
Revisite	0,66428571	0,55375	0,51

FIGURE B.1 – Indicateurs pour les 3 classes sur le Musée Carnavalet à T1

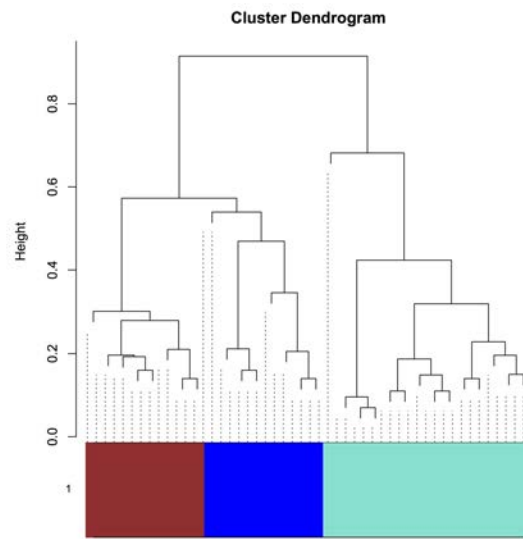


FIGURE B.2 – Dendrogramme sur le Musée Carnavalet à T1

INDICATEURS	1	2	3
NbFixation	84	28,2727273	15,2
DurTotale	18459,0833	6439,90909	4307,4
DurFixMoy	221,773333	230,382727	299,874
LongSp	13467,625	5737,19909	2788,878
AmplSacMoy	169,559167	203,813636	183,584
BarX	625,5	536,636364	455,2
BarY	499,083333	332,090909	228,2
Dispersion	359,560833	315,574545	284,788
DistanceBaryFirst	486,357221	337,206732	391,074495
Hull	622087,042	400622,182	240173,5
AoisDistinctes	18,5833333	11,5454545	7,8
Back	45,3333333	16,8181818	7,2
Revisite	0,55666667	0,38636364	0,15

FIGURE B.3 – Indicateurs pour les 3 classes sur le Musée Carnavalet à T2

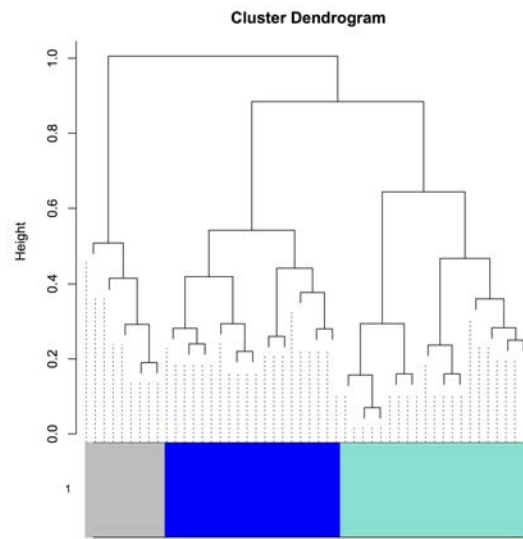


FIGURE B.4 – Dendrogramme sur le Musée Carnavalet à T2

INDICATEURS	1	2	3
NbFixation	57,6153846	20	16,4285714
DurTotale	12660,8462	5038,57143	4487,71429
DurFixMoy	220,11	250,375714	287,84
LongSp	10537,1708	3914,29	3206,84571
AmplSacMoy	200,188462	204,718571	196,078571
BarX	567	552,714286	481,142857
BarY	397,230769	309,571429	323,142857
Dispersion	341,471539	320,518571	288,297143
DistanceBaryFirst	380,522593	434,328605	316,852296
Hull	537019,923	364538,857	241756,714
AoisDistinctes	13	9,42857143	5,71428571
Back	30,1538462	11,5714286	6,14285714
Revisite	0,50846154	0,29714286	0,25

FIGURE B.5 – Indicateurs pour les 3 classes sur le Musée Carnavalet à T3

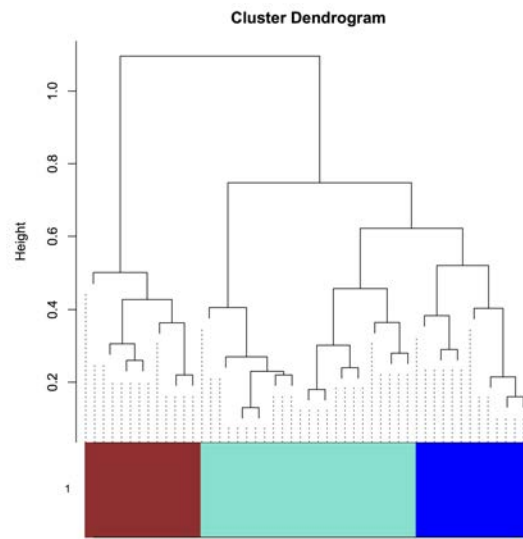


FIGURE B.6 – Dendrogramme sur le Musée Carnavalet à T3

B.2 France Diplomatie

INDICATEUR:	1	2	3
NbFixation	264,125	144,777778	52,2222222
DurTotale	61272,25	34514,6667	13530,1111
DurFixMoy	229,31625	241,631111	251,842222
LongSp	40014,4225	22107,1689	7707,95
AmplSacMoy	153,19125	157,051111	149,515556
BarX	553,625	574,111111	569,777778
BarY	479,125	455,444444	337,555556
Dispersion	357,48375	345,738889	296,86
DistanceBary	448,687747	488,998553	424,932792
Hull	792054,063	675024,056	415349,333
AoisDistincte	25,5	24,1111111	12,5555556
Back	129,625	74,2222222	23,7777778
Revisite	0,79625	0,67	0,49888889

FIGURE B.7 – Indicateurs pour les 3 classes sur France Diplomatie à T1

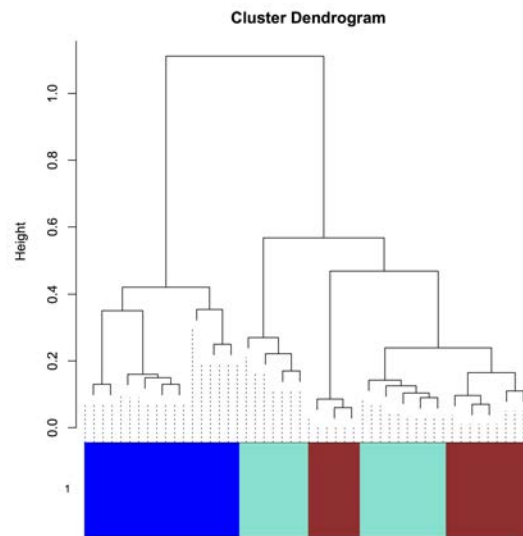


FIGURE B.8 – Dendrogramme sur le France Diplomatie à T1

Indicateurs	1	2	3
NbFixation	72,8571429	50,5714286	17,4545455
DurTotale	18226,1429	13909,8571	5041
DurFixMoy	257,542857	276,387143	298,167273
LongSp	9475,40429	7239,34286	2454,33727
AmplSacMoy	145,741429	143,532857	137,925455
BarX	493,428571	562,285714	535
BarY	442,857143	405,142857	300,454545
Dispersion	278,437143	301,365714	209,774545
DistanceBary	396,825169	376,600568	303,982474
Hull	406837,286	449828,643	145055,636
AoisDistincte	12,5714286	12,8571429	6,90909091
Back	28	23,4285714	8,90909091
Revisite	0,47857143	0,49142857	0,36181818

FIGURE B.9 – Indicateurs pour les 3 classes sur France Diplomatie à T2

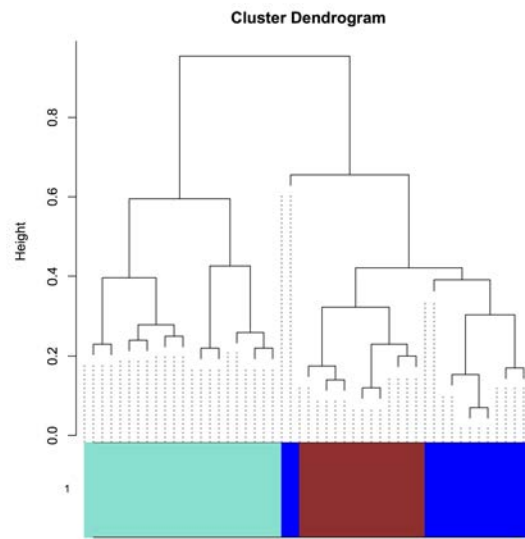


FIGURE B.10 – Dendrogramme sur le France Diplomatie à T2

INDICATEUR\	1	2	3
NbFixation	49,6363636	22,3636364	15,5
DurTotale	11914,9091	6389,72727	4818,5
DurFixMoy	284,36	289,17	327,3075
LongSp	7684,79	3659,31909	3012,085
AmpISacMoy	178,940909	161,362727	184,5375
BarX	579,454545	518,545455	465,25
BarY	430,454545	333,636364	372
Dispersion	271,886364	244,758182	254,76
DistanceBary	309,456937	343,389527	331,475789
Hull	343474,591	255539,773	226766
AoisDistincte	8,63636364	7,63636364	5,75
Back	18,3636364	10,1818182	6,5
Revisite	0,51181818	0,37272727	0,33

FIGURE B.11 – Indicateurs pour les 3 classes sur France Diplomatie à T3

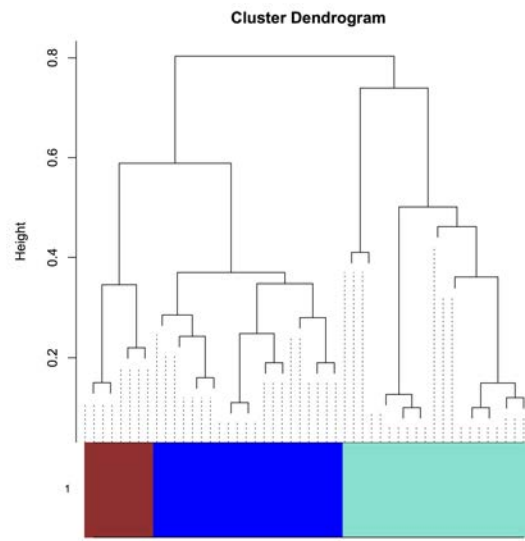


FIGURE B.12 – Dendrogramme sur le France Diplomatie à T3

Annexe C

Indicateurs et dendrogrammes de l'expérience 3

C.1 Session 1

Indicateurs	1	2	3	4
NbFixation	203,5	86,6666667	49,8888889	30,7142857
DurTotale	47156,8333	21299,5	11810,2222	8017,71429
DurFixMoy	234,1	246,256111	231,834444	244,915714
LongSp	28590,3883	11937,9322	8007,15778	4148,02
AmplSacMoy	140,636667	142,537778	161,4	147,97
BarX	601,666667	610,222222	619,666667	540,428571
BarY	366,166667	360,277778	354,666667	258,571429
Dispersion	405,5	367,018889	384,475556	275,117143
DistanceBary	270,245874	288,165764	257,644626	281,433714
Hull	901601,083	621196,028	549995,667	271127,5
AoisDistincte	19,5	15,2222222	14	9,28571429
Back	93,1666667	38,8888889	23,6666667	14
Revisite	0,79333333	0,61666667	0,45777778	0,40142857

FIGURE C.1 – Indicateurs pour les 3 classes pour la Session 1

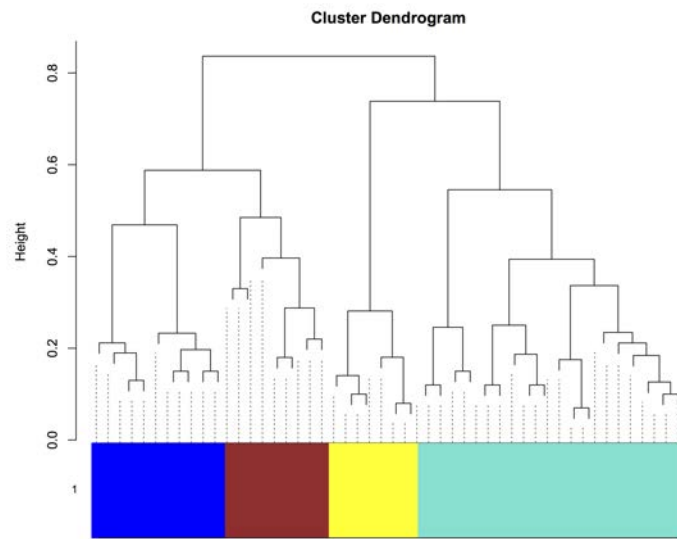


FIGURE C.2 – Dendrogramme sur la Session 1

C.2 Session 2

INDICATEUR:	1	2	3	4
NbFixation	183,222222	56	35,2142857	24
DurTotale	41495,6667	15979,2727	8374,07143	6099,09091
DurFixMoy	228,367778	285,584545	227,452143	254,685455
LongSp	26627,4	7797,66364	4169,25143	2601,41545
AmplSacMoy	155,684444	140,937273	124,944286	114,341818
BarX	622,222222	734,272727	391,571429	794,636364
BarY	438,666667	382,636364	271,285714	356,181818
Dispersion	386,204444	350,948182	228,596429	277,299091
DistanceBary	427,979179	326,415452	244,253714	323,562558
Hull	882422,944	494558,182	185074,893	174301,818
AoisDistincte	20,4444444	12,7272727	8,35714286	6,54545455
Back	85,5555556	23,8181818	15,7857143	7,72727273
Revisite	0,72666667	0,5	0,50142857	0,29181818

FIGURE C.3 – Indicateurs pour les 3 classes pour la Session 2

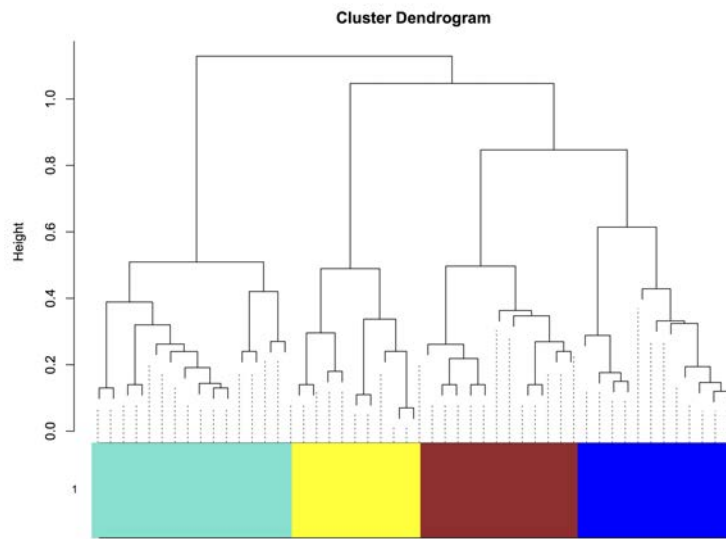


FIGURE C.4 – Dendrogramme sur la Session 2

C.3 Session 3

INDICATEUR:	1	2	3	4
NbFixation	104,636364	65,6666667	46,6666667	20,4117647
DurTotale	24403,0909	17745,8889	11739,1111	5488
DurFixMoy	220,72	277,868889	246,833333	278,407647
LongSp	14741,2591	8436,83	5078,59556	2398,28235
AmplSacMoy	146,141818	147,643333	119,49	121,961176
BarX	627,181818	477,333333	405	565,176471
BarY	399,272727	301,444444	278,555556	288,588235
Dispersion	370,384545	361,312222	332,84	215,046471
DistanceBary	256,255846	235,180385	327,2325	333,780582
Hull	603486,955	397868,778	270834,667	138918,294
AoisDistincte	11	10,4444444	7,44444444	4,82352941
Back	37	22,4444444	13,7777778	5,41176471
Revisite	0,64	0,57	0,52555556	0,28823529

FIGURE C.5 – Indicateurs pour les 3 classes pour la Session 3

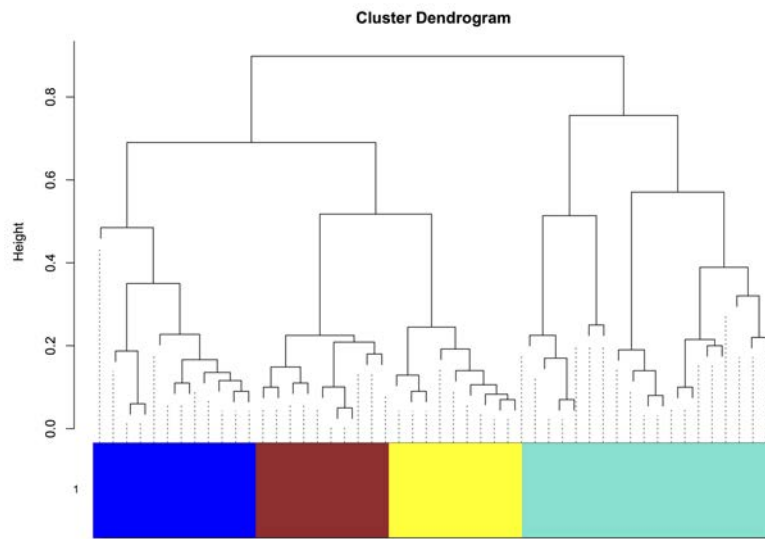


FIGURE C.6 – Dendrogramme sur la Session 3

C.4 Session 4

Indicateurs	1	2	3	4
NbFixation	86,3846154	28,1428571	17,3333333	9,6
DurTotale	20821,3846	8345,57143	5461	3003
DurFixMoy	229,134615	292,267143	316,051111	344,621
LongSp	13804,2023	4371,31786	2197,00222	1335,883
AmplSacMoy	165,588462	158,852143	128,585556	145,541
BarX	575,307692	771,928571	786,666667	768,9
BarY	314,076923	359,214286	334,444444	361
Dispersion	359,116154	341,247857	303,404444	206,645
DistanceBary	290,52308	380,107375	410,107747	286,381648
Hull	540549,385	341772,857	204622,056	69233
AoisDistincte	13,5384615	9,14285714	7,22222222	4
Back	39,1538462	12,5714286	6,66666667	2,3
Revisite	0,62307692	0,36714286	0,16777778	0,049

FIGURE C.7 – Indicateurs pour les 3 classes pour la Session 4

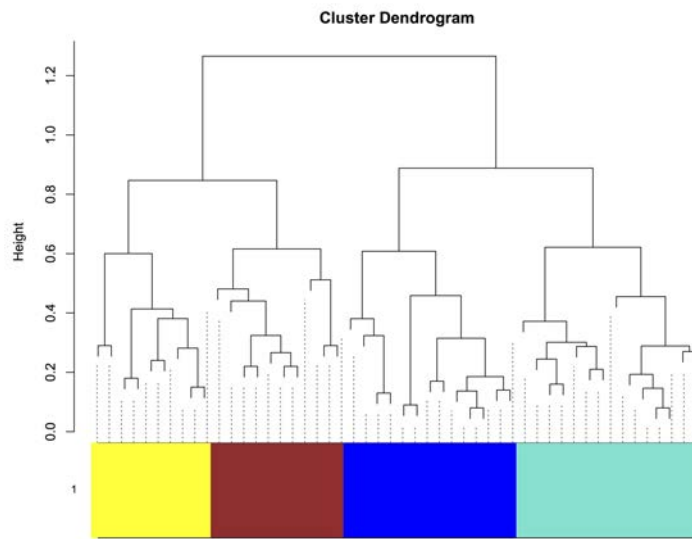


FIGURE C.8 – Dendrogramme sur la Session 4

C.5 Session 5

Indicateurs	1	2	3	4
NbFixation	88,83333333	29,33333333	20,8	10,3636364
DurTotale	20671,3333	8137,5	5255,3	3517,36364
DurFixMoy	216,181667	325,183889	242,243	379,212727
LongSp	11524,9267	4505,67278	3341,498	1167,04273
AmplSacMoy	141,335	154,534444	180,663	102,108182
BarX	468,166667	788,722222	485,9	590,181818
BarY	333,666667	361,833333	326,3	256,181818
Dispersion	284,13	283,720556	378,263	151,767273
DistanceBary	140,845764	332,917349	360,806341	241,514096
Hull	435764,75	287668,083	259872,7	73360,8182
AoisDistincte	7,33333333	6,33333333	5,9	2,54545455
Back	29,6666667	10,4444444	6,3	1,27272727
Revisite	0,745	0,44611111	0,285	0,16181818

FIGURE C.9 – Indicateurs pour les 3 classes pour la Session 5

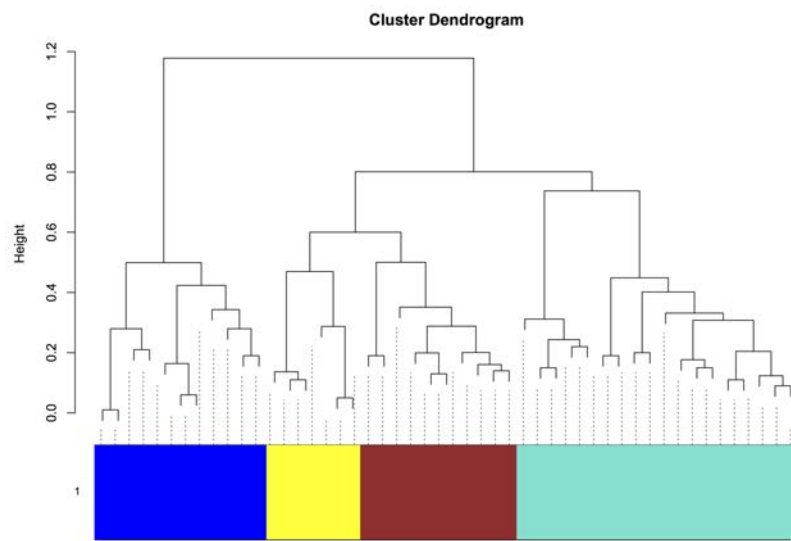


FIGURE C.10 – Dendrogramme sur la Session 5

Annexe D

Indicateurs et dendrogrammes de l'expérience 4

D.1 Education Nationale

INDICATEUR:	1	2	3
NbFixation	101,896552	34,1034483	18,5714286
DurTotale	22495,2414	7189,7931	4075,85714
DurFixMoy	222,268276	216,208276	224,683333
LongSp	15174,651	5793,93897	2610,68762
AmplSacMoy	156,571724	177,372069	156,509524
BarX	568,551724	419,827586	370,47619
BarY	491,862069	457,724138	449,428571
Dispersion	365,826207	317,577241	259,441429
DistanceBary	462,687656	414,200906	424,361757
Hull	626395,431	410364,155	229201,31
AoisDistincte	24,3103448	12,4827586	6,57142857
Back	55,2758621	16,9310345	6,52380952
Revisite	0,54172414	0,33	0,18285714

FIGURE D.1 – Indicateurs pour les 3 classes de la V1 du site Education Nationale

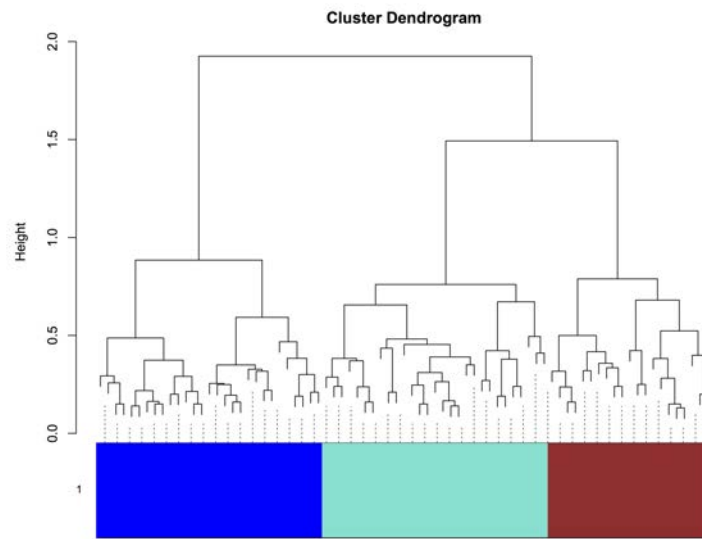


FIGURE D.2 – Dendrogramme sur la V1 du site Education Nationale

INDICATEUR:	1	2	3
NbFixation	128,363636	36,2380952	20,5
DurTotale	28570,8182	8008,09524	5331,77273
DurFixMoy	220,773636	220,803333	273,922727
LongSp	18481,0924	6274,55524	2825,19955
AmplSacMoy	151,996667	177,34381	146,58
BarX	605,727273	676,809524	737,090909
BarY	501,363636	499,047619	416,727273
Dispersion	356,882727	364,58381	320,135909
DistanceBary	471,24273	504,615122	559,994881
Hull	640535,939	514590	304307,159
AoisDistincte	19,7878788	12,3809524	7,18181818
Back	57,9090909	17	6,72727273
Revisite	0,62484848	0,33809524	0,17045455

FIGURE D.3 – Indicateurs pour les 3 classes de la V2 du site Education Nationale

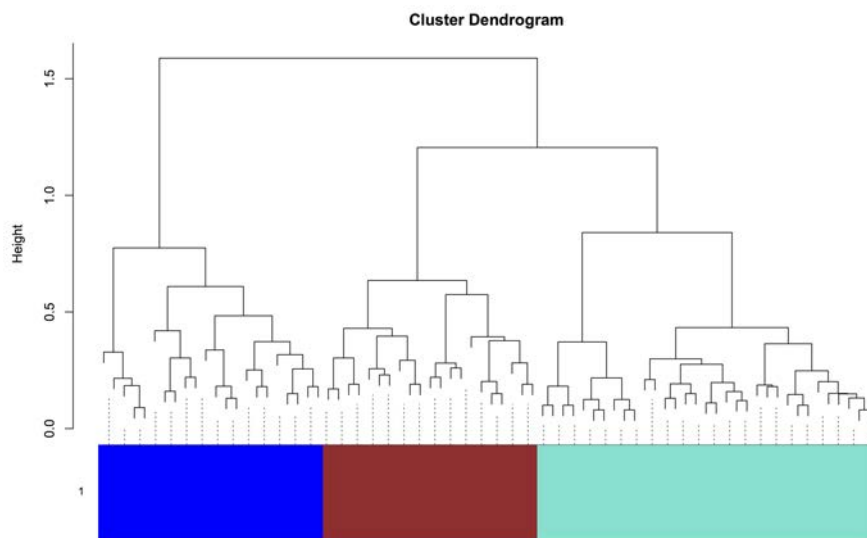


FIGURE D.4 – Dendrogramme sur la V2 du site Education Nationale

D.2 Service Public

INDICATEUR:	1	2	3
NbFixation	67,3928571	44,4137931	15,3888889
DurTotale	14890,3571	9883,37931	4334,11111
DurFixMoy	224,768929	228,757241	307,956667
LongSp	11124,3043	6885,61931	2328,6
AmplSacMoy	173,255357	159,468621	157,521111
BarX	559,892857	511,965517	446,611111
BarY	530,107143	544,137931	573,722222
Dispersion	305,145	295,211724	244,131667
DistanceBary	353,121576	441,567236	375,858825
Hull	453076,804	407578,448	178508,694
AoisDistincte	18,4642857	15,8275862	5,77777778
Back	40,1428571	22,6206897	6,05555556
Revisite	0,53535714	0,35137931	0,25555556

FIGURE D.5 – Indicateurs pour les 3 classes de la V1 du site Service Public

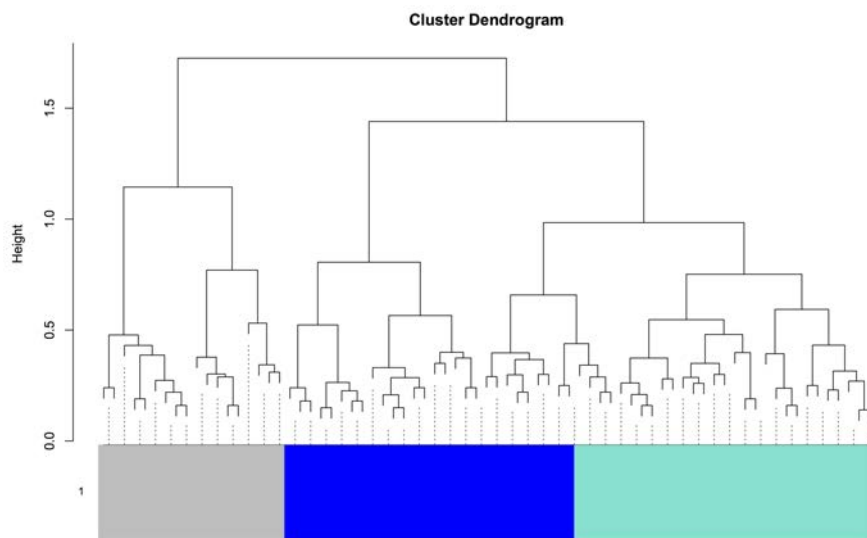


FIGURE D.6 – Dendrogramme sur la V1 du site Service Public

INDICATEUR:	1	2
NbFixation	69,2857143	21,1794872
DurTotale	16114,1429	5147,20513
DurFixMoy	234,583429	262,066154
LongSp	10654,57	3344,88462
AmplSacMoy	161,591143	163,884103
BarX	537,857143	514,307692
BarY	506,285714	568,692308
Dispersion	312,295429	256,576923
DistanceBary	398,32733	398,689481
Hull	499864,286	240233,641
AoisDistincte	19,1714286	9,05128205
Back	37,3714286	10,3076923
Revisite	0,48228571	0,24512821

FIGURE D.7 – Indicateurs pour les 2 classes de la V2 du site Service Public

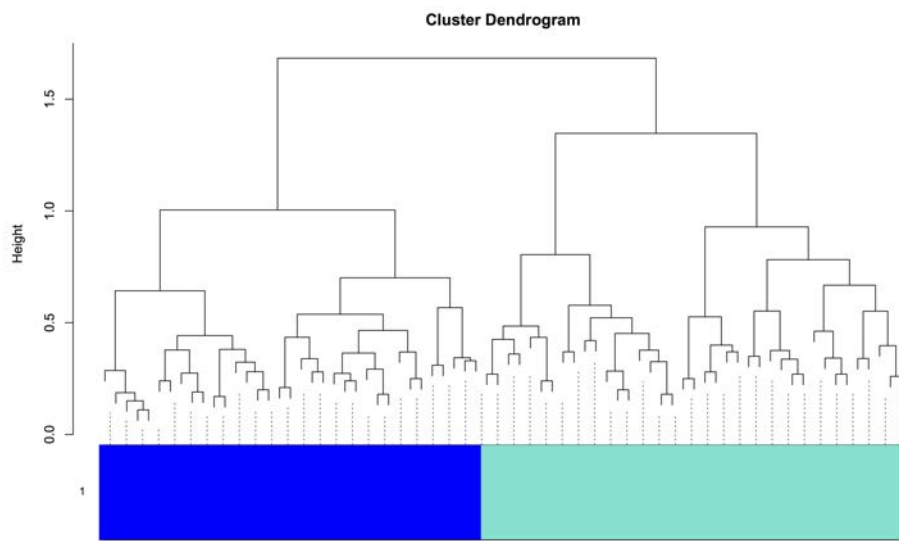


FIGURE D.8 – Dendrogramme sur la V2 du site Service Public

Annexe E

Captures d'écran de l'application Java développée pour la thèse

Cette application, développée en Java spécifiquement pour la thèse, permet d'importer les données de fixations pour chaque participant sous forme de fichier CSV ainsi que l'image du stimulus afin de calculer automatiquement les *AOIs*, de les visualiser puis de réaliser des exports de données traitées.

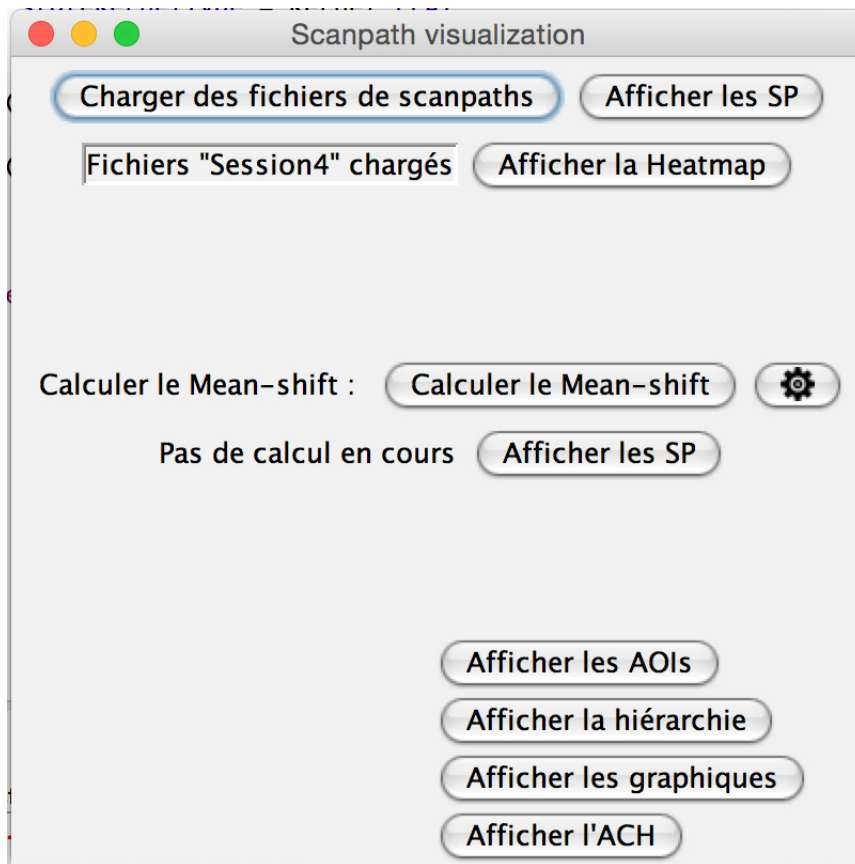


FIGURE E.1 – Fenêtre principale de l'application

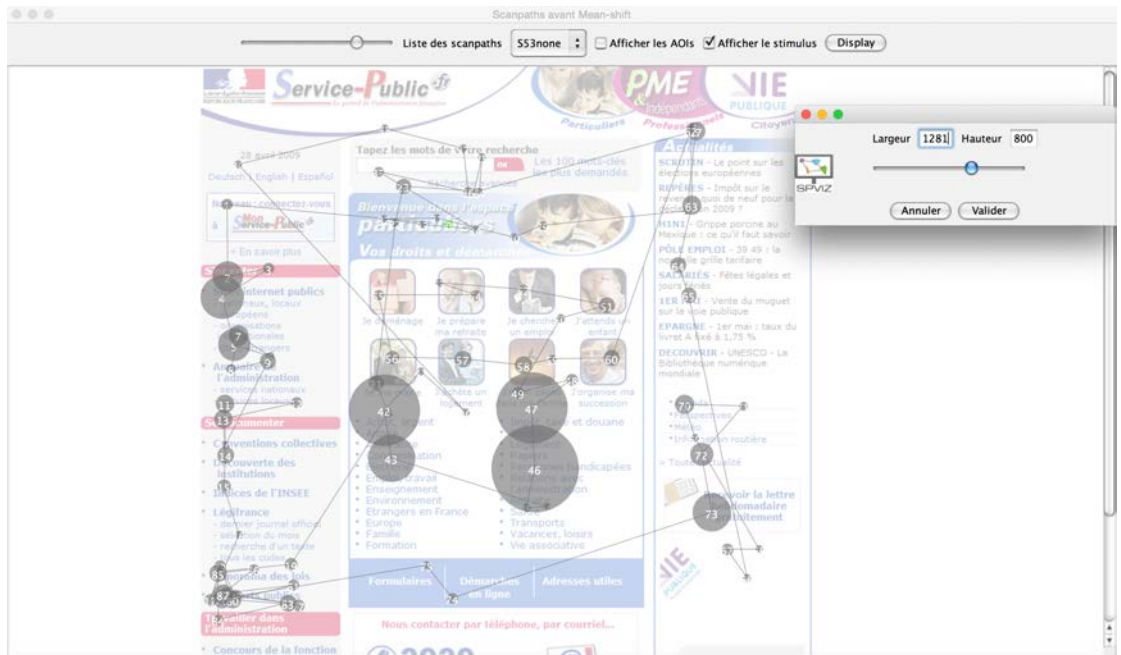


FIGURE E.2 – Fenêtre de visualisation d'un scanpath

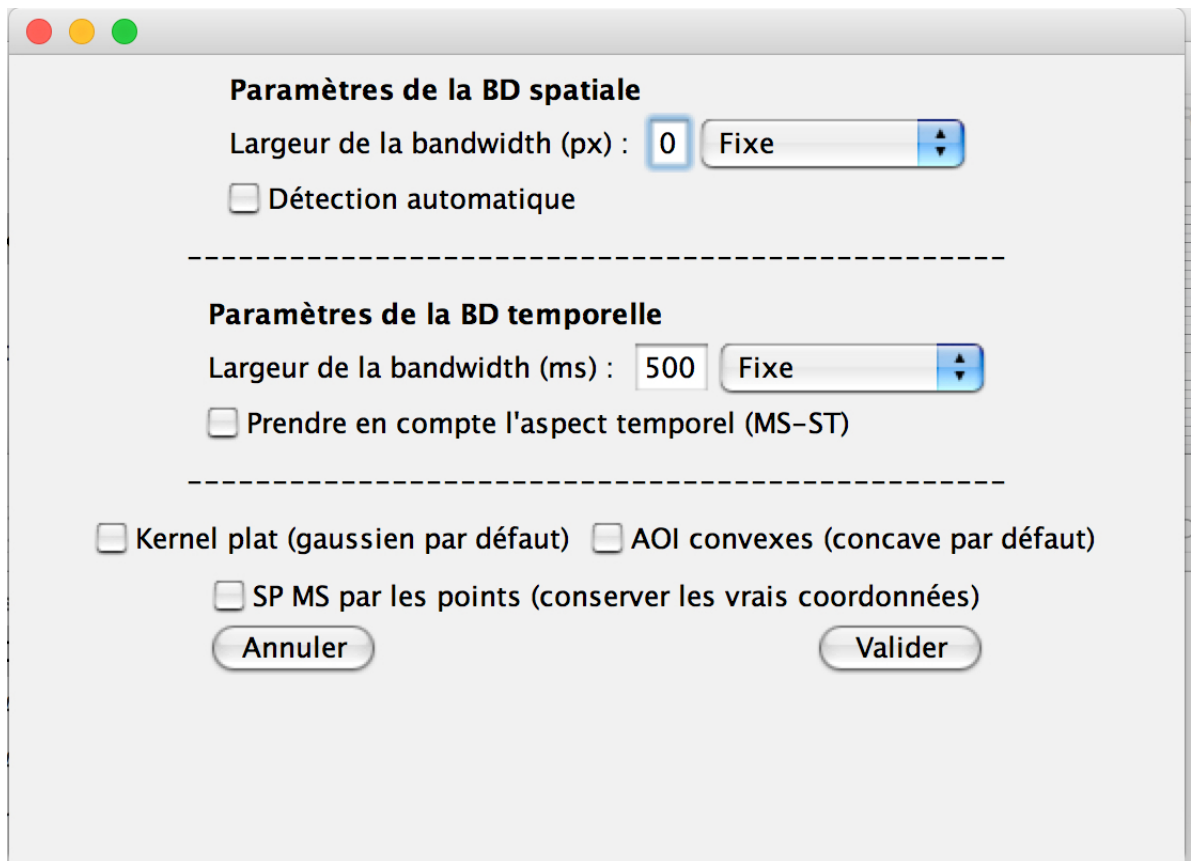


FIGURE E.3 – Fenêtre de sélection des options du Mean-Shift Clustering

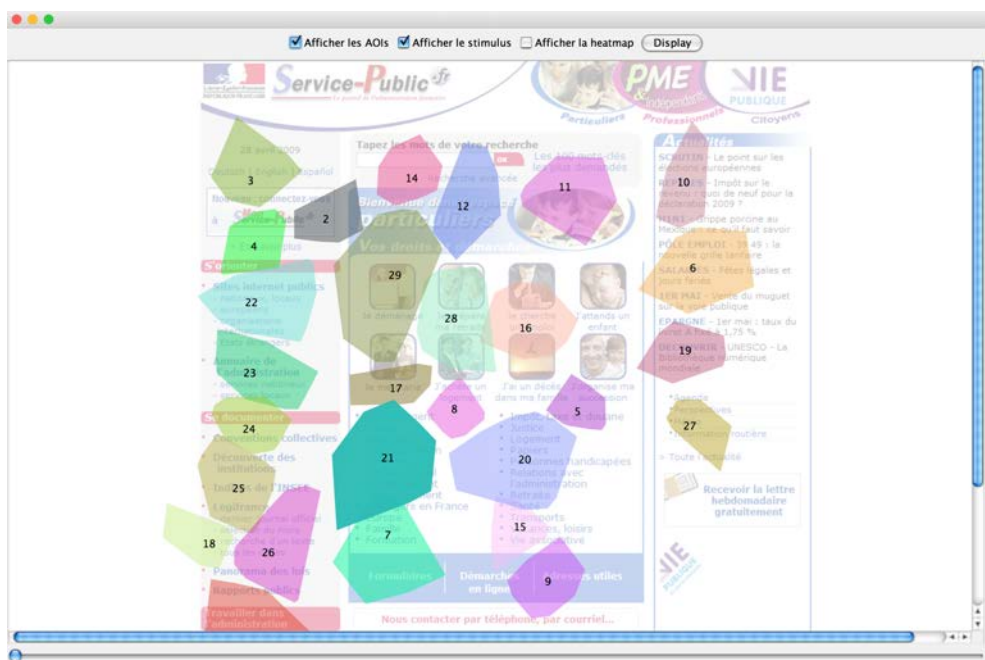


FIGURE E.4 – Fenêtre de visualisation des AOIs (le slider en bas permet de les visualiser temporellement)

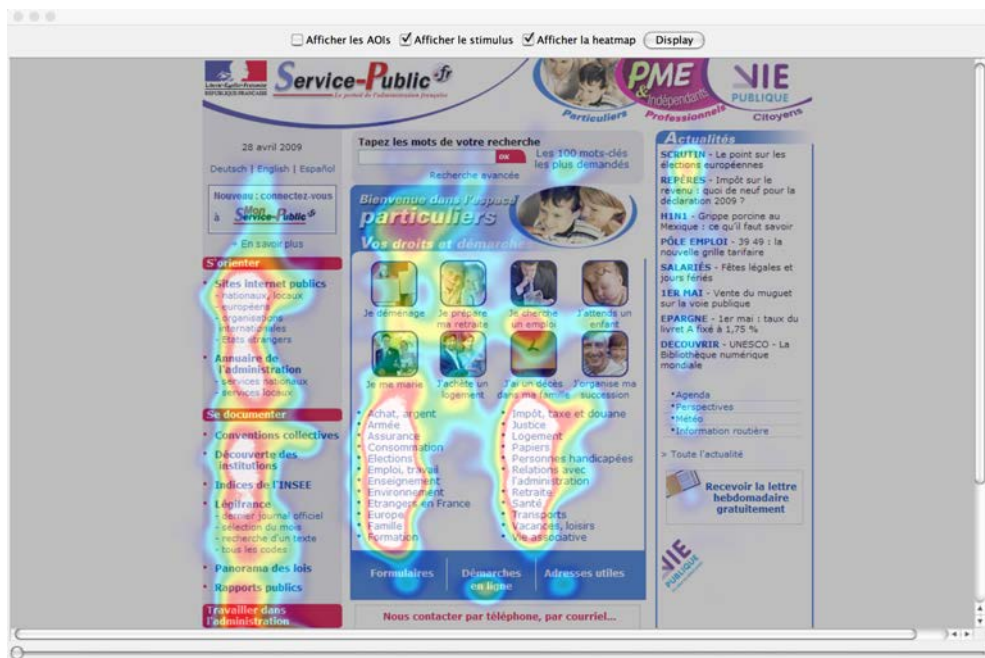


FIGURE E.5 – Fenêtre de visualisation de la heatmap

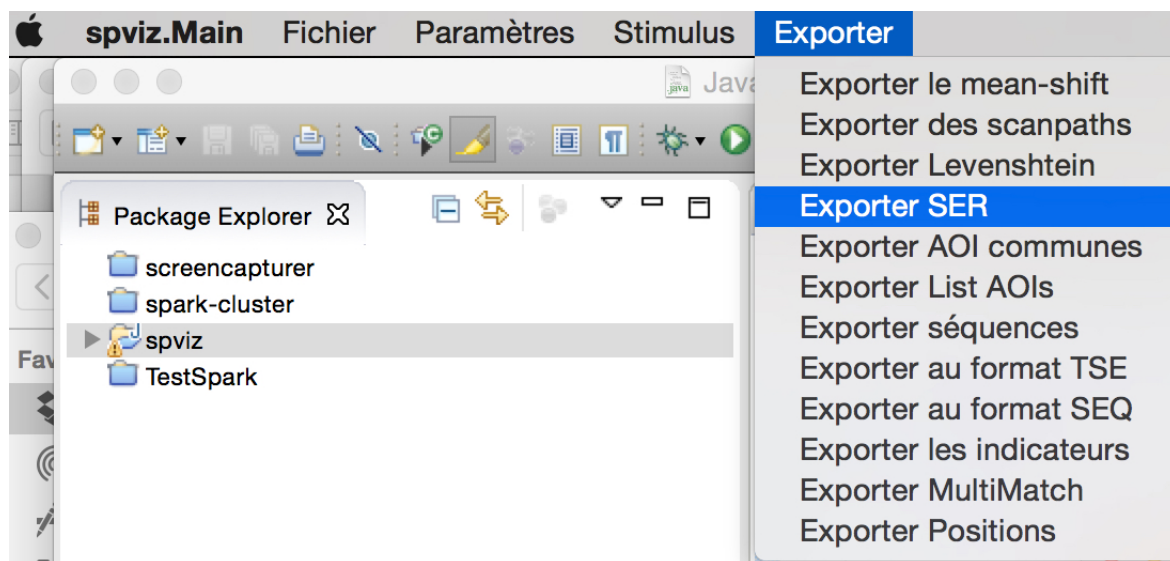


FIGURE E.6 – Menu d'export des données