



**HAL**  
open science

# L'exploration des pages web : de la caractérisation interindividuelle à l'identification de patterns comportementaux

Melanie Becker

► **To cite this version:**

Melanie Becker. L'exploration des pages web : de la caractérisation interindividuelle à l'identification de patterns comportementaux. Sociologie. Université de Lorraine, 2016. Français. NNT : 2016LORR0344 . tel-01661519

**HAL Id: tel-01661519**

**<https://theses.hal.science/tel-01661519>**

Submitted on 12 Dec 2017

**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](mailto: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.cfcopies.com/V2/leg/leg_droi.php)

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



Ecole doctorale Stanislas

# **L'exploration des pages web : de la caractérisation interindividuelle à l'identification de patterns comportementaux.**

Thèse de doctorat de l'Université de Lorraine  
en Ergonomie

Soutenue par Mélanie BECKER

Le 7 Décembre 2016

Sous la direction de J.M. Christian Bastien

Professeur en Ergonomie

Laboratoire PErSEUs (EA 7312)  
Psychologie Ergonomique et Sociale pour l'Expérience Utilisateurs

## **Composition du jury :**

Pr. BASTIEN Christian J.M, Université de Lorraine (directeur).  
Pr. BETRANCOURT Mireille, Université de Genève (rapporteur).  
Pr. DINET Jérôme, Université de Lorraine (examineur).  
Pr. VAN DE LEEMPUT Cécile, Université Libre de Bruxelles (rapporteur).

Années universitaires 2011-2016

# Remerciements

Je souhaite en premier lieu remercier chaleureusement toutes les personnes qui ont été liées de près ou de loin, à ces travaux de recherche. Cette thèse, c'est en partie, grâce à vous qu'elle existe.

Pour avoir contribué à la réalisation de cette thèse :

Merci à Christian qui fut une de ces rencontres inattendues qui changent le cours d'une vie. Je te dois un chemin que je n'aurais jamais osé envisager et il ne fait nul doute que ma carrière eût été beaucoup moins intéressante, stimulante et satisfaisante sans ton intervention.

Merci d'avoir cru en moi.

Merci à Axel, Gautier et Youssef qui ont développé les outils d'analyses nécessaires à cette étude. En plus de votre investissement sans limite, vos commentaires et vos questions m'ont beaucoup apporté.

Je souhaite également exprimer mes plus sincères remerciements aux membres du jury –Pr. Jérôme Dinet, Pr. Mireille Betrancourt et enfin Pr. Cécile Van De Leemput – qui me font l'honneur d'évaluer ce travail, et pour l'intérêt et l'attention qu'ils lui ont porté.

Pour finir, je remercie toutes les personnes qui ont accepté de donner de leur temps lors des expériences de cette étude.

Pour nos interactions riches et constructives à l'Université de Lorraine et en dehors, pour tout ce que vous m'avez enseigné :

Je souhaite remercier tous les membres de l'équipe PErSEUs et ses séminaires auprès desquels j'allais régulièrement me ressourcer.

Cette thèse a été une occasion unique de rencontrer des gens incroyables. Plus que des collègues, certains d'entre vous ont été de véritables compagnons de route et j'ai pu partager avec vous mes peines et mes joies.

Gautier, Sabrina, Jordan, Julie, et Anne Lorraine, vous avez permis à cette thèse de ne pas être une longue marche solitaire. Je remercie également Guillaume et Carine pour tous ces moments constructifs passés à Luxembourg.

Je n'oublie pas Audrey, ni tous mes collègues Serial qui ont su durant toute cette année trouver les mots justes pour me motiver à aller de l'avant.

Pour avoir été là tout au long de cette aventure :

Je remercie de tout mon cœur mes proches pour leur accompagnement, leur soutien sans faille, leurs conseils, leurs amitiés et bien plus encore : mes parents, ma sœur, Caroline, Laetitia, Stéphanie, Marina, Andrée et tous les autres.

J'exprime enfin toute mon affectueuse reconnaissance à Julien, pour sa bonne humeur, son soutien infailible et ses fous rires.

J'aurais réalisé ce travail avec beaucoup moins de plaisir sans votre présence et votre amour.

Merci infiniment à tous.



## Résumé

Une étude de Nielsen (2006), largement citée, indique que les internautes explorent les pages Web suivant un pattern en forme de "F". Ce résultat a amené les concepteurs à organiser les informations d'une page en fonction de ce comportement, même si aucune étude n'a permis de répliquer ces résultats. Bien que les conclusions de cette étude portent sur le comportement visuel, la question des patterns comportementaux permettant de décrire la navigation des internautes se pose de manière plus générale. L'objectif de cette thèse a donc été de déterminer si des patterns pouvaient être mis en lumière à partir de différents indicateurs.

Trois études ont été réalisées. Dans la première étude, 112 participants devaient réaliser quatre tâches de recherche d'information sur deux sites web différents. Le protocole impliquait une répétition immédiate de ces mêmes tâches. Une classification automatique a permis d'identifier 4 patterns qui se distinguent à la fois en termes de navigation sur la page d'accueil, mais aussi de performances. Lors des répétitions, la classification nous a permis d'identifier 3 des 4 patterns précédents. Ceci implique que les individus ne répètent pas forcément leur façon de rechercher l'information et ceci, peu importe la tâche, et le site.

La deuxième étude a porté sur 27 individus et impliquait, pour les participants, de se présenter trois fois consécutives à 48 heures d'intervalle, afin de refaire les mêmes tâches. La répétition des tâches, que ce soit à court ou à moyen terme, entraîne une augmentation des performances, c'est-à-dire que les tâches sont réalisées plus rapidement et de façon plus efficiente. Toutefois, les patterns identifiés diffèrent entre les répétitions à court et moyen terme. Un autre résultat observé est que les stratégies ou patterns ne sont pas propres aux individus. En d'autres termes, un individu peut présenter ou adopter plusieurs patterns d'une tâche à une autre, d'un site à un autre ou d'une répétition à l'autre.

Enfin, pour notre dernière étude, nous nous sommes demandé si l'homogénéité de nos échantillons pouvait influencer sur les patterns. Nous avons donc réalisé une expérimentation comptant 47 participants avec des profils variés. Les individus ont eu tendance à se distinguer selon les 4 mêmes patterns identifiés. Nous avons pu observer qu'en fonction des individus, une même stratégie pouvait conduire à la réussite ou à l'échec de la tâche. De plus, les styles d'apprentissage ne semblent pas liés aux patterns observés.

Les limites et les perspectives de ces travaux sont discutées.

**Mots-clés :** Comportements de navigation Web, performance, style de navigation, page d'accueil, classification, stabilité des patterns, variabilité

## **Abstract**

A study by Nielsen (2006), widely cited, indicates that Internet users explore web pages following a "F" shape pattern. This result brought the designers to organize the information of a page according to this behavior, even if no study replicated these results. Although the conclusions of this study concern the visual behavior, the question of the behavioral patterns allowing describing the navigation of the Internet users remains in a more general way. Thus the aim of this thesis was to determine if patterns could be revealed from various indicators.

Three studies were conducted. In the first study, 112 participants had to perform four information search tasks on two different websites. The experimental protocol involved an immediate repetition of the same tasks. A clustering method allowed us to identify 4 behavioral patterns, which distinguish themselves in terms of navigation on the homepage, but also in terms of performances. During the repetitions, the classification allowed us to identify 3 patterns out of the 4 previous ones. This implies that the individuals do not repeat necessarily the way they look for the information and this, no matter the task, and the Web site.

The second experiment involved 27 persons. They had to come three times, with 48 hour intervals to repeat the same tasks. The repetition of the tasks, whether in short or medium-term, increased the performances of the users, that is the tasks are more quickly realized and in a more efficient way. However, the identified patterns differ between the short and medium-term repetitions. Another observed result is that the strategies or patterns are not peculiar to the individuals. In other words, an individual can present or adopt several patterns from a task to another one, from a site to an other one or from a repetition to the other one.

Finally, in our last study, we wondered if the homogeneity of our previous samples could have influenced the patterns. So we conducted an experiment with 47 participants with varied profiles. The individuals tended to distinguish themselves according to 4 same identified patterns. We were able to observe that according to the individuals, the same strategy could lead to the success or to the failure of the task. Furthermore, the learning styles did not seem to be related to the observed patterns. Limits and prospects of this work are discussed.

**Keywords:** Web navigation behaviours, performance, behavioural patterns, navigation strategy, homepage, clustering, patterns stability, variability

*À Julien, à mes parents et à eux,  
avec tout mon amour.*



# Table des matières

<b>1. INTRODUCTION GENERALE.....</b>	<b>11</b>
<b>2. PARTIE THEORIQUE.....</b>	<b>13</b>
2.1 CARACTERISER LA RECHERCHE D'INFORMATION SUR LE WEB.....	13
2.1.1 <i>Les différentes tâches sur le web</i> .....	13
2.1.2 <i>Les principaux modèle de recherche d'information</i> .....	16
2.1.3 <i>Les principaux facteurs individuels impliquées dans la recherche d'information</i> .....	21
2.1.3.1 L'expertise .....	21
2.1.3.2 L'âge .....	23
2.1.3.3 Le genre.....	24
2.1.3.4 Le profil cognitif .....	25
2.2 LES METHODES DE RECUEIL DEEIEES A L'ETUDE DES COMPORTEMENTS SUR LE WEB.....	27
2.2.1 <i>Les techniques d'enregistrements directes</i> .....	28
2.2.1.1 L'analyse des données comportementales.....	28
Le test utilisateur.....	28
L'oculométrie.....	29
2.2.1.2 L'analyse des données déclaratives.....	30
Le think aloud .....	30
2.2.2 <i>Les techniques d'enregistrements différées</i> .....	30
2.2.2.1 L'analyse des données comportementales.....	30
Le Web mining.....	30
2.2.2.2 L'analyse des données déclaratives.....	34
Le questionnaire.....	34
L'entretien.....	35
2.3 LES NIVEAUX D'ANALYSES DES STYLES DE NAVIGATION ( <i>PATTERN</i> ) SUR LE WEB.....	36
2.3.1 <i>L'analyse des usages</i> .....	36
2.3.2 <i>L'analyse des comportements d'interaction et de navigation</i> .....	40
2.3.2.1 L'analyse des comportements pour la conception des navigateurs.....	40
2.3.2.2 L'analyse des comportements pour l'architecture des sites Web .....	47
2.3.2.4 L'analyse des comportements des internautes pour la conception de pages web .....	49
<b>3. REFLEXION METHODOLOGIQUE.....</b>	<b>54</b>
<b>4. ETUDES EMPIRIQUES.....</b>	<b>56</b>
4.1 METHODES ET ANALYSES POUR L'IDENTIFICATION DE PATTERNS COMPORTEMENTAUX SUR DES SITES WEB .....	56
4.1.1 <i>Problématique et contexte de l'étude</i> .....	56
4.1.2 <i>Méthodologie expérimentale</i> .....	57
4.1.2.1 Matériel et terrain.....	57
4.1.3.2 Procédures.....	58
4.1.3.3 Questionnaire de fin d'expérience.....	63
4.1.3.4 Échantillon .....	63
4.1.4 <i>Analyse des données exploratoires</i> .....	64
4.1.4.1 Modalités d'analyse des données de l'étude exploratoire.....	65

4.1.4.2	Résultats des analyses exploratoires et discussion.....	67
4.1.5	<i>Analyse des données approfondies</i> .....	71
4.1.5.1	Modalités d'analyse des données .....	71
4.1.5.2	Résultats des analyses approfondies.....	75
4.1.6	<i>Discussion</i> .....	90
4.2	EVOLUTION TEMPORELLE DES PATTERNS COMPORTEMENTAUX.....	94
4.2.1	<i>Problématique et contexte des études</i> .....	94
4.2.2	<i>L'étude des patterns comportementaux à court terme</i> .....	94
4.2.2.1	Méthodologie expérimentale .....	94
4.2.2.2	Analyse des données .....	95
4.2.2.3	Réflexion intermédiaire.....	108
4.2.3	<i>L'étude des patterns comportementaux à long terme</i> .....	109
4.2.3.1	Matériel et terrain.....	109
4.2.3.2	Procédures.....	109
4.2.3.3	Échantillon .....	113
4.2.3.4	Analyse des données .....	114
4.2.3.5	Réflexion intermédiaire.....	129
4.2.4	<i>Discussion</i> .....	130
4.3	INFLUENCE DES CARACTERISTIQUES DES PARTICIPANTS SUR LES PATTERNS .....	132
4.3.1	<i>Problématique et contexte de l'étude</i> .....	132
4.3.2	<i>Méthodologie expérimentale</i> .....	132
4.3.2.1	Matériel et terrain.....	132
4.3.2.2	Procédures.....	132
4.3.2.3	Echantillon .....	134
4.3.3	<i>Analyse des données</i> .....	135
4.3.4	<i>Discussion</i> .....	143
<b>5.</b>	<b>DISCUSSION</b> .....	<b>146</b>
5.1	SYNTHESE DE LA DEMARCHE .....	146
5.2	APPORTS ET LIMITES METHODOLOGIQUES.....	147
5.3	LES PATTERNS IDENTIFIES .....	149
5.3.1	<i>Description des patterns</i> .....	149
5.3.2	<i>Performance des patterns</i> .....	151
5.3.3	<i>Stabilité des patterns</i> .....	151
5.4	PERSPECTIVES .....	152
	<b>BIBLIOGRAPHIE</b> .....	<b>154</b>
	<b>TABLE DES FIGURES</b> .....	<b>167</b>
	<b>TABLE DES TABLEAUX</b> .....	<b>170</b>
	<b>SOMMAIRE DES ANNEXES</b> .....	<b>173</b>
	<b>ANNEXES A</b> .....	<b>175</b>
A.1	ANNEXES DE L'ETUDE « METHODES ET ANALYSES POUR L'IDENTIFICATION DE PATTERNS COMPORTEMENTAUX SUR DES SITES WEB ».....	175
A.1.1	<i>Questionnaire Post test</i> .....	175
A.1.2	<i>Résultats des analyses</i> .....	180

<b>ANNEXES B.....</b>	<b>184</b>
B.1 ANNEXES DE L'ETUDE « ÉVOLUTION TEMPORELLE DES PATTERNS COMPORTEMENTAUX ».	184
B.1.1 Résultats des analyses pour l'étude des patterns comportementaux à court terme.....	184
B.1.2 Résultats des analyses pour l'étude des patterns comportementaux à long terme.....	189
<b>ANNEXES C .....</b>	<b>213</b>
C.1 ANNEXES DE L'ETUDE « INFLUENCE DES CARACTERISTIQUES DES PARTICIPANTS SUR LES PATTERNS » .....	213
C.1.1 Questionnaire ISALEM Style d'apprentissage.....	213
C.1.2 Résultats des analyses.....	216



# 1. INTRODUCTION GENERALE

---

*« A man, viewed as a behaving system, is quite simple. The apparent complexity of his behaviour over time is largely a reflection of the complexity of the environment in which he finds himself. »*

*—Herbert A. Simon. The Science of the Artificial (1969)*

Ce travail de recherche s'inscrit dans le champ de la Psychologie Cognitive Ergonomique et se propose d'étudier les comportements de navigation sur le web. Internet est aujourd'hui devenu un moyen incontournable pour la distribution et l'échange d'informations, et nous sommes tous témoins de l'incroyable progression de cette masse d'information (Gerjets, Kammerer & Werner, 2011). Le site « internet live stats » nous renseigne, par exemple, qu'à la date du 27 Septembre 2016, plus de 1 milliard de sites existent sur la toile. Toutefois, plus que cette masse d'information mise à disposition des internautes, ce sont surtout les rapports entre la technologie et les utilisateurs qui ont beaucoup changé et changent encore. A cette même date, plus de 3 milliards d'utilisateurs du web sont recensés, chacun cherchant dans cet « environnement imposant » de quoi satisfaire ses besoins et sa curiosité (Holscher & Strube, 2000). Cette incroyable diversité met les concepteurs web face à une complexité sans précédent.

En effet, avec cette évolution, la conception de sites satisfaisants pour les internautes, et la compréhension des comportements de navigation constituent un défi pour la discipline. Concevoir des sites adaptés nécessite de mieux connaître les mécanismes impliqués dans l'activité de recherche, et de fournir aux concepteurs/designers de sites et services Web des lignes directrices et des recommandations pour la conception et l'évaluation. Ces travaux sont donc utiles à la fois d'un point de vue fondamental et pratique et ces recherches peuvent se résumer à la formule « connaître pour concevoir ».

A notre sens, deux postures s'opposent. D'une part, certaines études comme celle de Nielsen (2006) montrent que les internautes explorent souvent les pages Web selon un même pattern en forme de « F » (« F-shape »), ou encore qu'il faut préconiser un « design universel » adapté à tous les individus (Bevan, 2003; Vanderheiden, 2000). Ces études tendent donc à vouloir schématiser l'ensemble des comportements de navigation des internautes. Nous pensons que cette représentation de l'utilisateur « moyen » ne correspond en fait à aucun internaute.

D'autre part, certains chercheurs ont la volonté de prendre en compte la variabilité des comportements des internautes car aujourd'hui l'interface à « taille unique » ne peut plus satisfaire (Belk, Papatheocharous, Germanakos & Samaras 2013), et le web ne peut rester « statique » en traitant tous les utilisateurs de la même manière (Brusilovsky & Millan, 2007). Par exemple, tout le monde n'aurait pas besoin d'un plan de site pour trouver l'information recherchée. De même, un design spécifique

peut impacter un style de navigation plutôt qu'un autre. A l'instar de ces études, nous envisageons donc l'idée qu'il pourrait exister des patterns comportementaux correspondant à des stratégies de navigation plus ou moins similaires pour un ensemble d'individus. A cet égard, les comportements des internautes (pages vues, temps passé sur les pages etc.) font l'objet d'observations parfois approfondies au travers de métriques qui permettent de les décrire. Toutefois, ces études sont bien souvent réalisées à un niveau très macroscopique ne permettant pas vraiment de distinguer les gens entre eux et de décrire de façon précise la façon dont les usagers naviguent (Jiang, Tan, Phang & Weik, 2012, Belk et *al.*, 2013, Zhou, 2014). Le niveau d'analyse de la page pourrait permettre justement une description fine des comportements, mais les études sont rares. Nous pourrions penser, par exemple, que la page d'accueil n'a plus d'importance aujourd'hui à cause de l'utilisation généralisée des moteurs de recherche, mais ce n'est pas le cas pour tout le monde. Cette page est, d'ailleurs, souvent citée comme la page centrale d'un site, une clef de voûte, déterminante pour expliquer la navigation sur les autres pages. Cependant, à notre connaissance, peu d'études ce sont focalisées dessus pour expliquer des stratégies de navigation.

Nous proposons donc d'analyser plus finement la façon dont les individus naviguent sur le web dans des conditions similaires afin d'identifier ou non des patterns sur des sites web. L'objectif de cette thèse est double : (1) identifier une méthode d'analyse permettant d'étudier la variabilité, et de détecter des patterns comportementaux en mettant en relation le style comportemental sur la page d'accueil avec les performances de navigation (2) découvrir si ces patterns sont stables de manière intra et interindividuelles et d'en connaître les déterminants.

Ce travail débutera par une présentation des notions théoriques nécessaires à l'introduction de notre problématique. Les aspects liés aux méthodes de recueil de données sur le web et les différents niveaux d'analyses des comportements d'interactions seront aussi abordés.

# 2. PARTIE THEORIQUE

---

## 2.1 CARACTERISER LA RECHERCHE D'INFORMATION SUR LE WEB

---

### 2.1.1 Les différentes tâches sur le web

Depuis les 25 dernières années, de nombreuses recherches sont consacrées à l'optimisation du World Wide Web. Une partie de ces efforts se concentrent sur les aspects techniques de ces nouvelles technologies et cela concerne surtout l'utilisabilité des hypertextes, hypermédias et l'usage des navigateurs.

Toutefois, avant même d'aller dans ce détail et comme Byrne, Wehrle et Crow (1999) le font remarquer, il est nécessaire de tirer une compréhension plus claire des principales activités que peut réaliser l'utilisateur sur le web, et notamment mieux identifier les différents objectifs d'une recherche d'information (RI). Il est donc essentiel de comprendre ce que les gens font sur internet, comment ils le font, mais il est tout aussi important de comprendre pourquoi ils le font.

Dans cette approche globale et dans l'idée de découvrir les tâches des utilisateurs lorsqu'ils naviguent, pendant une journée, Byrne et son équipe ont demandé à 8 volontaires, issus de la communauté universitaire (enseignants, personnels de secrétariat ou encore étudiants) de naviguer sur le web, comme « ils le font en temps normal » et de décrire verbalement ce qu'ils faisaient au cours de la navigation. Des caméras étaient mises en place sur les ordinateurs. En combinant les données issues des fichiers logs et du protocole verbal, Byrne et ses collaborateurs ont démontré l'existence d'une taxonomie regroupant 6 activités :

- L'utilisation de l'information : Il s'agit ici de toutes les activités ou séries d'activités pour lesquelles l'utilisateur va utiliser une information trouvée sur le web. Ces informations peuvent alors être lues, écoutées, visionnées (animations, vidéo), télécharger ou encore imprimer.
- La localisation de l'information : Byrne définit ici la recherche visuelle de l'information ou d'un lien lorsque l'individu pense avoir trouvé quelque chose « d'intéressant ».
- Aller d'un point A à B : Toutes activités qui impliquent que l'utilisateur se dirige vers une autre page. Cela peut se faire, par exemple, via le bouton retour arrière, taper une adresse URL, accéder à son historique ou ses favoris
- Fournir des informations : Les utilisateurs utilisent le web pour obtenir de l'information mais également pour en envoyer. Il peut s'agir par exemple de compléter un filtre pour sélectionner un objet, compléter des critères de recherche ou encore remplir un formulaire.
- Configurer le navigateur : Il s'agit là de toutes les activités permettant de configurer le navigateur (taille et nombre de fenêtres, etc.).
- Réagir à l'environnement : La plupart des activités sont guidées par l'utilisateur mais parfois le navigateur exige quelque chose de l'utilisateur.

Ces situations ont été considérées comme des tâches « réactives » exigeant un feedback (valider ou annuler un téléchargement, fermer une pop-up, etc.).

Les principales critiques faite à cette étude concernent son périmètre trop large ne tenant pas compte de la recherche sur un site en particulier, et son échantillon trop restreint (Holscher & Strube, 2000).

En 2001, Morrison, Pirolli et Card ont réalisé une enquête et ont demandé à 2188 utilisateurs du Web de décrire un épisode récent dans lequel ils ont trouvé des informations sur le Web et qui les a conduits à prendre une décision ou à réaliser une action significative. Quatre objectifs principaux sont ressortis ; collecter de l'information, trouver de l'information, explorer le web et monitorer l'information en surveillant le changement sur des pages précises.

Se basant sur ces études et sur les travaux de Broder (2002), Rose et Levinson (2004) ont donc proposé une nouvelle classification.

Les chercheurs ont dans un premier temps établi un cadre conceptuel dans lequel ils ont répertorié un ensemble de buts et ont ensuite fait correspondre manuellement plus d'une centaine de requêtes effectuées sur le moteur de recherche AltaVista aux buts identifiés par ailleurs. Une typologie comptant trois types de recherches différente a été observée :

- Les recherches de navigation où l'on se destine à retrouver un site Web spécifique, cela implique que l'utilisateur ait déjà un site à l'esprit.
- Les recherches informationnelles centrées sur le but de l'utilisateur qui est celui d'obtenir des informations sur sa question. Cette catégorie inclut des questions ouvertes ou fermées, des besoins de conseils ou encore de la curiosité pour en apprendre plus sur certains sujets. La requête informationnelle peut aussi inclure le besoin de localiser des éléments dans l'environnement réel (acheter tel produit), ou simplement obtenir une liste de suggestions pour de futures recherches.
- La recherche ressource qui a pour but d'obtenir autre chose que de l'information et donc d'accéder à une ressource telle que le téléchargement d'une chanson ou d'une recette de cuisine par exemple. Broder (2002) avait nommé cette catégorie sous le terme de requête transactionnelle.

Jansen, Booth et Spink (2008, 2009) émettent tout de même une réserve sur cette classification car selon eux 25% des recherches pourraient avoir plusieurs intentions. Cette catégorisation reste, néanmoins l'une des plus référencées lorsqu'il s'agit de classer les différents types de recherches d'informations (Weber & James, 2011).

Plus récemment, Cordova et Jansen (2014) ont donc souhaité approfondir ces travaux en identifiant des types de tâches que l'on peut réaliser sur le web et qui combinent des recherches que l'on peut effectuer sur un moteur de recherche et un site web en particulier. Cette étude va donc déjà un peu plus loin que les précédentes en identifiant les besoins de recherche et en commençant à décrire « comment » ils le font.



Pendant 5 mois, 295 271 requêtes effectuées à partir d'un moteur de recherche espagnole ont été classifiées. Selon cette équipe, on distinguerait six types de recherche d'information (RI) qui seraient motivées par différents besoins :

- L'exploration: sans but précis, il s'agit d'effectuer une requête large sur le moteur de recherche puis de multiples requêtes sur le site (par exemple, quand l'utilisateur veut découvrir une musique)
- La recherche pour la navigation : requête large sur le moteur de recherche mais précise sur le site (rechercher une musique)
- La recherche pour l'acquisition : correspond à une recherche précise sur le moteur de recherche et sur le site (télécharger une musique)
- La recherche orientée : l'utilisateur effectue la même recherche sur le moteur de recherche et le site, mais précise sa question (recherche d'un artiste et ensuite de sa biographie)
- La recherche persistante : l'utilisateur recherche et surveille la même information sur le moteur de recherche et sur le site (information à propos d'un artiste, voir s'il y a des nouveautés)
- Et enfin la recherche inconsistante : requête différente entre celle qui est faite sur le moteur de recherche et le site (rechercher une information sur le rock et ensuite la musique latine).

Même si ces études diffèrent en termes de méthodologie et d'objectifs de recherches, d'importantes similarités sont quand même à relever entre ces catégorisations. La première concerne la recherche informelle et la simple consultation. L'objectif de l'utilisateur est de récupérer des informations assez brèves et pas forcément détaillées. La recherche peut se faire sur une seule page. La seconde est la recherche formelle qui pourrait être comparée à une recherche bibliographique afin de collecter un maximum d'information sur un sujet. L'utilisateur brasse de nombreuses pages afin de recueillir toutes ces données qui se chevauchent. La troisième catégorie concerne l'activité « ludique » de navigation où l'utilisateur est engagé dans une recherche fortuite. Une quatrième comprend le fait de surveiller l'information et voir les changements s'il y en a. La dernière comprend les tâches permettant d'obtenir autre chose que de la ressource (transactions, configuration, chat, etc.)

Enfin, d'autres chercheurs ont tenté d'aller plus loin afin d'étudier l'effet de la tâche sur la navigation. Kellar, Watters et Shepherd (2006, 2007) ont porté leur attention sur 21 étudiants en informatique et leurs ont demandé d'installer un logiciel afin de pouvoir récupérer les fichiers d'évènements. L'objectif était de voir ce que les gens faisaient en fonction du type de tâches réalisées et pour cela les usagers devaient définir de quelles tâches il s'agissait. En se basant, en partie, sur la catégorisation de Byrne *et al.*, (1999) citée plus haut, les chercheurs ont observé que peu importe la tâche, toutes étaient initiées par le fait de taper une requête. Toutefois des différences marquées ont été observées. Par exemple, l'utilisation des favoris était surtout observée dans les tâches de transactions, tout comme l'ouverture d'une nouvelle fenêtre. Ou encore, l'utilisation du retour arrière se retrouvait surtout dans des tâches de collecte d'informations. Les chercheurs ont pu conclure à un effet de la tâche sur le comportement de navigation et à l'importance de connaître le contexte d'une recherche pour comprendre ce que les gens font.

Internet offre donc une multitude de possibilités pour les utilisateurs afin de répondre à leurs besoins et constituer ainsi un véritable progrès social et technologique.

La partie suivante présente les principaux modèles de recherche d'information dans le domaine des hypermédias et tentera de mettre en exergue les principaux facteurs individuels qui peuvent influencer sur la recherche.

## **2.1.2 Les principaux modèles de recherche d'information**

La recherche d'information part d'une prise de conscience d'un besoin information pour pouvoir réaliser une tâche ou en optimiser la réalisation. Il s'agit donc point d'entrée dans l'activité de recherche et à partir des informations qu'il souhaite compléter, l'utilisateur va se construire une représentation mentale du but à atteindre.

Guthrie et ses collègues (Guthrie, 1988; Guthrie & Mosenthal, 1987 ; Guthrie, Briten & Barker, 1991) sont souvent considérés comme les premiers à avoir étudié de manière expérimentale les processus mentaux impliqués en recherche d'information (Tricot, Demarcy & Boussarghini, 1998).

Leurs études concernaient alors les environnements « traditionnels » papier (manuels, notices). Selon ces auteurs, rechercher des informations passe par cinq étapes successives :

- La formation d'un but : construction d'une représentation mentale du but à atteindre
- La sélection d'une source d'information : celle paraissant la plus adaptée pour accéder à l'information recherchée
- L'extraction de l'information : l'individu traite l'information contenue dans le document
- L'intégration aux connaissances antérieures des informations pertinentes trouvées dans le dit texte ;
- Le recyclage : évaluation de l'objectif et si elle n'est pas satisfaisante, l'individu recommence à la première étape jusqu'à ce que le but soit atteint.

Par la suite, ce modèle a été enrichi par Rouet et Tricot (1995, 1998). Ces auteurs soulignaient une trop grande linéarité du modèle initial qui ne permettait pas de considérer les activités de recherche dans des documents complexes ou à but moins précis. L'objectif de ce modèle était donc de décrire à la fois des recherches d'informations simples mais aussi des recherches plus complexes dans lesquelles le but est complété au fur et à mesure de la recherche.

Le cycle de traitement est constitué de trois phases principales : la sélection de l'information, le traitement de l'information sélectionnée et l'évaluation de la pertinence en fonction du but visé par la personne. Par ailleurs, dans ce modèle, les activités de sélection et d'évaluation vont avoir un rôle particulier car elles représentent un contrôle de l'activité (planification de la recherche et mesure de la distance état actuel / but) et un processus de traitement des informations « interactionnelles » (liens, menus, boutons). Ce double processus est en lien avec la représentation que se fait le sujet de la tâche, cette représentation pourra être modifiée dynamiquement tout au long de l'activité. Par ailleurs, dans ce modèle la

représentation du but va être comparée à la représentation du contenu traité. Cette comparaison entraînera un jugement de concordance entre ces deux représentations et entraînera une décision selon trois modalités :

- L'information trouvée correspond exactement à la représentation du but, le but semble atteint donc l'activité est arrêtée
- L'information ne correspond que partiellement à la représentation du but, le but est partiellement atteint donc l'activité de sélection est poursuivie
- L'information traitée ne correspond pas à la représentation du but, le but n'est pas atteint donc un changement de stratégie de sélection sera envisagé.

Les chercheurs précisent qu'à cette prise de décision « rationnelle » une variabilité liée à l'humain pourra se rajouter et elle se traduit par une évolution de la représentation du but. En effet, avec les nouvelles informations traitées, la représentation pourra s'enrichir et permettre au but de devenir plus précis. Cependant, cette caractéristique peut amener une difficulté à maintenir une représentation stable du but, et par conséquent une désorientation de l'utilisateur. Foss (1988) parle de phénomène de noyade en digression. En étant plus généraliste (Dinet & Tricot, 2008) ce modèle permet d'expliquer la recherche d'information en tenant compte des processus de régulation de l'activité en permettant ainsi de présenter un plus large panel de RI que le modèle sur lequel il se base.

Ces deux modèles présentés qui n'ont pas la prétention d'être exhaustif, tentent de décrire les comportements et processus mentaux sous-jacents impliqués lors d'une RI. Parallèlement à ces modèles plutôt descriptifs, on retrouve d'autres auteurs qui élaborent des modèles visant à simuler et prédire ces processus mentaux et comportementaux.

Kitajima et ses collaborateurs (Blackmon, Mandalia, Polson & Kitajima, 2007 ; Kitajima, 2003, 2008 ; Kitajima, Blackmon & Polson, 2005) ont développé l'un de ces modèles prédictifs ; le modèle CoLiDeS (Comprehension-based Linked model of Deliberate Search ).

CoLiDeS est un modèle qui s'intéresse plus particulièrement à la manière dont les pages des sites web peuvent être traitées. Les chercheurs expliquent que toutes les actions qui peuvent être réalisées par un individu à partir d'une page Web sont sous l'influence de deux processus :

- Les processus attentionnels : l'utilisateur va partitionner la page en zone
- Les processus de sélection d'action : l'utilisateur se concentre sur certaines de ces zones.

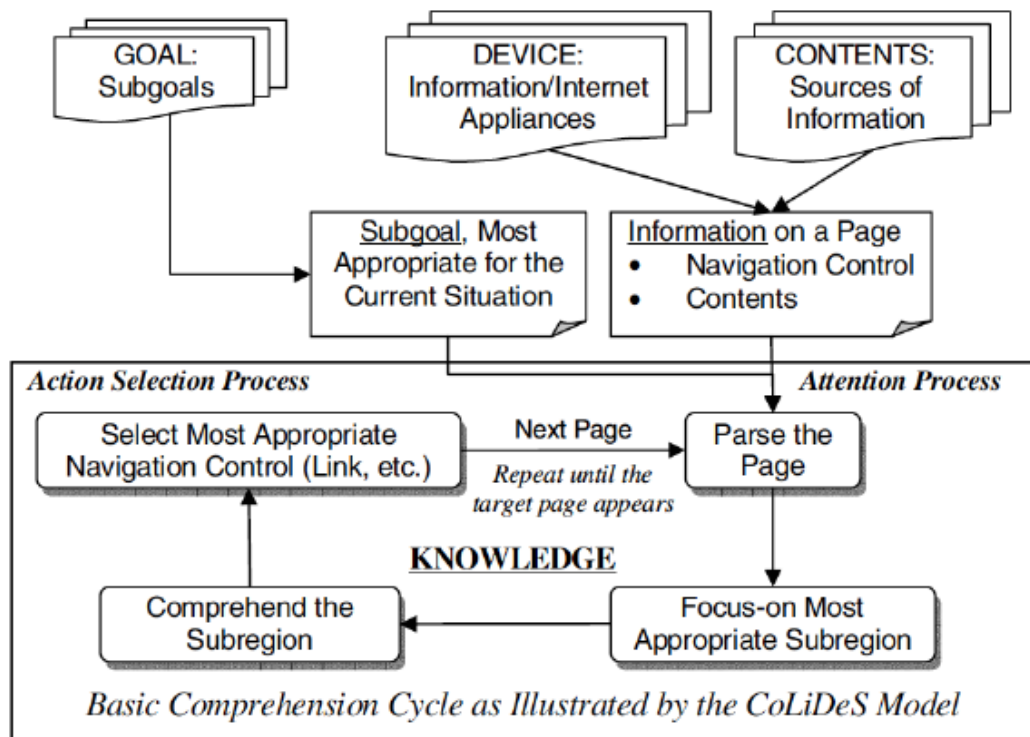


Figure 1 : Schéma du modèle CoLiDeS (Kitajima et al., 2005)

À partir de ces deux processus, la recherche d'information peut être décrite en actions :

- Sur la base de ces connaissances et de la question ou consigne, l'individu se construit une représentation mentale du but.
- Il analyse la structure de la page web et les caractéristiques des contenus (emplacement, typographie, etc.). Il s'agit d'une analyse de forme.
- L'individu segmente la page en zone de contenu en regroupant les différents éléments sémantiquement liés. Cette analyse se fait à partir d'un apprentissage explicite ou par expérience. Il s'agit d'une analyse de fond.
- Après avoir focalisé son attention sur la partie de la page qui lui semble la plus pertinente, l'individu va tenter de comprendre le contenu de cette zone.
- Si la zone est jugée pertinente et qu'elle correspond à la représentation mentale de son but, l'individu va sélectionner le lien qui lui paraît le plus adéquat en cliquant dessus.

Tout comme les deux précédents modèles, le CoLiDeS suppose la présence conjointe de processus ascendants et descendants. D'une part, l'individu va se baser sur les caractéristiques de la source d'information, et d'autre part utiliser ses connaissances (procédurales et du domaine) pour prendre des décisions. Avec ce modèle, la recherche d'information et la navigation sont donc étroitement liées.

Le modèle SNIF-ACT (Scent-based Navigation and Information Foraging in the Act cognitive architecture) établi par Pirolli et Fu (Pirolli, Fu, Reeder & Card 2002 ; Pirolli & Fu 2003 ; Fu & Pirolli 2007) vise également à prédire les comportements de recherches d'information, et les auteurs précisent qu'il peut s'agir de RI non

familiales. Ce modèle relève d'une approche écologique qui se fonde sur la théorie de la fouille d'information (« *information foraging* ») (Pirolli & Card, 1999) et la notion de parfum ou trace d'information (« *information scent* »). Cette notion est basée sur un principe économique de rapport du coût de la recherche et du traitement de l'information par rapport au bénéfice escompté de cette information pour la réalisation d'une tâche. L'engagement dans une RI est donc fonction de la valeur de l'information recherchée. Les chercheurs se basent ici sur la métaphore de la recherche alimentaire.

Lors d'une recherche d'information, l'utilisateur doit donc faire des choix concernant les interfaces qui donnent accès à l'information et juger de la pertinence des options de parcours qui s'offrent à lui par rapport à la concordance estimée avec son but. Il estimerait de manière probabiliste quels sont les liens ayant le meilleur « parfum », c'est à dire ceux qu'ils estiment les plus riches et les plus accessibles en terme de traitement de l'information. Cette analyse rationnelle est une extension de la théorie ACT-R (Anderson, 1990). Le schéma suivant présente le modèle SNIF-ACT.

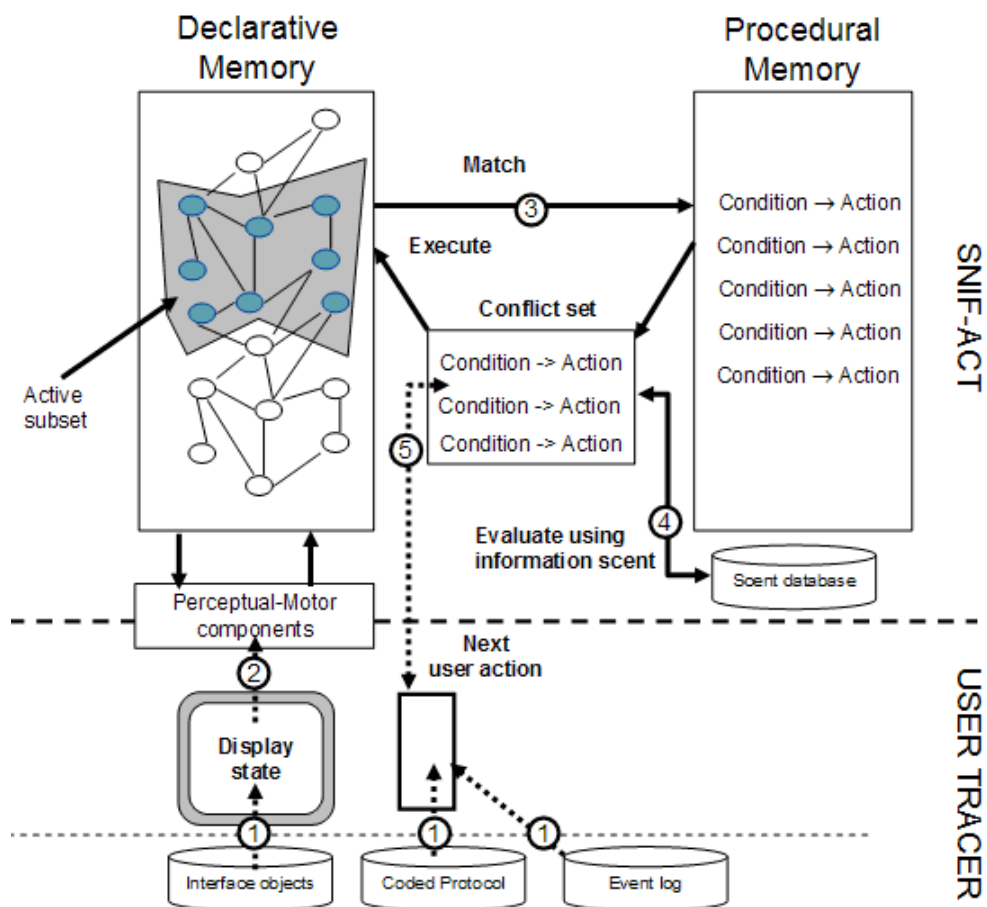


Figure 2 : Schéma du modèle SNIF-ACT (Fu & Pirolli, 2007)

La majorité des modèles s'intéressant aux processus mentaux impliqués lors d'une recherche d'information dans les environnements numériques sont censés s'appliquer à tous les types d'environnements électroniques. Cependant, certains chercheurs ont décidé de s'intéresser en priorité aux comportements des utilisateurs face à l'utilisation d'un moteur de recherche (Pak & Price, 2008 ; Sharit, Hernandez, Czaja & Pirolli, 2008).

Sharit et ses collègues ont mené une étude dans laquelle des participants de trois groupes d'âge (jeunes adultes, seniors jeunes et seniors âgés) devaient réaliser des tâches de recherche d'information de complexité différente. Les résultats ont souligné que les connaissances antérieures avaient un effet sur la performance, mais que la variance obtenue ne s'expliquait pas qu'au travers de ce facteur. Plus précisément, pour les RI complexes où le participant devait formuler correctement différents termes de recherche, les chercheurs ont souligné l'effet combiné des connaissances et de certaines capacités cognitives (mémoire de travail, raisonnement). L'argument soutenant les recherches de cette équipe met en exergue que les comportements et performances liés à l'utilisation d'un moteur de recherche vont déterminer, en partie, la réussite ou l'échec d'une recherche.

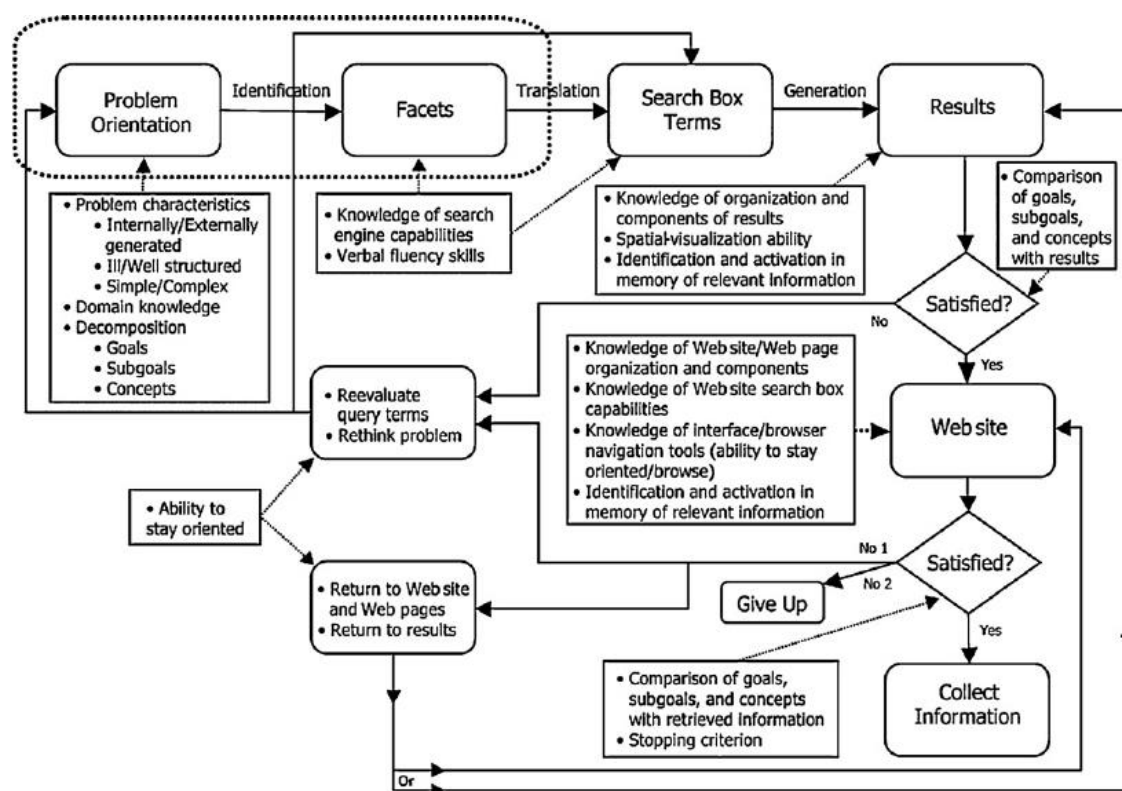


Figure 3 : Schéma du modèle de recherche d'information par moteur de recherche ( Sharit et al., 2008)

Les auteurs ont proposé à la suite de cette étude un modèle dans lequel une recherche d'information est assimilée à une « résolution de problème flou » pour laquelle l'individu va rencontrer trois étapes :

- Représentation mentale du problème à résoudre ; cette représentation est sous l'influence des connaissances du domaine et de la nature du besoin informationnel
- Planification en vue d'atteindre l'objectif
- Exécution ; la personne opérationnalise la procédure qu'elle a planifiée. Il s'agit ici d'interroger le moteur de recherche avec des mots clefs.

Ce cycle est itératif jusqu'à ce que le but final soit atteint ou si l'individu décide de reformuler une autre requête. Les chercheurs précisent également ici que, comme pour les autres modèles de RI, d'autres connaissances peuvent influencer sur la navigation dans les systèmes d'information.

### **2.1.3 Les principaux facteurs individuels impliqués dans la recherche d'information**

Selon Tricot et Rouet (2004), ces principaux modèles de recherche d'information nécessitent pour l'utilisateur d'avoir certaines connaissances.

Principalement, on recense trois grands types de connaissances :

- Des connaissances sur le contenu recherché. En effet, pour se représenter l'objectif de la recherche et interpréter les informations recherchées, il est important d'avoir des connaissances dans le domaine en question.
- Des connaissances dans le domaine de l'information ou « connaissances méta-documentaires ». Il s'agit là de connaître les ressources les plus efficaces pour trouver les informations souhaitées.
- Des connaissances plus spécifiques relatives à l'utilisation du système d'information et à son mode opératoire (index, table des matières, moteur de recherche, par exemple).

La partie suivante présente plusieurs études ayant tenté de faire la lumière sur ces questions relatives à l'expertise. Cependant, nous verrons aussi que le facteur humain ne peut se limiter à cette variable et que d'autres caractéristiques individuelles sont impliquées (Chen & Macredi, 2010).

#### **2.1.3.1 L'expertise**

Navarro-Pietro, Scaife et Rogers (1999) font partie des premiers chercheurs à s'être intéressés au lien pouvant exister entre l'expertise des utilisateurs et la recherche d'information sur le web. Leur attention s'est portée sur 23 volontaires dont 10 étudiants en Sciences Informatique avec deux ans d'expérience sur le web et 13 étudiants en Psychologie avec un an d'expérience. L'objectif de cette étude exploratoire était de voir si des différences s'observaient entre ces deux groupes. En combinant des questionnaires concernant la recherche d'information sur le web, des observations lors de leur navigation et des interviews, les auteurs ont identifié trois stratégies de navigation.

- Stratégie « *top down* », les utilisateurs recherchent l'information dans une zone générale puis affinent en fonction des liens sur lesquels ils tombent, jusqu'à trouver l'information spécifique désirée.
- Stratégie « *bottom-up* », en contraste avec la stratégie citée juste au-dessus, les usagers utilisent directement les bons mots clés dans le moteur de recherche et scrollent afin de faire défiler les résultats. Ils consultent les liens et reviennent sur la page de résultat si l'information n'est pas la bonne. Cette stratégie a été le plus souvent utilisée par les participants expérimentés.
- Stratégie « mixte », une majorité des participants combinent les deux stratégies, il s'agit ici d'ouvrir de multiples fenêtres afin de faire plusieurs recherches en parallèle.

Chevalier et Kicka (2006) ont également réalisé une étude afin d'étudier l'impact du niveau d'expertise et du type de site web consulté (avec une qualité ergonomique différente) sur trois activités distinctes lors de la navigation. Comme pour l'étude de Holscher et Strube (2000), un individu était considéré comme un expert s'il surfait en moyenne pendant une heure sur le web. Les principaux résultats ont montré que les utilisateurs considérés comme experts recherchaient et utilisaient plus le menu de navigation et ceci d'autant plus si le site web souffrait d'une mauvaise qualité ergonomique. Au contraire, les participants novices effectuaient d'avantage une lecture linéaire du site en utilisant plus les hyperliens. Enfin, l'étude a relevé une différence significative entre les experts et novices concernant l'exploration des pages web. Les participants experts (comprenant des concepteurs web) exploraient plus fréquemment les pages que les utilisateurs novices et ceci indépendamment de la qualité du site web.

Ces résultats étaient en accord avec ceux effectués auparavant par Holscher et Strube (2000) qui avaient mis en relation deux types d'expertises ; expertise dans le Web et expertise dans le domaine de la recherche (l'économie). En comparant 12 participants répartis chacun dans un groupe en fonction de leurs compétences ; groupe double experts, groupe double novices, groupe novice-web expert, groupe expert-web novice, les résultats ont montré que les doubles experts résolvaient plus de tâches sur le web que les autres tout en passant moins de temps sur les pages consultées. Ce groupe était également moins enclin à revenir en arrière. Par ailleurs, les chercheurs ont montré que les experts dans le domaine effectuaient des requêtes moins longues laissant l'idée que plus on a de connaissances sur le contenu de notre recherche et plus nous serions capable de choisir les termes qui correspondent le mieux à ce que l'on recherche. Thatcher (2008) comparera les novices et leur façon de faire à des « touristes virtuels ». Aula (2003) et Aula, Jhaveri et Kaki (2005) préciseront également que les experts dans le Web font une plus grande utilisation des opérateurs booléens et qu'ils réitèrent plus souvent leurs requêtes lorsque les informations souhaitées ne sont pas trouvées (Aula & Kaki, 2004).

Plus récemment, White, Dumais et Teevan (2008, 2009) ont approfondi ces conclusions pour différencier les types de requête faites sur un moteur de recherche en étudiant uniquement l'effet de l'expertise dans le domaine du médical, de la finance, du droit et de l'informatique. Leurs conclusions sont similaires aux précédents



travaux effectués dans le domaine et les chercheurs proposent un modèle afin de prédire automatiquement l'expertise dans un domaine.

Nous pourrions citer encore d'autres travaux comme ceux de Lazonder, Biemans et Worpeis (2000) ou encore Palmquist et Kim (2000) qui ont abouti aux mêmes résultats en mettant en avant que les experts sont bien souvent plus efficaces, et rapides. Ils effectuent des requêtes plus ciblées et réussissent plus de tâches que les novices. Toutefois, comme Aula et Norhaussen (2006) le soulignent, il convient d'être assez prudent avec ces études car certaines présentent des résultats controversés qui pourraient s'expliquer en partie par la notion « d'expert ». Chacune des études citées aborde à sa manière cette notion, et il ne paraît pas y avoir de consensus, ce qui en fait une limite à prendre en compte. Brand-Gruwel, Wopereis et Vermetten (2005) retrouvent peu de différences entre ces deux populations qui ont été constituées par des étudiants en dernière année de doctorat de Technologie (groupe appelé les experts) et des étudiants en psychologie (groupe considéré comme les novices). Afin de contourner cette limite, Aula et Norhaussen recommandent d'utiliser le TCS (Task Completion Speed) qui calcule un score d'efficacité ((nombre de tâches réalisées / temps total) x 3600). Les chercheurs partent du postulat que l'expérience ne fait pas nécessairement un expert et donc que la seule prise en compte de l'expérience ne doit pas être utilisée pour diviser les gens en experts et novices. L'expertise est donc ici traitée comme une qualité évolutive et non plus comme un état binaire à deux valeurs (expert / novice).

Certaines recherches ont été plus loin en suggérant que les individus avec différents niveaux d'expertises n'avaient pas les mêmes préférences concernant la structure de l'information. Calisir et Gurel (2003) ont étudié l'impact de trois types de structure de texte ; linéaire, hiérarchique et mixte (structure hiérarchique combinée avec des liens référencés) en relation avec les connaissances ou non dans le domaine de la productivité. Les résultats montrent que la structure hiérarchique est plus appropriée pour les usagers qui n'ont pas de connaissances car elle permet de fournir une vision claire du contenu et de l'organisation du sujet. Enfin, d'autres études comme celle de Potelle et Rouet (2003) ont voulu affiner les résultats et tenté de savoir si l'expertise pouvait aussi impacter l'usage de certaines aides à la navigation. Trois versions d'un même système d'information ont été testées ; avec une carte hiérarchique, une liste alphabétique ou une carte conceptuelle. 47 élèves avec différents niveaux de connaissances du domaine ont pris part à l'expérience et ont été assignés à utiliser l'une des trois versions. Les résultats ont mis en évidence la supériorité de la carte hiérarchique permettant de favoriser une meilleure compréhension du contenu, y compris par les individus n'ayant aucune connaissance au préalable. Plus récemment, Minetou, Chen et Liu (2008) ont abouti à des résultats similaires.

### **2.1.3.2**

#### **L'âge**

À ces facteurs classiquement cités, peuvent venir s'en ajouter d'autres comme l'âge. Chin, Fu et Kannampallil (2009) et plus récemment Chin et Fu (2010) ont observés que des stratégies de recherche différentes peuvent apparaître entre des individus jeunes (entre 18 et 28 ans) et âgés (entre 60 et 77 ans). En demandant aux participants de réaliser des tâches sur deux versions d'un site d'information médical, les chercheurs

ont remarqué que deux stratégies influençaient les performances de recherche en fonction du type de tâche demandée (consigne avec un but précis et consigne avec un but moins bien défini).

- La stratégie « *bottom up* » (interface driven strategies) guidée par l'interface optimiserait les usagers dans des tâches de recherche bien définies. Ce pattern se décrit par une certaine rapidité pour parcourir les pages web et prendre une décision avant de cliquer sur un lien, de cliquer sur plus de liens et de naviguer davantage entre les liens d'une même catégorie (par exemple les rubriques douleur musculaire et douleur articulaire et osseuse constituaient une même catégorie).

Selon Chin, ce type de stratégie demande une bonne maîtrise du système d'information et de son mode opératoire et serait plus souvent observé chez les usagers jeunes.

- La stratégie « *top down* » (knowledge driven strategies) guidée par les connaissances serait plus efficace pour les individus dans le cas où les tâches de recherche seraient plus floues. Ce type de stratégie conduit les utilisateurs à prendre un long moment pour cliquer sur un lien et parcourir une page web, de naviguer davantage entre les liens de rubriques différentes et de se centrer uniquement sur un ensemble restreint d'hyperlien.

### 2.1.3.3

#### **Le genre**

Depuis le début des années 1980, la littérature dans le domaine des IHM s'intéresse aux différences de genre et a reconnu le sexe comme une variable influençant les compétences informatiques. De façon générale, les recherches suggèrent que les hommes déclarent des niveaux d'anxiété inférieurs que les femmes, et que les hommes atteignent plus facilement leurs objectifs de recherche que les femmes (Karavidas, Lim & Katsikas, 2004)

Large, Behesthi et Rahman (2002) ont réalisé une étude afin d'observer s'il existe des différences en termes de styles de navigation en fonction du genre. Leur échantillon était composé de 23 garçons et 30 filles répartis dans 6 niveaux de classes différents. Les résultats ont montré qu'en général, les garçons, comparativement aux filles, sont plus actifs sur le Web, formulent plus de requêtes et cliquent sur plus de liens hypertextes par minute. De plus, l'étude montre que ce sont eux qui passent le moins de temps à regarder les pages et qui, par conséquent, en voient le plus.

Ces résultats sont similaires avec ceux de Roy, Taylor et Chi (2003), cités par Chen et Macredie (2010) qui ont réalisé une étude sur un échantillon de 14 étudiants comportant autant de filles que de garçons. Les résultats ont montré qu'en général les garçons voient plus de pages en une minute, effectuent plus d'aller-retour entre les pages, effectuent plus de requêtes sur le moteur de recherche, sauvegardent plus d'informations tout en passant moins de temps sur les pages que les filles. Les chercheurs en ont conclu que les garçons naviguent de façon non linéaire. À l'inverse, les filles auraient tendance à parcourir un ensemble de documents liés entre eux et à adopter une approche plus linéaire. Zhou (2014) est aussi parvenu à cette conclusion.

À l'inverse, Protosaltis et Bouki (2008) ont mené une étude afin d'étudier un aspect bien précis de la navigation et qui concerne l'utilisation des liens hypertextes. Cette recherche comptait 99 participants dont 27 femmes, tous issus de la filière informatique. L'expérimentation avait lieu en laboratoire. Les sujets devaient naviguer sur un site qui contenait 136 liens, lire le texte autant de fois qu'ils le souhaitent et une fois qu'ils se sentaient prêts, devaient quitter le poste informatique pour répondre à des questions. Plusieurs variables dépendantes étaient prises en compte ; un score concernant la compréhension du document, le temps de lecture, le nombre d'hyperliens visités, la cohérence de ces hyperliens et précisément si les deux liens faisaient partie du même contexte, s'il s'agissait ou non d'un lien parent vers un lien enfant ? Les chercheurs avaient émis l'hypothèse que ces variables allaient être affectées par le genre. Après analyse, les résultats n'ont montré aucune différence significative. Les chercheurs ont donc conclu que le genre n'avait pas d'effet sur les variables étudiées.

Selon ces mêmes chercheurs, les études essayant de refléter des différences entre hommes et femmes dans la navigation sont souvent controversées. Les résultats issus de ces recherches peuvent dépendre en partie de la population étudiée et notamment du niveau d'expérience avec l'outil informatique et du niveau d'étude de la population.

#### **2.1.3.4 Le profil cognitif**

Au même titre que les autres facteurs, la littérature relève que le style cognitif joue également un rôle dans la recherche d'information.

Le concept de style cognitif a été défini, au début des années 1990, comme étant « une tendance pour un individu d'adopter un type particulier de stratégie » (Ford, Wood, & Walsh, 1994) et ce, « lors du traitement d'une ou plusieurs informations » (Ford, Wilson, Foster, Ellis, & Spink, 2002). En d'autres termes, il s'agit de l'approche qu'un individu adopte, habituellement, pour organiser et se représenter les informations lui parvenant (Riding & Rayner, 1998). La littérature précise qu'il n'existe pas de bon ou mauvais profil cognitif. Toutefois certains styles semblent permettre de résoudre des tâches plus facilement que pour d'autres (Wang, Hawk & Tenopir, 2000).

Dans ce contexte, le style « *Field Dépendence* » (FD), ou « *Field Independence* » (FI) émerge comme l'un des facteurs les plus étudiés. Il se réfère au degré auquel la perception ou la compréhension de l'information par l'utilisateur est dépendant ou non du champ perceptif ou contextuel environnant (Chen & Macredie, 2010).

Les études portant sur les préférences d'apprentissage ont soulevé des différences concernant ces deux styles cognitifs. Il s'agit ici des préférences de l'apprenant lorsqu'il se retrouve dans un environnement ou des activités de formation (Lee, McCool & Napieraski, 2000).

Par exemple, Ford et Chen (2000) ont examiné les effets du style cognitif sur les préférences de navigation de 65 étudiants. Afin d'identifier les styles cognitifs, le Riding Cognitive Style (CSA, Riding, 1991) a été utilisé. Les résultats ont montré que ceux ayant un profil « *Field Independence* » (FI) préfèrent utiliser l'index

alphabétique pour localiser des informations spécifiques alors que les FD utilisent plutôt la carte hiérarchique pour avoir une vision globale et complète. Plus récemment, Lin et Chen (2008) ont observé le même résultat. Avec un échantillon plus important comptant 101 individus, les chercheurs ont fait passer le CSA, cité plus haut, afin d'identifier les styles cognitifs. Les résultats montrent que les individus présentant un style FI utilisent davantage l'index alphabétique et moteur de recherche, tandis que les participants FD préfèrent l'utilisation de la carte du site.

Par ailleurs, d'autres travaux montrent des relations entre la façon de rechercher de l'information et le style FI/FD. Wang, Hawk et Tenopir (2000) ont réalisé une étude en laboratoire avec 24 étudiants qui devaient répondre à deux questions en cherchant l'information sur le web. Afin d'identifier à quel type de profil cognitif chaque individu dépendait, les chercheurs ont fait passer le « *Embedded Figure Test* » (EFT, Witkin, 1971). Il s'agit dans un premier temps de présenter au participant une figure simple à retenir, puis lui représenter mais dans une figure plus complexe. La tâche est de localiser la figure simple dans la figure complexe. Le test compte 12 figures et il est chronométré afin d'obtenir un score. Les résultats montrent une corrélation significative entre le score au test, le temps passé pour faire les tâches et le nombre de pages visitées. Les participants avec le score EFT (FD) le plus élevé mènent des sessions de recherches plus longues, et visitent plus de pages que ceux avec un score plus faible. Une autre étude pertinente est celle de Palmquist et Kim (2000) dans laquelle les chercheurs ont étudié les effets du style cognitif et de l'expérience sur les performances de recherche. L'EFT a été administré aux 48 étudiants afin de définir leur style cognitif. La performance a été étudiée à travers deux variables ; le temps mis pour trouver l'information demandée sur le site de l'université et le nombre de nœuds consultés avant de trouver les informations pertinentes. Les chercheurs observent que les individus avec un style FD présentent une meilleure performance. Frias-Martinez, Chen et Liu (2008) ont observé comment 50 individus naviguaient sur une librairie digitale. Ils ont également identifié des différences. Les résultats ont montré que les FD utilisent plus souvent les boutons retour et avancer, alors que les FI utilisent plus les fonctions de recherches avancées.

Enfin, plus récemment nous pouvons citer l'étude de Belk (2013) qui s'est intéressé au lien entre le style cognitif des individus et à leur façon de chercher de l'information sur le site de Wikipedia. 106 participants âgés de 17 à 25 ans ont participé à cette recherche réalisée en laboratoire. En se basant sur le test CSA de Riding, les chercheurs ont distingué les participants plutôt « analytiques » (FI) et « holistiques » (FD). Concernant l'interaction avec le site, deux indicateurs ont intéressé les chercheurs ; l'interaction avec le menu et l'interaction avec les liens. Dans un premier temps, une classification a été réalisée afin d'identifier les groupes d'utilisateurs présentant les mêmes stratégies de navigation ; linéaire, non linéaire. Dans un second temps, les chercheurs ont regardé si les individus étant identifiés dans le même groupe présentaient le même profil cognitif. Les résultats ont infirmés cette idée, les profils cognitifs étaient variables. Belk et son équipe présentent cette étude comme exploratoire et recommandent de continuer les travaux dans ce sens en affinant l'analyse des comportements de navigation.

Toutefois, comme Huang (2003) et Bousbia, Rebia, Labat et Balla (2010) le précisent, il serait pertinent d'étendre ces travaux et de ne pas se limiter uniquement à l'analyse du style (FI/FD) car en fonction des tâches et du protocole d'autres déterminants personnels seraient aussi susceptibles de nous expliquer la navigation.

## 2.2

### **LES METHODES DE RECUEIL DEDIEES A L'ETUDE DES COMPORTEMENTS SUR LE WEB**

---

La littérature présente de nombreuses recherches qui ont voulu étudier la façon dont les internautes naviguent sur le Web et quels comportements ils adoptent en conséquence (Milic-Frayling, Jones, Rodden, Smyth, Blackwell & Sommerer, 2004). En effet, l'utilisation des sites web peut s'avérer difficile (Chevalier et Kicka 2006, Ling et van Schaik 2006, van Deursen et van Dijk, 2009) et certains chercheurs estiment que la moitié des recherches d'informations sur le web restent infructueuses (Nielsen 2000, Gwidzka et Spence, 2007).

Dans le but d'appréhender ces comportements et stratégies de navigation, différentes méthodes de recueils de données existent telles que les questionnaires, les tests utilisateurs, ou l'oculométrie par exemple (Hsieh-Yee, 2001).

En se basant sur Jamet, Lemercier et Février (2007) et Jamet et Février (2008) distinguant les techniques « off-line » utilisées à posteriori dans le but d'évaluer la charge mentale ou les performances d'apprentissage sous forme de questionnaire, et les méthodes dites « online » qui permettent de recueillir les traces des comportements des utilisateurs de manière précise (Jamet, Lemercier & Février, 2008), le tableau suivant met en exergue les principales méthodes qui ont pour objectif de comprendre la navigation.

Nous les avons distingué en fonction de deux variables :

- Le type de données ; s'agit-il de données déclaratives ou comportementales,
- Et, ce que nous avons appelé la temporalité ; est ce que l'observation est directe et permet d'avoir un feedback en temps réel ou s'agit-il d'une technique d'observation « statique ». (Cockburn, McKenzie & JasonSMith, 2002).

Par ailleurs, précisons que certaines de ces méthodes peuvent être utilisées simultanément.

		Type de données	
		Données comportementales	Données déclaratives
Temporalité	Observation dynamique (en direct)	- Oculométrie - Test utilisateur	- Think aloud
	Observation statique (en différée)	- Web mining	- Questionnaire - Entretien

Tableau 1 : Principales méthodes pour le recueil de données sur la navigation

## 2.2.1

### Les techniques d'enregistrements directes

#### 2.2.1.1

#### L'analyse des données comportementales

##### ***Le test utilisateur***

Le test utilisateur appelé aussi « *user testing* » fait partie des méthodes d'analyses requérant la participation directe des utilisateurs.

Cette approche empirique et expérimentale consiste à observer et enregistrer les comportements d'un ou plusieurs utilisateurs dans des contextes « contrôlés » dans lesquels les participants sont invités à réaliser des tâches réelles ou encore explorer librement un site avec une passation standardisée et supervisée en temps réel. (Tullis & Albert, 2008 ; Albert, Tullis & Tedesco, 2010).

Dans ce type de situation, on s'intéresse aux performances et aux comportements des utilisateurs lors d'interaction avec le site Web. On mesurera notamment, le temps requis pour l'exécution d'une tâche, le nombre d'erreurs, les liens sélectionnés, le lien emprunté, etc. Il s'agit donc, à partir d'indices comportementaux d'identifier des choix de conception (par exemple le positionnement et l'intitulé des liens, l'existence ou non de plans de sites) qui pourraient entraîner des erreurs d'utilisation, susciter des interrogations ou encore alourdir la recherche d'information. Les résultats de ces tests peuvent aussi permettre de comparer plusieurs versions d'un même site en termes d'efficacité ou de préférences. Dans tous les cas, cette façon permet de mieux comprendre ce que font les internautes et de mieux les connaître (Bastien, Drusch & Dinet, 2010).

Les sessions issues des tests centrés utilisateurs sont capturées au moyen d'enregistrements vidéo qui permettent de visualiser et de coder les comportements des utilisateurs en terme de temps de réalisation des tâches par exemple. Afin d'effectuer le codage, il est nécessaire d'utiliser certains équipements spécifiques (logiciel « Morae », « Evalyzer » par exemple). Ces équipements permettent

d'identifier le commencement et la fin des tâches mais également des comportements que présentent les utilisateurs. Des statistiques descriptives ou des analyses de séquences sur les comportements des utilisateurs peuvent aussi être obtenues par le biais de ces outils. Dans d'autres cas, l'enregistrement des comportements utilisateurs est automatisé et peut se traduire, par exemple, par la sélection de certaines options dans les menus. De plus, ce type de comportements est défini de manière temporelle.

Toutefois, précisons que depuis le début des années 1990 avec l'émergence du Web, l'idée de réaliser des tests utilisateurs à distance est apparue. Andreasen, Nielsen, Schroder et Stage (2007) définissent les tests utilisateurs à distance comme une « situation dans laquelle les évaluateurs sont séparés spatialement et/ou temporellement de l'utilisateur » (Drusch, Bastien, Becker & Rodriguez, 2012).

Même si ces tests sont réalisés à distance, ils sont perçus comme plus écologiques car les modalités de passations se rapprochent des conditions réelles d'utilisation en termes de matériel et contexte d'usage. Toutefois, dans ce contexte, étant donné qu'aucun contact téléphonique ou visuel ne peut être établi au moment de la passation, il est plus difficile de pouvoir recueillir des informations subjectives sur le ressenti des utilisateurs au cours de la tâche. Cependant, ce type de méthodologie a l'avantage d'être réalisé dans le contexte naturel des personnes, c'est-à-dire à leur domicile et ainsi éviter de les perturber. Lors de ces tests des oculomètres peuvent aussi permettre d'appréhender différemment les comportements des internautes.

### ***L'oculométrie***

L'oculométrie (appelé aussi « *eye tracking* ») est une technique de recueil de données visuelles qui est habituellement utilisée en laboratoire. Le dispositif appelé oculomètre permet de suivre en temps réel ce que regarde un internaute sur un écran d'ordinateur ou une page Web, et ainsi de connaître la position d'une fixation visuelle sur un espace donné. Pour se faire, la technique la plus utilisée est celle du reflet cornéen. Elle consiste à envoyer un faisceau de lumière infrarouge invisible à l'œil, filmé par une caméra. Ce faisceau se reflète sur la rétine rendant la pupille lumineuse et permettant d'en délimiter les contours. 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. En effet, sur Internet, la finesse des données recueillies par cette méthode permet de savoir, d'une part, où regardent exactement les participants, et par conséquent de savoir quelles sont les zones les plus (ou les moins) explorées, et d'autre part, d'identifier les comportements oculaires erratiques, indicateurs d'une désorientation de l'utilisateur.

Dans le cadre de l'étude des comportements sur le web, différentes problématiques existent : l'étude de l'impact de l'organisation des contenus Web sur les comportements oculaires (Lorigo, Haridasan, Brynjarsdottir, Xia, Joachims, Gay, Granka, Pellacini & Pan 2008), l'impact du type de site Web (Habuchi, Takeuchi & Kitajima, 2007), l'impact de variables individuelles (genre, âge, tâches, etc.) (Pan 2004).

## 2.2.1.2

### L'analyse des données déclaratives

#### ***Le think aloud***

Les protocoles de « *think aloud* », appelés également protocole de pensée à voix haute sont l'un des principaux outils utilisés par les professionnels de l'utilisabilité afin d'observer et d'analyser ce que l'individu pense au cours de la navigation et de sa recherche d'information (Hsieh-Yee, 2001). Cette technique permet d'accéder aux informations qui ne sont pas directement observables, telles que les difficultés ressenties (Hawala, Murphy, Ashenfelter, & Kathleen 2010). Selon Ericsson et Simon (1980) qui sont à l'origine de cette technique, les verbalisations que fait le participant peuvent être catégorisées en deux niveaux. Le premier niveau fait appel à des informations contenues dans la mémoire à court terme. Le deuxième niveau concerne les informations dites « nulles ». Elles sont contenues dans la mémoire à long terme et selon les chercheurs ces verbalisations ne sont pas naturelles mais le résultat d'une redirection par rapport à ce que le participant faisait et pensait à l'origine.

Un environnement expérimental et contrôlé permet donc d'observer les utilisateurs, dans le cadre de tâches de recherches d'informations prédéfinies, sur un ou plusieurs sites donnés. Néanmoins, comme le souligne Kim et Allen (2002), il convient d'être prudent avec ces types de méthodes qu'il qualifie « d'artificiel » car elles sont susceptibles de biaiser les comportements de navigation des participants. En effet, le fait que des tâches soient dictées par un expérimentateur peut influencer sur sa motivation en l'amenant à être plus performant, par exemple, pour finir sa tâche.

Toutefois, il n'est pas toujours possible d'étudier et d'analyser l'interaction web en laboratoire de façon directe et de recueillir un feedback immédiat au moment même où l'utilisateur réalise sa tâche. Certaines méthodes permettent donc d'enregistrer et d'étudier le comportement de navigation de façon différée.

## 2.2.2

### Les techniques d'enregistrements différées

#### 2.2.2.1

#### L'analyse des données comportementales

##### ***Le Web mining***

À ce niveau, nous parlons d'observations statiques, les chercheurs analysent les traces et les actions d'exploration libre des internautes lorsqu'ils naviguent sur le web. Le terme plus communément utilisé depuis quelques années est celui de « Web mining » c'est à dire fouille du web (Liu & Zhang, 2004 ; Khoury, Dawborn & Huang, 2011, Cooley, Mobasher & Srivastava, 1997). Il s'agit ici d'extraire des informations pertinentes à partir de ressources disponibles sur le web, une ressource pouvant être un document ou un service Web.



Selon l'objectif visé, trois types d'études peuvent être réalisées :

- L'analyse du contenu des pages web (web content mining), c'est le processus d'extraction des connaissances à partir du contenu des pages web.
- L'analyse des liens entre les pages web (web structure mining), il s'agit de l'analyse de la structure du web, de l'architecture et des liens qui existent entre les différents documents ou pages sur le web.

Sujatha et Iyakutty (2010) et Sujatha et Punithavalli (2012) précisent que l'e-marketing et l'e-commerce sont très friands de ce type d'études. La plupart du temps ces recherches sont réalisées dans le domaine de l'informatique et l'objectif est d'élaborer des algorithmes poussés dans l'objectif de simuler et toujours mieux prédire les intentions et comportements d'achats des utilisateurs.

- L'analyse de l'usage des pages web (web usage mining), cette branche consiste à analyser le comportement de l'utilisateur à travers l'analyse de son interaction avec les sites et pages web. Compte tenu de notre problématique, notre attention s'est surtout portée sur les études portant sur ce domaine.

De façon générale, le processus du web mining se déroule en trois étapes passant par la collecte des données, la transformation de ces données et l'analyse.

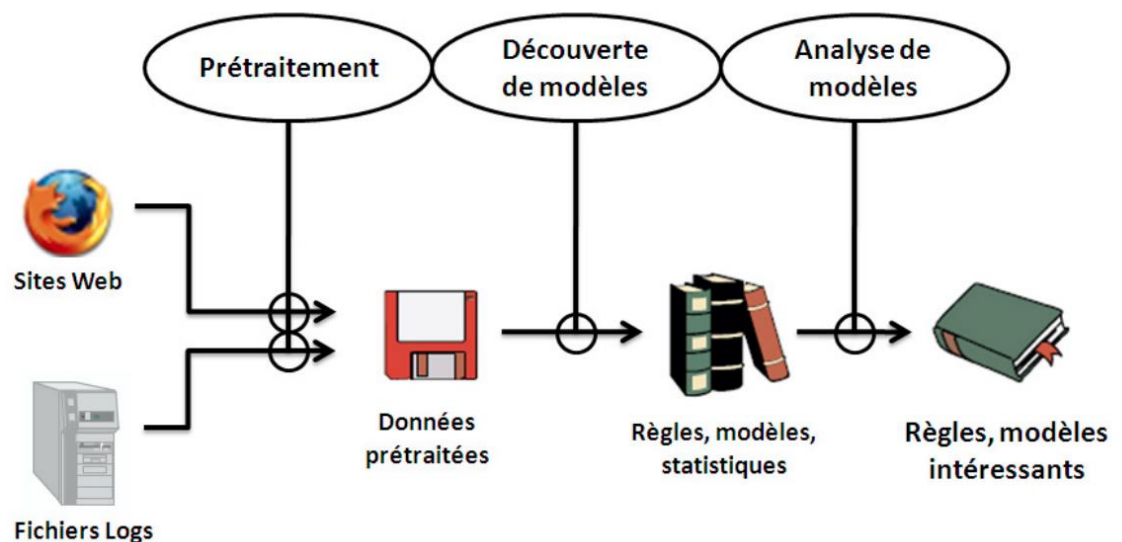


Figure 4 : Schématisation du processus de web mining

Les études sur lesquelles nous avons portées notre attention récoltent les données à l'aide de différents moyens d'enregistrements. L'une de ces principales méthodes est l'analyse des logs. Ils font référence à des journaux d'événements pour lesquels les URLs des sites consultés ainsi que les événements qui ont eu lieu sont enregistrés de manière chronologique. On parle aussi d'analyse des flux de clics ou clickstream study. La structure de ces fichiers a été définie par le World Wide Web Consortium (cité par Bastien, Drusch & Dinet, 2010).

Ce type de fichier peut être récupéré grâce à une technologie serveur ou client.

- D'un point de vue serveur, les données récupérées sont assez limitées. Ce type de méthode ne fournit pas d'informations sur les activités et l'interaction avec le navigateur, l'analyse se limite aux sites Web visités. Plus précisément, le serveur Web joue le rôle d'intermédiaire. En effet, il récupère les informations de navigation chez le participant, les traite et les envoie chez l'expérimentateur.
- Lorsque les technologies « *client-side* » sont utilisées, les chercheurs s'intéressent aux parcours et interactions réalisés par chacun des participants sur les sites consultés. L'enregistrement automatique réalisé sur les postes des participants permet d'analyser les parcours qu'ils réalisent sur plusieurs sites, ceux-ci pouvant être différents d'un utilisateur à l'autre. Avec ce type de technologie, un plug in peut être téléchargé et exécuté afin d'enregistrer l'historique de navigation, les actions réalisées ainsi que la fréquence d'utilisation des composants de l'interface. Dans ce cadre, les données sont exécutées sur le poste. Toutefois, si certains internautes conservent leur historique, d'autres s'assurent, pour des raisons diverses, que l'historique s'efface lors de la fermeture de leur navigateur Web. Par conséquent, dans ce genre d'études, on demande aux participants de conserver leurs historiques, ou on leur fournit des navigateurs instrumentés qui récupèrent automatiquement ces données (Catledge & Pitkow, 1995, Tauscher & Greenberg, 1997).

Champs	Description	Valeurs (exemples)
Date	Date à laquelle la transaction a eu lieu	2010-07-15
Time	Durée de la transaction	00:00:54
c-ip	Adresse IP du client	195.220.225.25
cs-username	Nom d'utilisateur pour les sites Web nécessitant une authentification	
s-sitename	Nom du serveur de site	W3svc3
s-computername	Nom de l'ordinateur	Business
s-ip	Adresse IP du serveur	195.220.225.22
cs-method	Commande http envoyée au serveur	GET
cs-uri-stem	Adresse de la page demandée (sans les attributs supplémentaires)	/Department/dept_footnote.idc
cs-uri-query	Tous les attributs supplémentaires de l'URL	FootNoteID=6&DegreeID=11
sc-status	L'état est un code de 3 chiffres renvoyé par le serveur au client (ex., 200 pour Ok, 404 pour « <i>Not found</i> »)	200
sc-win32-status	État du serveur win 32	0
sc-bytes	Le nombre d'octets retourné au client	16240
cs-bytes	Le nombre d'octets envoyé au serveur par le client	409
time-taken	Temps de réalisation de la transaction en secondes	32
s-port	Port du serveur (le port par défaut est 80)	80
cs-version	Version du protocole http utilisée	http/1.1
cs (user agent)	Navigateur Web utilisé par le client et système d'opération	Mozilla/4.0+ (compatible; +MSIE+5.5; +Windows+NT+4.0...
cs (Cookies)	Informations transmises par les cookies	- (ce signe indique que le cookie n'a pas été accepté ou qu'il n'a pas été utilisé)
cs (Referrer)	Indique la page Web à partir de laquelle le lien a été sélectionné	Business.tamu.edu/mgmt/Curric_BBA_MGMT.html

Figure 5 : Information pouvant être récupérées des fichiers d'évènements selon le World Wide Web Consortium

Toutefois, comme Berendt et Spiliopoulou (2000) le soulignent, derrière ce type de méthodologie, qui est pourtant la plus utilisée pour étudier les comportements de navigation, se dissimule de nombreuses limites qu'il est important de considérer.

La première concerne les motivations inhérentes à la navigation. En effet bien que ce type de protocole ait l'avantage de ne pas perturber les individus car les recherches sont réalisées dans le contexte naturel des personnes, c'est-à-dire à leur domicile, les chercheurs n'ont pas d'informations sur les objectifs qui ont motivé les individus à naviguer sur le web. Les usagers cherchent une information, mais les chercheurs ne savent pas laquelle. Zhou (2014) met en avant le delta qu'il peut y avoir entre la perception qu'à le chercheur sur la façon dont l'individu a fini sa tâche et la façon dont effectivement elle a été réalisée.

Par ailleurs, l'utilisateur est bien souvent identifié derrière un numéro associé à son adresse IP. Mais dans ce contexte, il n'est alors pas possible de savoir s'il s'agit toujours du même utilisateur et s'il a fait sa recherche seule par exemple (White & Horvitz, 2012). La littérature identifie certaines heuristiques qui peuvent être utilisées afin de contrer cette limite et d'en déduire si derrière le journal d'événement se cache plusieurs usagers. Cooley, Mobasher et Srivastava (1999) proposent, notamment, de vérifier combien de navigateurs différents ont été utilisés à partir du même poste. S'il y en a deux, il faut en déduire que plusieurs personnes ont utilisé l'ordinateur.

Par ailleurs, même si certains chercheurs sont passés par l'envoi d'invitation à des participants directement issus de leurs universités et ainsi avoir connaissances de certaines caractéristiques propres à leurs échantillons, dans la majorité des études ils n'ont pas de connaissance plus précises sur leurs échantillons. Les chercheurs n'obtiennent donc que peu d'informations sur les caractéristiques individuelles et le profil des sujets (Cockburn et McKenzie, 2001 ; Weinreich, Obendorf, Herder & Mayer, 2008).

Une autre limite très souvent relevée concerne la notion de session. En effet, il n'existe pas vraiment de consensus concernant ce terme et il convient au chercheur de définir quand la recherche de l'utilisateur prend fin. Berendt et Spiouloupoulou (2000) considèrent une session comme une séquence de pages qui conduit à l'obtention d'un but spécifique. Pour d'autres, une session peut s'établir sous la forme d'une durée définie par un seuil. Si le seuil est dépassé sans aucune autre action sur la page, alors on peut supposer que l'utilisateur a abandonné (Sullivan 1997). Enfin selon Chen, Park et Yu (1996) cité par Weber et Jaimes (2011) une session se compose uniquement d'actions de navigation vers l'avant, si l'utilisateur fait un retour arrière alors cela est à considérer comme une autre tâche. Cependant, comme Tauscher et Greenberg (1997) le mettent en lumière, la moitié des actions réalisées par un utilisateur sont des « back » et cela ne signifie pas nécessairement qu'il a fini ou changé son objectif.

Ces différents éléments montrent que le contrôle de certains facteurs parasites devient, dans ce cas, plus difficile puisqu'on ne peut pas être certain, notamment, que chaque sujet navigue dans des conditions similaires, et que les motivations sous-jacentes sont similaires. De cette façon, il semble compliqué de parler de standardisation des procédures (Aula & Jhaveri, 2005).

## **2.2.2.2 L'analyse des données déclaratives**

### ***Le questionnaire***

Les questionnaires peuvent être utilisés afin de compléter des mesures objectives recueillies par ailleurs. Jamet, Lemercier et Février (2007) cité plus haut, considèrent cette technique comme une méthode off-line permettant a posteriori d'évaluer la charge mentale, ou les performances d'apprentissage après interaction avec une interface. Ce type d'outil donne accès à des données plus subjectives relatives au ressenti et au vécu de l'utilisateur par rapport à la tâche qui lui a été demandé de réaliser. Par exemple, Cockburn, McKenzie et JasonSmith en 2002 ont étudié les différences qui existent entre deux types de retour en arrière ayant un fonctionnement différent. À la fin de la tâche, les chercheurs demandaient aux sujets de répondre, sur une échelle de Likert, à des questions relatives à la qualité de l'interface mais également par rapport à la facilité d'usage des boutons de retour en arrière. Plusieurs questionnaires ont été développés à l'intention des concepteurs, certains permettent d'évaluer la qualité du site web, alors que d'autres mesurent l'utilisabilité du site avec lequel le sujet a interagit.

### ***L'entretien***

Les techniques d'analyses qualitatives telles que l'entretien semi-directif font parties des méthodes permettant d'étudier les comportements de navigation des utilisateurs (Ackerman & Halverson, 1998 cité par Teevan, 2004). En effet, Teevan a employé cette technique afin de comprendre les comportements d'orientation lors d'une recherche d'information. Le chercheur demandait aux sujets de décrire leurs dernières activités de recherche d'information en explicitant leurs stratégies et leurs choix de navigation. Ce type de méthodologie peut permettre d'obtenir des données démographiques sur les participants telles que les habitudes de navigation (Holscher & Strube, 2000) mais également des données subjectives sur le ressenti de l'utilisateur par rapport à son interaction (Cockburn & McKenzie, 2001). Cette technique peut donc compléter des données objectives. Toutefois, les recherches ayant pour but d'identifier des patterns et comportements de navigation ne peuvent se résoudre à utiliser uniquement une méthodologie qualitative.

## 2.3

# LES NIVEAUX D'ANALYSES DES STYLES DE NAVIGATION (*PATTERN*) SUR LE WEB

---

Pour étudier l'activité des internautes, les auteurs tentent parfois de classifier les différents niveaux d'analyses (Le Guel, 2004).

La littérature présente donc différents niveaux d'extractions permettant d'observer l'interaction entre l'individu et le web tout en considérant les objectifs de recherche. Beauvisage (2004), propose une classification permettant de « décomposer l'analyse d'un parcours en cinq échelles distinctes ». Le niveau micro porte sur une suite de pages vues, l'objectif étant d'étudier le contenu des pages. Le niveau mini décrit le parcours d'un utilisateur au sein d'un site donné. Le niveau méso s'attache à l'assemblage de pages sur un ou plusieurs sites. L'avant dernier niveau macro correspond à l'étude d'un ou plusieurs cours d'actions au sein d'une session. Et enfin, le niveau méga prend en compte un parcours de l'utilisateur pour le comparer avec d'autres. Herder et Juvina (2000) ont quant à eux évoqué trois types de niveaux d'informations qui permettent d'identifier et d'appréhender la navigation sur le web. Les comportements sur Internet peuvent être alors étudiés par l'intermédiaire de l'information syntaxique qui identifie les liens suivis, les parcours de navigation, ainsi que le temps passé sur chaque page. L'information sémantique permet, elle, d'analyser les différents contenus des pages que va rencontrer l'utilisateur durant sa navigation. Et pour terminer l'information pragmatique étudie le comportement au travers des buts et des tâches de l'utilisateur. Différents paliers existent donc pour appréhender les comportements des individus sur le web. Il est important alors de prendre en compte les objectifs de sa recherche afin de mettre en exergue le niveau d'analyse approprié. Bucklin et Sismeiro (2003) distinguent quant à eux deux niveaux principaux sur lesquels portent les études s'intéressant à la modélisation des comportements de navigation ; le niveau intra site portant sur un site en particulier et le niveau inter site s'intéressant à la navigation à travers une multitude de sites web.

Pour notre part, nous les avons catégorisé en fonction de leurs usages ou de leurs comportements de navigation. Cela conduit les recherches à se concentrer sur deux niveaux d'analyses différents : l'analyse des usages, qui permet généralement de fournir des recommandations aux spécialistes des « analytics », c'est-à-dire ceux qui cherchent à mieux cibler les visiteurs sites Web ; et l'analyse des comportements de navigation, qui permet quant à elle, de concevoir de manière adaptée les navigateurs, sites Web et pages Web.

### 2.3.1

#### L'analyse des usages

Cette première approche est la plus générale. À ce niveau, on s'emploie en effet à identifier les sites Web consultés, leurs caractéristiques et quand c'est possible de les mettre en relation avec des groupes d'utilisateurs afin d'analyser ce que certains chercheurs appellent « le trafic » (Chi, Rosien & Heer, 2003). L'un des objectifs de ces études descriptives est donc de faire correspondre des profils d'utilisateurs aux sites Web visités.

Chau, Fang et Sheng (2005) parlent aussi d'analyses des usages. Cette démarche est semblable à celle que l'on trouve en marketing lorsqu'on étudie le contenu des paniers d'achat des consommateurs : quels sont les produits qu'achètent différents segments de consommateurs. Sur le Web, la question devient : qui consulte quoi et comment ?

Plus précisément, les chercheurs s'attachent à décrire, par exemple, les parcours de navigation à travers différents sites ou encore les sites le plus souvent consultés (Obendorf, Weinreich, Herder & Mayer, 2007 ; Weinreich, Obendorf, Herder & Mayer 2008) en relation avec le profil des internautes ou du type de site. Obendorf et al. (2007) ont notamment étudié le comportement de « revisite », autrement dit, le nombre de fois où les internautes revisitent un même site. Les chercheurs ont réalisé une étude portant sur 25 participants pour lesquels ils ont enregistré les événements de navigation sur le navigateur instrumenté Firefox entre 52 et 195 jours grâce à l'installation d'un proxy. Cette étude réalisée sur une période de 2 à 4 mois demandait également aux participants s'ils étaient d'accord pour réaliser 90 minutes d'entretiens à l'issue de l'étude afin de comprendre plus en détail leur comportement de revisite. Bien que les résultats issus de cette étude portent en grande partie sur les interactions avec le navigateur, les chercheurs se sont aussi concentrés en parallèle sur les 10 sites les plus visités par les participants.

L'une de leurs conclusions est que le type de sites a une influence sur le nombre de revisites. Ainsi, les moteurs de recherche ou les dictionnaires en ligne sont des sites pour lesquels la page d'accueil présente de nombreuses revisites puisque c'est à partir de celle-ci qu'il est possible de faire une recherche d'information. Ces sites ont donc une page très populaire et une liste d'autres pages qui le sont moins. Au contraire, les sites institutionnels ou de news ont une « page d'atterrissage » sur laquelle on identifie de nombreuses visites, mais aussi une succession d'autres pages présentant autant de revisites. Dans ce cas, il s'agit de pages présentant de l'information sur un domaine ou sujet particulier (sport, législation, etc.). Même si les chercheurs n'ont pas identifié de lien entre les revisites et certaines caractéristiques personnelles de l'échantillon (âge, genre), les chercheurs recommandent au concepteur de tenir compte du type de site afin de savoir si toutes les pages ont la même probabilité d'être visitées ou non.

Avec un échantillon plus conséquent, Adar (2008) s'est intéressé au comportement de revisite sur le web en lien avec le type de site. L'étude réalisée depuis le domicile des utilisateurs a porté sur 5 semaines et a compté 612000 participants. Chacun d'entre eux s'est vu attribué un numéro, et en analysant l'historique de navigation, les chercheurs ont porté leur attention, notamment, sur l'URL visité, le nombre de visites pour chacune de ces pages, le délai moyen entre chaque revisite. En utilisant la notion de « *revisitation curve* », ils ont pu schématiser pour tous les utilisateurs le temps de revisite sur chacune des pages revisitées. Adar et son équipe considèrent ainsi une courbe comme une signature du comportement de revisite.

Ils ont identifié 4 catégories de revisite chacune associées à des types de sites :

- Les pages du groupe de revisites rapides sont revisitées dans un intervalle de moins d'une heure. Il s'agit de sites d'e-commerce, catalogues en ligne, Wikipédia ou encore pornographiques.
- Les pages du groupe médium sont revisitées dans un intervalle s'étalant d'une heure jusqu'à une journée. Les auteurs ont identifié ici des pages issues de sites bancaire, sport, news, ou boites mail. Dans cette catégorie, on retrouve également les pages d'accueil des moteurs de recherche Google et Yahoo.
- Les pages du pattern lent sont des pages visitées dans un délai de plus d'un jour et qui ne semblent pas représenter un besoin urgent et important pour l'utilisateur. Ces pages font majoritairement référence aux activités du weekend (film, religion, astrologie, téléchargement logiciel).
- Enfin, les pages du groupe hybride correspondent à des revisites lentes et rapides. Elles comprennent des pages web sur la musique et les jeux en ligne par exemple.

Les résultats de leur étude montrent aussi qu'une grande majorité des pages (1 million environ) n'ont eu qu'un unique visiteur pour la durée de l'étude.

L'étude plus récente de Jiang, Tan, Phang et Wei (2012) analyse les séquences des sites Web qui se définissent comme une suite de pages Web consultées (sites e-commerce, d'informations, social, etc.) par un utilisateur donné et en tenant compte de l'ordre des visites. En effet, selon ces auteurs il est plus pertinent d'analyser des séquences de navigation au détriment de données disparates et ponctuelles. Les données de 200 utilisateurs Chinois recueillis via un panéliste ont été enregistrées pendant 30 jours. Les participants devaient installer un plug in sur leur navigateur afin que les chercheurs puissent récupérer les logs.

À l'issue de la récupération de ces logs, tous les sites sur lesquels les sujets ont navigué ont été répertoriés et ils se sont vus assigner à une catégorie. Des séquences de navigation ont ensuite été identifiées et une classification a été réalisée afin d'identifier les patterns similaires.

Ils identifient 4 types de séquences de navigation.

- La première correspond à une navigation pour la recherche et l'acquisition d'informations (Search and information retrieving). Il s'agit ici de consultations de portails d'informations permettant à l'utilisateur d'acquérir de nouvelles connaissances (santé, politique, technologie, etc.).
- Le deuxième type regroupe des visites de sites Web qui présentent une navigation sociale, centrée sur les blogs, forums, réseaux sociaux (Social navigation).
- La troisième catégorie regroupe des navigations sur des sites d'e-commerce (Commerce navigation).
- La quatrième catégorie concerne les utilisateurs qui se connectent directement sur des sites Web précis.

En parallèle, Jiang et son équipe ont analysées ces classes en fonction des caractéristiques démographiques de l'échantillon et notamment du genre. Les résultats étaient cohérents avec la littérature. En effet, plus de femmes présentaient une



navigation axée sur les réseaux sociaux (Weiser, 2000) et sur les sites d'e-commerce (Van Slyke, Comunale & Belanger, 2002). Concernant les hommes et en concordance avec les travaux de Odell, Korgen, Schumacher et Delucchini (2000), ils étaient identifiés majoritairement à travers la séquence de navigation concernant la recherche et l'acquisition d'information. Les chercheurs précisent que l'un des buts de cette recherche est l'étude des comportements d'exploration à des fins marketings et publicitaires.

Weber et Jaimés (2010, 2011) abordent cette même problématique en se référant aux requêtes émises sur un moteur de recherche et aux pages de résultats consultées. Les chercheurs considèrent que la recherche sur le net constitue pour les industriels un marché équivalent à un milliard de dollars et qu'aujourd'hui l'un des plus gros challenges est de caractériser les recherches en terme de profil démographique. Le web mining constitue alors la base de nombreuses décisions marketing relatives au référencement ou encore à la publicité. Les chercheurs ont analysé 2,3 Millions de requêtes issus du moteur de recherche Yahoo. Grâce à un cookie, ils ont pu recueillir le sexe de l'utilisateur, son âge et sa localisation. En mettant ces données en rapport avec le dernier recensement fait sur la population, les chercheurs ont pu mettre en relation plusieurs caractéristiques démographiques des utilisateurs, le type de recherche d'information (requête informationnelle, navigationnelle ou transactionnelle) et la nature de la session de navigation (notamment, nombre de sessions, durée de la session).

Les résultats montrent d'importantes différences entre les requêtes. Par exemple, l'utilisateur identifié comme un « blanc » conservateur et ayant voté pour le parti républicain chercherait surtout des informations sur la maison et le jardinage. Les baby-boomers seraient plutôt intéressés par la finance et la plupart de leurs requêtes seraient destinées à réaliser des opérations bancaires donc transactionnelles. Par ailleurs, Weber et son équipe ont pu remarquer que les requêtes liées au jeu seraient plus répandues sur Las Vegas que Salt Lake City, et que les recherches portant sur les acteurs plutôt à Los Angeles. Enfin, les chercheurs mettent aussi en évidence des différences de genre et d'âge. Les hommes chercheraient davantage sur le thème du sport et du bricolage, et les femmes sur la santé, la famille et les arts. Concernant l'âge, les individus ayant autour de 20 ans effectueraient plutôt des requêtes sur le fitness, la reproduction, et l'éducation, tandis que les personnes âgées sur la santé, le casino, la maladie et les voyages.

Ce type d'étude permet notamment aux spécialistes en marketing d'avoir une meilleure compréhension des différents segments d'internautes à partir des sessions de navigation observées dans les outils « analytics ». Les recherches réalisées à ce niveau mettent donc en exergue la volonté de prédire et de segmenter ce que les individus explorent librement sur le Web. Ce niveau d'analyse reste donc très général et descriptif, il s'agit d'une approche que l'on peut qualifier de macroscopique. L'intérêt de cette approche dans le cadre de l'analyse de patterns comportementaux à un niveau plus précis reste limité.

## 2.3.2

### **L'analyse des comportements d'interaction et de navigation**

L'analyse des comportements de navigation a pour objectif d'améliorer les interfaces Web que ce soit au niveau des navigateurs, des sites ou des pages Web. Dans cette optique, certaines études ont choisi d'analyser la variabilité comportementale au travers des styles, stratégies, ou patterns de navigation selon la terminologie employée. Cette approche permet à la fois d'affiner les recommandations génériques pour la conception mais aussi de mieux décrire la façon dont les internautes explorent le Web. Précisons que c'est à ce niveau que certaines études citées dans la première partie ont tenté de décrire les différences de stratégies qui pouvaient exister entre novices et experts.

### 2.3.2.1

#### **L'analyse des comportements pour la conception des navigateurs**

Pour fournir des recommandations sur la conception des navigateurs, certaines études se centrent :

- D'une part, sur des indices comportementaux comme la durée de visite des pages Web (Obendorf et al. 2007), le nombre de pages revisitées (Danielson, 2002), le nombre total de pages visitées pendant une session, ou encore l'utilisation du bouton retour arrière « back » du navigateur Web. Qu'il s'agisse du design des historiques de recherche (Tauscher e& Greenberg, 1997), des fonctionnalités du bouton « back » (Milic-Frayling, et al. 2004), ou encore de l'usage des onglets multiples (Dubroy & Balakrishnan, 2010), ces recommandations sont établies pour le plus grand nombre d'utilisateurs possibles.
- Et d'autre part, sur la découverte de patterns reposant sur ces mêmes indices.

Nous présenterons en ce début de partie, les principaux résultats issus de ce premier type d'études.

L'utilisation et la gestion des « bookmarks », plus connu sous le terme favoris est un comportement parfois analysé afin d'étudier le comportement de navigation des utilisateurs. Choo, Detlor et Turnbull (1999) font partis des pionniers à s'être intéressés à ce comportement. Cette étude portant sur 34 utilisateurs tous spécialistes en informatique avait combiné l'enregistrement de données au domicile des personnes et des entretiens. Les chercheurs avaient pu observer que la majorité des pages étaient consultées à partir des favoris. Par la suite, Cockburn et McKenzie (2001) ont réalisé une étude empirique sur une durée de quatre mois afin de caractériser les actions des utilisateurs avec le navigateur Netscape Navigator. Les chercheurs ont sauvegardé quotidiennement l'historique du navigateur des 17 participants ainsi que leurs fichiers de favoris. L'étude a montré que les participants réorganisent périodiquement leurs signets afin de ne pas avoir une liste trop longue à dérouler (Abrams, Baecker & Chignell, 1998). Ainsi sur les 17 sujets, 12 ont, sans le savoir, des onglets copiés plusieurs fois dans des dossiers différents. Les chercheurs ont observé que 25% des bookmarks n'étaient plus valides après deux mois car l'adresse web a changé.

Cockburn et McKenzie ont conclu que même si en général les utilisateurs ajoutent beaucoup de pages à leurs bookmarks, l'option « favoris » reste peu utilisée. Weinreich, Obendorf, Herder et Mayer (2008) que nous avons cité plus haut ont également observé que pour accéder aux pages web, 10% des actions réalisées correspondent soit à la sélection d'un favori ou à la recherche d'un favori dans la barre du navigateur.

Ainsi en s'intéressant au taux d'effacement et d'adjonction des favoris, les chercheurs relèvent que ce comportement fait appel à une caractéristique du navigateur qui est peu utilisée et cela pour deux raisons. D'une part, cette option demande à l'utilisateur d'enregistrer à l'avance les pages qu'ils envisageraient de revisiter à l'avenir. Et d'autre part, Comme Won, Jin et Hong (2009) le soulignent, si l'utilisateur souhaite rapidement s'y trouver, cela lui demanderait de trier et de classer quotidiennement ces favoris. Ainsi, afin d'encourager et d'optimiser l'option des bookmarks, Cockburn et McKenzie (2001) recommandent, d'une part, de les présenter avec un design qui faciliterait la lecture de grandes listes et, d'autre part, assister l'utilisateur en l'aidant à classer ces favoris en identifiant les pages qui n'existent plus.

L'étude de la navigation parallèle, permise grâce à l'utilisation des onglets du navigateur, a également fait l'objet de plusieurs recherches. Selon Dubroy et Balakrishnan (2010), le phénomène de navigation parallèle est de plus en plus populaire et il est nécessaire de l'étudier si l'on souhaite s'intéresser aux comportements sur le web. Les onglets ont pour rôle de faciliter la navigation de l'individu en termes d'organisation. Ils ont deux fonctions principales.

- La première permet de mettre une page de côté afin d'y retourner plus tard et de jongler entre plusieurs onglets.
- La deuxième fonction est celle d'être une « fourche » dans la navigation, c'est à dire qu'un lien peut être ouvert dans un onglet sans perturber la page qui est en train d'être consultée.

Dubroy et Balakrishnan (2010) ont conduit une étude sur le navigateur Mozilla avec 21 sujets afin de comprendre comment et pourquoi les utilisateurs utilisent les onglets. Les chercheurs ont demandé aux participants d'installer une extension de Firefox et afin de garantir une navigation la plus « naturelle » possible, ils ont assuré aux utilisateurs qu'ils ne prendraient pas en compte la nature des pages visitées. Pour se faire les URLs n'étaient pas accessibles aux chercheurs, un code était associé à chaque page et ils avaient connaissance si oui ou non cette page avait été précédemment ouverte. Par ailleurs, grâce à des interviews, et à la tenue d'un journal de bord dans lequel les utilisateurs expliquaient leurs activités de navigation, les auteurs ont observé que les onglets étaient une alternative à l'utilisation du retour en arrière. En effet, selon les participants, les onglets sont perçus comme plus pratiques, plus rapides et moins fastidieux que le bouton back. Les individus expliquent leurs motivations à utiliser les onglets par le fait que c'est un bon moyen de rappel, qu'ils permettent de faire des comparaisons entre deux pages, et qu'à court terme ils sont équivalents à l'usage des favoris. Afin d'optimiser d'avantage l'usage des onglets, les chercheurs recommandent de fournir un feedback direct afin de lister l'ensemble des onglets ouverts et de montrer la relation entre eux afin de voir les derniers consultés.

Concernant la préférence de l'usage des onglets par rapport au bouton back, Weinreich et al., (2008) avaient déjà observé que les changements de comportements les plus importants concernaient le nombre croissant de pages ouvertes dans une nouvelle fenêtre du navigateur ou dans de nouveaux onglets, et l'usage décroissant du bouton « Back ». L'ouverture de nouvelles fenêtres ne constituait qu'environ 1% des stratégies de navigation dans les années '90 alors que les auteurs constatent aujourd'hui un pourcentage d'environ 10%. Quant à l'usage du bouton « Back », le pourcentage d'action de navigation est passé de 30 à environ 15%. L'utilisation des onglets et des nouvelles fenêtres pour afficher des pages Web explique pour une bonne part l'usage décroissant du bouton « Back ». Les utilisateurs reviennent aux pages précédentes en sélectionnant les onglets correspondant ou en sélectionnant les fenêtres correspondantes.

Huang et White (2010) ont réalisé une étude sur l'usage des onglets, mais à une plus grande échelle. À travers 60 milliards de pages vues, 50 millions d'utilisateurs ont participé à cette étude où les données étaient enregistrées à partir du moment où l'utilisateur en consentait l'envoi. Trois indicateurs ont été étudiés : le nombre de pages vu par onglet, le nombre d'onglet ouvert par session et le nombre de clics réalisés à partir d'une page vue. L'étude a révélé qu'en moyenne 5 à 10 pages sont vues à partir d'un onglet, et 4 onglets sont ouverts par fenêtre. De plus 88,7% des pages vues sont ouvertes à partir d'un onglet existant et les utilisateurs passent 57,4% de leur temps à passer d'un onglet à l'autre. Bien que les chercheurs concluent à une navigation parallèle de plus en plus prédominante, ils soulignent que la non prise en compte des caractéristiques individuelles est une limite à considérer. Par ailleurs, comme pour l'étude de Dubroy et Balakrishnan (2010) la notion de tâches ou sessions n'est pas prise en compte, il est donc difficile de savoir quel est le comportement de navigation parallèle sur une période précise.

Concernant les liens hypertextes, de nombreuses recherches ont montrés qu'il s'agissait du moyen préféré des utilisateurs afin de se déplacer d'une page à une autre lors de la navigation. La fréquence d'utilisation varie de 44% à 50% (Catledge & Pitkow, 1995 ; Obendorf & Weinreich, 2007 ; White, 2007). Notons, que les auteurs comptabilisent souvent le nombre de fois où ce moyen de navigation a été utilisé pour tout l'échantillon, sans distinguer les pages ou sites visités. Les résultats restent donc généraux.

Le nombre total de pages visitées, toutes URL confondues, par les individus est aussi pris en compte mais le chiffre varie en fonction de certains paramètres tels que le nombre de sujets ou la durée de l'expérience. Catledge et Pitkow (1995) ont relevé ainsi un total de 1222 pages vues sur une durée de 3 semaines pour 107 participants, tandis que Cockburn et MacKenzie (2002) ont, eux, un total de 84841 pages sur une durée de 4 mois pour un nombre de 25 personnes. Certains de ces auteurs vont plus loin et calculent une moyenne ou nombre de pages vus par sujet et par jour (Tauscher & Greenberg, 1997 ; Cockburn & MacKenzie, 2002 ; Obendorf & Weinreich, 2007). En fonction des études, les moyennes varient de 14 pages à 283 pages visitées en moyenne par certains sujets.

Ce type de description pose quand même la question de sa pertinence car sans connaître les motivations sous-jacentes des individus, ni même si les sites étaient les mêmes, cela permet difficilement de comparer ces moyennes entre elles.

Par ailleurs, l'indicateur temps et précisément la moyenne du temps passé sur les différentes pages visitées (Obendorf & Weinreich, 2007) sont aussi étudiés pour caractériser la navigation des internautes. Il peut également s'agir d'analyser le temps de latence entre la visite de deux pages (Cockburn & MacKenzie, 2002).

L'utilisation du bouton retour en arrière « back » est également une métrique qui est analysée afin de répondre à trois objectifs :

- Le premier est de mieux saisir la façon qu'ont les individus de naviguer (Cockburn & McKenzie, 2002 ; Milic-Frayling, Jones, Rodden, Smyth, Blackwell, & Sommerer, 2004 ; Obendorf & Weinreich, 2007).
- Le deuxième objectif est de différencier les individus pour identifier des stratégies de navigation (White & Drucker, 2007).
- Le troisième objectif est celui d'étudier l'impact de différentes organisations de sites web sur le comportement de navigation (Danielson, 2002).

Les chiffres concernant l'utilisation du bouton retour en arrière sont d'ailleurs très variables d'une étude à l'autre. En effet, dans l'étude de Catledge & Pitkow publiée en 1995 et portant sur 107 utilisateurs pendant 21 jours, les chercheurs ont rapporté environ 31000 actions de navigation, l'utilisation du retour en arrière représentait 41% des actions, alors que l'utilisation des liens hypertextes ne représentait que 52% des actions. En se basant sur cet indicateur, les chercheurs ont identifié la stratégie du « *hub and spoke* » qui consiste à effectuer des allers-retours entre les pages consultées et la page de départ à l'aide du bouton de retour arrière. Les chercheurs précisent que cette stratégie se différencie du « *looping back* » défini par le retour au point de départ à partir d'une fonction d'aide à la navigation tel que l'historique par exemple.

Tauscher et Greenberg en 1997 publient une étude portant sur 23 sujets pendant six semaines et rapporte que l'utilisation du retour en arrière constitue 30% des actions de navigation. Ces deux études réalisées dans le contexte naturel des sujets, avec un navigateur instrumenté, sont celles rapportant un pourcentage d'utilisation du « back » le plus élevé. Greenberg et Cockburn (1999) justifient cette importante utilisation en expliquant que les individus cliquent sur le bouton retour en arrière jusqu'à reconnaître la page qui les intéresse. Ces derniers se baseraient plutôt sur des indices visuels que textuels. Les individus préfèrent alors le retour en arrière plutôt que l'historique ou taper l'URL. D'autres recherches ont mis en exergue une plus faible utilisation de ce bouton. Toutefois, l'analyse de l'utilisation de ce bouton comporte certaines limites. En effet, le pourcentage d'utilisation de cette option est donné la plupart du temps de façon « brute ». Le nombre de pages visitées et revisitées au cours de la session de navigation, ou encore le temps de latence entre la visite et la revisite d'une page ne sont pas considérés.

Concernant les revisites, la plupart des études qui se sont intéressées à ce type de parcours ont été réalisées à la fin des années 1990 et début 2000 (Zhang, & Zhao, 2011).

La revisite est définie comme le chargement répété d'une page web identifiée par son URL, par ailleurs la littérature assimile le taux de revisite à la lettre « R ». Le taux de revisite correspondant au nombre total d'URLs visitées auquel on soustrait les différents URLs visités, puis le tout divisé par le nombre total d'URLs visités. La majorité des études présentent les résultats en termes de fréquence, mais certaines vont plus loin en étudiant les motivations qui poussent un individu à revenir sur une page (Adar, Teevan & Dumais, 2009). Deux grands types de méthodologies sont donc utilisées pour étudier le comportement de revisite ; l'étude de logs avec proxy et navigateurs instrumentés (Cockburn, et *al.*, 2001 ; Herder, 2005, Obendorf, et *al.*, 2007 ; Tauscher, et *al.*, 1997) et les études qualitatives avec passation d'entretiens ou de questionnaires (Aula, 2005, Kellar, Watters & Inkpen, 2007).

Tauscher et Greenberg (1997) rapportent ainsi dans leur étude un « R » de 58 %, il s'agit ici d'une moyenne pour toutes les pages et sujets confondus. Les auteurs indiquent également que 60% des pages ont été visitées une fois, 8% ont été visitées 3 fois et 4% ont été visitées 4 fois. Les pages qui ont été fréquemment visitées ont tendance à faire partie de la même catégorie, et ceci s'explique par leur popularité ; page d'accueil du navigateur ou moteur de recherche.

Toutefois, il est important de noter que même si la formule pour calculer le taux de revisite est commune à la majorité des études citées (Won, Jin & Hong, 2009), les résultats des recherches ont montré une certaine hétérogénéité. Cockburn et McKenzie (2000) ont, par exemple, trouvé un taux de revisite à 81 %. Ce résultat s'explique par le fait que les chercheurs tronquaient l'adresse URI sans considérer tous les paramètres de la requête, et que par conséquent, le nombre de revisite était plus élevé.

Obendorf et Weinreich (2007) apportent également une nuance en calculant un taux de revisite individuel pour les 25 participants. Les fréquences variaient entre 17,4% et 61,4%. Dans ce cas, le calcul d'une moyenne individuelle diffère totalement d'une moyenne générale. Pour de futures études, les chercheurs recommandent de privilégier le calcul d'un taux de revisite individuel et de l'étudier en rapport avec les tâches effectuées.

L'analyse des logs ainsi que les entretiens réalisés leurs ont permis de différencier trois intentions de revisites et les actions leur permettant d'accéder à la page :

- Les revisites à court terme sont celle qui se font dans l'heure juste après avoir visité la page qui intéresse le sujet. Ce type de revisite est prépondérant à 72,6% et les utilisateurs utilisent à la fois le bouton back et aussi la navigation parallèle.
- Les revisites à moyen terme sont réalisées dans la journée (12%) ou dans la semaine (7,8%). Une part importante de ces revisites a pour but de revoir une page web afin de trouver de nouvelles ressources pour affiner une recherche, ou simplement observer les mises à jour du site. L'accès à ces pages se fait à la fois par les favoris mais aussi en tapant directement l'URL.
- Et enfin, les revisites à long terme qui sont réalisées plus d'une semaine après la visite initiale (7,6%). Elles sont motivées afin de redécouvrir un site et de ré accéder à des informations vues bien auparavant. Bien souvent, ce type de revisite se fait grâce aux hyperliens (58%) et dans une moindre mesure via les favoris (16%).

Adar (2008, 2009) précisera que le fait de revisiter une page peut dépendre de l'intention de la personne, du contenu de la page et de la structure du site.

À notre connaissance, la dernière étude en date s'intéressant au comportement de revisite est celle de Zhang et Zao (2011). Cette étude client side a porté sur 20 participants et a enregistré 235 707 événements de navigation. Comme pour l'étude de Catledge et Pitkow, Tauscher et Greenberg, McKenzie et Cockurn et Obendorf et Weinreich, les informations sur l'échantillon étaient limitées et on demandait aux utilisateurs de naviguer comme à leur habitude. Selon eux, la définition actuelle de la revisite considère uniquement les pages initiées par le navigateur et néglige l'activité de revisite par les onglets. Les chercheurs soulignent ici une limite relative aux études précédentes. Les résultats montrent que le taux de revisite effectif (avec prise en compte des onglets) est de 59,6%, alors qu'il est de 39,3% pour le taux conventionnel.

La littérature révèle donc que cet indicateur est largement utilisé dans de nombreuses études pour définir et étudier le comportement de navigation des internautes lors de l'exploration de pages Web. Toutefois, la façon de le calculer et de l'interpréter est souvent remise en question (Won, Jin, & Hong, 2009). Les différences de résultats s'expliquent par le manque d'homogénéité dans la méthodologie employée (Zhang & Zhao, 2011). À l'issue de ces travaux, certaines recommandations ont été établies. Par exemple, une idée pourrait être de garder en photo une capture des 2,3 dernières pages vues afin que l'utilisateur puisse plus facilement y accéder et identifier les pages. Par ailleurs, il est recommandé également de mettre en couleur les éléments de la page qui ont évolué depuis la dernière visite afin que l'utilisateur identifie en un coup d'œil les nouvelles informations. Enfin, l'historique pourrait présenter les pages les plus souvent revisitées.

Weinreich et son équipe (2008) ont également étudié un autre aspect intéressant de la navigation inter page-inter site il s'agit des aspects temporels relatifs au temps passé sur les pages web. Les chercheurs estiment que les individus passent peu de temps sur les pages web en général, 25% des pages sont visionnées moins de 4 secondes, 52% des visites sont inférieures à 10 secondes, et 10% des pages sont visionnées plus de 2 minutes. Les chercheurs ont voulu préciser les résultats en comparant le nombre de mots des pages web et le temps passé sur ces pages. Les résultats basés sur 60000 pages visitées montrent que le temps passé sur une page dépend de son contenu. Les pages visitées moins de 12 secondes comportent en moyenne 430 mots, ces résultats sont significativement plus faibles que pour les pages de taille moyenne ayant un temps de visite plus important. Dans ce type d'analyse, plusieurs limites sont à soulever. D'une part, il n'est pas possible de savoir si les utilisateurs ont lu la page en question, ni même ce qu'ils y ont cherché. D'autre part, le type de page n'est pas pris en compte, s'agit-il d'une page de résultats Google, une publicité, etc.

D'autres études se basent sur ces mêmes métriques pour identifier des « stratégies » de navigation pour affiner les recommandations. Catledge et Pitkow (1995) ont été les pionniers concernant l'analyse de données sur la navigation des utilisateurs du Web. L'étude a porté sur 107 sujets pour une durée de 3 semaines. Les données étaient recueillies sur les ordinateurs des participants grâce à un navigateur instrumenté. Selon cette étude, il existerait trois styles de navigation :

- Le premier style est celui du « chercheur » (*searcher*). L'internaute faisant partie de ce groupe est décrit comme ayant un objectif précis. Sa navigation se caractérise par une combinaison de séquences courtes de pages avec des visites peu fréquentes, et des séquences plus longues et fréquentes. Les données indiquent que plus on visite les mêmes sites sur Internet, moins on y consacre de temps les fois suivantes.
- Le second style de navigation est celui du « navigateur aux buts généraux » (*general purpose browser*). Il est défini comme celui dont les séquences de navigation ne se répètent que rarement.
- Enfin, le « navigateur fortuit » (*serendipitous browser*) explore quant à lui des sites qui sont peu liés les uns aux autres. Il navigue donc de façon fortuite sans jamais s'attarder sur un site particulier.

Trois métriques ont permis d'aboutir à ces patterns : les sites visités, la fréquence de ces visites, et les chemins parcourus. Catledge et Pitkow ne donnent pas de recommandations spécifiques en fonction des patterns de navigation, mais concluent toutefois que les concepteurs doivent prendre en compte les stratégies de navigation des utilisateurs.

Plus récemment, en vue de différencier les parcours des individus et d'étudier la variabilité comportementale lors d'une recherche d'information, White et Drucker (2007) ont mené une étude utilisant des outils de captures d'événements utilisateurs. Leur étude a porté sur la navigation de 2527 individus. Pour chaque participant, les adresses des pages Web consultées ont été enregistrées. Ces suites d'adresses sont ensuite analysées comme des chaînes de caractères, ce qui permet à White et Drucker de calculer une distance entre les sessions. Pour cela ils utilisent la Distance de Levensthein (LD) qui consiste à calculer le nombre minimal de caractères qu'il faut supprimer, insérer ou remplacer d'une chaîne pour la rendre identique à une autre.

Les auteurs ont alors pu observer que certains participants sont cohérents avec eux-mêmes entre différentes sessions de navigation : ce sont les « navigateurs » (la LD intra individu est faible) alors que d'autres, qualifiés « d'explorateurs », ne le sont pas (la LD intra individu est forte).

- Les « navigateurs » visitent souvent les mêmes pages, atteignent leurs objectifs de façon directe, c'est-à-dire en empruntant les chemins les plus courts. Leur mode de navigation ne varie pas.
- Dans le cas contraire, les individus sont définis comme des « explorateurs ». Ils explorent des sources diverses et variées et ont tendance à revisiter et à visiter de nouvelles pages très fréquemment. Ils ne parviennent pas directement à leurs objectifs et « perdent du temps » en s'éparpillant au cours de leur navigation.

En tenant compte de cette variabilité, les chercheurs ont proposé quelques recommandations pour la conception web. Ils recommandent, par exemple, de fournir un historique de navigation sur les pages d'accueil des moteurs de recherche pour les « navigateurs », et une liste des pages populaires pour les « explorateurs ».



Par ailleurs, dans cette approche d'autres recherches se sont appuyées sur des méthodes d'analyses déclaratives pour concevoir le « navigateur parfait ». Teevan, Alvarado, Ackerman et Karger (2004) ont réalisé, avec 15 participants tous diplômés en sciences informatiques, deux fois par jours et pendant 5 jours, 151 entretiens semi-directifs de cinq minutes chacun. Les participants devaient relater leurs dernières activités de recherche d'information sur le Web. Les chercheurs ont également mené des observations directes afin d'étudier les patterns de navigation des individus lors de leurs recherches d'informations. Deux stratégies de navigation ont été identifiées.

- La stratégie du « comportement d'orientation » (*orienteeing behavior*) concerne les individus qui utilisent leurs connaissances a priori combinées à des informations contextuelles pour trouver une information. Ils procèdent ainsi par étape sans tenter d'accéder directement au contenu recherché.
- L'autre stratégie est décrite comme de la « téléportation » (*teleporting*), où les individus ont tendance à formuler des requêtes permettant d'accéder directement à l'information.

À l'issu de leur étude, Teevan et son équipe ont suggéré de nouvelles pistes pour la conception de la page du moteur de recherche, dont l'idée d'un navigateur capable de redéfinir et reformuler automatiquement les requêtes afin de réduire les étapes de recherche.

Card, Pirolli, Wege, Morrison, Reeder, Schraedley, & Boshart, (2001) ont, quant à eux, réussi à vérifier l'impact que pouvait avoir l'information scent sur le comportement et le parcours des utilisateurs. Ces chercheurs ont donc établi un protocole original en demandant à 14 individus de décrire à haute-voix ce qu'ils faisaient au cours des tâches qui leurs étaient assignées. Ainsi, pour chaque participant, et pour chaque page visitée, les auteurs rendent compte du niveau « d'information scent » en se basant sur les commentaires des sujets. Les résultats montrent que les sujets passent à autre chose lorsque le parfum de l'information ne semble pas pertinent. Des remarques du type « cela semble prometteur », « c'est ce que je veux » indiquent un parfum d'information pertinent, alors que « mais il n'y a rien sur cette page » ou « je vais voir ailleurs » révèlent un parfum d'information faible. De cette façon, le nombre de liens suivis ou de pages visitées dépendrait du niveau de parfum des informations perçus par l'utilisateur tout au long de sa recherche (Card, et *al.*, 2001).

### **2.3.2.2**

#### **L'analyse des comportements pour l'architecture des sites Web**

L'objectif de cette approche est d'identifier et de comparer les parcours des internautes lorsqu'ils naviguent sur un site Web afin de formuler des recommandations pour la conception.

Dans cette optique, Danielson (2002) a évalué la mise en page de deux sites Web différents : l'un avec un plan du site et l'autre sans. 20 participants ayant entre 18 et 40 ans ont été recrutés. Cette étude ayant lieu en laboratoire s'est déroulé en plusieurs phases. Une première pour laquelle le chercheur a expliqué la consigne, une deuxième phase où les utilisateurs devaient naviguer sur l'un des deux sites (avec ou sans plan)

et trouver les informations demandées. L'étude prenait fin après le remplissage d'un questionnaire.

Le chercheur s'est appuyé, notamment, sur le nombre de pages visitées pour chacun des sites, la réussite ou non aux tâches, l'usage du retour en arrière, et sur la capacité du sujet à se souvenir du nombre de pages vues pour évaluer l'utilisabilité de ces deux sites. Avec moins de retour arrière, de pages revisitées, moins de temps passé sur les pages, et une meilleure représentation du nombre de pages visitées, les résultats indiquaient la supériorité de la mise en page associée à un plan du site.

Dans le but de concevoir de nouveaux systèmes d'aides adaptatifs pour les sites Web, Herder et Juvina (2000) ont réalisé une étude en laboratoire où 30 participants devaient réaliser des tâches sur des sites précis (financiers et e-commerce) permettant d'identifier des patterns de navigation en lien avec des problèmes de désorientation.

Deux styles comportementaux ont été observés : les « fragiles » (flimsy) et les « laborieux » (laborious).

- Ainsi, les « fragiles » se servent fréquemment du bouton « back » pour naviguer et ont tendance à se concentrer sur le contenu d'une page. Ces individus visitent peu de pages et ils y passent un certain temps. Leur navigation gravite essentiellement autour de la page d'accueil, ils ne s'éloignent pas trop et reviennent régulièrement dessus. Ils utilisent rarement les hyperliens présents sur les pages visitées. Herder et Juvina considèrent cette façon de naviguer comme un signe de désorientation. Dans ce cas, des aides visuelles à la navigation telles que l'historique de navigation, et plans de sites leur seraient utiles.
- A contrario, les « laborieux » explorent les sites plus profondément et revisitent des pages déjà consultées lorsque cela leur paraît utile. Leurs revisites sont nombreuses mais non redondantes, quand une page est revisitée, ils vont sélectionner un autre lien. Ces individus, s'aperçoivent rapidement de la pertinence ou non des liens. Ils semblent agir par essais erreurs en visitant les liens qu'ils rencontrent afin de mesurer leurs degrés de pertinence par rapport à leurs requêtes initiales. Pour ce pattern, les chercheurs recommandent de présenter, un plan de site local c'est-à-dire par pages avec un index des liens.

De nombreux indicateurs tels que, la complexité du chemin parcouru, le taux de revisites, le nombre de retours arrières, le nombre de pages visitées, le temps passé et le nombre de visites sur les pages d'accueil ont été pris en compte dans l'identification de ces deux stratégies. Ces résultats pourraient permettre de concevoir des aides adaptées en fonction du style de navigation.

Gwidzka et Spence (2007) ont poursuivi ces recherches en étudiant les facteurs pouvant prédire le succès et l'échec à une tâche de recherche d'information pour améliorer les aides en ligne. L'expérience qui a eu lieu en laboratoire avec 48 participants ayant 8 tâches de recherche d'informations à effectuer sur deux sites Canadiens distincts d'e-gouvernement, a permis d'identifier les indicateurs prédisant le mieux la réussite ou l'échec aux tâches. Les participants avaient 3 minutes pour répondre à chaque tâche, au-delà, cela était considéré comme un échec. Les analyses ont montré que la similarité entre le parcours utilisateur et le parcours optimal pourrait

prédire la réussite ou l'échec à une tâche. La linéarité du chemin de navigation serait également un bon indicateur.

Par ailleurs, afin d'expliquer la variabilité entre les stratégies de navigation, trois styles de navigation ont été observés :

- La navigation inefficace (nombreuses pages visitées, temps important pour réaliser la tâche, faible similarité avec le chemin optimal et plus important taux d'échecs)
- La navigation laborieuse (taux de revisites élevé et chemins non linéaires). Ce type de navigation ressemble à celui décrit par Herder et Juvina (2000).
- La navigation rapide (beaucoup de clics, et surtout une navigation très rapide sur les pages visitées).

Les auteurs précisent que cette étude est un premier pas vers la détection automatique de la désorientation des utilisateurs pour concevoir des systèmes d'aides adaptatifs.

Enfin, nous pouvons également citer l'étude de Chin et Fu (2010) décrite plus haut qui a permis de mettre en lumière des stratégies différentes en comparant deux populations réalisant des tâches sur deux versions d'un même site d'information médical ou encore l'étude de Potelle et Rouet (2003) qui a porté son attention sur les différents types de dispositifs de représentation de contenu (carte hiérarchique, liste alphabétique ou carte conceptuelle) sur la compréhension de documents hypertextes. Dans les deux cas, ces études avaient comme objectif de s'intéresser aux principales particularités (âge, expertise) de leurs échantillons au cours de la navigation afin de fournir des recommandations adaptées.

#### **2.3.2.4**

#### **L'analyse des comportements des internautes pour la conception de pages web**

Cette approche consiste à étudier de façon précise l'interaction sur une page donnée, ce qu'elle doit contenir et à identifier les effets de différentes organisations de l'information sur les performances d'exploration et de recherche afin d'en évaluer l'utilisabilité et d'établir des recommandations (Nielsen, 2003)

Weinreich et *al.* (2008) ont étudié notamment la façon dont les Internautes explorent de longues pages Web et plus précisément l'utilisation de la barre d'ascenseur du navigateur (les scrolls) pouvant désorienter l'individu. Ce dernier peut alors perdre le fil de sa navigation car les éléments présentés sur une page ne sont plus disponibles au même endroit après un scroll. Face à une page longue, le scrolling augmente la charge cognitive. De plus, si le sujet scroll sur une durée assez conséquente cela exige qu'il se rappelle des informations qui ont été vues précédemment, tandis que des pages plus courtes permettent de comparer toutes les informations disponibles au premier coup d'œil (Lynch & Horton 2002).

Leur attention s'est portée sur la page de résultat du moteur de recherche Google. Pour obtenir une vue d'ensemble de l'activité de navigation dans les différentes régions de l'écran, les chercheurs ont créé des cartes de clics. Celles-ci sont réalisées à l'aide de « heatmaps » appelées aussi cartes de chaleur. Elles illustrent la densité des clics afin de comparer leurs différents emplacements et selon Shrestha et Lentz (2007) permettent de dépeindre les régions d'attention de l'utilisateur sur l'écran. Les résultats montrent que 76,5% des liens étaient choisis au niveau de la région visible de la page et que seulement 23% des clics étaient fait en dessous de la région initialement visible, c'est-à-dire la partie nécessitant du scroll. L'étude a par ailleurs révélé une ressemblance entre la distribution des clics et des fixations oculaires qui permettrait d'envisager une corrélation entre l'attention que l'on porterait à la page et l'endroit où l'on positionne le curseur. En outre, ces auteurs ont observé que les internautes sélectionnent les liens essentiellement dans les parties visibles de la page web (ne nécessitant pas de scrolls). Ces conclusions rejoignent celle de Bernard, Baker et Fernandez (2002) qui montrent que les pages courtes sont préférées et offrent de meilleures performances de navigation. Ils ont conclu que la structure des pages Web devrait être flexible en fonction de la taille de la fenêtre du navigateur.

Buscher, White, Dumais et Huang (2012) ont poursuivi ces travaux sur les pages de résultats du moteur de recherche « Bing » pour en améliorer l'affichage. Les chercheurs se sont cette fois intéressés notamment aux comportements d'interactions (scrolls, clics et déplacement du curseur de la souris). En effet, selon Buscher et *al.*, l'analyse des logs du curseur de la souris constitue une opportunité pour évaluer l'usage du Web. De récentes recherches ont montré que les mouvements de la souris sont corrélés avec le suivi du regard (Guan, & Cutrell, 2007) et il serait ainsi un indicateur pertinent pour analyser l'attention du participant. Ces études réalisées en laboratoire avec de petits panels de sujets et portant sur les pages de résultats de moteurs de recherches ont mis en exergue un comportement de « unclicked hover » (Huang, White & Dumais, 2011, 2012). Il s'agit ici d'un temps de latence pendant lequel le sujet va déplacer son curseur de souris autour d'un résultat de recherche mais sans cliquer. Ces déplacements peuvent permettre d'aider à la lecture, de marquer un résultat intéressant ou encore de positionner le curseur sur un endroit de l'écran où il ne gênera pas.

L'étude a été réalisée en contexte écologique auprès de 22084 utilisateurs pendant 13 jours. Par l'utilisation d'indicateurs tels que la moyenne de la distance parcourue par la souris, la vitesse de déplacement de la souris ou encore le nombre de scrolls, les chercheurs ont pu identifier 3 classes d'utilisateurs pour les tâches informationnelles :

- les « exhaustifs-passifs » (*exhaustive-passive*) qui cliquent et scrollent beaucoup, passent très peu de temps sur les pages de résultats de recherche mais qui abandonnent très souvent.
- Les « exhaustifs-actifs » (*exhaustive-active*) qui cliquent beaucoup sur les liens hypertextes, font des mouvements de souris assez lent, scrollent beaucoup également mais qui abandonnent rarement leurs tâches.
- Enfin, les « économiques » (*economic*) qui passent très peu de temps sur les pages de résultats de recherche, font des mouvements de souris très rapides et

qui en moyenne cliquent sur un seul résultat (ils correspondent à la stratégie de « *teleporting* » identifiée par Teevan et son équipe).

Les auteurs concluent par exemple que les utilisateurs « économiques » bénéficieraient de réponses plus riches à leurs requêtes directement sur la page de résultats alors que les « exhaustifs-actifs » auraient besoin d'outils permettant de comparer les résultats. Concernant les « exhaustifs passifs », il serait intéressant de leur présenter plus de résultats afin d'augmenter la probabilité qu'ils trouvent quelque chose leur correspondant. Par ailleurs, et comme nous l'avons souligné avec l'étude de Kellar (2006), les chercheurs concluent également à un effet de la tâche. En effet, en se basant sur les requêtes effectuées, ils ont distingué les tâches de navigation et d'informations. De futures publications devront définir plus précisément la méthode utilisée pour les classer. Les analyses ont montré que dans le cadre de tâches informationnelles, les participants faisaient plus de clics sur des hyperliens, mettaient plus de temps avant d'en sélectionner un, réalisaient plus de scrolls sur la page de résultat.

Au-delà des études portant leur attention sur la page de résultat d'un moteur de recherche, Kim et Fesenmaier (2008) mettent en lumière l'importance de s'attarder sur les pages d'accueil ou « home page ». En effet, selon les chercheurs, c'est essentiellement la page d'accueil du site, sur lequel atterrit l'utilisateur, qui va l'inciter à entrer ou non dans le site, et ceci d'autant plus s'il n'est pas connu de la personne. Les acteurs dans le domaine de l'e-commerce s'intéressent de près à cet aspect particulier d'un site. Pour les entreprises de biens et services, l'un des enjeux majeurs, est que l'utilisateur reste suffisamment longtemps sur la page d'accueil de leur site pour leur donner envie d'y entrer.

Récemment, une équipe de Microsoft Research (Liu, White, & Dumais, 2010) a proposé une modélisation mathématique pour expliquer les comportements d'abandon des usagers. Après avoir analysé le comportement de plus d'un million d'utilisateurs à travers 200000 sites web, les chercheurs ont déterminé que le temps passé par un individu sur une page Web suivait la distribution de Weibull (Cohen, 1965). Selon cette distribution, la probabilité qu'un individu quitte une page diminue à mesure que le temps passé dessus augmente.

Plus précisément, ce modèle a au départ été pensé pour rendre compte de la fiabilité de systèmes techniques. Cette distribution permet d'estimer la probabilité qu'un composant (d'un système technique donné) dysfonctionne à un instant  $t$  alors que ce composant a correctement fonctionné jusqu'à ce moment  $t$ . Cette distribution mathématique permet ainsi de prédire le moment auquel le composant doit être remplacé.

Dans le contexte de la navigation Web, la défaillance du composant correspond à l'abandon de la page Web par l'utilisateur. Selon la distribution de Weibull, deux vieillissements d'un dispositif peuvent être distingués pour expliquer une défaillance:

- le « vieillissement positif »: il s'agit du vieillissement le plus logique puisqu'il prédit que le risque de défaillance d'un composant est directement fonction de son temps d'utilisation ( $t$ ). En d'autres termes, plus un composant est utilisé, plus il a de risque de faire défaut;

- le « vieillissement négatif » : il prédit que le risque de défaillance d'un composant est inversement proportionnel à son temps d'utilisation (t). En d'autres termes, plus un composant est utilisé, moins il a de risque de dysfonctionner. Ce résultat s'expliquerait par le fait que, dans ce cas, le composant aurait prouvé sa robustesse.

Liu et ses collègues précisent que la grande majorité des pages Web présenteraient un « vieillissement négatif ». Ainsi, plus un individu reste sur une page Web, plus la probabilité qu'il la quitte diminue. Et il apparaît que les 10 premières secondes sont déterminantes (Liu, White, & Dumais, 2010; Nielsen, 2011). En d'autres termes, si un individu reste plus de 10 secondes sur la page d'accueil d'un site Web, alors la probabilité qu'il entre et explore le site augmente considérablement (Dinet, Vivian, La Mantia, Kitajima & Bertolo, 2013).

Nielsen conseille justement de mettre en évidence, sur la page d'accueil d'un site, sa valeur ajoutée, pour qu'en moins de 10 secondes l'utilisateur ait perçu l'information. Nielsen (2003) préconise qu'une page d'accueil doit contenir les zones d'informations suivantes :

- Outils d'aide à la navigation (20%)
- Le contenu (20%)
- De l'espace et des zones vides (20%)
- Le menu (19%)
- Du contenu de promotion (témoignages et feedback des usagers, clients) (9%)
- Du contenu interactif (vidéos, photos) (5%)
- Le logo de l'enseigne, mots de bienvenue (5%)
- Publicités (2%)

Il est intéressant de préciser que depuis quelques années, certaines recommandations portant sur les pages d'accueil accentuent d'avantage l'impact de l'esthétique sur l'affect, les émotions et l'expérience utilisateur. (Yu, Kong, 2011 ; Liu, Gao, Ye, Liang, 2016).

Enfin, à ce niveau d'analyse, certaines études se centrent sur les comportements visuels. Drusch, Bastien et Dinet (2011) ont ainsi réalisé une étude en laboratoire sur 113 participants afin de caractériser et de comparer les comportements oculaires des utilisateurs afin d'identifier d'éventuels patterns visuels intra-page. Pour se faire, ils ont appliqué une classification basée sur la distance de Hausdorff permettant de comparer la forme géométrique des scanpaths pour tous les participants. Des analyses complémentaires ont notamment été réalisées sur la base de comparaisons de données oculométriques inter-participants à l'aide de méthodes classification.

Ces chercheurs ont pu identifier pour deux tâches différentes, quatre classes d'utilisateurs selon leurs comportements oculaires. Les analyses portant sur les indicateurs oculométriques ont permis de mettre en exergue les différences inter-classes (Drusch, Bastien & Dinet, 2011 ; Drusch & Bastien, 2012).

### 3. REFLEXION METHODOLOGIQUE

---

L'analyse des comportements de navigation permet de mieux connaître les comportements des internautes et de formuler, en se basant sur les données empiriques, des recommandations pour la conception des navigateurs, sites ou pages Web. Une partie de ces études tente donc de décrire globalement les comportements de navigation des internautes. L'autre partie va plus loin et suggère qu'il existerait des patterns de navigation permettant d'affiner les recommandations pour la conception. A ce niveau certains chercheurs comparent des populations avec des caractéristiques différentes (âge, sexe, expertise, etc.).

A cet égard, les comportements des internautes (pages vues, temps passé sur les pages etc.) font l'objet d'observations parfois approfondies au travers de métriques qui permettent de les décrire. Toutefois, ces études sont bien souvent réalisées à un niveau très macroscopique ne permettant pas vraiment de distinguer les gens entre eux et de décrire de façon précise la façon dont les usagers naviguent (Jiang, Tan, Phang & Weik, 2012, Belk et *al.*, 2013, Zhou, 2014). Ce constat nous a amené à penser que le comportement des internautes devrait être analysé plus finement avant de formuler des recommandations. Le niveau d'analyse de la page permet justement une description fine des comportements, mais les études sont assez rares. La page d'accueil est, d'ailleurs, souvent citée comme la page centrale d'un site, déterminante pour expliquer la navigation sur les autres pages du site. Cependant, à notre connaissance, peu d'études se sont focalisées dessus pour expliquer des stratégies de navigation (Herder & Juvina, 2004).

Par ailleurs, d'un point de vue méthodologique, ces études sont souvent réalisées dans le contexte naturel des sujets alors que les buts de recherche des participants ne sont pas connus. Dans cette approche de web mining, ce type d'environnement, familier aux individus, a l'avantage de ne pas les perturber, mais le contrôle de certains facteurs parasites devient dans ce cas plus difficile puisqu'on ne peut pas être certain que chaque sujet navigue sur le web de façon similaire. Ce constat amène donc les chercheurs à comparer des parcours qui n'ont pas eu lieu sur les mêmes sites, pages, et qui découlent de tâches différentes (White, et *al.*, 2007) afin d'en dégager des stratégies de navigation.

Nous proposons donc d'analyser plus finement la façon dont les individus naviguent sur le web dans des conditions similaires afin d'identifier ou non des patterns sur des sites web. Cette approche pourrait, d'une part, nous permettre d'affiner nos connaissances car il est possible qu'indépendamment des caractéristiques des individus, des « styles » ou « patterns » de navigation existent. D'autre part, les recommandations actuelles concernant la conception des pages Web, ou encore la formation à la recherche d'informations pourraient être optimisées.

Cette thèse tentera donc de répondre à plusieurs questions :



- Quelle méthode d'analyse permettrait d'étudier la variabilité et de détecter des patterns comportementaux ?
- Existe-t-il des styles ou patterns comportementaux lors de la consultation de la première page vue qui entraineraient des niveaux de performance de navigation différents ?
- Par ailleurs, s'il existe bien des patterns, c'est à dire des façons de faire commune à un ensemble d'individus, est ce que ces patterns sont stables d'une tâche et d'un site à l'autre ?
- Quels sont les déterminants de ces patterns ?

# 4. ETUDES EMPIRIQUES

---

« Le comportement est un objet d'étude difficile, non parce qu'inaccessible, mais parce qu'il est extrêmement complexe »  
*Burrhus Frederic Skinner (1969).*

## 4.1 METHODES ET ANALYSES POUR L'IDENTIFICATION DE PATTERNS COMPORTEMENTAUX SUR DES SITES WEB

---

Certains points de cette partie reprennent des éléments publiés dans :

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(2), 159-171.

Becker, M., Bastien, J.M.C. & Drusch, G. (2013) Navigation sur le web : de nouveaux indicateurs pour l'identification de patterns comportementaux. EPIQUE 2013, Bruxelles, Belgique.

### 4.1.1 Problématique et contexte de l'étude

Nous avons vu précédemment que les auteurs conduisent des études pour améliorer certains éléments des sites et pages Web comme le plan, les aides en ligne, ou encore leur structure. Il est étonnant de voir que les interactions avec la page d'accueil n'ont pas été étudiées en elles-mêmes, surtout lorsque les patterns de navigation sont analysés dans l'optique d'aboutir à des recommandations. Il est pourtant possible que les comportements sur cette page aient une influence sur les performances de navigation globale. Nous pourrions penser que la page d'accueil n'a plus d'importance aujourd'hui à cause de l'utilisation généralisée des moteurs de recherche, mais ce n'est pas toujours le cas. En effet, les styles de navigation identifiés par Teevan, Alvarado, Ackerman et Karger (2004) puis Buscher, White, Dumais et Huang (2012) montrent que la recherche directe n'est l'apanage que de certains utilisateurs. Par ailleurs, pour la plupart des sites web, la page d'accueil est celle sur laquelle on compte le plus de visites (Nielsen & Loranger, 2006), et elle constitue le lieu de départ le plus commun pour les utilisateurs (Fang, Hu, Chau, Hu, Yang & Sheng, 2012). L'étude de la navigation sur cette page représente donc un enjeu capital pour la conception (Scott, 2015).

Dans cette optique, nous proposons d'analyser plus finement le comportement de participants sur la page d'accueil et du rapport de ces comportements avec la navigation dans le site.

- En d'autres termes, existe-t-il des styles ou patterns comportementaux lors de la consultation de la première page vue qui entraîneraient des niveaux de performance de navigation différents ?
- Par ailleurs, s'il existe bien des patterns, c'est à dire des façons de faire commune à un ensemble d'individus, est ce que ces patterns sont stables d'une tâche et d'un site à l'autre ?

Pour répondre à notre objectif, nous avons réalisé une expérimentation de laboratoire où les participants effectuaient tous les mêmes tâches dans les mêmes conditions, contrairement à la plupart des études précitées où les internautes consultaient librement le Web. Les conditions mises en place permettent donc d'explicitier les objectifs poursuivis et d'attribuer des variations comportementales à des styles et non à des objectifs différents.

Cette étude a été conduite en deux temps :

- Une première phase exploratoire pour laquelle une seule tâche du protocole a été traitée afin d'éprouver la méthode envisagée.
- À l'issue de cette première étude nos outils d'analyses et nos indicateurs ont été affinés, permettant ainsi d'analyser l'ensemble du protocole de façon plus approfondie.

## **4.1.2 Méthodologie expérimentale**

### **4.1.2.1 Matériel et terrain**

Nous avons utilisé le système d'oculométrie déporté Tobii T120. Le logiciel *Tobii Studio* (v.1.3.14) permet de recueillir des fichiers logs pour chaque participant et chaque tâche et ainsi de contrôler le protocole expérimental.

Il s'agissait d'une procédure expérimentale de laboratoire, réalisée dans Pergolab (la salle d'observation du laboratoire PErSEUs) pendant 25 jours. Les participants étaient installés face au système Tobii. Ce dernier était posé sur une stable stabilisée.



*Figure 6 : le dispositif expérimental*

La navigation sur les sites Web était rendue possible par la mise à disposition de la souris, les participants n'ayant pas le droit au clavier. Ce dernier était contrôlé par l'expérimentateur pour lancer les différentes phases de l'expérimentation.

### 4.1.3.2

### Procédures

Lors de l'accueil des participants, nous leur avons précisé qu'ils auraient à répondre à des questions en naviguant sur un site Internet particulier sans avoir l'usage du clavier. Nous leur avons expliqué que les questions ne comportaient pas de pièges et qu'il n'y avait pas de limite temporelle afférente aux tâches. L'expérimentateur était assis en léger retrait à côté du sujet et un observateur était installé à une table derrière eux. Ce dispositif permettait un double contrôle du bon déroulement de l'expérience car l'observateur notait systématiquement tout événement susceptible de constituer un biais expérimental.

Tous les participants réalisaient ensuite une tâche d'entraînement (exploration libre) sur deux sites Web, puis différentes tâches répétées sur deux sites Web.

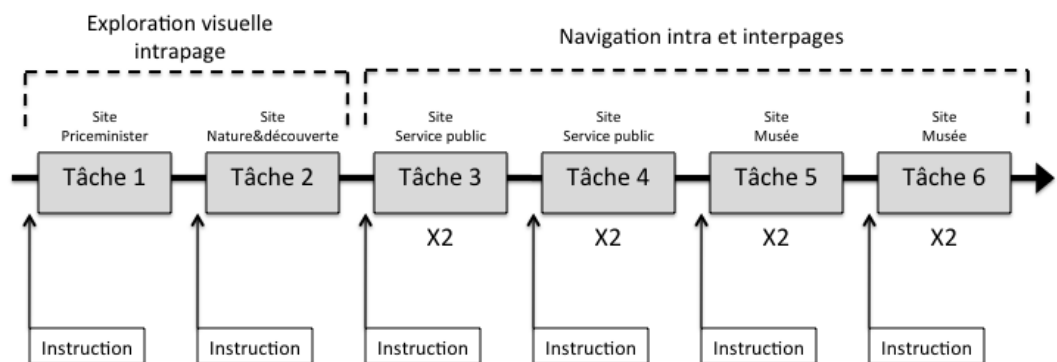


Figure 7 : Schéma récapitulatif de la procédure expérimentale

Les tâches étaient les suivantes.

1. « Explorez librement pendant 15 secondes la page Web qui va apparaître ». Page Web : [www.priceminister.com](http://www.priceminister.com) (page d'accueil enregistrée).

Figure 8 : Page d'accueil du site Priceminister

2. « Explorez librement pendant 15 secondes la page Web qui va apparaître ».  
 Page Web : [www.natureetdecouverte.com](http://www.natureetdecouverte.com) (page d'accueil enregistrée).



Figure 9 : Page d'accueil du site Nature et Découverte

Les tâches 3 à 6 ont été répétées deux fois de suite.

3. « À 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](http://www.service-public.fr) (page d'accueil)

4. « À 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 fonctionnaire) ? Quand vous aurez trouvé les réponses, citez-les à haute voix. » Page Web de départ : [www.service-public.fr](http://www.service-public.fr) (page d'accueil)

28 avril 2009

Deutsch | English | Español

Nouveau : connectez-vous à **Mon Service-Public.fr**

→ En savoir plus

**S'orienter**

- Sites internet publics**
  - nationaux, locaux
  - européens
  - organisations internationales
  - Etats étrangers
- Annuaire de l'administration**
  - services nationaux
  - services locaux

**Se documenter**

- Conventions collectives**
- Découverte des institutions**
- Indices de l'INSEE**
- Légifrance**
  - dernier journal officiel
  - sélection du mois
  - recherche d'un texte
  - tous les codes
- Panorama des lois**
- Rapports publics**

**Travailler dans l'administration**

- Concours de la fonction publique, bourses de l'emploi...**

Tapez les mots de votre recherche

Les 100 mots-clés les plus demandés

Recherche avancée

**Bienvenue dans l'espace particuliers**

**Vos droits et démarches**

- Je déménage
- Je prépare ma retraite
- Je cherche un emploi
- J'attends un enfant
- Je me marie
- J'achète un logement
- J'ai un décès dans ma famille
- J'organise une succession

- Achat, argent
- Armée
- Assurance
- Consommation
- Elections
- Emploi, travail
- Enseignement
- Environnement
- Etrangers en France
- Europe
- Famille
- Formation
- Impôt, taxe et douane
- Justice
- Logement
- Papiers
- Personnes handicapées
- Relations avec l'administration
- Retraite
- Santé
- Transports
- Vacances, loisirs
- Vie associative

Formulaires | Démarches en ligne | Adresses utiles

Nous contacter par téléphone, par courriel...

**3939**  
Allô, service public

Posez votre question

**Actualités**

**SCRUTIN** - Le point sur les élections européennes

**REPÈRES** - Impôt sur le revenu : quoi de neuf pour la déclaration 2009 ?

**H1N1** - Grippe porcine au Mexique : ce qu'il faut savoir

**PÔLE EMPLOI** - 39 49 : la nouvelle grille tarifaire

**SALARIES** - Fêtes légales et jours fériés

**1ER MAI** - Vente du muguet sur la voie publique

**EPARGNE** - 1er mai : taux du livret A fixé à 1,75 %

**DECOUVRIR** - UNESCO - La Bibliothèque numérique mondiale

\*Agenda  
\*Perspectives  
\*Météo  
\*Information routière

> Toute l'actualité

Recevoir la lettre hebdomadaire gratuitement

Mentions légales | Presse | Avis et suggestions | Statistiques | Plan du site | Syndication RSS  
Informations sur service-public.fr | Faire des liens vers service-public.fr | Co-marquage | Contact

© La Documentation française  
Autres sites gérés par la Documentation française :  
ladocumentationfrancaise.fr | pms.service-public.fr | vie-publique.fr | formation-publique.fr

Figure 10 : Page d'accueil du site Service Public

5. « À l'aide du site Internet qui va s'afficher, répondez à la question suivante : quelle est l'adresse du théâtre municipal du 15<sup>ème</sup> arrondissement parisien ? Quand vous aurez trouvé la réponse, citez-la à haute voix ». Page web de départ : [www.carnavalet.paris.fr](http://www.carnavalet.paris.fr) (page d'accueil).

6. « À 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](http://www.carnavalet.paris.fr) (page d'accueil).

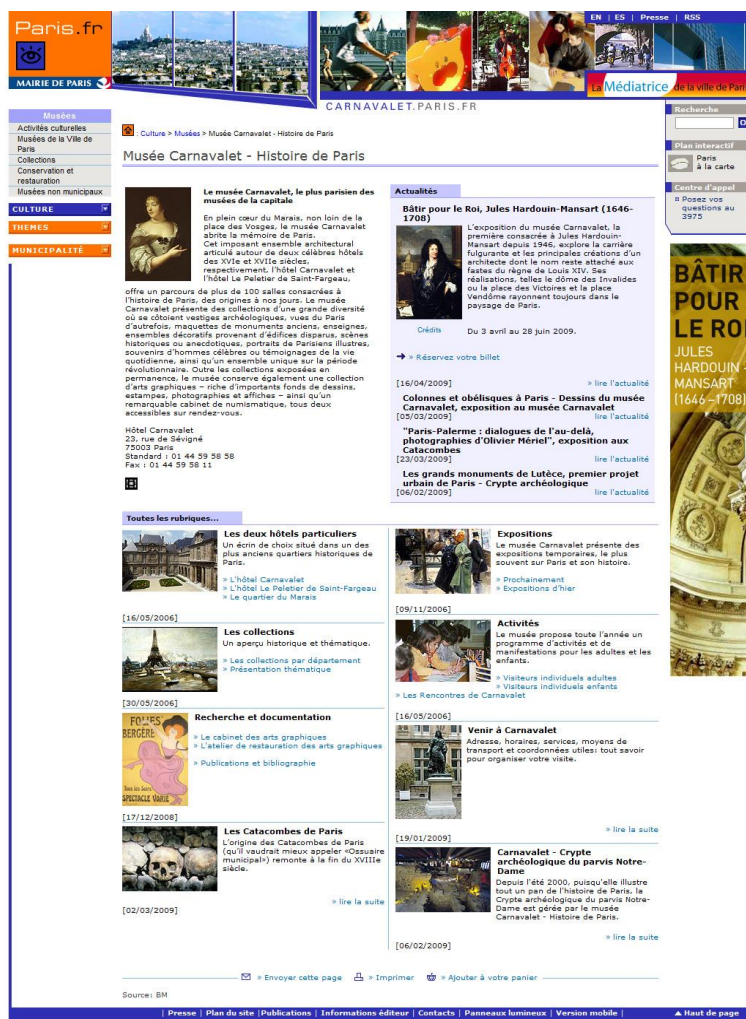


Figure 11 : Page d'accueil du site Musée Carnavalet

Les différentes tâches n'ont pas fait l'objet d'un contrebalancement car, étant dans l'optique d'une identification de patterns, l'objectif était que tous les participants soient placés dans des conditions similaires. Chaque tâche était précédée 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 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 post test était rempli par les participants. (Annexe A)



Par ailleurs, afin que ces sites correspondent à nos objectifs, ils devaient répondre à plusieurs critères. Les stimuli devaient être peu connus et faiblement mis à jour, ils devaient permettre d'effectuer des tâches plutôt simples tout en permettant d'accéder aux réponses par plusieurs chemins (2 minimums, le deuxième étant bien souvent le plan du site). Ces sites nécessitaient d'être relativement riche en informations afin de ne pas rendre les tâches trop simples. Il était également important de sélectionner des sites Web pour lesquels les participants ne risquaient pas d'être perturbés par des publicités ou pop-up intempestives. Enfin, un autre critère concernait l'organisation du contenu. L'un des sites Web devait se présenter de façon structurée (Service Public), tandis que l'autre (Musée Carnavalet) devait paraître comparativement moins organisé.

Site Web Service Public	Site Web Musée Carnavalet
Page d'accueil structurée en trois parties et présentant distinctement le menu principal au milieu	Page d'accueil agencée en plusieurs parties assez dispersées (une partie des menus sur la gauche, une autre en bas de pages et blocs d'informations au milieu)
Les items du menu principal sont directement accessibles et présentés sous format vignette ou mots-clefs	Nécessité de cliquer sur les titres du menu pour faire apparaître la fenêtre déroulante et les sous-menus
Le contenu principal de la page est affiché sans avoir besoin de scroller	Nécessité de scroller pour voir la page s'afficher en entier

#### 4.1.3.3

#### Questionnaire de fin d'expérience

Ce questionnaire a permis de contrôler, à posteriori, l'homogénéité de notre échantillon selon différents critères.

Il est découpé en plusieurs parties :

- Questions concernant l'étude : Nous avons évalué le degré de connaissance antérieure des sites Web parcourus durant les passations et le degré de difficulté perçue des différentes tâches. Nous avons aussi vérifié que les participants n'étaient pas conscients de l'objectif de notre expérimentation.
- Fiche signalétique : Recueil de données démographiques, 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 devaient se positionner en termes d'expertise par rapport à des affirmations. Cette partie du questionnaire a été inspirée et traduite de Lim et Lee (2000).

#### 4.1.3.4

#### Échantillon

Dans le cadre de cette expérimentation, 112 participants ont été recrutés. Par rapport à nos objectifs, et pour chacune des tâches respectives, seul les participants ayant réussi ont été retenus. Par ailleurs, pour certaines tâches, les données de plusieurs

participants ont été écartées à la suite de difficultés techniques (mauvais chargement des pages, problème d'enregistrement des données).

Le tableau ci-dessous présente les effectifs étudiés pour chaque tâche.

Tableau 2 : Effectifs pour chaque tâche, étude 1

Tâches	Effectifs
Tâche 3 Service Public	93 participants
Tâche 4 Service Public	101 participants
Tâche 5 Musée Carnavalet	99 participants
Tâche 6 Musée Carnavalet	86 participants

La moyenne d'âge était de 23 ans ( $\sigma = 3,72$ ) avec 47,7% d'hommes et 52,3% de femmes. La plupart des participants (93,8%) utilisaient Internet depuis plus de 3 ans et 87,7% étaient connectés sur le Web au moins une heure par jour. Le tableau ci-dessous présente les caractéristiques de notre échantillon.

Tableau 3 : Caractéristiques de l'échantillon, étude 1

Variables	Données sur les variables
Âge moyen (écart type)	23,64 ans (3,72)
Proportion d'hommes et de femmes	47,7% d'hommes et 52,3% de femmes
Nombre d'années d'utilisation d'Internet	Plus de 3 ans pour 93,8%
Fréquence d'utilisation d'Internet	Au moins 1 heure par jour pour 87,7%
Moyenne du niveau d'expertise sur le Web perçu par les participants	En moyenne, score de 3,82 (un niveau égal ou supérieur à 4 correspond à une connaissance experte) ( $\sigma = 0,65$ )

#### 4.1.4 Analyse des données exploratoires

Pour analyser les données de cette expérimentation, nous avons procédé en deux étapes ; une analyse exploratoire suivie d'une analyse plus approfondie.

La première a donc consisté à investiguer une seule tâche choisit au hasard. Il s'agit de la tâche 5 concernant le site Internet du Musée Carnavalet. C'est ce qui suit dans la prochaine partie.

La deuxième étape nous a permis d'étudier le reste de l'étude en affinant nos outils, nos analyses et en proposant une nouvelle méthode.

#### 4.1.4.1

### Modalités d'analyse des données de l'étude exploratoire

La préparation des données pour l'analyse exploratoire a nécessité plusieurs étapes. La première consistait à délimiter les sessions de navigation des participants. Le débit Internet ayant impacté le temps de téléchargement et d'affichage du site, et donc le temps mis pour réaliser la tâche, des segments ont été créés à partir du moment où la page du site apparaissait entièrement jusqu'à l'événement clavier indiquant que la tâche prenait fin.

Les données de ces segments ont été exportées vers Excel. Les fichiers logs issus de *Tobii Studio* nous fournissaient millisecondes par millisecondes le récit des événements ayant eu lieu pendant la navigation, le nombre et la nature des clics (clic sur un lien hypertexte ou un autre élément de la page, ou sur le navigateur), l'Url de chaque pages visitées, les coordonnées X et Y de la souris lors d'un clic, et la taille de la fenêtre de navigation (pouvant dépasser la taille de l'écran et justifier le scrolling). Il s'agissait d'un fichier par participant.

Un premier traitement des données a été réalisé sur Excel, ce qui a permis de calculer certaines métriques. Par exemple, les retours arrière ont pu être identifiés à partir des positions précises du curseur X et Y de la souris. Il a été nécessaire, dans un premier temps, de visionner les vidéos d'enregistrement pour repérer les clics correspondant à un retour en arrière, et voir, dans un second temps, comment cela était codé dans les fichiers logs. La figure XX illustre l'ensemble des étapes nécessaires à cette analyse.



Figure 12 : Principales étapes pour l'analyse des données

Au final, 9 indicateurs ont été étudiés.

Pour évaluer les performances de navigation, deux métriques d'efficacité à la tâche ont été calculées:

- Temps total pour accomplir la tâche (en millisecondes) ;
- Nombre total de pages visitées pour réaliser la tâche ;

Combinées à deux indicateurs de désorientation (Gwizdka & Spence, 2007) :

- Nombre de retours en arrière ;
- Taux de revisite (calculé selon la formule de Tauscher & Greenberg, 1997) :

$$\frac{\text{nombre total d'URLs visitées} - \text{nombre d'URLs différentes visitées}}{\text{nombre total d'URLs visitées}}$$

Il a été décidé également d'analyser le plus finement possible les comportements sur la page d'accueil, et de ne pas s'attarder uniquement aux métriques globales. Les indicateurs concernant les comportements de navigation sur les pages Web ayant été encore rarement utilisés dans ce cadre, nous avons calculé :

- Le temps passé sur la page d'accueil (PA) au début de la tâche, c'est à dire lors de la première visite ;
- Le nombre de clics gauche réalisés sur la page d'accueil au début de la tâche. Il s'agit des clics permettant d'accéder à une nouvelle page, de faire apparaître le contenu du menu déroulant ou encore de clics sur des éléments non cliquables ;
- Le nombre de scrolls montants et descendants. Ce comportement est comptabilisé dès que le défilement du scroll s'arrête plus de 500 millisecondes, ou qu'il change d'orientation.

Enfin, la plupart des participants ayant visité plusieurs fois la page d'accueil, nous avons également calculé des indicateurs permettant d'étudier la répartition des comportements en fonction du nombre total de visites sur cette page :

- Temps total passé sur la PA (calculé dès que la page d'accueil s'affiche jusqu'au clic sur un lien mettant fin à la navigation sur cette page) pondéré par le nombre de visites sur cette page ;
- Le nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page

Un tableau contenant, en ligne, la liste des participants et en colonne les différentes métriques a été obtenu.

Afin de caractériser les comportements de navigation sur le Web et d'identifier des patterns comportementaux, les indices présentés ci-dessus ont été traités. L'analyse a consisté à réaliser une classification ascendante hiérarchique (CAH) à l'aide du logiciel SPSS. Cette classification, a été calculée sur la base du carré de la distance euclidienne. Cette méthode partitionne la population en groupes homogènes ayant des caractéristiques communes (faible variabilité intra-groupe) et sépare les individus

ayant des caractéristiques différentes (forte variabilité inter-groupes). Les indicateurs ont été centrés-réduits afin d'éviter une sur- ou sous-pondération d'un indicateur. L'indice d'agrégation choisi pour regrouper les individus les plus proches est la méthode de Ward. Cette procédure consiste à minimiser l'inertie (ou variance) intra classe et de maximiser l'inertie inter classe. Chaque individu était ainsi associé à sa classe la plus proche (Murtagh & Legendre, 2011).

Cette agrégation d'individus a permis d'aboutir à un arbre de classification ou dendrogramme. En se basant sur la chute des classes, le dendrogramme (appelé aussi arbre hiérarchique) a été coupé en prêtant attention à garder des groupes les plus homogènes possibles. Ces analyses ont permis d'aboutir à trois classes de comportements.

#### 4.1.4.2 Résultats des analyses exploratoires et discussion

Comme nous l'avons indiqué plus haut, cette étude exploratoire se centre exclusivement sur l'une des 4 tâches qui consistait à trouver l'adresse du théâtre municipal du 15<sup>ème</sup> arrondissement de Paris sur le site du musée Carnavalet.

Les analyses réalisées ont permis d'aboutir à trois classes de comportements. Reprécisons que tous les participants de toutes les classes avaient réussi la tâche qui était assignée.

Les moyennes pour chacun de ces indicateurs sont présentées dans le tableau 4. Les valeurs les plus élevées apparaissent en gris.

Tableau 4 : Moyenne des indicateurs relevées pour chaque pattern

Métriques Analysées	Pattern 1 (n=41)	Pattern 2 (n=31)	Pattern 3 (n=25)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	59 399	122 147	65 606
Nombre total de pages visitées pour réaliser la tâche	3,37	6,06	3,92
Nombre de retours en arrière	0,10	0,55	0,12
Taux de revisite (en pourcentage)	2,97	8,01	2,17
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	29 386	15 747	19 177
Nombre de clics gauche réalisés	2,95	1,52	2
Nombre de scroll descendant	2,07	0,81	1,48
Nombre de scroll montant	1,63	0,39	0,92
<b>Comportements sur la page d'accueil pondéré par le nombre total de visites sur cette page</b>			

Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	28 037	15 160	16 728
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	2,91	1,62	1,97

Le graphique suivant, réalisé avec l'application Web Bertifier (Perin, Dragicevic & Fekete, 2014) permet d'illustrer rapidement les classes d'individus. Inspiré de la méthode d'analyse des matrices de Jacques Bertin, les valeurs du tableau 4 ont été codées afin de regrouper visuellement les lignes et les colonnes similaires.

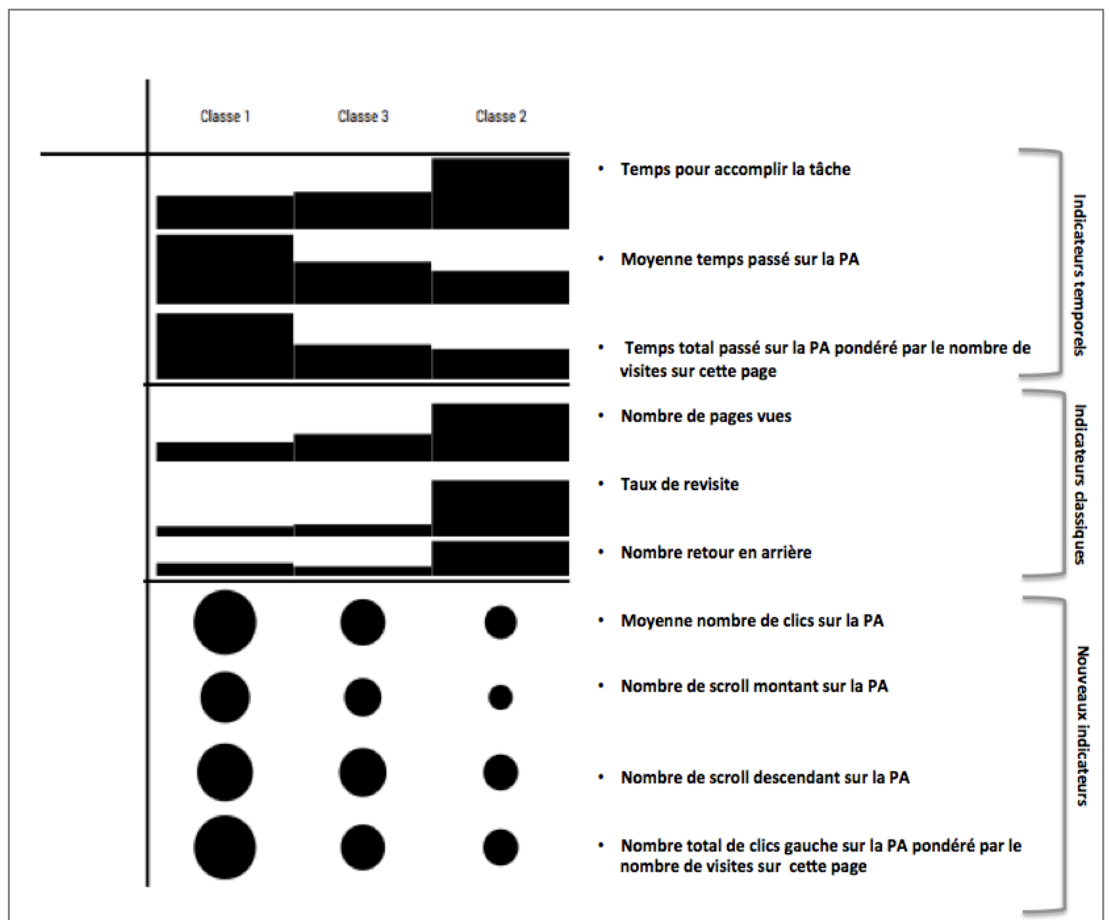


Figure 13 : Visualisation des classes avec la méthode Bertifier

Comme précisé par la tableau 4 et la figure 13, la première classe regroupe les individus qui réalisent la tâche le plus rapidement (59,3 secondes), et visitant en moyenne le moins de pages (3,37). Ce sont eux également qui effectuent le moins de retours en arrière (0,10). Cette classe d'individus est celle qui passe le plus de temps sur la page d'accueil au début de la tâche (29,3 secondes), et effectue le plus de scrolls montant (1,63) et descendant (2,07). Les individus effectuent aussi en moyenne plus de clics sur la page d'accueil pondérés par le nombre de visites (2,91) que les autres groupes.

Le deuxième groupe se différencie nettement du premier. Ces individus passent plus de temps pour réaliser la tâche (122,1 secondes), visitent (6,06) et revisitent plus de pages (80,1%), et font aussi plus de back (0,55). Ce sont eux qui cliquent et scrollent le moins sur la page d'accueil au cours de la première visite, tout en y passant le moins de temps (15,7 secondes).

Le dernier groupe identifié se caractérise par le plus faible taux de revisites (2,17%) et un temps assez faible pour réaliser la tâche (65,6 secondes). C'est un groupe intermédiaire : les moyennes des autres indicateurs se situent entre les classes 1 et 3.

D'après ces résultats, il semblerait que les comportements sur la page d'accueil contribuent le plus à la définition des classes. Les figures 13 et 14 illustrent clairement ce phénomène : on constate que les trois classes ont des ratios différents entre le temps passé sur la page d'accueil lors de la première visite (Temps Page d'Accueil, PA), et le temps total pour réaliser la tâche (Temps total). Bien que la classe 1 mette le plus de temps à naviguer sur la page d'accueil au début de la tâche, ces participants la réalisent le plus rapidement. La classe 2 au contraire passe en moyenne le moins de temps sur la page d'accueil mais passe beaucoup plus de temps pour réaliser la tâche.

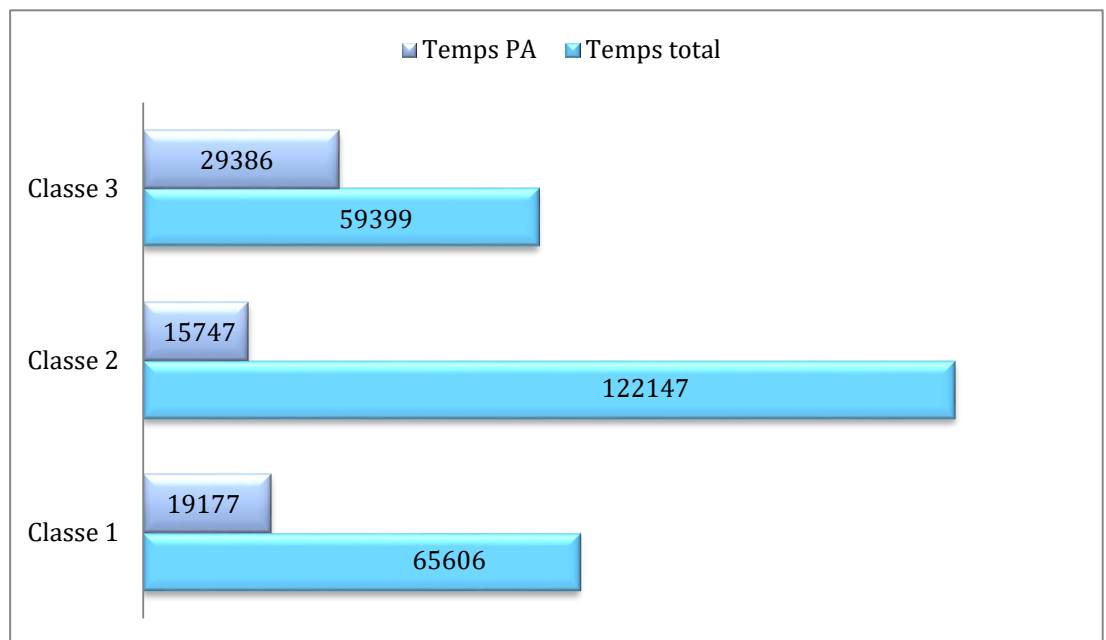


Figure 14 : Illustration des classes en fonction du temps passé sur la PA (en ms) au début de la tâche par rapport au temps total (ms) pour réaliser la tâche

Ces analyses préliminaires avaient donc pour objectif de savoir si les indicateurs choisis permettaient de faire apparaître des styles ou patterns comportementaux lors de la consultation de la page d'accueil qui entraîneraient des niveaux de performance de navigation différents. Des travaux similaires existent dans la littérature mais très rares sont ceux qui prennent en compte les interactions avec la page d'accueil. La prise en compte des interactions avec cette page nous a donc apporté de nouveaux éléments de réflexion concernant les comportements de navigation sur le Web. Des patterns

comportementaux ont pu être identifiés dans des conditions de passations strictement contrôlées.

En effet, la première classe identifiée à l'aide de la CAH rassemble 41 participants. Ce sont ces individus qui réalisent le plus de scrolls sur la page d'accueil et qui y passent le plus de temps tout en réussissant la tâche le plus rapidement. Ce groupe a une stratégie efficace et dynamique. Ces individus ont un taux de revisites très faible et réalisent très peu de retour en arrière. L'utilisation importante du scroll sur la page d'accueil lors de la première visite laisse supposer que ce groupe d'individus a exploré largement cette page afin de pouvoir s'orienter au mieux pour la suite de la navigation. Nous pouvons supposer que si ces participants effectuent la tâche de manière rapide et efficace c'est parce que le temps passé sur la page d'accueil conduit à la sélection correcte d'un lien permettant d'accéder à la bonne réponse. Ce pattern présente des similitudes avec les navigateurs décrits par White et Drucker (2007) bien que les conditions de passations de l'étude diffèrent.

Le deuxième groupe comprend 31 participants. Il est constitué des individus ayant passé le plus de temps pour réaliser la tâche et qui visitent le plus de pages différentes. Ces individus naviguent très peu sur la page d'accueil et n'y ont presque pas scrollé. En effet, c'est dans ce groupe que l'utilisation du scroll sur la page d'accueil est la plus faible. Cela suppose donc qu'ils n'ont pas pris le temps d'explorer toute la page d'accueil et qu'ils n'avaient pas en leur possession l'ensemble des informations permettant d'identifier rapidement le lien le plus pertinent. Ce sont eux également qui ont le taux de revisites le plus élevé, indiquant une certaine désorientation. Ainsi, bien qu'ils réussissent la tâche, il semblerait qu'ils se perdent sur le site.

Enfin, la troisième classe est composée d'individus qui réalisent la tâche assez rapidement et qui explorent peu la page d'accueil. Ces individus visitent un nombre de pages assez restreint et ont un faible taux de revisites, ce qui signifie qu'ils ne reviennent que très rarement sur leurs traces. Dans leur cas, à partir du moment où le choix de parcours est fait, les individus s'y tiennent, et il semblerait que les liens qu'ils suivent leur permettent d'accéder rapidement à l'information. Nous pouvons supposer que ces individus ont la capacité à collecter rapidement les informations pertinentes pour la réussite de la tâche.

Les patterns identifiés semblent donc indiquer que les interactions avec la page représentant le point de départ (dans notre expérimentation, la page d'accueil) à une tâche déterminent les performances de navigation. En effet, les principales différences concernant ces patterns s'expliquent par leurs façons d'explorer la page d'accueil. Ce résultat est particulièrement intéressant car il laisse envisager la possibilité que cette page influe sur toute la navigation, mais d'une façon différente selon le pattern. Cependant, cette analyse exploratoire présente certaines limites que nous avons pris en compte afin de poursuivre nos travaux.

En effet, le traitement et l'analyse non automatisé des fichiers logs issus de *Tobii Studio* ont constitué des difficultés auxquelles il a fallu remédier. De plus, sans outil d'analyse automatisé, la distance de scroll ou encore les chemins de navigation n'ont pas pu être analysés. En tenant compte de ces métriques, il est probable que certains patterns auraient pu être davantage explicités ou même découverts. Par ailleurs, la classification ascendante hiérarchique réalisée pour l'analyse exploratoire présente



également une limite. La découpe des classes n'étant pas automatisée, il n'est pas toujours évident de savoir si le nombre de groupe identifié est bien optimal (Bittner, Meltzer & Trent, 1999).

Ainsi, en considérant de nouvelles modalités d'analyse des données, le protocole de recherche devrait permettre de répondre à plusieurs questions. Qu'en est-il de la stabilité de ces patterns comportementaux ? D'autres tâches sur d'autres sites Web permettraient-elles de classer les individus dans les mêmes groupes ? Ces patterns dépendent-ils du contexte, c'est à dire du site ? La partie suivante présente ces analyses approfondies et répond à ces questions.

## **4.1.5 Analyse des données approfondies**

### **4.1.5.1 Modalités d'analyse des données**

Pour pouvoir analyser l'ensemble des tâches sur les différents sites Web, il a été nécessaire de développer un nouvel outil permettant d'automatiser le calcul des indicateurs. Nos besoins bien précis, et le format de fichiers logs de *Tobii Studio* étant bien spécifiques, il n'existait rien pouvant répondre à nos attentes. L'outil Web « *Application Analyzer* » pour le traitement de données comportementales sur Internet a été développé.

Le fonctionnement de l'outil était simple. Une fois les fichiers des participants sélectionnés, il ne restait plus qu'à choisir les indicateurs que l'on souhaitait analyser. Deux exports étaient possibles : par page (permettant de recueillir des données pour une page précise, la page d'accueil par exemple) et par tâche (afin d'obtenir des données générale sur la tâche).

Au-delà de la rapidité d'analyse de l'outil, l'intérêt résidait également dans sa précision et sa fiabilité pour l'analyse des métriques. Il a été possible, par exemple, de consolider l'analyse du scroll en incluant la notion de distance et ainsi rendre l'indicateur plus fin. En effet, les principales différences concernant les patterns identifiés pour l'analyse exploratoire s'expliquaient par leurs façons d'explorer la page d'accueil. Il était donc important de poursuivre notre recherche en incluant des analyses plus fines et complètes sur le comportement ayant lieu sur cette page.

Au total, 16 indicateurs ont été étudiés. Les métriques associées à un astérisque (\*) correspondent à ceux ayant été rajoutés ou modifiés par rapport à l'analyse exploratoire.

Pour évaluer les performances de navigation, les mêmes métriques d'efficacité à la tâche que pour les analyses exploratoires ont été sélectionnées :

- Temps total pour accomplir la tâche (en millisecondes) ; Catledge et Pitkow (1995) ; Obendorf et Weireinch (2007) ; Cockburn et Mackenzie (2002)
- Nombre total de pages visitées pour réaliser la tâche ;

Combinées à deux indicateurs de désorientation (Gwizdka & Spence, 2007) :

- Nombre de retours en arrière ; Gwizdka et Spence (2007), Danielson (2002)
- Taux de revisite (calculé selon la formule de Tauscher & Greenberg, 1997) ;

L'analyse des comportements sur la page d'accueil au moment de la première visite a été affinée :

- Le temps passé sur la page d'accueil (PA) lors de la première visite ; Herder et Juvina, (2004)
- Le nombre de clics gauche réalisés sur la page d'accueil lors de la première visite ;
- \* Le temps de latence avant la première action. Cet indicateur définit le temps pris par l'individu avant toute action sur la page (clic sur un menu déroulant, clic sur un lien hypertexte afin de quitter la page ou scroll).
- \* La distance moyenne de scroll montant (en pixel). Il s'agit ici de pondérer la distance de scroll montant par le nombre de scroll montant.
- \* La distance moyenne de scroll descendant (en pixel). Comme pour l'indicateur de scroll montant, il s'agit de pondérer la distance de scroll descendant par le nombre de scroll descendant.
  - Comme pour l'étude de Buscher, White et Dumais (2012), les coordonnées relatives au scroll sont vérifiées plusieurs fois par seconde. Le scroll est comptabilisé dès que son défilement s'arrête plus de 500 millisecondes minimum ou qu'il change de sens.
  - Ces deux indicateurs permettent d'obtenir une moyenne de la distance de scroll plus précise que l'indicateur nombre de scroll. En effet, prenons l'exemple de 2 participants S1 et S2 ayant réalisés chacun une distance de scroll descendant de 200 pixels. S1 a utilisé la barre d'ascenseur 2 fois tandis que S2 4 fois. S1 aura donc une distance moyenne de scroll descendant de 100 pixels tandis que S2 50 pixels. Par rapport à S1, S2 réalise donc de plus petits scrolls par « à coups ».
  - Ainsi, tout en tenant compte de la distance de scroll, cette nouvelle métrique permet de mieux préciser et différencier les comportements sur la page d'accueil.

La plupart des participants ayant visité plusieurs fois la page d'accueil, nous avons également calculé des indicateurs permettant d'étudier la répartition des comportements en tenant compte du nombre de visites sur cette page :

- \* Le nombre de visites sur la PA pendant la tâche ;
- Temps moyen passé sur la PA en fonction du nombre de visites sur cette page ;
- Le nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page.

À ceux-ci ont été rajoutés les indicateurs relatifs au scrolling (Buscher, White & Dumais, 2012) :

- \* Distance totale de scroll descendant en fonction du nombre total de scroll descendant sur la PA ;
- \* Distance totale de scroll montant en fonction du nombre total de scroll montant sur la PA.
  - Ces deux indicateurs indiquent si les participants en revenant sur cette page, l'ont davantage exploré ou non.

Afin d'avoir des données sur le comportement de navigation des participants sur les autres pages visitées, nous avons calculé deux indicateurs :

- \* Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées ;
- \* Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues.
  - Il nous semblait surtout pertinent d'étudier cette métrique car de nombreuses pages sur les sites comportaient des menus déroulant et cela pouvait nous renseigner sur l'exploration du site.

Enfin, pour chaque participant nous avons observé son parcours de navigation :

- \* Enchaînement des URLs visitées.
  - Afin de compléter les analyses et d'étayer les éventuels patterns qui pourraient être trouvés, il était pertinent de différencier et d'estimer la variabilité des parcours de navigation entre tous les participants. Tout comme White et Drucker (2007), nous avons utilisé la Distance de Levensthein (LD) (Levensthein, 1966). Cette méthode permet de comparer les sessions de navigation 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. L'indice varie de 0 (similarité parfaite) à 1 (dissimilarité parfaite). Précisons qu'une chaîne représente une session de navigation. Chaque lettre de la chaîne représente une URL.

Par exemple pour les parcours des sujets S1 et S2 :

S1 = ABCDE	} Opérations à effectuer : remplacement de « C » et « D »
S2 = ABBCE	

La distance d'édition sera ici de 2 et l'indice de similarité (ou LD normalisée) de 0,4. La chaîne comportant 5 caractères, on divise par ce nombre.

Trois autres indicateurs ont été considérés ; scroll horizontal droit, scroll horizontal gauche et clic droit, mais ces comportements n'ont jamais été observés dans l'échantillon.

Comme pour l'analyse exploratoire nous avons souhaité faire une classification. Toutefois, comme nous l'avons dit précédemment, les classes étant définies par les branches du dendrogramme, l'une des principales difficultés peut concerner l'interprétation et la qualité des classes obtenues afin d'extraire des inférences pour comprendre la navigation des utilisateurs (Bittner, Meltzer & Trent, 1999 ; Chen & Liu, 2003). Palis, Angelis, et Vakali (2007) recommandent donc d'opter pour une approche qui sélectionne automatiquement le nombre de classes « optimales ».

La méthode du *Dynamic Hybrid Cut* (Langfelder, Zhang & Horvath, 2007) a été utilisée avec le logiciel R. Elle distingue les classes très proches qui peuvent parfois paraître imbriquées, détecte les « outsiders », découpe automatiquement les classes et elle combine les avantages de la classification hiérarchique et du partitionnement autour des médoïdes (PAM) (Kaufman & Rousseeuw, 2005). Un médoïde est le représentant le plus central d'une classe. Le PAM réduit la distance entre les points d'une classe et le médoïde désigné pour être le meilleur représentant de ce groupe.

Comme le PAM, la méthode du *Dynamic Hybrid Cut* s'applique sur les matrices de dissimilarités, puis sur les informations du dendrogramme.

Le *Dynamic Hybrid Cut* procède en deux étapes :

- 1<sup>ère</sup> étape : les branches considérées comme des classes sont détectées
- 2<sup>ème</sup> étape : tous les objets précédemment non affectés à un groupe sont testés afin de voir s'ils ont une proximité suffisante avec les classes détectées dans la première étape, ou s'ils sont considérés comme des « outsiders ».

Au final, les classes détectées par l'algorithme se voient chacune attribuer une couleur et les objets non affectés (« outsiders ») sont assignés à la couleur grise (Figure 10).

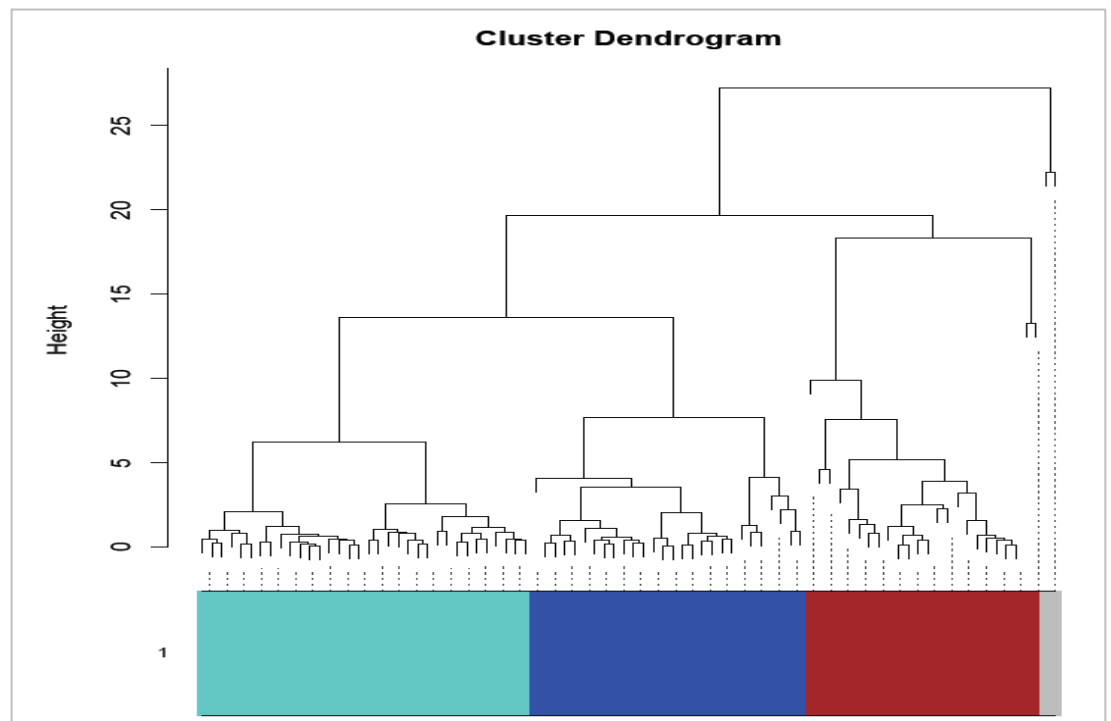


Figure 15 : Exemple de dendrogramme avec la méthode *Dynamic Hybrid Cut*

## 4.1.5.2

### Résultats des analyses approfondies

Rappelons ici que tous les participants réalisaient deux tâches différentes sur deux sites Web (T3-T4 sur le site du Service Public et T5-T6 sur le site du Musée Carnavalet). Chacune de ces tâches été répétées tout de suite, avant même d'en commencer une nouvelle.

Les analyses sont présentées en suivant l'ordre dans lequel les participants ont réalisé les tâches.

#### ➤ Site Service Public – Tâche 3, conditions de dissolution du PACS

L'application de la procédure que nous avons décrite précédemment permet d'identifier deux patterns (ou classes) différents. Pour cette tâche, une des classes est composée de 63 participants et la seconde de 30 participants.

Les moyennes des indicateurs sont présentées en Annexe A. L'illustration suivante permet de présenter graphiquement et rapidement les 2 classes d'individus et leurs comportements.

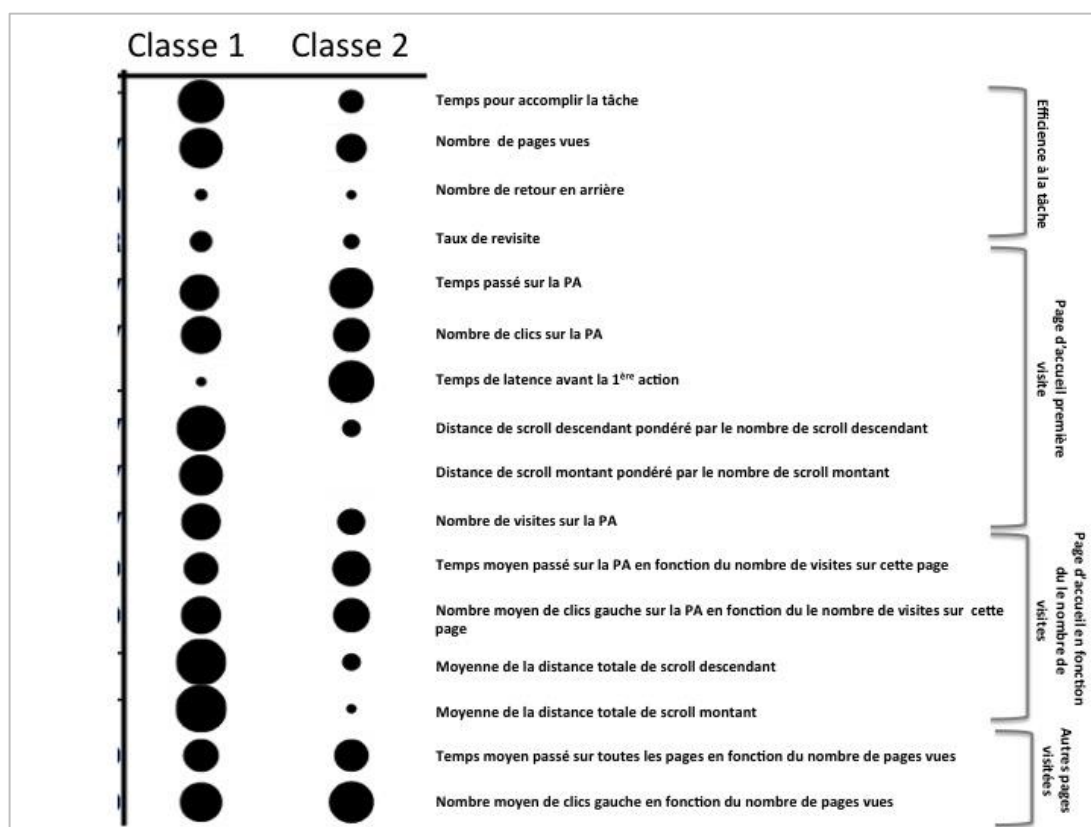


Figure 16 : Illustration des patterns pour la Tâche 3 – Service Public

Comme l'illustre la représentation Bertier, le pattern 1 regroupe la majorité des participants :

- En visitant en moyenne 13 pages (deux fois plus que le pattern 2), et en revenant presque 3 fois sur la page d'accueil du site, ces individus mettent près de deux fois plus de temps pour effectuer la tâche (219,9 secondes).

- Concernant le comportement sur la page d'accueil lors de la première visite, ils l'explorent avec une distance moyenne de scroll descendant de 115 pixels et de 106 pixels pour les scrolls montant. Enfin, ces participants effectuent très rapidement une première action sur la page (au bout de 11 secondes).
- **Malgré une exploration assez approfondie de la page d'accueil, ces individus semblent être désorientés.**

Concernant le pattern 2 :

- Les participants accomplissent, en moyenne, la tâche une fois et demie plus vite (135,4 secondes) et pourtant ils passent un peu plus de temps que les individus du pattern 1 sur la page d'accueil. Durant cet intervalle de temps, les participants scrollent sur des distances moyennes très faible, et ils n'effectuent qu'un clic ciblé sur la page afin de sélectionner un lien. Précisons que sur cette page, le scroll n'était pas nécessaire car le contenu principal de la page était entièrement affiché.
- Par ailleurs, en passant en moyenne un peu plus de temps sur la page d'accueil, ces individus prennent aussi trois fois plus de temps avant de réaliser une première action sur la page (33 secondes).
- Les individus du pattern 2 semblent donc aussi « **exhaustif** » dans leur façon de naviguer mais pour eux cette stratégie leur permet d'être **plus rapide et efficace en les conduisant à emprunter les parcours les plus optimaux.**

En effet, afin de mieux décrire nos patterns et compléter ces données, nous avons étudié la variabilité des parcours en calculant la Distance de Levensthein (expliquée plus haut page). La matrice de similarité obtenue, nous a permis de réaliser une classification reposant sur la méthode du *Dynamic Hybrid Cut*. Cette classification a ainsi permis de regrouper les individus ayant emprunté les parcours de navigation les plus similaires entre eux.

Par ailleurs, pour mieux illustrer chacune des classes de parcours, nous avons retenu le parcours avec la distance de Levensthein (DL) la plus représentative, c'est à dire celle qui a la plus faible valeur. En ayant ainsi, en moyenne, la plus forte similarité avec les autres chemins de navigation des participants au sein de la classe, nous avons pu représenter visuellement ce parcours type.

Pour cette première tâche réalisée sur le site du Service Public, la classification des parcours basée sur la DL a permis de classer les participants en fonction de 3 types de parcours différent.

Le tableau 5 permet d'étudier la répartition des parcours au sein de chaque pattern. Pour chaque pattern, la classe de parcours la plus empruntée est en surbrillance.

Tableau 5 Répartition des parcours pour chaque pattern pour la Tâche 3 - Service Public

	Pattern 1	Pattern 2
Parcours classe 1	<b>33,3%</b> (n=21)	<b>70%</b> (n=21)
Parcours classe 2	<b>27%</b> (n=17)	<b>16,7%</b> (n=5)
Parcours classe 3	<b>39,7%</b> (n=25)	<b>13,3%</b> (n=4)
Nombre total de participants	<b>100%</b> (n=63)	<b>100%</b> (n=30)

En général, dans l'échantillon, les parcours de la classe 1 sont les plus empruntés. Le parcours le plus représentatif de cette classe comprend 5 pages et il s'agit du chemin optimal permettant d'accéder le plus rapidement à la réponse.

Au sein du pattern 2 regroupant les individus ayant une stratégie plus « posée », cette classe de parcours a été empruntée majoritairement à 70%. Le parcours 1 le plus représentatif est représenté ci dessous. La représentation des chemins s'inspire de l'étude de White et Drucker (2007) et White et Horvitz (2012).



Figure 17 : Chemin le plus représentatif pour le parcours n°1 Tâche 3 - Service Public (LD = 0,12)

Pour la seconde classe de parcours, le chemin le plus représentatif est constitué en moyenne de 7 pages. 27% des participants du pattern 1 et 16,5% du pattern 2 ont emprunté ces parcours afin d'accéder à la réponse.

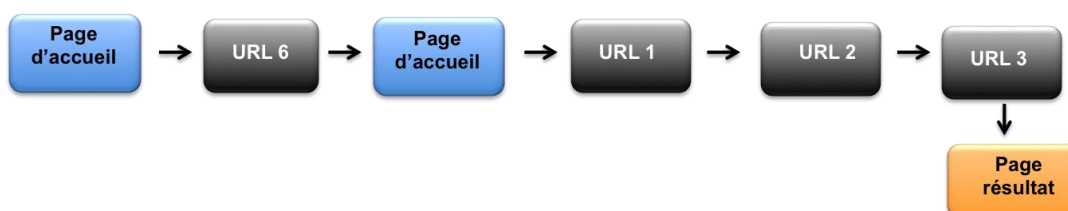


Figure 18 : Chemin le plus représentatif pour le parcours n°2 Tâche 3 - Service Public (LD = 0,22)

Enfin, les parcours de la classe 3 regroupent 29 participants. Au sein du pattern 1, il s'agit de la classe de parcours la plus empruntée à 39,7%. Rappelons que ce pattern regroupe les individus ayant mis le plus de temps pour réaliser la tâche, visité le plus de pages et ayant rencontré le plus de difficultés à sélectionner le lien optimal pour accéder rapidement à la réponse. La distance de Levensthein est très élevée pour cette classe car les individus ayant vu en moyenne 13 pages, les parcours de navigation sont très hétérogènes.

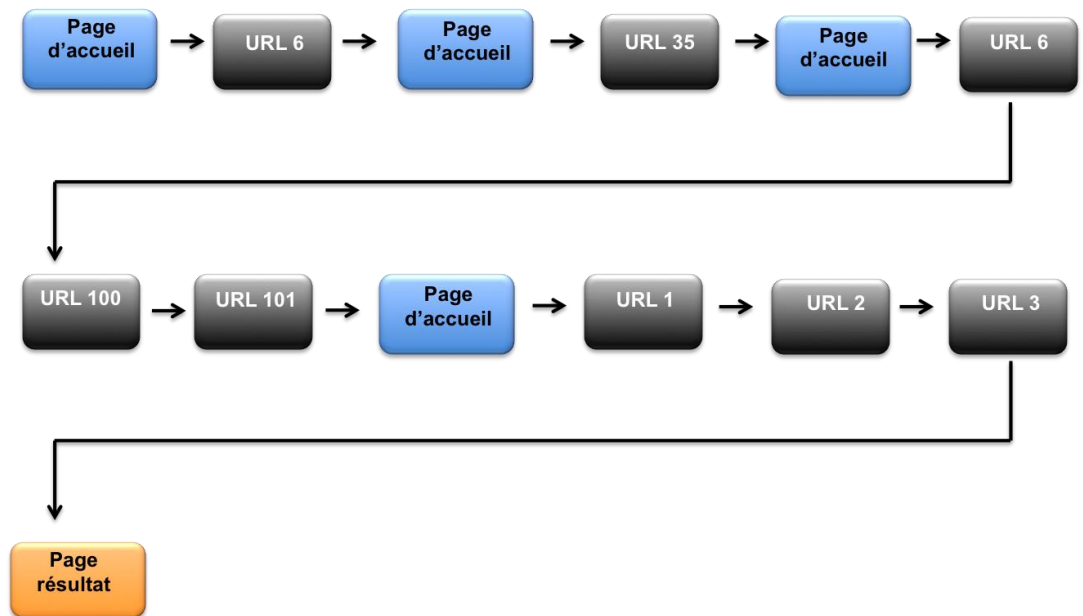


Figure 19 : Chemin le plus représentatif pour le parcours n°3 Tâche 3 - Service Public LD = 0,65

- *Site Service Public – Tâche 4, deux premières conditions générales de recrutement pour accéder à la fonction publique*

L'analyse de la tâche suivante réalisée sur le même site met en exergue trois patterns. Le premier est constitué de 47 participants, et les deux autres respectivement de 24 individus chacun. Six participants ont été considérés comme des outsiders pour la classification.

Les moyennes pour chaque pattern et chacun des indicateurs sont présentées en Annexe A. La visualisation réalisée avec l'application Web Bertifier permet d'illustrer rapidement ces 3 classes d'individus.



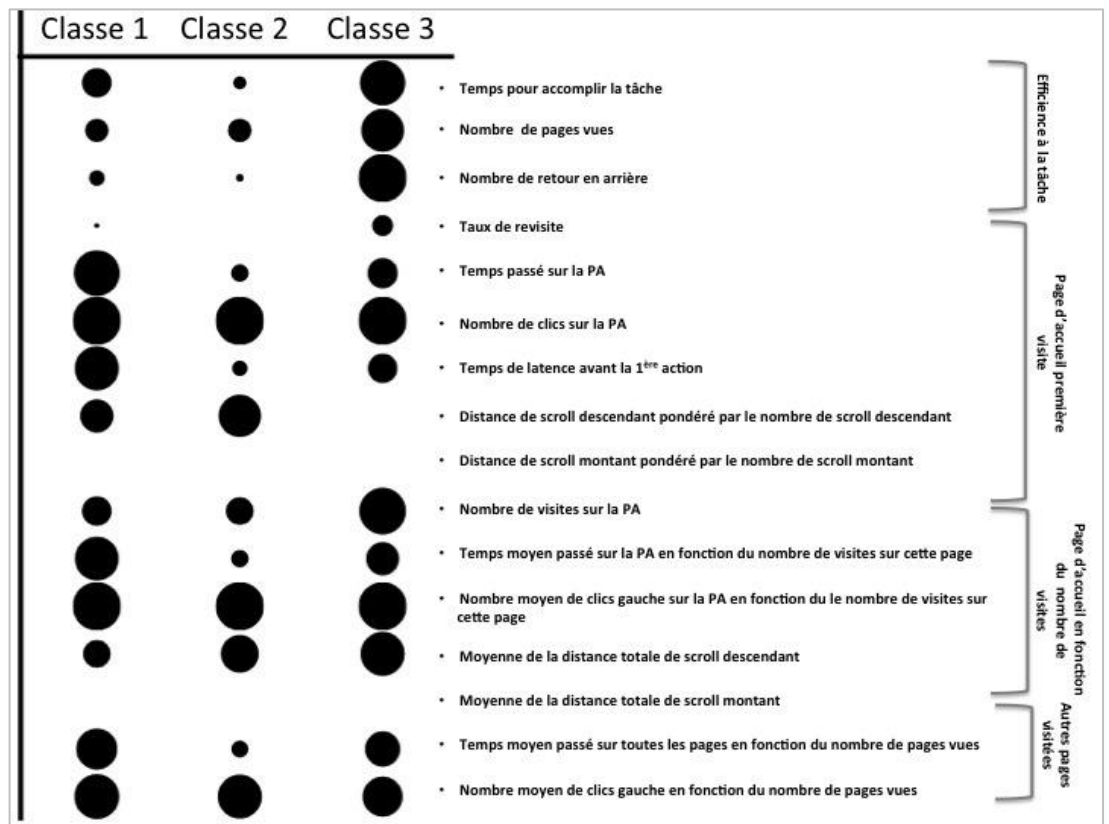


Figure 20 : Illustration des patterns pour la Tâche 4 - Service Public

La classe 1 rassemble la majorité des participants avec 47 individus :

- De façon générale, ils mettent trois fois plus de temps que les individus du pattern 2 et presque deux fois plus de temps que les individus du pattern 3 sur la page d'accueil à la première visite, c'est à dire 18,6 secondes.
- De plus, ils sont ceux prenant, en moyenne, le plus de temps avant de réaliser une action sur la page d'accueil et à naviguer sur l'ensemble des pages du site. **Cette façon de naviguer peut également être caractérisée comme « consciencieuse ».**
- Toutefois, bien qu'ils ne soient pas les plus rapides pour réaliser la tâche (71,2 secondes), ils explorent le moins de pages pendant la tâche (5,82) et n'effectuent presque aucun retour arrière.

La classe 2 est la plus rapide pour réaliser la tâche en 38 secondes (presque trois fois moins de temps que les participants du pattern 3 et, une fois et demi moins de temps que ceux du pattern 1) :

- Pourtant elle ne passe en moyenne que 6,9 secondes sur la page d'accueil lors de sa première et unique visite. Au sein de cette classe, les participants sont ceux également qui passent le moins de temps à naviguer sur les pages visitées.
- Ils sont ceux réalisant le plus rapidement une première action sur la première page vue. **En effectuant cette deuxième tâche, il semblerait que ces participants aient retenu la structure du site.**

Enfin la troisième classe est celle ayant mis en moyenne le plus de temps pour réaliser la tâche (127,5 secondes) :

- Ils passent cependant peu de temps sur la page d'accueil lors de la première visite (4 secondes de moins que la classe 2 qui elle a mis 3 fois moins de temps à réaliser la tâche) et ils passent relativement peu de temps à explorer la page avant de réaliser une première action (10,6 secondes).
- Au cours de cette première visite, les participants ne scrollent pas du tout sur la page d'accueil. Ce sont eux également qui font le moins de clics sur l'ensemble des pages visitées.
- Par ailleurs, ce groupe visite beaucoup de pages (12,5) et revient presque 3 fois sur la page d'accueil du site. Le taux de revisite de 29% et les précédents éléments cités laissent penser que **ces participants se perdent au cours de la navigation**. Ici contrairement aux participants du pattern 2, **cette rapidité de navigation sur la page d'accueil à la première visite semble plus infructueuse**.
- Comme pour le pattern 1 Tâche 3, il semblerait donc qu'une stratégie trop rapide peut être moins fructueuse. **Certains individus ont besoin de prendre plus de temps sur les pages visitées, y compris la première page vue, pour prélever les bonnes informations.**

Concernant les chemins de navigation, la classification des chemins catégorise les participants en fonction de trois groupes de parcours. Le tableau suivant permet d'étudier la répartition des parcours au sein de chaque pattern.

Tableau 6 : Répartition des parcours pour chaque Pattern pour la Tâche 4 - Service Public

	Pattern 1	Pattern 2	Pattern 3
Parcours classe 1	<b>36,1%</b> (n=17)	<b>12,5%</b> (n=3)	<b>0%</b> (n=0)
Parcours classe 2	<b>46,9%</b> (n=22)	<b>83,3%</b> (n=20)	<b>29,2%</b> (n=7)
Parcours classe 3	<b>17%</b> (n=8)	<b>4,2%</b> (n=1)	<b>70,8%</b> (n=17)
Nombre total de participants	<b>100%</b> (n=47)	<b>100%</b> (n=24)	<b>100%</b> (n=24)

La première classe de parcours regroupe 20 participants. Le chemin le plus représentatif avec la LD moyenne la plus faible implique de passer par le plan du site. C'est au sein du Pattern 1 que cette classe de chemin a été plus suivie. Précisons que c'est dans ce pattern que les participants ont passé le plus de temps sur la page d'accueil et qu'ils ont pris le plus de temps avant de réaliser une action sur la page. Ce temps de latence élevé leur a sans doute été profitable pour identifier ce lien.

Le parcours 1 le plus représentatif de la classe est représenté ci-dessous.



Figure 21 : Chemin le plus représentatif pour le parcours n°1 Tâche 4 - Service Public LD = 0,06

49 individus ont emprunté les parcours de la classe 2. Le chemin le plus caractéristique de la classe montre que dans ce cas les participants passent par le chemin « classique » et non via le plan du site. Nous considérons que ce parcours de navigation est tout aussi optimal. Respectivement, au sein du Pattern 1 les parcours de cette classe ont été suivis à 46,9% et au sein du Pattern 2 à 83,3%.



Figure 22 : Chemin le plus représentatif pour le parcours n°2 Tâche 4 - Service Public LD = 0,11

Enfin, la dernière classe regroupe 26 participants. Au sein de ce groupe, le chemin le plus représentatif fait en moyenne 6 pages, il s'agit donc ici des parcours les plus laborieux. Au sein du Pattern 3, 71% des individus ont emprunté ce type de parcours. Si l'on s'intéresse à la description des participants du pattern, nous pouvons envisager que la rapidité et la faiblesse d'exploration de la première page vue peuvent être déterminantes pour la performance de la navigation chez certains individus.



Figure 23 : Chemin le plus représentatif pour le parcours n°3 Tâche 4 - Service Public LD = 0,55

### Résumé :

Les analyses réalisées sur le site du Service Public mettent en exergue plusieurs patterns dont un présentant de fortes similitudes entre les deux tâches. Le style de navigation que l'on a caractérisé de « consciencieux, exhaustif » regroupe des individus prenant leur temps pour naviguer sur l'ensemble des pages du site, y compris sur la page d'accueil, et ici prenant plus de temps que les autres avant de réaliser une première action. Ces individus semblent avoir ainsi en leur possession les informations permettant de s'orienter au mieux sur le site. Toutefois, en fonction de la tâche, cette façon de naviguer n'engendre pas toujours la meilleure performance. En effet, nos résultats semblent indiquer que, d'une tâche à l'autre, certains participants ont plus de facilités à mémoriser la structure du site que d'autres. Ainsi en **étant les plus rapides, ces individus sont aussi les plus efficaces.**

Par ailleurs, le pattern 1 Tâche 3 et le pattern 3 Tâche 4 laisse envisager l'idée que même si tous les individus accèdent à la réponse, **certains individus auraient besoin de prendre plus de temps sur les pages visitées, y compris la première page vue, pour prélever les bonnes informations** et accéder à la réponse.

➤ *Site Musée Carnavalet – Tâche 5, Adresse du théâtre municipal du 15<sup>ème</sup> arrondissement de Paris*

Cette tâche est la première sur le site du Musée Carnavalet. Comme pour la première tâche du Service Public, deux classes d'individus sont observées. La première regroupe 62 participants et la deuxième 37. Les moyennes pour chaque patterns et indicateurs sont présentées en Annexe A.

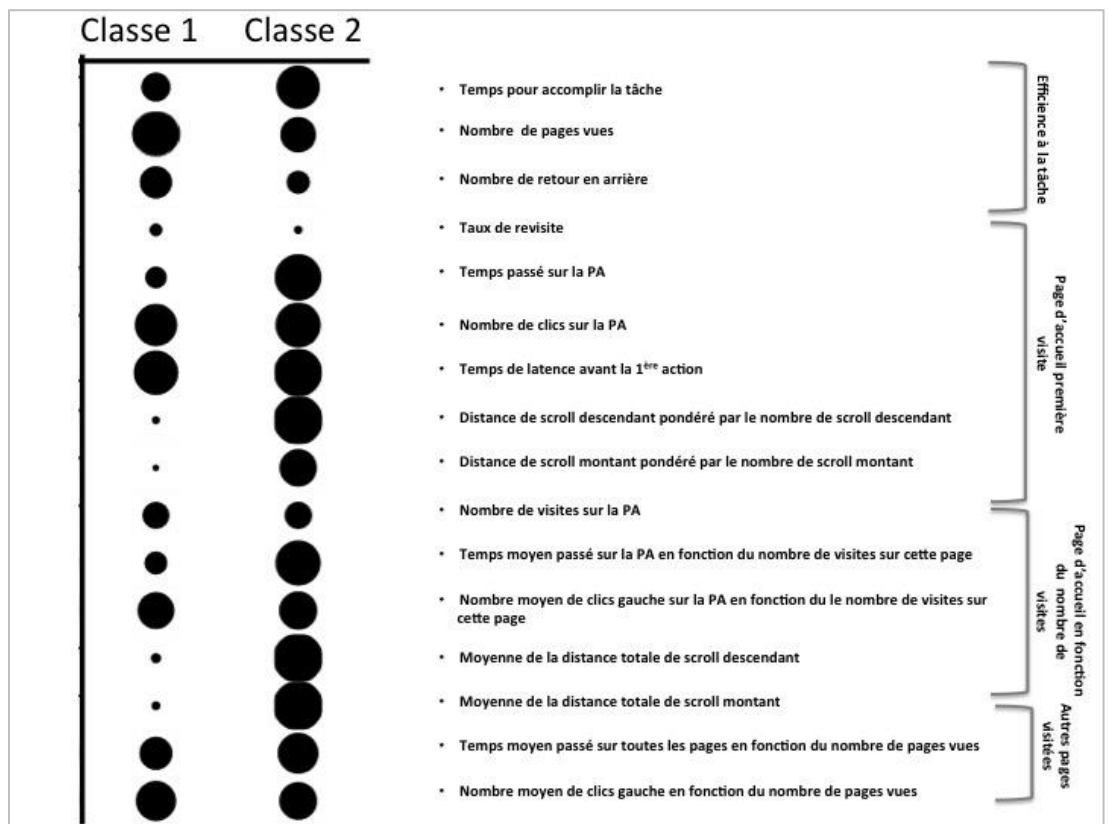


Figure 24 : Illustration des patterns pour la Tâche 5 – Musée Carnavalet

Les participants du pattern 1 sont les plus rapides plus réaliser la tâche (65,5 secondes) :

- Ils sont également ceux passant le moins de temps sur la page d'accueil lors de la première visite (15 secondes) et ont le temps de latence le plus faible (8,5 secondes).
- Ces individus ont, en moyenne, peu scrollé sur la page d'accueil et sur des distances très courtes. De plus le faible temps qu'ils passent, en moyenne, sur les pages visitées semblent indiquer qu'ils naviguent assez vite. **Avec cette stratégie de navigation moins exhaustive, ce sont eux qui revisitent trois fois plus de pages, et qui effectuent, en moyenne, deux fois plus de retour arrière.**

Le pattern 2 regroupe 30 participants :

- Ce groupe passe deux fois plus de temps pour réaliser la tâche et sur la page d'accueil au cours de la première visite (31 secondes). Même s'ils ont mis plus de temps à finir la tâche, ils ont visité en moyenne une page de moins que les individus du pattern 1, et ils n'ont que 2% de revisites.
- De plus, c'est au sein de ce groupe que les participants ont effectué, en moyenne, le nombre de clics et les distances moyennes par scroll les plus élevés, ce qui peut expliquer ce temps plus important passé sur la page d'accueil. Dans ce contexte, il était nécessaire de scroller pour afficher le contenu du site dans son intégralité. De plus, plusieurs menus déroulants été affichés, ces individus ont donc pris le temps d'en explorer le contenu.
- Ils sont finalement ceux prenant le plus de temps à naviguer sur les pages visitées. **Cette stratégie ressemble donc à celle identifié pour la tâche 3 et 4 c'est à dire une navigation plus « exhaustive » pour prélever les informations et les amenant à plus facilement trouver la réponse.**

La classification des chemins a permis de classer les participants en fonction de 2 groupes de parcours. Le tableau 7 présente la répartition des parcours au sein de chaque pattern.

Tableau 7 : Répartition des parcours pour chaque Pattern pour la Tâche 5 - Musée Carnavalet

	Pattern 1	Pattern 2
Parcours classe 1	53,2% (n=33)	59,5% (n=22)
Parcours classe 2	46,8% (n=29)	40,5% (n=15)
Nombre total de participants	100% (n=62)	100% (n=37)

Avec 3 pages (Figure 24), le parcours le plus représentatif de la classe est aussi le chemin optimal pour accéder à la réponse. Au sein des deux patterns, c'est cette classe de parcours qui a été majoritairement empruntée.



Figure 25 : Chemin le plus représentatif pour le parcours n°1 Tâche 5 – Musée Carnavalet LD = 0,12

Le chemin le plus représentatif de la classe 2 comptabilise minimum 5 pages.



Figure 26 : Chemin le plus représentatif pour le parcours n°1 Tâche 5 – Musée Carnavalet LD = 0,38

➤ *Site Musée Carnavalet – Tâche 6, Le plein tarif d'une visite conférence au Musée*

Cette tâche était la deuxième à réaliser sur le site du Musée Carnavalet. L'application de la procédure que nous avons décrite permet d'identifier trois classes de comportements différents. La première classe est composée de 36 participants, la seconde de 30 et la dernière de 20 participants. Les moyennes pour chacun de ces indicateurs sont présentées en Annexe A.

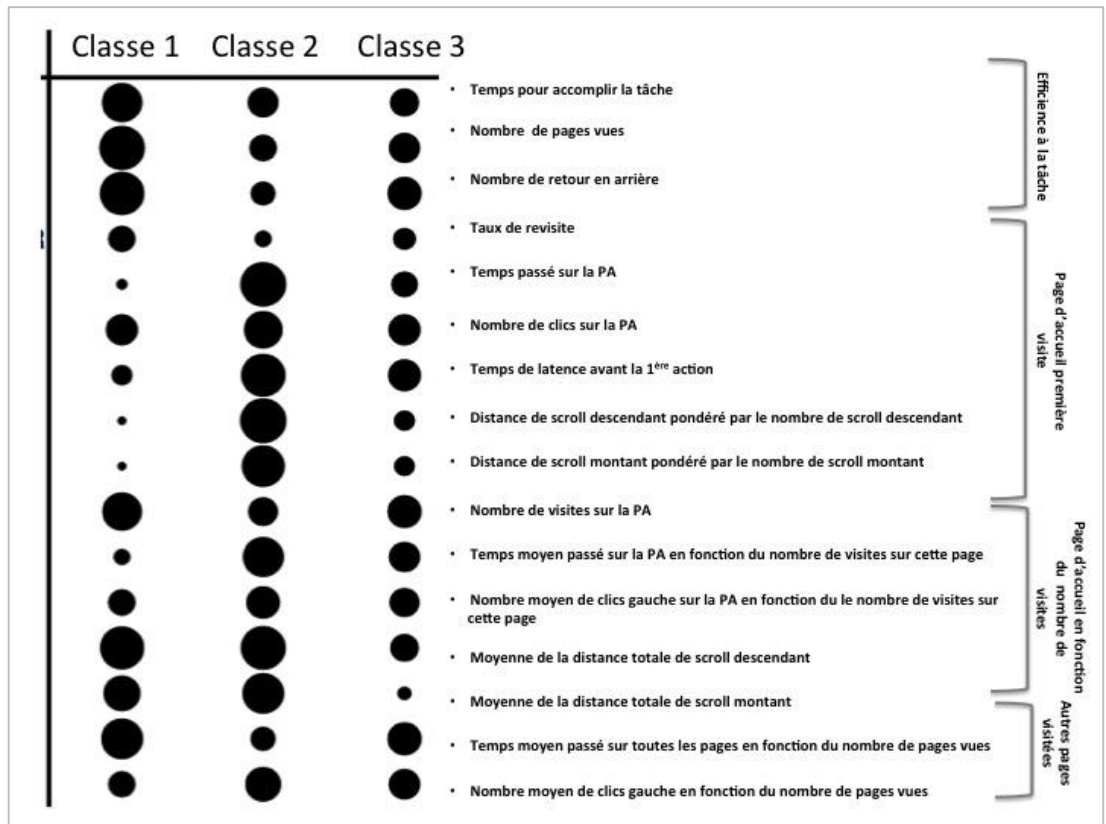


Figure 27 : Illustration des patterns pour la Tâche 6 – Musée Carnavalet

Le premier pattern est constitué par la majorité des participants :

- Il est celui dans lequel les individus passent le plus de temps pour réaliser la tâche (151,4 secondes), visitent le plus de pages (11,25), et font aussi le plus de retour arrière.
- Ce sont eux qui cliquent et qui ont les distances moyennes de scroll les plus faibles sur la page d'accueil au cours de la première visite, tout en y passant le moins de temps (9,8 secondes). Leur temps de latence est le plus faible (7,6 secondes).
- En revenant sur la page d'accueil, leurs distances moyennes par scroll sont bien plus élevées. Les données montrent qu'ils ont cette fois affiché la page dans son intégralité et ils l'ont exploré plus en profondeur. Par ailleurs, les nombreuses revisites semblent corrélées avec le nombre de visites sur la page d'accueil car ce sont eux qui la fréquentent le plus. Enfin, l'analyse des parcours montre **d'importantes difficultés pour s'orienter et accéder à l'information.**

Le deuxième pattern se distingue nettement du premier :

- Les individus visitent en moyenne le moins de pages (4,16) et ce sont eux également qui effectuent le moins de retours en arrière (0,56). Ils réalisent cette tâche assez rapidement (130 secondes), et pourtant ils sont ceux qui passent le plus de temps sur la page d'accueil lors de la première visite (38,3 secondes) et qui effectuent la distance moyenne par scroll verticale la plus importante.

- En moyenne, cette classe réalise le plus de clics sur les pages et y passe le plus de temps (31 secondes). Tout comme pour le pattern 2 Tâche 3, pattern 1 Tâche 4, et le pattern 2 Tâche 5, ces individus explorent longuement les pages, notamment la page d'accueil avant de naviguer sur une autre et ils accèdent sans trop de difficultés à la réponse.

Enfin, le dernier pattern est celui qui regroupe le moins de participants :

- Néanmoins, ces participants terminent la tâche le plus rapidement (126,4 secondes). C'est un groupe intermédiaire, ces moyennes se situent toutes entre les classes 1 et 3. Ces individus prennent leur temps pour explorer les pages visitées mais cela ne les empêche pas de revenir sur leurs traces.

Le tableau suivant présente la répartition des parcours au sein de chaque pattern.

Tableau 8 : Répartition des parcours pour chaque Pattern pour la Tâche 6 - Musée Carnavalet

	Pattern 1	Pattern 2	Pattern 3
Parcours classe 1	0% (n=0)	46,7% (n=14)	5% (n=1)
Parcours classe 2	36,1% (n=13)	16,6% (n=5)	50%(n=10)
Parcours classe 3	63,9% (n=23)	36,7% (n=11)	45%(n=9)
Nombre total de participants	100% (n=36)	100% (n=30)	100% (n=20)

De façon générale, les parcours de la classe 1 sont ceux ayant été les moins empruntés dans l'échantillon (15 participants au total l'ont suivi) alors qu'ils sont les plus direct pour accéder à la réponse. En effet, le chemin le plus représentatif de la classe fait 2 pages (Figure 28). Au sein du pattern 2, les participants ont emprunté le plus les parcours de cette classe. Le fait que ces individus soient ceux prenant le plus de temps à explorer la page d'accueil à la première visite peut expliquer cette différence de navigation. Ils ont ainsi pu avoir en leur possession l'ensemble des informations permettant d'identifier rapidement le lien le plus pertinent.



Figure 28 : Chemin le plus représentatif pour le parcours n°1 Tâche 6 – Musée Carnavalet LD = 0,2

Comme l'illustre le chemin le plus représentatif de la classe 2, le second type de parcours implique de passer par 4 pages minimum dont une revisite sur la page d'accueil (Figure 29). En effet, il est important de souligner que c'est à l'issue de cette revisite que les participants ayant suivi ce type de chemin ont remarqué le lien pour



accéder à l'information. Les participants du Pattern 3 ont emprunté cette classe de parcours à 50%. Ce résultat est cohérent avec les caractéristiques du pattern. En effet, ces individus ont été décrits comme ayant eu minimum 2 visites sur la page d'accueil du site.

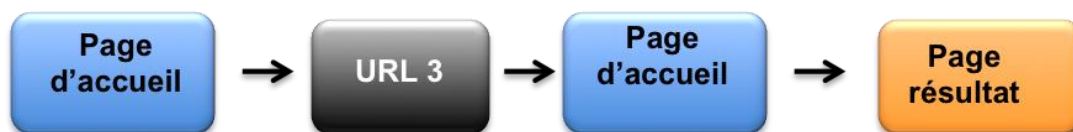


Figure 29 : Chemin le plus représentatif pour le parcours n°2 Tâche 6 – Musée Carnavalet LD = 0,12

Enfin, 44 participants ont suivi les parcours de la classe 3. Le chemin le plus représentatif de la classe fait en moyenne 6 pages (Figure 30). Il s'agit donc ici des parcours les plus laborieux. 63,9% des individus du Pattern 1 ont emprunté ces types de parcours. En s'intéressant aux comportements de navigation de ces participants, on s'aperçoit que malgré le fait qu'ils aient revisité plusieurs fois la page d'accueil du site Carnavalet, paradoxalement ils y ont passé le moins de temps au total. Par rapport aux autres patterns, il semblerait que le faible temps passé sur la page d'accueil soit donc une conséquence de leur dispersion.



Figure 30 : Chemin le plus représentatif pour le parcours n°3 Tâche 6 – Musée Carnavalet LD = 0,52

### Résumé :

À l'issue des analyses réalisées sur ce second site, et comme pour l'étude exploratoire, **le pattern faisant référence au style de navigation « consciencieux » a été retrouvé.** Il est intéressant d'observer ici que, pour les deux tâches, ce sont les individus ayant adopté cette façon de faire qui ont cette fois, de meilleures performances de navigation. Au contraire, pour le site du service Public, ces conclusions sont à nuancer car pour la Tâche 4, il s'agit des individus ayant été les plus rapides pour l'ensemble de la navigation à avoir été les plus performants.

Nous remarquons également que le pattern 1 tâche 5 et pattern 1 tâche 6 se ressemblent. En effet, ici il s'agit **d'une même stratégie de navigation amenant les individus à naviguer plus rapidement sur les pages, à faiblement les explorer pour au final revenir un peu plus sur leurs traces.**

Par ailleurs, nous avons l'idée que l'interaction avec la page d'accueil influe sur les performances de navigation. Toutefois si l'on regarde le pattern 3 Tâche 6 correspondant à une stratégie de navigation assez « consciencieuse » amenant les individus à prendre un certain temps sur les pages visitées, y compris la page d'accueil,

il semblerait qu'au fur et à mesure de la navigation, ils éprouvent quand certaines difficultés à s'orienter. **Ce temps d'exploration pourrait donc plutôt faire référence à une difficulté pour prélever des informations pertinentes.**

Nous avons également souhaité nous intéresser à la façon dont les participants évoluent à travers les patterns et tâches. En effet, l'une des questions que nous nous posions était de savoir si d'une tâche et d'un site à l'autre, les individus restaient cohérents dans leur façon de naviguer. 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 ?

Nous avons donc développé un outil de visualisation sous la forme de coordonnées parallèles permettant de suivre l'évolution des participants à travers les différentes tâches et patterns (Figure 31).

Chaque axe représente une tâche (T3, T4, T5 et T6) et dessus les patterns qui y sont associés.

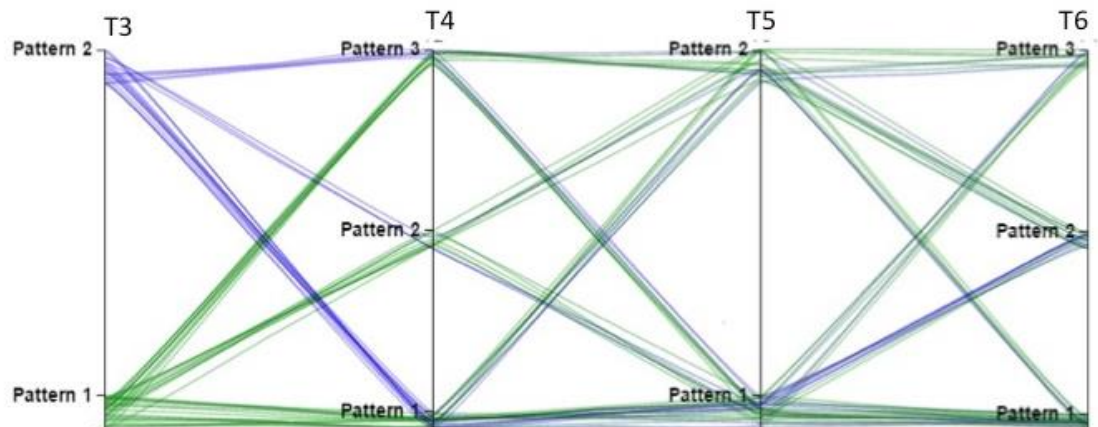


Figure 31 : Visualisation de l'évolution des participants à travers les différentes tâches et patterns

Intéressons-nous d'abord au passage de la première à la deuxième tâche sur le site du Service Public (T3 vers T4) :

- Pour 46,6% des participants du pattern 2 Tâche 3 **on observe que la façon de naviguer ne semble pas changer.** En effet, ils gardent cette façon « consciencieuse » de naviguer dans le pattern 1 Tâche 4 : Les autres participants de ce pattern ne semblent pas présenter de stabilité puisqu'on les retrouve autant dans la classe 2 ou 3.
- Concernant les individus ayant réussi la tâche 3 avec plus de difficultés (pattern 1), la majorité d'entre eux (46%) migrent également vers ce pattern qualifié de « consciencieux ». **Ce changement dans la façon de naviguer leur permet ainsi d'être plus performant** pour réaliser la seconde tâche et il semblerait donc qu'ils aient appris de leurs « maladresses ».

Le passage entre les deux sites engendre une distribution aléatoire des participants. Nous n'observons pas de proportion plus forte à migrer vers le pattern 1 ou 2 pour la tâche 5.

Concernant la transition entre la première et la deuxième tâche sur le site du Musée Carnavalet (T5 vers T6) :

- 52% des participants issus du pattern 1 tâche 5 gardent sensiblement cette même façon rapide de naviguer sur les pages visitées. Au contraire, pour les 48% restant **la façon de naviguer à évoluer et pour cette dernière tâche, ils prennent plus de temps pour explorer le contenu des pages visitées.**
- Concernant les individus du Pattern 2 Tâche 5 pour qui on a qualifié la navigation de « consciencieuse», que deviennent-ils à la Tâche 6 ? La plus grande proportion (40,7%) se retrouve dans le Pattern 2 présentant les mêmes caractéristiques. Les autres participants de ce pattern ne semblent pas présenter de stabilité puisqu'on les retrouve autant dans la classe 1 ou 3.

Enfin, pour aller plus loin, bien **que pour chaque sessions on ne retrouve pas le même nombre de patterns à chaque fois**, certains présentent des similitudes et il y en a un bien commun aux 4 tâches. Le pattern qualifié de « consciencieux» (Pattern 2 à T3, Pattern 1 à T4, Pattern 2 à T5 et Pattern 2 à T6) présente de fortes ressemblances par rapport aux comportements sur la page d'accueil et de façon générale pendant la navigation. Toutefois, nous ne savons pas si dans ce pattern nous trouvons toujours les mêmes participants. Certes, de la Tâche 3 à 4 46,6% des individus gardent la même façon de naviguer et 40,7% de la Tâche 5 à 6, mais s'agit-il toujours des mêmes qui gardent tout au long de l'étude la même façon de faire ?

Dans ce cas, par rapport à ce pattern, il semblerait que non. Comme le montre la figure 23, au fur et à mesure des tâches, les individus changent de façon de naviguer jusqu'à ne plus identifier aucun sujet du départ à la dernière tâche. Ainsi, **concernant ce pattern, les sujets ne gardent pas pour toutes les tâches, toujours cette même façon de faire.**

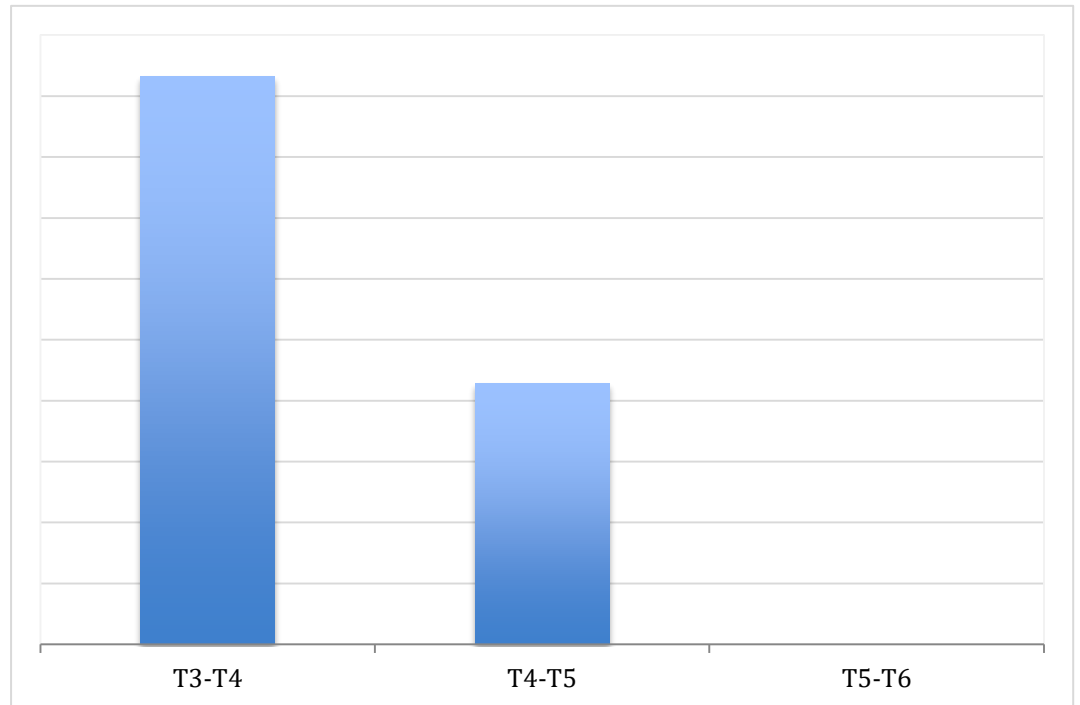


Figure 32 : Évolution des participants caractérisés comme ayant une « navigation consciencieuse » de la première tâche et jusqu'à la fin de l'étude

#### Résumé :

Nos résultats tendent à montrer qu'il n'y a pas vraiment de tendance concernant l'évolution des stratégies d'une tâche et d'un site à l'autre. En effet, même si 4 patterns ressortent de notre analyse, on observe qu'ils ne sont pas stables d'une session à l'autre, nous n'avons pas obtenu le même nombre de classes pour chaque session.

Les données indiquent donc que les individus ne semblent pas garder pour toutes les tâches, toujours la même façon de faire. Nos résultats tendent donc à montrer que **les individus ont tendance à « adopter » une façon différente de naviguer qui s'établit en fonction du contexte du site et de la tâche**. Ainsi, même si nous avons pu caractériser certains styles de navigation, il n'est pas possible de le faire pour les participants.

## 4.1.6

### Discussion

Deux objectifs sous tendaient cette première étude. Le premier a été d'identifier une méthode permettant de rassembler les individus les plus proches sur leur façon de naviguer, et pour se faire de choisir des indicateurs pertinents. Le deuxième a été de voir si, à partir de cette démarche, il existait des styles ou patterns comportementaux lors de la consultation de la page d'accueil qui entraîneraient des niveaux de performance de navigation différents.

En effet, nous avons vu dans la littérature que de nombreuses études ont cherché à analyser et définir des stratégies de navigation (Danielson, 2002, White & Drucker 2007, Obendorf & Weinreich, 2007, Tauscher & Buscher 2012) en vue d'améliorer

certain éléments des sites et pages Web comme le plan, les aides en ligne, le moteur de recherche, etc. Toutefois, à notre connaissance, il a été surprenant de voir que lorsque l'on s'intéresse aux patterns de navigation, les interactions avec la page d'accueil n'ont pas été étudiées et mis en lien avec les performances de navigation globale au sein de ce site. Compte tenu du fait que l'objectif principal des études s'intéressant aux patterns concerne la définition de recommandations, et que la majorité de ces recherches étudient ces comportements à travers la façon dont les internautes explorent différents sites Web et pages consultées, il nous a semblé pertinent d'analyser plus finement le comportement des internautes (au niveau d'un site et d'une page en particulier), avant de formuler, à terme, des recommandations.

Une première étude exploratoire se basant sur une tâche de notre protocole a donc été réalisée. La classification ascendante hiérarchique basée sur nos 10 indicateurs a permis d'identifier trois patterns de navigation. Les principales différences entre les classes montraient que la façon d'explorer la page d'accueil entraînaient des niveaux de performance de recherche d'informations différents. Pour la suite de nos analyses, nous avons décidé d'affiner et de préciser nos indicateurs afin de pouvoir mieux différencier les classes entre elles, mais également afin de mieux décrire les comportements des individus.

Ainsi, au regard des patterns identifiés, la prise en compte des interactions avec, notamment, la page d'accueil a apporté de nouveaux éléments de réflexion sur les comportements de navigation sur Internet. Les conditions de passations étant strictement contrôlées, et tous les individus ayant réussi les tâches, il ne paraissait pas si évident de trouver des patterns différents, et qui s'expliquent, notamment, par rapport à leur façon de faire sur une page en particulier.

Au cours des 4 tâches, quatre patterns se sont exprimés :

- Un pattern reflétant une façon **consciencieuse, prudente et linéaire** de naviguer. Ces individus mettent plus de temps avant de réaliser une première action sur la page d'accueil et pour en explorer le contenu. On observe également cette même façon de faire sur les autres pages visitées. De façon générale, ces participants reviennent moins sur leurs traces, ils ne sont pas toujours les plus rapides pour finir la tâche (comme cela a été le cas pour l'étude exploratoire) mais ils semblent éprouver moins de difficultés pour s'orienter en empruntant souvent les chemins optimaux.
- Le second pattern présente des similarités avec celui décrit juste au-dessus. Ces individus prennent également du temps pour naviguer sur les pages visitées (y compris la page d'accueil), mais malgré cette **exploration consciencieuse ils semblent éprouver des difficultés pour s'orienter**. Leur taux de revisite, le nombre de retour en arrière et le chemin emprunté montrent qu'ils se dispersent en cherchant l'information. Ils sont souvent ceux qui mettent le plus de temps pour finir la tâche.
- Un pattern dans lequel les individus passent peu de temps sur les pages visitées, y compris la page d'accueil, réalisent très vite une action sur cette page, revisitent très peu de pages et ce sont eux qui de façon générale passent le moins de temps sur les pages visitées. Ce pattern **rapide et linéaire** n'a été observé qu'une fois lorsque les gens reviennent pour la deuxième fois sur le

- site du Service Public (Pattern 2). Dans ce cas, il semblerait qu'ils aient parfaitement assimilé la structure du site et qu'ils sachent précisément où aller.
- Un pattern dans lequel les **gens naviguent relativement vite** sur les pages visitées sans prendre le temps de les explorer en détail. Ces individus reviennent sur leurs traces, ils paraissent se perdre et rencontrer des **difficultés d'orientation**. Leur stratégie paraît infructueuse car ces personnes semblent vouloir rapidement effectuer la tâche, mais en explorant peu les pages elles sont obligées d'y revenir afin de collecter les informations manquantes. Ces personnes visitent et revisitent donc plus de pages. Par ailleurs, cette volonté de rapidement naviguer et de faiblement explorer les pages les amènent à emprunter le parcours le plus long. Il est possible que les « téléportés » décrits par Teevan *et al.* (2004) et les « économiques » décrits par Buscher *et al.* (2012) correspondent à ce pattern. À la différence de ces auteurs, notre expérimentation ne permettait pas aux participants de recourir à une navigation directe (l'utilisation du moteur de recherche était défendue). Dans ce cas, il semblerait qu'une page qui ne permet pas d'être accessible par un moteur de recherche performant (page Web non référencée, intranets, etc.) rendrait leur style de navigation inadapté. Par ailleurs, ce pattern présente d'autres points communs avec les « explorateurs » de White et Drucker (2007) décrits comme des utilisateurs faisant fréquemment des retours en arrière, visitant de nombreuses pages et ne parvenant pas directement à leurs objectifs.

Au regard de ces patterns, l'idée que nous avons et qui suggérait que l'interaction avec la page d'accueil influe sur les performances de navigation ne semble plus être fondée. En effet, si l'on regarde les patterns 1 Tâche 3 et 3 Tâche 6 correspondant à cette stratégie de navigation « d'exploration prudente » amenant les individus à prendre un certain temps sur les pages visitées, y compris la page d'accueil, il semblerait qu'au fur et à mesure de la navigation, ils finissent par éprouver certaines difficultés à s'orienter. Pour certains individus, ce temps d'exploration serait plutôt synonyme d'une difficulté pour prélever des informations pertinentes et les conduirait à avoir une navigation dispersée.

Par ailleurs, ce pattern se retrouvant à la fois sur le site du Musée Carnavalet et du Service Public, nous ne pouvons attribuer cette façon de naviguer à une certaine disposition de l'information. Toutefois, il serait intéressant d'approfondir la question en comparant un même site mais décliné en deux designs différents. En effet, une question subsiste, le pattern décrit comme rapide et linéaire n'est observé qu'une fois lorsque les gens réalisent la tâche une seconde fois. Est-ce que cela est dû à un apprentissage ou ce site présente des caractéristiques optimisant son assimilation ?

De plus, à la question de savoir si les individus ont des stratégies de navigation stables dans le temps sur le même site, nos résultats laissent apparaître qu'il n'y a pas vraiment de tendance concernant la stabilité des patterns en fonction du changement de tâche ou de site. D'une part nous n'avons pas obtenu le même nombre de classes pour chaque session, d'autres part comme nous l'avons démontré à travers l'analyse de l'évolution des participants ayant démontré une façon de naviguer « consciencieuse », les résultats indiquent que les individus ne semblent pas garder pour toutes les tâches, toujours la même façon de faire. On observe plutôt une stabilité à court terme c'est à

dire d'une tâche à la suivante. Il apparaît que les individus ont donc tendance à adopter un comportement différent selon les tâches et les sites. Ainsi, même si nous avons pu caractériser certains styles de navigation, il n'est pas possible de le faire pour les participants.

Plusieurs questions et perspectives émanent donc de ces conclusions. Compte tenu de soucis techniques, nous avons dû porter notre attention uniquement sur les individus ayant réussi les tâches indépendamment les unes des autres. Est-ce que l'analyse des comportements pour des échecs ferait émerger de nouveaux patterns ? Il sera intéressant de prendre en compte cette variable pour nos futurs travaux.

De plus, comme nous l'avons précisé au début de notre protocole, nous avons souhaité prendre deux sites avec une disposition de l'information différente. Cependant, les individus cherchant des informations bien distinctes, il est difficile de justifier les patterns à travers le design des stimuli. Il serait intéressant d'approfondir la question du contexte mais en comparant cette fois un même site décliné en deux designs différents.

Par ailleurs, l'objectif principal des recherches portant sur les stratégies et patterns de navigation est de proposer, à terme, des recommandations pour le web, l'évolution des patterns comportementaux est donc une question qui reste à explorer. Comment évoluent les comportements de navigation si les gens reviennent à court terme ou à long terme sur le même site pour rechercher l'information ? Notre prochaine étude portera sur cette problématique.

## 4.2

# EVOLUTION TEMPORELLE DES PATTERNS COMPORTEMENTAUX

---

### 4.2.1

## Problématique et contexte des études

La littérature met en exergue de nombreuses études s'étant attardées sur les comportements de revisites sur Internet (Zhang & Zhao, 2011). Comme nous l'avons dit précédemment, parmi celles-ci se distinguent les recherches portant sur :

- La façon dont les gens ré-accèdent et recherchent une information sur un moteur de recherche ou site web familier en vue d'améliorer les interfaces (Aula, Jhaveri & Kaki, 2005 ; Teevan, Adar, Jones, & Potts, 2007 ; Adar, Teevan & Dumais, 2009)
- Les motivations de l'individu l'encourageant à revisiter un site (Adar, Teevan & Dumais, 2008 ;)
- La quantification des changements de contenu sur un site web

Nous concernant, nous avons souhaité voir si les répétitions de tâches à court et long terme feraient émerger de nouveaux patterns. En effet, l'évolution temporelle des patterns est une question pertinente dans la mesure où l'on sait que les individus peuvent être amenés à revenir plusieurs fois sur un même site. La littérature nous renseigne sur l'utilisation de l'option des favoris, du retour en arrière, etc. mais quand est-il concernant les stratégies de navigation ?

Par ailleurs, dans l'hypothèse que les individus soient plus familiers avec la structure du site et qu'ils ont tous réussi la tâche, est ce que leur façon de naviguer est optimisée ? Si oui comment ?

Deux expérimentations en laboratoire ont été réalisées. La première permet d'étudier les comportements de navigation lors de la revisite d'un site à court terme (répétition immédiate), il s'agit de l'expérimentation citée dans la partie précédente.

La seconde permet d'observer les comportements lors d'une revisite à plus long terme (répétition à 48h d'intervalle).

### 4.2.2

## L'étude des patterns comportementaux à court terme

#### 4.2.2.1

### Méthodologie expérimentale

La première expérimentation demandait à tous les participants de réaliser deux tâches différentes sur deux sites Web (T3-T4 sur le site du Service Public et T5-T6 sur le site du Musée Carnavalet). Chacune des tâches été répétées tout de suite avant même d'en commencer une nouvelle (T3R-T4R, T5R-T6R).

Pour chacune des tâches, seul les participants ayant réussi ont été retenus. Par ailleurs, pour certaines tâches, les données de nombreux participants ont été écartées à la suite



de difficultés techniques (mauvais chargement des pages, problème d'enregistrement des données).

Tableau 9 : Effectifs pour chaque tâche

Tâches	Effectifs
Tâche 3R Service Public	91 participants
Tâche 4R Service Public	105 participants
Tâche 5R Musée Carnavalet	101 participants
Tâche 6R Musée Carnavalet	88 participants

#### 4.2.2.2

#### **Analyse des données**

Concernant les modalités d'analyse des données, elles restent inchangées. Nous avons appliqué la procédure décrite précédemment.

Deux niveaux d'interprétation des données sont envisagés :

- Description du ou des patterns identifiés pour la tâche répétée
- Comparaison du ou de ces patterns avec la tâche initiale

➤ *Site Service Public – Tâche 3 Répétition, conditions de dissolution du PACS*

Trois patterns comportementaux différents ont été identifiés, et comptent respectivement 41, 41 et 9 participants.

Les moyennes des indicateurs sont présentées dans l'annexe A. L'illustration 1 présente visuellement les 3 patterns et leurs comportements.

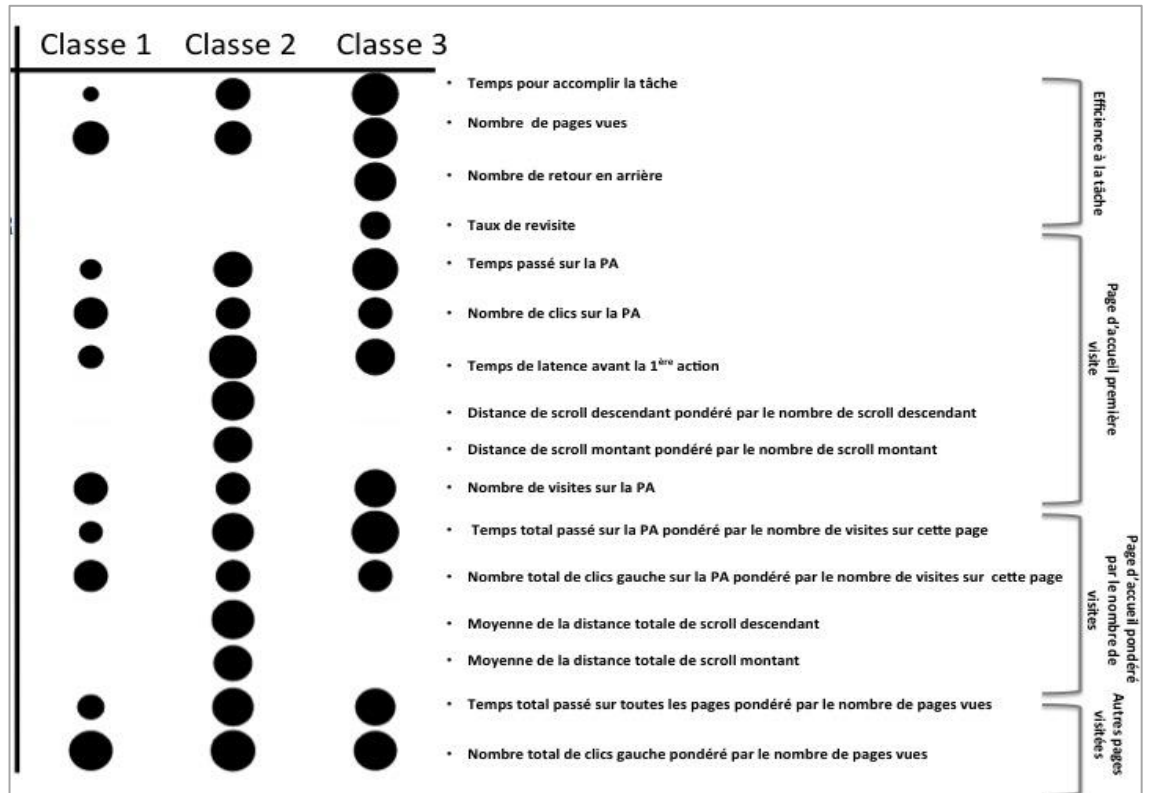


Figure 33 : Illustration des patterns pour la Tâche 3R - Service Public

Trois patterns sont identifiés, le premier se distingue nettement des deux autres car la **stratégie est rapide et linéaire**:

- Ces individus sont les plus rapides pour réaliser la tâche en 18 secondes (10 secondes de moins que le pattern 2 et 20 secondes de moins que le pattern 3) et passant le moins de temps sur la page d'accueil (3,4 secondes) et sur les autres pages visitées.
- Ils empruntent le chemin le plus court, et naviguent de façon très linéaire, sans retour arrière et sans faire de revisites. Il semblerait ici qu'ils aient bien assimilé la structure du site pour rechercher l'information.

Le second pattern se montre aussi performant que le premier mais **ils naviguent plus exhaustivement sur les pages visitées** :

- Ces participants visitent autant de pages, n'effectuent aucun retour arrière et ne revisitent aucune page.
- Concernant le comportement sur la page d'accueil, c'est uniquement au sein de ce pattern que l'on trouve des individus ayant scrollé. Par ailleurs, avec les individus du Pattern 3 ils ont sensiblement mis autant de temps sur la page d'accueil et avant de réaliser une première action. C'est au sein de ce pattern que l'on retrouve la plus forte proportion de participants à avoir identifié un nouveau lien pour accéder à la réponse (via « Les 100 mots clés les plus utilisés »).

Concernant les individus du Pattern 3, bien qu'ils aient tous réussi la tâche, ils sont ceux **éprouvant le plus de difficultés**:

- Ils sont ceux mettant le plus de temps (39,3 secondes), visitant (7,3 pages) et revisitant le plus de pages. Pourtant même en mettant presque autant de temps que les participants du pattern 2 à surfer sur la page d'accueil, **ils ne semblent pas avoir aussi facilement assimilé la structure du site même s'ils prennent leur temps sur les pages visitées.**

Concernant les chemins de navigation, la classification des chemins basée sur la Distance de Levensthein a permis de classer les participants en 2 groupes. Le tableau suivant permet d'étudier la répartition des parcours au sein de chaque pattern.

Tableau 10 : Répartition des parcours pour chaque pattern pour la Tâche 3R - Service Public

	Pattern 1	Pattern 2	Pattern 3
Parcours type 1	<b>95,1%</b> (n=37)	<b>75,6%</b> (n=31)	<b>100%</b> (n=9)
Parcours type 2	<b>4,9%</b> (n=4)	<b>24,4%</b> (n=10)	<b>0%</b> (n=0)
Nombre total de participants	<b>100%</b> (n=41)	<b>100%</b> (n=41)	<b>100%</b> (n=9)

De façon générale, 79 participants ont emprunté le parcours de type 1. Et pour les 3 patterns, c'est celui étant le plus suivi. Le parcours le plus représentatif est représenté ci dessous. Il s'agit ici du même chemin que pour la tâche initiale.



Figure 34 : Chemin le plus représentatif pour le parcours n°1 Tâche 3R - Service Public (LD = 0,03)

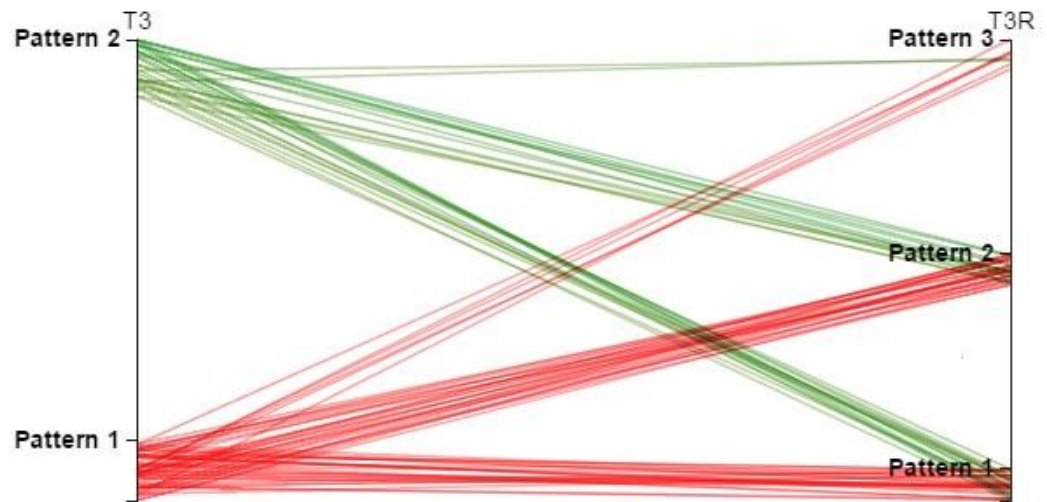
Le deuxième parcours regroupe 12 participants dont 8 issus du pattern 2. Comme nous l'avons dit, ce parcours impliquait de passer via « Les 100 mots clés les plus utilisés ». La première fois que les individus ont du réaliser cette tâche, aucun n'avait identifié cet hyperlien. La figure présente le parcours le plus représentatif.



Figure 35 : Chemin le plus représentatif pour le parcours n°2 Tâche 3R- Service Public (LD = 0,13)

Bien que l'on observe à travers ces trois patterns de meilleures performances pour la répétition de la tâche, il est intéressant de se demander comment les individus ont évolué entre la tâche initiale et la répétition (Figure 35).

*Evolution dans le temps:*



*Figure 36 : Visualisation de l'évolution des participants à travers la répétition de la première tâche sur le site du Service Public*

On observe que 10,4% des participants qui présentaient une stratégie exhaustive mais désorientée reste dans ce même pattern après avoir répété la tâche. Dans ce cas, il semblerait que **certains individus aient plus de mal à assimiler la structure du site** et à retrouver facilement l'information.

Les autres participants se retrouvent respectivement à 44,8% dans le Pattern 1 et 2. Ces **89,6% ont pu optimiser leur façon de naviguer et certains ont ainsi pu même découvrir un nouveau chemin pour accéder à l'information.**

Concernant les individus que l'on a qualifié de « consciencieux » pour la tâche initiale, on observe dans la majorité que leur façon de naviguer a évolué. Dans ce cas, la structure du site semble avoir été apprise sans aucune difficulté et **ils ne ressentent plus le besoin de passer autant de temps sur la page d'accueil.** Leur stratégie a donc évolué et ces participants se trouvent être alors les plus rapides à naviguer sur le site.

➤ *Site Service Public – Tâche 4 Répétition, deux premières conditions générales de recrutement pour accéder à la fonction publique*

L'analyse de la répétition de la tâche 4 met en exergue deux patterns. L'un constitué de 59 participants et l'autre de 41. Cinq participants ont été exclus des analyses considérés comme des outsiders pour la classification.

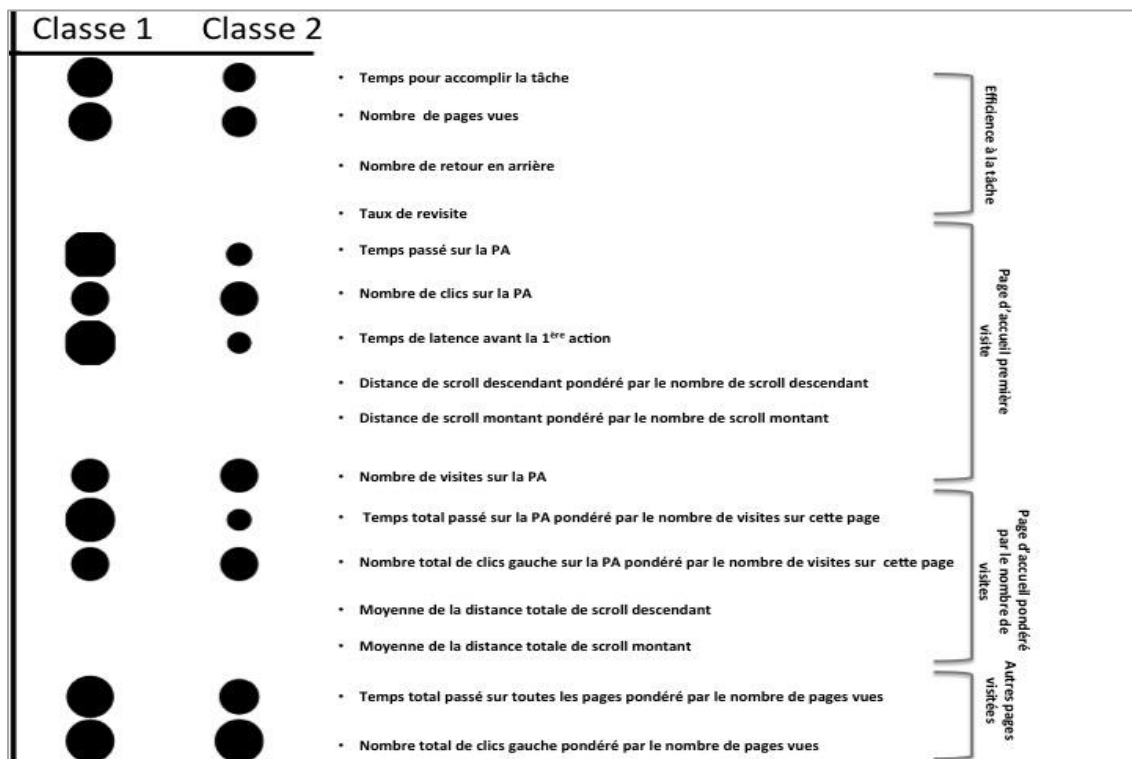


Figure 37 : Illustration des patterns pour la Tâche 4R – Service Public

Comme l'illustre la représentation Bertifier ci-dessus, les deux patterns identifiés se distinguent surtout par le fait que :

- Les participants du pattern 2 sont **les plus rapides pour réaliser la tâche** (16 secondes), ils mettent en moyenne moins de 2 secondes pour quitter la page d'accueil
- Comme l'illustre l'analyse des parcours, ces mêmes participants ont privilégié l'accès à la réponse via le lien du site et impliquant en l'occurrence moins de pages.
- Les participants du pattern 1 mettent 6 secondes de plus à accéder à la réponse, et à naviguer sur les autres pages du site.
- Dans leur cas, ils empruntent en majorité le chemin « classique » afin d'accéder à la réponse. **Cette façon de prendre leur temps sur les pages, et avant de réaliser une action peut être qualifiée de « consciencieuse ».**

Bien qu'il s'agisse des deux mêmes types de parcours que pour la tâche initiale, le tableau 3 montre qu'en fonction du pattern, le parcours emprunté est différent.

Tableau 11 : Répartition des parcours pour chaque Pattern pour la Tâche 4R – Service Public

	Pattern 1	Pattern 2
Parcours type 1	<b>69,5%</b> (n=41)	<b>39%</b> (n=16)
Parcours type 2	<b>30,5%</b> (n=18)	<b>61%</b> (n=25)
Nombre total de participants	<b>100%</b> (n=59)	<b>100%</b> (n=41)

Au sein du pattern 1, 69,5% ont suivi le parcours de type 1 impliquant de passer par le chemin « classique ».



Figure 38 : Chemin le plus représentatif pour le parcours n°1 Tâche 4R – Service Public (LD = 0,07)

Au contraire, pour le pattern 2, 61% ont emprunté le parcours de type 2 qui fait 1 page de moins et qui consiste à passer via le plan du site. La figure présente le parcours le plus représentatif.



Figure 39 : Chemin le plus représentatif pour le parcours n°2 Tâche 4R – Service Public (LD = 0,02)

### Evolution dans le temps:

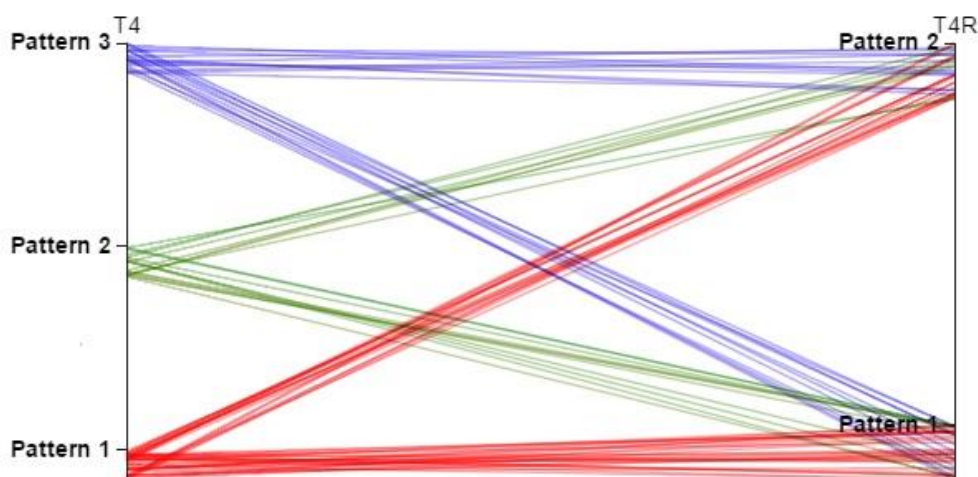


Figure 40 : Visualisation de l'évolution des participants à travers la répétition de la deuxième tâche sur le site du Service Public

Contrairement à ce que l'on aurait pu penser, **la majorité des participants faisant partie du pattern décrit comme rapide et linéaire n'est comparativement plus aussi rapide pour la répétition de la tâche.** À 59% ils font désormais partis du pattern décrit comme plus « exhaustif » dans la façon de naviguer.

Les individus que l'on avait qualifiés de « consciencieux » pour la tâche initiale gardent à 60,9% sensiblement la même façon de faire, ils sont toujours ceux prenant le plus de temps pour naviguer sur la page d'accueil et sur le site.

Concernant les individus du Pattern 3 qui avaient éprouvé le plus de difficultés à la tâche initiale nous n'observons pas de proportion plus forte à migrer vers le pattern 1 ou 2. La répétition de la tâche engendre une distribution aléatoire des participants mais qui montre que tous ont pu cette fois trouver l'information plus rapidement et facilement. **Même si l'on observe une optimisation dans leur façon de naviguer, en majorité, tous les participants ont changé de pattern.**

➤ *Site Musée Carnavalet – Tâche 5, Adresse du théâtre municipal du 15<sup>ème</sup> arrondissement de Paris*

Pour la première répétition de tâche sur le site du Musée Carnavalet, deux patterns sont observés. L'un de 21 participants et l'autre de 80.

Le tableau en annexe A et la figure suivante présentent les caractéristiques et les comportements des deux classes d'individus

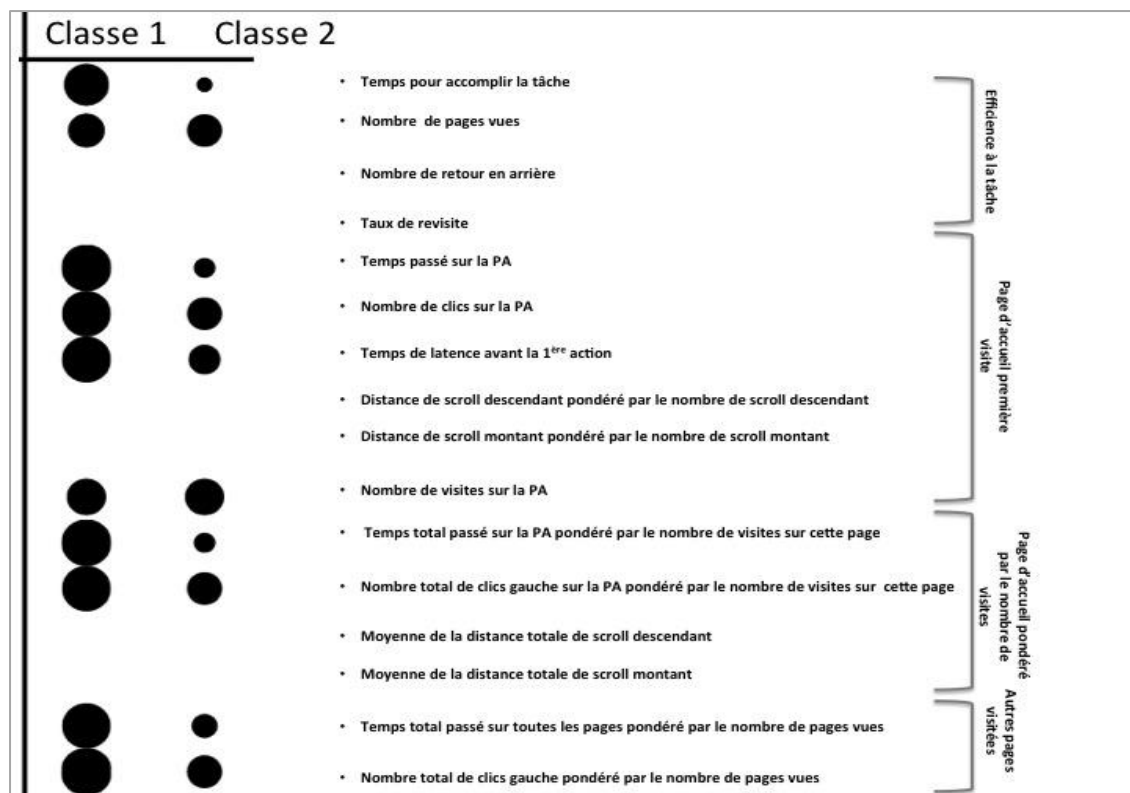


Figure 41 : Illustration des patterns pour la Tâche 5R – Musée Carnavalet

Tout comme pour la tâche précédente, les deux patterns observés se distinguent en grande partie par la rapidité dans la navigation:

- Les participants du pattern 2 mettent **deux fois moins de temps à réaliser la tâche** (18 secondes) et **trois fois moins de temps sur la page d'accueil** (5 secondes).
- Concernant les autres participants, ce temps plus important passé sur la page d'accueil peut s'expliquer par le fait qu'ils l'explorent via les menus déroulants. Ils réalisent, en moyenne, presque deux fois plus de clics. À travers cette façon de naviguer, nous pouvons attribuer à ce pattern l'étiquette de « **conscientieux ou exhaustif** ».



Le tableau 12 présente la répartition des parcours au sein de chaque pattern.

Tableau 12 : Répartition des parcours pour chaque Pattern pour la Tâche 5R - Musée Carnavalet

	Pattern 1	Pattern 2
Parcours type 1	76,2% (n=16)	86,3% (n=69)
Parcours type 2	23,8% (n=5)	13,7% (n=11)
Nombre total de participants	100% (n=21)	100% (n=80)

Bien qu'il s'agisse des deux mêmes type de parcours que pour la tâche initiale, dans ce cas, tous ont majoritairement emprunté les parcours de type 1. Ce parcours de 3 pages est optimal pour accéder à la réponse. Cette fois ce sont les participants les plus rapides sur l'ensemble du site qui ont emprunté en majorité ce chemin.

Le parcours 1 le plus représentatif de la classe est représenté ci dessous.



Figure 42 : Chemin le plus représentatif pour le parcours n°1 Tâche 5R - Musée Carnavalet (LD = 0,01)

Le second chemin compte minimum 4 pages. Le parcours le plus représentatif de la classe est représenté ci dessous.



Figure 43 : Chemin le plus représentatif pour le parcours n° 2 Tâche 5R - Musée Carnavalet (LD = 0,16)

### Evolution dans le temps:

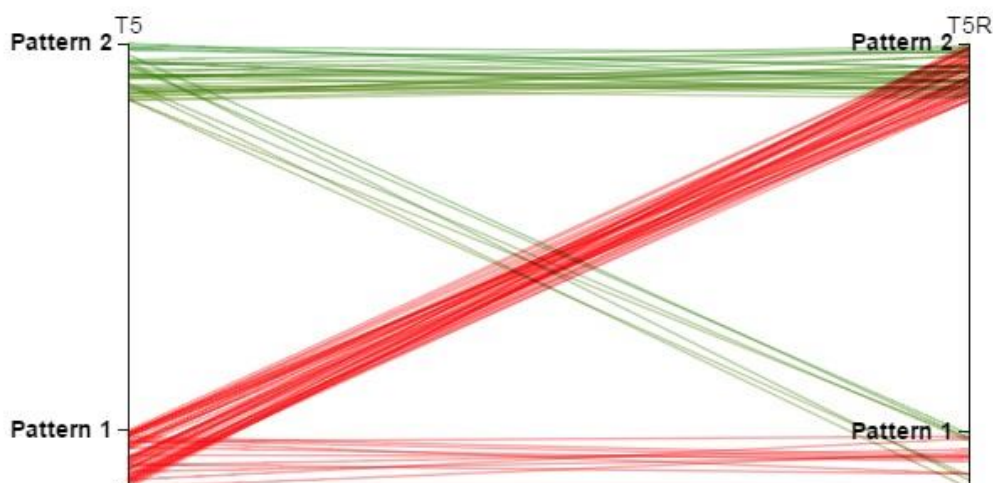


Figure 44 : Visualisation de l'évolution des participants à travers la répétition de la première tâche sur le site du Musée Carnavalet

L'analyse des parcours met clairement en évidence **qu'à la suite de la répétition de la tâche, tous les participants ont eu beaucoup plus de facilités à accéder à la réponse**. Toutefois, l'analyse des patterns nous montre tout de même des différences au niveau des stratégies de navigation.

Pour la tâche initiale, les individus que l'on avait qualifié d' « exhaustifs » sont-ils toujours ceux prenant plus de temps sur l'ensemble des pages du site et avant de réaliser une première action ? Il semblerait que non, la grande majorité d'entre eux (81,1%) ont évolué dans leur façon de faire. Ils font maintenant **partis du pattern décrit comme étant le plus rapide**.

Concernant les individus qui étaient les plus rapides à la tâche initiale mais rencontrant des difficultés d'orientation, on remarque que 80,6% d'entre eux sont encore plus rapides pour cette répétition de tâche, mais **cette fois cette rapidité ne leur porte pas défaut. Ils se sont appropriés la structure du site et le chemin d'accès à l'information**.

➤ Site Musée Carnavalet – Tâche 6 Répétition, Le plein tarif d'une visite conférence au Musée

Pour cette dernière tâche du protocole, trois patterns sont observés. À la suite de l'analyse, deux participants ont été considérés comme des outsiders. La représentation Bertier présente les caractéristiques et les comportements des deux classes d'individus

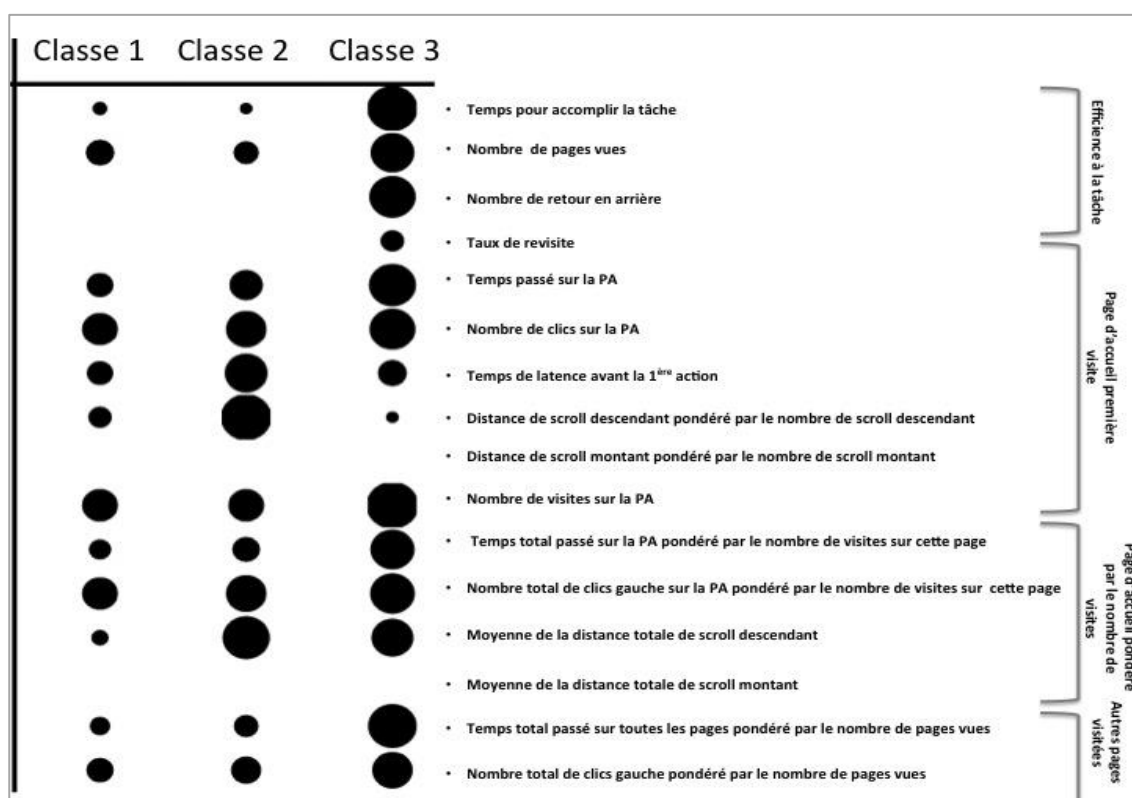


Figure 45 : Illustration des patterns pour la Tâche 6R - Musée Carnavalet

En ayant réussi la tâche avec une seconde de différence, visité quasiment le même nombre de pages, et navigué de façon linéaire sans retour arrière, les patterns 1 et 2 présentent de fortes similitudes. Les comportements sur la page d'accueil permettent de les distinguer :

- En moyenne, les individus du premier pattern ont été **plus rapides** pour naviguer sur la page d'accueil et ont effectué une distance moyenne par scroll descendante plus petite.
- Le pattern 2 présente une **stratégie d'exploration plus exhaustive** en ayant la distance moyenne par scroll la plus grande, à effectuer plus de clics sur la page d'accueil et à avoir un temps de latence un peu plus élevée.
- Enfin, pour les participants du pattern 3, la **recherche de l'information semble avoir été plus compliquée**. En effet, avec en moyenne un taux de

revisite de 22%, mis approximativement trois fois plus de temps à réaliser la tâche, et visité presque 4 pages, ils sont les moins performants. En effectuant le plus de clics sur les pages visitées, ils font partie du pattern explorant le plus les contenus. L'analyse des parcours montre qu'au sein de ce pattern la grande majorité des participants ont emprunté les parcours les moins rapides.

La classification des parcours permet de classer les participants en deux groupes. Le tableau suivant montre la répartition de ces parcours au sein de chaque pattern.

Tableau 13 : Répartition des parcours pour chaque Pattern pour la Tâche 6R - Musée Carnavalet

	Pattern 1	Pattern 2	Pattern 3
Parcours type 1	<b>79,4%</b> (n=27)	<b>100%</b> (n=28)	<b>12,5%</b> (n=3)
Parcours type 2	<b>20,6%</b> (n=7)	<b>0%</b> (n=0)	<b>87,5%</b> (n=21)
Nombre total de participants	<b>100%</b> (n=34)	<b>100%</b> (n=28)	<b>100%</b> (n=24)

Le premier type de parcours identifié est le même que pour la tâche initiale. Il est cette fois-ci le plus emprunté dans l'échantillon et 100% des participants du pattern 2 l'ont suivi.



Figure 46 : Chemin le plus représentatif pour le parcours n°1 Tâche 6R - Musée Carnavalet (LD = 0,05)

Le second type de chemin implique minimum 4 pages, la quasi-totalité des participants du pattern 3 a emprunté cette classe de parcours.



Figure 47 : Chemin le plus représentatif pour le parcours n°2 Tâche 6R - Musée Carnavalet

### Evolution dans le temps:

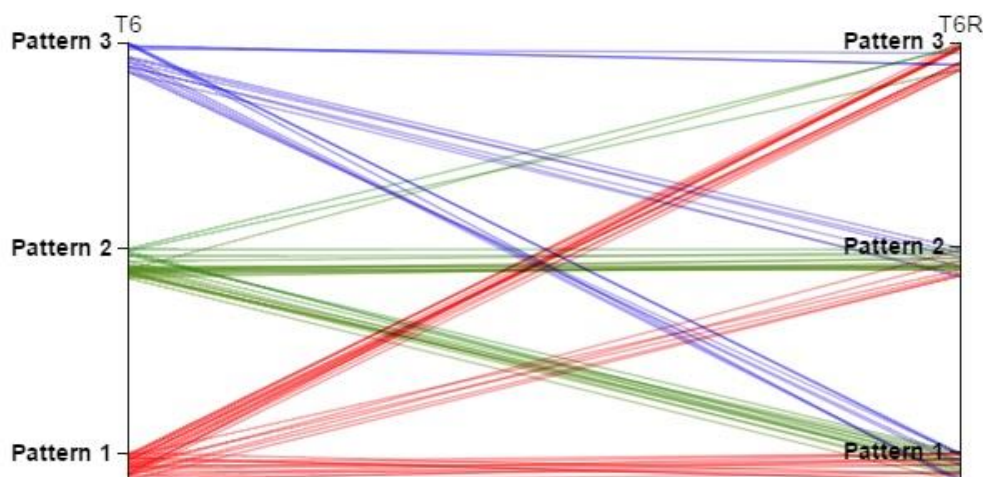


Figure 48 : Visualisation de l'évolution des participants à travers la répétition de la deuxième tâche sur le site du Musée Carnavalet

Les analyses réalisées sur cette dernière tâche montre que **même si certains individus ont gardé quelques difficultés à naviguer sur le site, peu importe le pattern, tous ont progressé pour rechercher l'information** (moins de pages vues, de temps pour réaliser la tâche, de retour arrière).

Concernant les individus du Pattern 1 ayant éprouvé le plus de difficultés à la tâche initiale on observe que la plus grande proportion (42,9%) d'entre eux se retrouve à nouveau dans un pattern présentant des difficultés, pourtant leur stratégie a changé, ils prennent plus de temps à naviguer sur les pages visitées et à cliquer sur le contenu. Ils ne sont que 20% à faire partie du pattern 2 ayant emprunté le parcours le plus court. De façon générale, **même si la recherche de l'information a été optimisée, les individus ayant eu des difficultés semblent donc les garder mais leur façon d'explorer les pages est plus exhaustive.**

À proportion égale, on retrouve les participants présentant une façon « consciencieuse » de naviguer dans les patterns 1 et 2. Dans tous les cas, ils restent ceux ayant identifié la réponse avec le moins de difficultés. Même en ayant passé le moins de temps sur la page d'accueil, ils ont les distances moyennes de scroll les plus élevées. **Ces participants savaient donc où trouver le bon hyperlien et la structure du site semble donc assimilée.** Le constat est identique pour les individus qui à la tâche initiale naviguaient de façon assez posée sur les pages tout en étant désorientés. Dans cette répétition, ils ne sont que 14,2% à conserver des difficultés pour accéder à la réponse.

### 4.2.2.3

## Réflexion intermédiaire

Plusieurs conclusions émanent de ces analyses. La première porte sur l'évolution dans la façon de naviguer. En effet, en revenant sur les sites pour les mêmes tâches, **la grande majorité des individus ont optimisé leur façon de naviguer** et ils accèdent avec beaucoup plus de facilités à la réponse. On observe tout de même que **certains individus qui rencontraient des difficultés aux tâches initiales les gardent**.

La seconde concerne **l'homogénéité des patterns**. Nous constatons que la représentation Bertier aboutit à **des similitudes importantes entre les sessions de répétition** et que l'on retrouve trois patterns identifiés dans l'étude initiale.

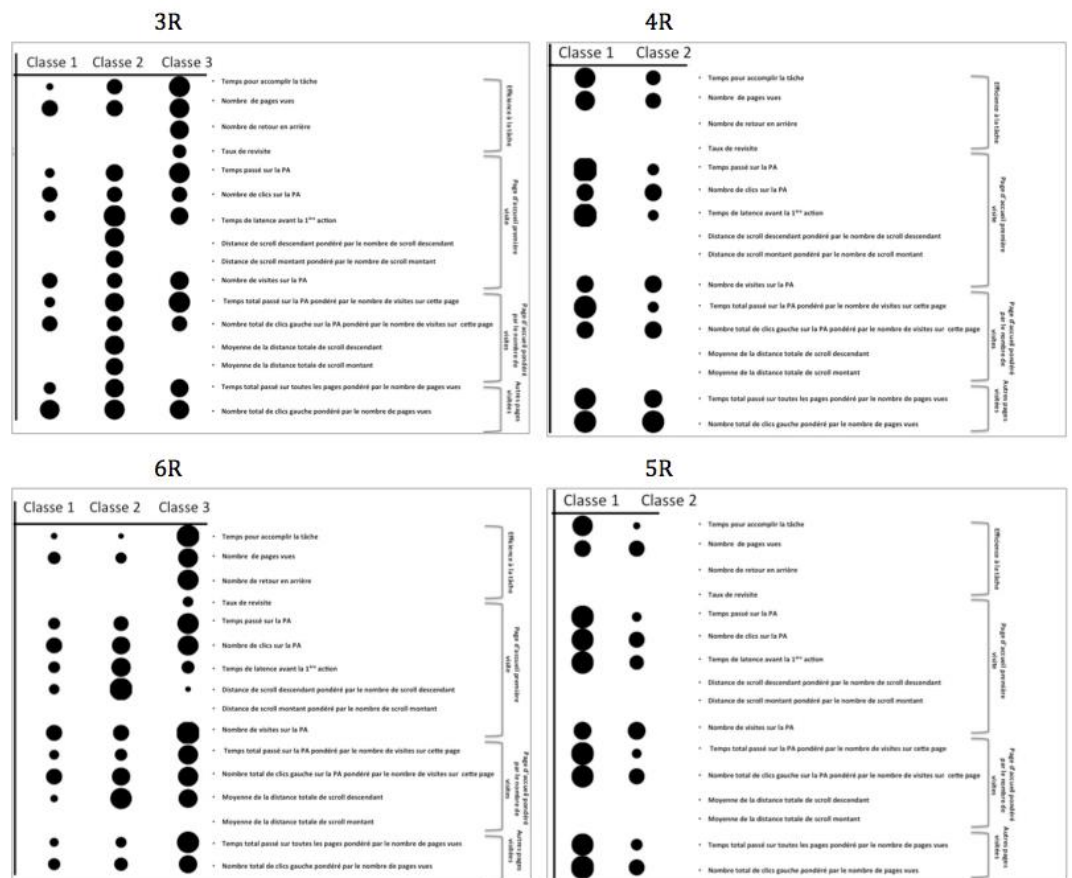


Figure 49 : Représentation Bertier des sessions 3R, 4R, 5R et 6R

- Un pattern dans lequel les individus naviguent de façon très rapide sur le site, passent peu de temps sur la page d'accueil et avant de réaliser une action, visitent très peu de pages et ne reviennent jamais en arrière. **Dans ce cas la structure du site et le chemin d'accès à la réponse semblent bien assimilés.**
- Un second pattern dans lequel on identifie de fortes similitudes avec la stratégie de navigation « exhaustive, consciencieuse » décrite par ailleurs. Ils prennent plus de temps sur la page d'accueil et avant de réaliser une action, plus de temps à naviguer sur les pages en général, et

visitent très peu de pages également. **La structure du site et le chemin d'accès à la réponse semblent avoir été retenu.**

- Le dernier pattern est celui regroupant **les participants désorientés et ayant éprouvés le plus de difficultés.** Dans ce cas il se caractérise par un temps plus important pour réaliser la tâche, plus de pages vues et revisitées et c'est uniquement au sein de ce pattern que l'on identifie des individus ayant recours au bouton « back ». Malgré une navigation dispersée, ils accèdent à l'information.

Le pattern que nous avons décrit comme rapide mais amenant à une certaine désorientation pour la tâche initiale (identifié pour 3 tâches sur 4), n'a pas été retrouvé. **Cette observation amène à penser que de façon générale les individus optimisent l'accès à la réponse et leurs stratégies évoluent.**

Par ailleurs pour l'ensemble de ces répétitions, et contrairement aux tâches initiales, **le pattern le plus performant reste à chaque fois celui regroupant les individus ayant été les plus rapides à naviguer.** Dans ce contexte, l'exploration approfondie de la page d'accueil et des autres pages du site ne semblent plus avoir autant d'influence sur les performances à la tâche. Mais est-ce que cela est dû à une répétition immédiate ? Il est possible qu'un phénomène d'apprentissage à court terme explique ce résultat.

La prochaine étude présente une étude longitudinale où les participants devaient répéter les tâches mais dans un délai de plusieurs jours.

## **4.2.3 L'étude des patterns comportementaux à long terme**

### **4.2.3.1 Matériel et terrain**

Comme pour la première expérimentation, nous avons utilisé le système d'oculométrie déporté Tobii T120. Le logiciel *Tobii Studio* (v.2.3.2.0) permettait de recueillir des fichiers logs pour chaque participant et chaque tâche.

Il s'agissait d'une procédure expérimentale de laboratoire, réalisée dans la salle d'observation du laboratoire Pergolab pendant 21 jours. Les participants étaient installés face au système Tobii. Ce dernier était posé sur une table stabilisée.

La navigation sur les sites Web était rendue possible par la mise à disposition de la souris. Le clavier était contrôlé par l'expérimentateur pour lancer les différentes phases de l'expérimentation.

### **4.2.3.2 Procédures**

Nous avons mis en place un protocole expérimental standardisé pour lequel tous les participants devaient se présenter trois fois en une semaine au laboratoire, avec 48 heures d'intervalle entre chaque passation. À savoir :

Lundi = temps 1 (T1), Mercredi = temps 2 (T2) et Vendredi = temps 3 (T3).

Dès l'étape de recrutement, les participants étaient informés que l'expérimentation se déroulait en trois temps et qu'ils devaient s'engager à venir trois fois.

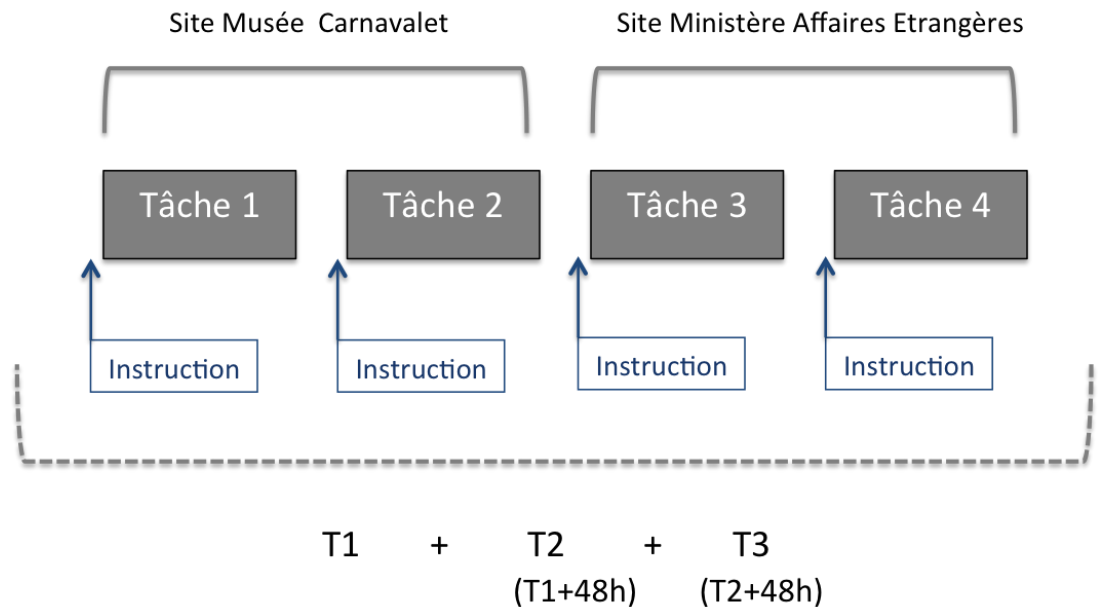


Figure 50 : Schéma récapitulatif de la procédure expérimentale

Comme pour la première expérimentation, lors de l'accueil des participants à la première passation, nous leur avons précisé qu'ils auraient à répondre à des questions en naviguant sur un site Internet particulier sans avoir l'usage du clavier. Nous avons expliqué que les questions ne comportaient pas de pièges et qu'il n'y avait pas de limite temporelle afférente aux tâches. L'expérimentateur était assis en léger retrait à côté du sujet et notait systématiquement tout événement susceptible de constituer un biais expérimental.

Tous les participants ont réalisé les tâches suivantes dans le même ordre :

1. « À partir du site web qui va s'afficher, trouvez qui fut le premier donateur du Musée ? Quand vous aurez trouvé la réponse, citez-la à haute voix ». Page Web de départ : [www.carnavalet.paris.fr](http://www.carnavalet.paris.fr) (page d'accueil).

2. « Vous êtes gérant d'une entreprise et vous souhaitez organiser un cocktail au Musée avec vos collaborateurs. À partir du site web qui va s'afficher, trouvez le nom de la personne à contacter ? Quand vous aurez trouvé la réponse, citez-la à haute voix. » Page Web de départ : [www.carnavalet.paris.fr](http://www.carnavalet.paris.fr) (page d'accueil).



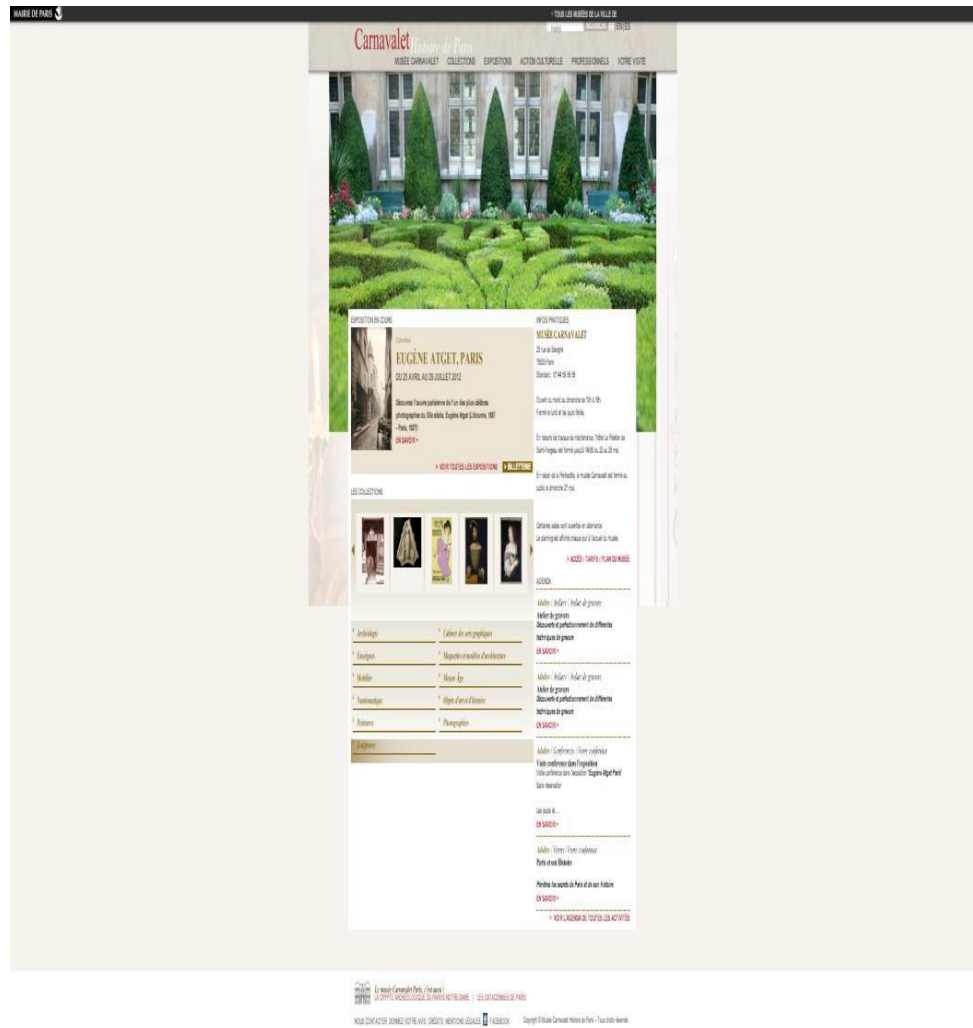


Figure 51 : Page d'accueil du site Musée Carnavalet

3. « À partir du site web qui va s'afficher, trouvez en quelle année fut créé le CNED (Centre National d'Enseignement à Distance ?) Quand vous aurez trouvé la réponse, citez-la à haute voix. » Page Web de départ : [www.diplomatie.gouv.fr](http://www.diplomatie.gouv.fr) (page d'accueil).

4. « À partir du site web qui va s'afficher, trouvez les informations relatives à la sécurité à Andorre si vous vous déplacez dans ce pays ? Quand vous aurez trouvé la réponse, citez-la à haute voix. » Page Web de départ : [www.diplomatie.gouv.fr](http://www.diplomatie.gouv.fr) (page d'accueil).



Figure 52 : Page d'accueil du site Diplomatie Française

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 le temps de latence d'affichage. Une fois que les sujets avaient lu une consigne, nous leur demandions de nous dire ce qu'ils devaient faire et nous nous assurons de leur bonne compréhension.

Chaque tâche était précédée de l'effacement de l'historique de navigation d'Internet Explorer.

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. Par ailleurs, pour ne pas biaiser l'étude, les participants n'ont pas été informés qu'ils devaient effectuer les mêmes tâches aux prochaines sessions.

Nous avons ainsi évité qu'une fois à leur domicile, ils retournent sur ces sites web pour se familiariser avec et / ou trouver les réponses aux tâches à effectuer pendant l'expérience. La procédure expérimentale fut identique pour la seconde et la troisième passation. Les différentes tâches n'ont pas fait l'objet d'un contrebalancement.

Un questionnaire post test était rempli par les participants. Celui cité dans la première étude a été réutilisé afin de contrôler, à posteriori, l'homogénéité de notre échantillon selon différents critères (Annexe A). Rappelons qu'il évaluait le niveau d'expertise avec Internet (inspiré et traduit de Lim et Lee, 2000), l'usage d'Internet des participants et y était inclut une fiche signalétique

Enfin, pour sélectionner ces sites, nos critères sont restés inchangés par rapport à la première expérimentation. La structure du site du Musée Carnavalet ayant été entre temps modifié afin de présenter plus clairement le contenu, nous avons choisi un autre site que celui du Service public de façon à ce que l'un paraisse comparativement moins organisé.

Site Web Musée Carnavalet	Site Web Diplomatie Française
Page d'accueil structurée en trois parties et présentant distinctement le menu principal en haut de la page et un second au milieu de page	Page d'accueil agencée en plusieurs parties assez dispersées (une partie des menus en haut, à droite et à gauche de la page, blocs d'informations au milieu avec plusieurs menus déroulants )
Pas de dispersion vers d'autres liens, sites ou publicités	De nombreux liens, logos, etc. renvoyant vers des sites associés ou d'autres focus
Présentation des actualités du Musée dans un seul et même encadré	Dispersion des actualités dans de nombreux encadrés à plusieurs endroits de la page

### 4.2.3.3

#### Échantillon

Trente et un volontaires ont participé à cette étude, mais seul les données pour 27 participants ont pu être exploitées. En effet, 4 personnes ne se sont pas présentées à toutes les sessions du protocole. Au total, 108 passations ont donc eu lieu.

Même si peu d'individus ont échoué aux tâches, cette fois ci, sans perte de données dans l'enregistrement des sessions, les individus ayant échoué ont été pris en compte.

Le tableau ci-dessous présente les effectifs étudiés pour chaque tâche.

Tableau 14 : Effectifs pour chaque tâche, étude 2

Tâches	Effectifs
Tâche 3 Musée Carnavalet	27 participants
Tâche 4 Musée Carnavalet	27 participants
Tâche 5 Diplomatie Française	27 participants
Tâche 6 Diplomatie Française	27 participants

La moyenne d'âge de l'échantillon était de 21 ans ( $\sigma = 2,24$ ) avec une proportion plus importante de femmes (65,4%). Tous les participants utilisaient Internet depuis plus de 3 ans et 92,3% étaient connectés sur le Web au moins une heure par jour. Le tableau ci-dessous présente les caractéristiques de notre échantillon.

Tableau 15 : Caractéristiques de l'échantillon, étude 2

Variables	Données sur les variables
Âge moyen (écart type)	21,46 ans ( $\sigma = 2,24$ )
Proportion d'hommes et de femmes	34,6% d'hommes et 65,4% de femmes
Nombre d'années d'utilisation d'Internet	Plus de 3 ans pour tout l'échantillon
Fréquence d'utilisation d'Internet	Au moins 1 heure par semaine pour 92,3%

#### 4.2.3.4

#### Analyse des données

Concernant les modalités d'analyse des données, elles restent inchangées, nous avons appliqué la procédure décrite précédemment. Cependant, un paramétrage sur le logiciel R a été nécessaire. Le *Dynamic Tree Cut* étant initialement paramétré pour privilégier des classes avec 20 éléments minimum, pour cette étude comptant 27 participants, nous avons dû l'ajuster à un niveau de granularité plus fin.

L'objectif étant ici de suivre l'évolution des patterns dans le temps, les analyses sont présentées en fonction des tâches.

➤ *Site Musée Carnavalet – Tâche 1, premier donateur du Musée*

Les analyses nous révèlent deux ou trois patterns en fonction du moment auquel a eu lieu la tâche.

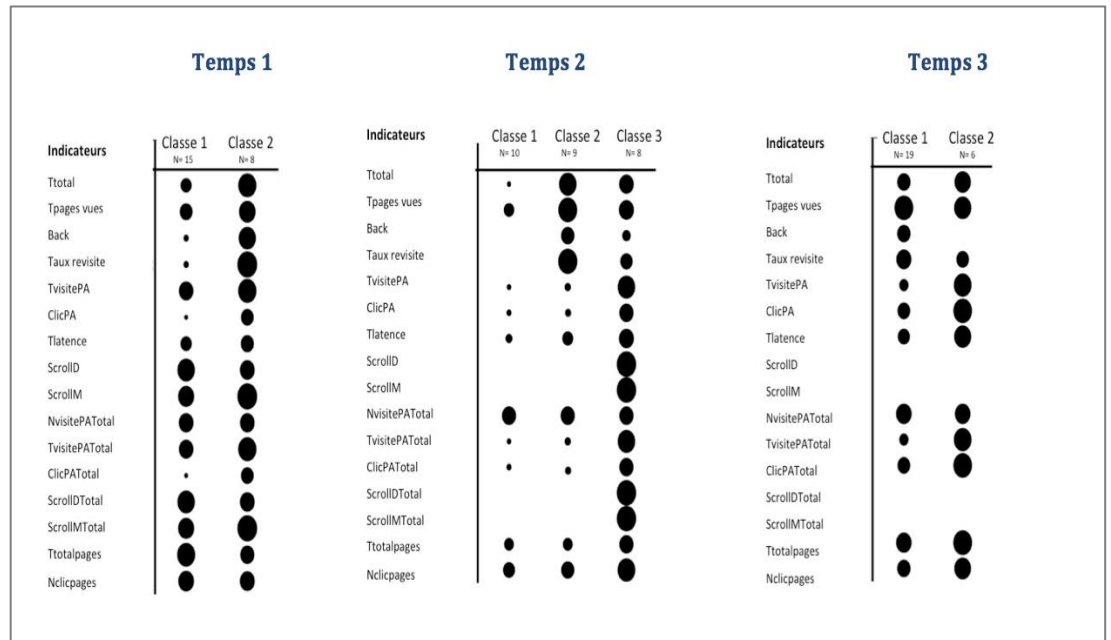


Figure 53 : Illustration des patterns pour la première Tâche à Temps 1, 2 et 3 – Musée Carnavalet

*Description des patterns :*

À Temps 1, deux patterns (ou classes) sont identifiés. Quatre individus ont été considérés comme des *outsiders*.

Le pattern 1 se différencie surtout par sa navigation plus linéaire :

- Presque **deux fois moins de pages visitées**, 50 secondes de moins pour réaliser la tâche, et **très peu de retour en arrière** ;
- Ces individus sont ceux à avoir passé, en moyenne, moins de temps sur la page d'accueil, même s'ils y ont scrollé.
- Enfin, concernant les parcours, le chemin le plus court fait deux pages. Les participants du pattern 1 sont les seuls à avoir emprunté cette classe de parcours. La représentation des chemins est en annexe B.

Au contraire, le pattern 2 présente une navigation plus laborieuse avec deux fois plus de pages visitées et revisitées, mais aussi presque deux fois plus de temps passé sur la page d'accueil. Ils y ont largement scrollé.

Tableau 16 : Répartition des parcours pour chaque Pattern Tâche 1-Temps 1 - Musée Carnavalet

	Pattern 1	Pattern 2
Parcours classe 1	46,7% (n=7)	25% (n=2)
Parcours classe 2	0% (n=0)	75% (n=6)
Parcours classe 3	53,3% (n=8)	0% (n=0)
Nombre total de participants	100% (n=15)	100% (n=8)

À temps 2, (48 heures plus tard), trois patterns se distinguent, comptant sensiblement le même nombre d'individus. La majorité des participants (patterns 2 et 3) éprouvent autant voir plus de difficultés à accéder à la réponse qu'à Temps 1:

- Par rapport aux participants du pattern 1, ils mettent trois ou quatre fois plus de temps pour accéder à la réponse, visitent et revisitent beaucoup plus de pages, et ont des taux de revisite compris entre 11 et 27%.
- Les participants des patterns 2 et 3 semblent se perdre au fur et à mesure de la navigation. En effet, les parcours de la classe 1 avec un minimum de 5 pages sont les plus laborieux et ils rassemblent les participants de ces deux patterns.
- Cependant même si les participants du pattern 2 et 3 rencontrent des difficultés, ceux de P2 sont plus rapides à naviguer sur les pages visitées.
- Pour 63% de l'échantillon, les difficultés pour accéder à la réponse ne se sont donc pas atténuées.

Tableau 17 : Répartition des parcours pour chaque Pattern Tâche 1-Temps 2 - Musée Carnavalet

	Pattern 1	Pattern 2	Pattern 3
Parcours classe 1	0% (n=0)	100% (n=9)	62,5% (n=5)
Parcours classe 2	100% (n=10)	0% (n=0)	37,5%(n=3)
Nombre total de participants	100% (n=10)	100% (n=9)	100% (n=8)

Enfin, à temps 3, deux patterns se différencient. L'un comptant 19 participants et l'autre 6. Deux sujets ont été considérés comme *outsider*.

- Bien que les individus réalisent la tâche pour la troisième fois, cette fois-ci, **le pattern ayant eu le moins de difficultés pour accéder à la réponse** est celui qui présente comparativement une **navigation plus « exhaustive »**.

De plus, chacun des patterns a emprunté majoritairement une classe de parcours différente. Ce sont les participants du pattern 2 qui ont emprunté les chemins les plus optimaux.

Tableau 18 : Répartition des parcours pour chaque Pattern Tâche 1-Temps 3 - Musée Carnavalet

	Pattern 1	Pattern 2
Parcours classe 1	22% (n=4)	67% (n=4)
Parcours classe 2	78% (n=15)	33% (n=2)
Nombre total de participants	100% (n=19)	100% (n=6)

#### À retenir sur ces patterns :

Pour cette première tâche, on observe qu'**une même façon de naviguer sur les pages du site ne conduit pas toujours à une même performance pour réussir la tâche**. En effet, la première fois que les individus réalisent la tâche on observe qu'ils prennent tous un temps assez conséquent pour naviguer sur les pages, y compris la page d'accueil, à y scroller et pourtant même avec cette exploration certains individus reviennent sur leurs traces et empruntent des chemins plus laborieux.

#### Evolution dans le temps:

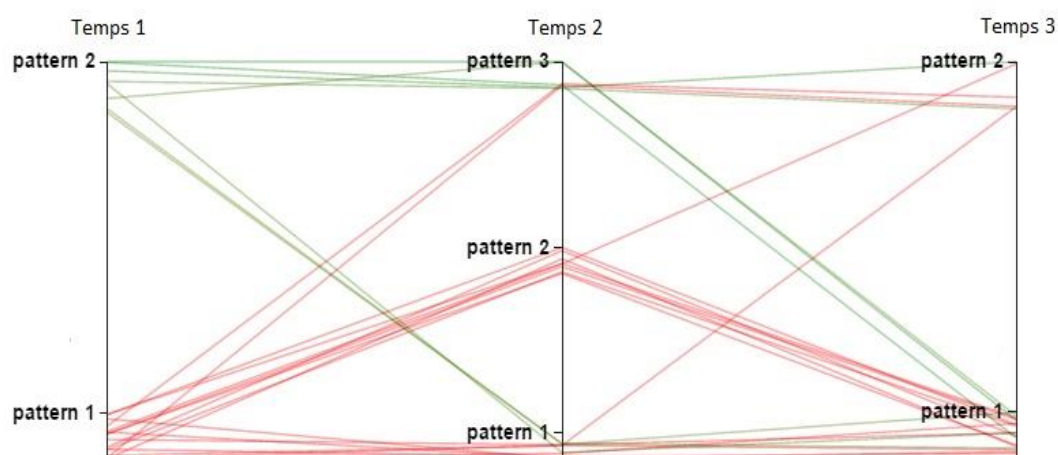


Figure 54 : Visualisation de l'évolution des participants à travers les répétitions de la première tâche sur le site du Musée Carnavalet

Comme l'illustre la représentation graphique, parmi les 15 participants naviguant de façon assez linéaire, 46,6% deviennent beaucoup plus rapide à naviguer sur les pages visitées et ne reviennent plus sur leurs traces. Cela signifie donc qu'**après avoir réalisé la tâche une première fois, la plus grande majorité des individus change de patterns et se retrouve avec une stratégie présentant autant de difficultés**. Nous envisagerons plus tard l'idée que le design de l'expérience puisse être à l'origine de ce résultat.

Le constat est similaire entre T2 et T3. 76% des participants de l'échantillon se retrouvent dans le pattern 1 présenté comme rapide mais provoquant tout de même

une certaine désorientation. **Pour cette tâche, il semblerait que la répétition à 48h et 72h d'intervalle impacte et fragilise les performances des individus.**

➤ *Site Musée Carnavalet – Tâche 2, organisation cocktail*

En fonction du moment auquel a eu lieu la tâche, deux ou trois patterns se distinguent.

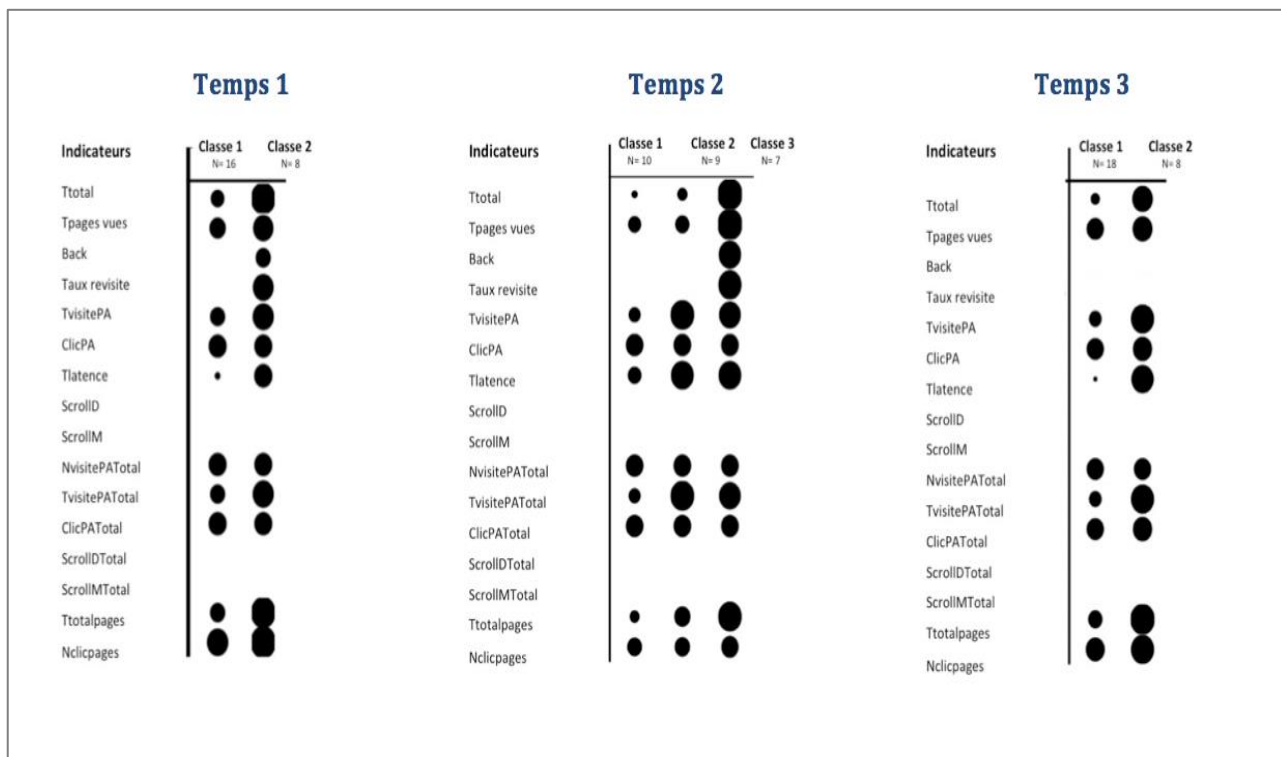


Figure 55 : Illustration des patterns pour la deuxième Tâche à Temps 1, 2 et 3 – Musée Carnavalet

*Description des patterns :*

La première fois que les individus ont réalisé cette tâche (Temps 1), les analyses révèlent deux patterns (trois outsiders ont été exclus des analyses).

Deux classes d'individus se distinguent, l'une rencontrant plus de difficultés pour accéder à l'information que l'autre :

- Avec une **navigation rapide et linéaire**, les individus du pattern 1 semblent n'éprouver aucune difficulté pour réaliser la tâche. En majorité, ils empruntent le chemin optimal afin d'accéder à la réponse (Tableau 12).
- À l'inverse, dans le second pattern, les participants ont un taux de revisite de 22% ce qui laisse paraître une **désorientation**. En moyenne, ils prennent également plus de temps à naviguer sur la page d'accueil et sur les autres pages visitées. C'est également au sein de ce pattern qu'ils empruntent en majorité les chemins de la classe 1 correspondant aux parcours les plus fastidieux. La représentation des chemins est en annexe.



- Ces éléments montrent que malgré le fait que tous les individus finissent par trouver la réponse, **certains ont besoin de plus de temps pour explorer les pages visitées que d'autres.**

Tableau 19 : Répartition des parcours pour chaque Pattern Tâche 2-Temps 1 - Musée Carnavalet

	Pattern 1	Pattern 2
Parcours classe 1	<b>31,3%</b> (n=5)	<b>75%</b> (n=6)
Parcours classe 2	<b>56,2%</b> (n=9)	<b>12,5%</b> (n=1)
Parcours classe 3	<b>12,5%</b> (n=2)	<b>12,5%</b> (n=1)
Nombre total de participants	<b>100%</b> (n=16)	<b>100%</b> (n=8)

À Temps 2, trois patterns se différencient (un outsider identifié). On identifie deux patterns pour lesquels les individus ont optimisé la recherche d'information :

- On retrouve pour les participants des patterns 1 et 2 une navigation rapide et directe. **Ces individus semblent avoir facilement mémorisé la structure du site et le chemin d'accès à la réponse.** Tous ces individus empruntent les parcours de la classe 1 regroupant les chemins les plus courts.
- Cependant, les participants du pattern 2 ont une stratégie de navigation un peu différente. En effet, on observe ici qu'ils mettent presque deux fois plus de temps à naviguer sur la page d'accueil, avant d'effectuer un clic sur un hyperlien, et que de façon générale ils prennent plus de temps sur l'ensemble des pages visitées. Cette manière de faire s'apparente à cette façon **exhaustive et consciencieuse de naviguer** que nous avons décrite par ailleurs, et qui est ici fructueuse.

Au sein du pattern 3, les individus présentent plus de difficultés pour accéder à la réponse qu'à Temps 1 :

- En visitant plus de 10 pages dont 26% sont des revisites, en mettant entre 3 et 6 fois plus de temps que les individus des autres patterns pour accéder à la réponse, **ces individus démontrent une navigation dispersée.**
- Par ailleurs, leur façon de naviguer sur la page d'accueil et sur le site présente des similitudes avec le pattern 2. À nouveau, il semblerait que pour certains individus **une façon assez exhaustive de naviguer sur les pages s'apparente plutôt à une difficulté pour prélever la bonne information et n'engendre pas une meilleure performance.**
- **Malgré le fait qu'ils répètent la tâche pour la seconde fois, la performance est encore moins bonne qu'à Temps 1.**

Tableau 20 : Répartition des parcours pour chaque Pattern Tâche 2-Temps 2 - Musée Carnavalet

	Pattern 1	Pattern 2	Pattern 3
Parcours classe 1	100% (n=10)	100% (n=9)	14,3% (n=1)
Parcours classe 2	0% (n=0)	0% (n=0)	85,7%(n=6)
Nombre total de participants	100% (n=10)	100% (n=9)	100% (n=7)

Enfin, à Temps 3, la dernière fois que les individus réalisent la tâche, on distingue deux patterns dans lesquels tous les participants ont amélioré la recherche d'information :

- Le pattern 1 regroupe les participants avec une **stratégie rapide et directe**, tous ont également emprunté les types de chemins les plus courts (3 pages). **Cette dernière répétition de tâche permet à ces participants d'être encore plus efficaces.**
- Dans le deuxième pattern, ils mettent deux fois plus de temps pour terminer la tâche et pour naviguer sur la page d'accueil. Ils visitent une page de plus mais nous **n'identifions plus de participants s'étant égaré sur le site.**
- Indépendamment du pattern, et comme le montre le tableau suivant, tous les participants ont majoritairement emprunté les parcours optimaux de la classe 1. La représentation des parcours est en annexe.

Tableau 21 : Répartition des parcours pour chaque Pattern Tâche 2-Temps 3 - Musée Carnavalet

	Pattern 1	Pattern 2
Parcours classe 1	100% (n=18)	62,5% (n=5)
Parcours classe 2	0% (n=0)	37,5% (n=3)
Nombre total de participants	100% (n=18)	100% (n=8)

#### À retenir sur ces patterns :

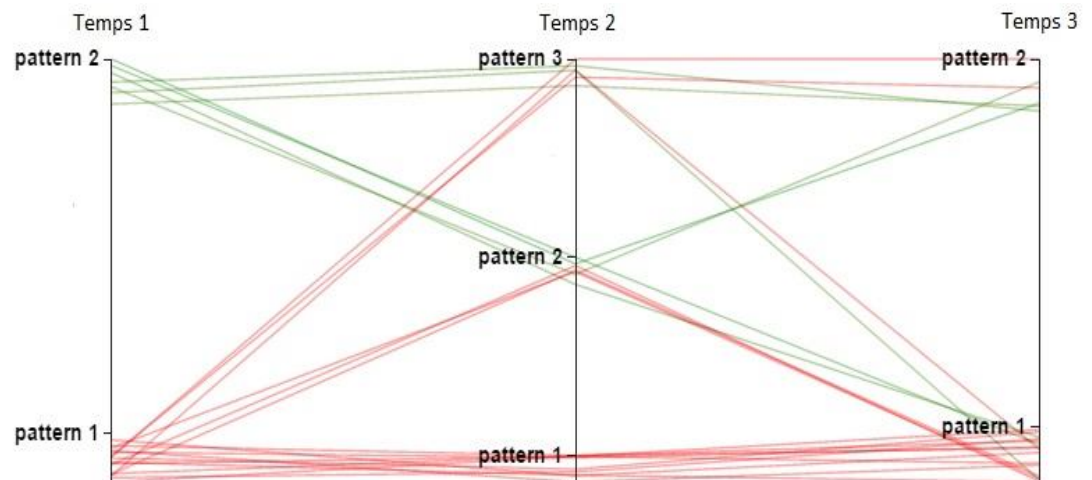
Pour les trois sessions, le pattern le **plus performant regroupe les individus les plus rapides** à naviguer sur le site et qui ne reviennent pas sur leurs traces.

Un second pattern se caractérise par **une façon plus exhaustive et consciencieuse de naviguer sur le site**, mais ces participants sont tout aussi performants.

Enfin, le troisième rassemble les individus éprouvant le plus de difficultés à accéder à la réponse. L'élément intéressant est qu'ici la façon de naviguer s'apparente également à une certaine « exhaustivité » car ils explorent les pages visitées, mais dans ce cas cela s'interprète comme **une difficulté pour prélever les bonnes informations.**

Indépendamment du moment de la répétition de la tâche, il semblerait donc, **qu'en fonction des individus, une stratégie de navigation peut être fructueuse pour certains et synonyme de difficultés pour d'autres.**

*Evolution dans le temps:*



*Figure 56 : Visualisation de l'évolution des participants à travers les répétitions de la deuxième tâche sur le site du Musée Carnavalet*

Entre T1 et T2, **62,5 % des participants identifiés dans le pattern rapide et direct conservent la même stratégie**, et tous l'a garde aussi pour T3.

Par ailleurs, ceux ayant éprouvé **le plus de difficultés à T1 sont ceux qui continuent à en ressentir le plus** pour la répétition 48 heures plus tard (57,1%). Il est difficile de parler de stabilité pour ce pattern car on ne le retrouve pas à T3. **Tous ont optimisé la recherche et ils ne sont plus désorientés.**

Contrairement à la première tâche, et de façon générale, les répétitions renforcent ici les performances des individus.

➤ Site Diplomatie Française – Tâche 3, création CNED

Nous identifions ici deux patterns pour chaque session.

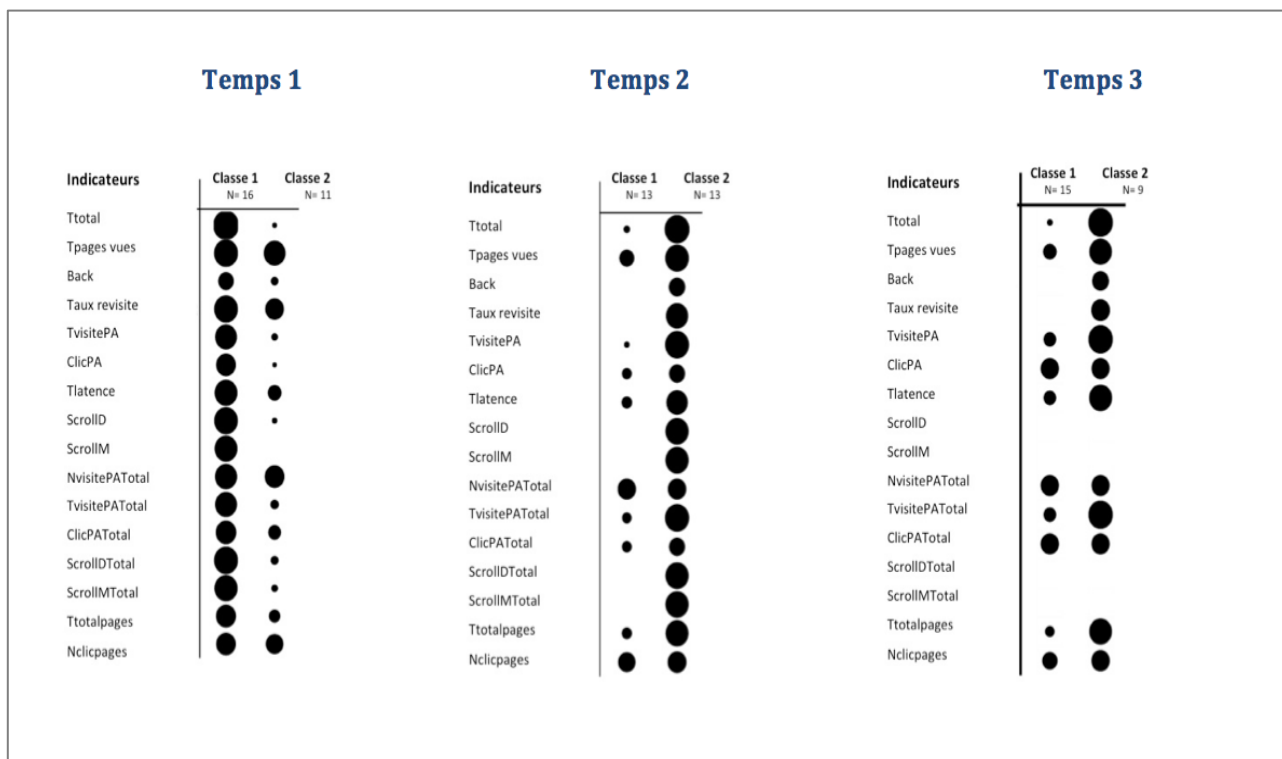


Figure 57 : Illustration des patterns pour la troisième Tâche à Temps 1, 2 et 3 – Diplomatie Française

*Description des patterns :*

La première fois que les 27 individus réalisent cette tâche, un tiers d’entre eux ne réussit pas la tâche. Ces participants se retrouvent dans deux patterns :

- **Au sein du pattern 1, 37,5% n’ont pas réussi à accéder à la réponse, et 27,2% dans le pattern 2.**
- Les individus du pattern 1 se distinguent surtout par le fait qu’ils **explorent plus exhaustivement le contenu des pages visitées**. Ils mettent deux fois plus de temps pour finir la tâche (265 secondes), 4 fois plus de temps sur la page d’accueil à la première visite (61 secondes) et que les distances de scroll verticales soient plus élevées sur cette page. Malgré cette exploration, ils empruntent en majorité les parcours les plus fastidieux.
- Au contraire, au sein du pattern 2, les individus naviguent de façon plus rapide sur l’ensemble des pages, y compris la page d’accueil. C’est d’ailleurs dans ce pattern que les individus ont en majorité emprunté les parcours les plus optimaux.

- En identifiant deux patterns différents comptant tous les deux des individus en échec, **il n'est pas possible ici de définir les caractéristiques propres à la réussite de la tâche.**

Tableau 22 : Répartition des parcours pour chaque Pattern Tâche 3-Temps 1 – Diplomatie Française

	Pattern 1	Pattern 2
Parcours classe 1	68,7% (n=11)	45,4% (n=5)
Parcours classe 2	31,3% (n=5)	54,6% (n=6)
Nombre total de participants	100% (n=16)	100% (n=11)

Quarante-huit heures plus tard, lorsque les individus réalisent cette tâche pour la seconde fois, nous identifions deux patterns, dont l'un comptant 46,1% de participants en échec:

- Comme pour la session précédente, les participants ayant échoué à la tâche se retrouve dans le pattern regroupant les individus ayant passé le plus de temps à explorer la page d'accueil et, de façon générale, le site. L'étude des parcours montre qu'ils se perdent au fur et à mesure de la navigation.
- Dans ce pattern même si 46 % des personnes échouent, les 54% restants réussissent la tâche.
- Pour le pattern 1, tous les individus ont réussi la tâche en étant très rapide et ils empruntent les parcours les plus efficaces.

Tableau 23 : Répartition des parcours pour chaque Pattern Tâche 3-Temps 2 – Diplomatie Française

	Pattern 1	Pattern 2
Parcours classe 1	38,5% (n=5)	92,3% (n=12)
Parcours classe 2	61,5% (n=8)	7,7% (n=1)
Nombre total de participants	100% (n=13)	100% (n=13)

Enfin, la dernière fois que les individus réalisent nous distinguons deux patterns dont l'un comptant 13,3% des individus en échec. Quatre individus ont été considérés comme des outsiders.

Comme pour les sessions précédentes, l'un des patterns éprouve des difficultés pour accéder à la réponse :

- Au sein du pattern 1, malgré **l'exploration des pages visitées, ces individus se perdent** encore au fur et à mesure de la navigation.

- Au contraire, dans le second pattern les participants ont optimisé la recherche de l'information et ils empruntent majoritairement les parcours de la classe 1 regroupant les parcours les plus courts.

Tableau 24 : Répartition des parcours pour chaque Pattern Tâche 3-Temps 3 – Diplomatie Française

	Pattern 1	Pattern 2
Parcours classe 1	73,3% (n=11)	22,2% (n=2)
Parcours classe 2	26,7% (n=4)	77,8% (n=7)
Nombre total de participants	100% (n=15)	100% (n=9)

#### À retenir sur ces patterns :

Nous observons des similitudes assez importantes entre les trois sessions. Les participants se retrouvent soit dans :

- une stratégie « exhaustive » d'exploration des pages mais impliquant une désorientation,
- ou à l'inverse une stratégie plus rapide et fluide ne laissant percevoir que de faibles difficultés, voir aucune, pour accéder à l'information.

Sur ce site l'exploration approfondie, des pages, y compris celle de l'accueil, **n'engendre pas une meilleure performance dans la navigation**. À l'inverse, cette façon de naviguer s'apparente d'avantage à **une difficulté pour prélever les informations sur le site**.

Par ailleurs, en observant, le comportement de navigation des individus ayant échoué à la tâche, deux points sont à souligner:

- Il n'est **pas possible** ici de définir **des caractéristiques propres à la réussite de la tâche**.
- En fonction des individus, **une même façon de naviguer peut conduire à la réussite ou l'échec à une tâche**.

### Evolution dans le temps:

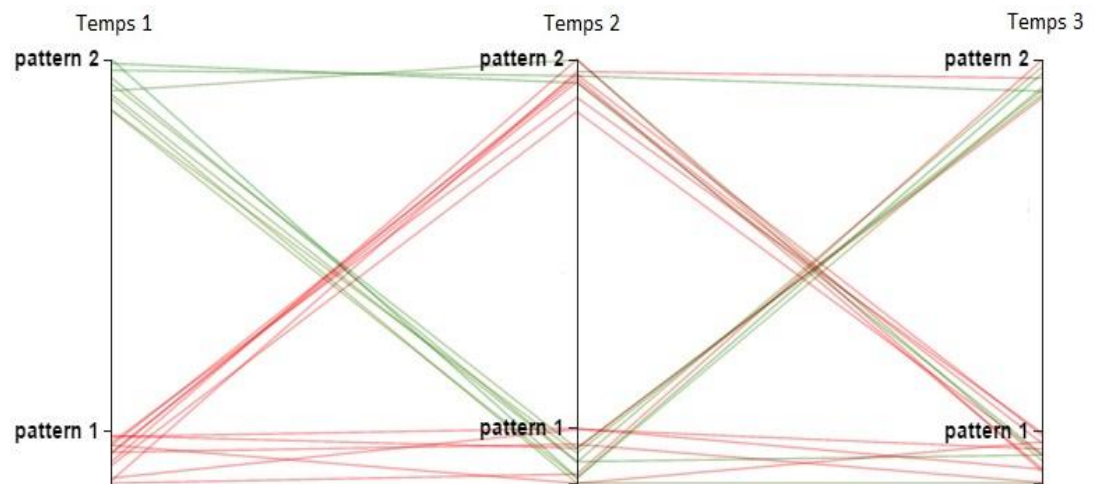


Figure 58 : Visualisation de l'évolution des participants à travers les répétitions de la troisième tâche sur le site de Diplomatie Française

Les participants avec une stratégie « plus consciencieuse » mais désorientée présentent une certaine stabilité entre T1 et T2 : 56,2% des individus restent dans ce même pattern. Au contraire entre T2 et T3 ce résultat baisse à 11,1%.

Le constat est identique pour les participants avec une navigation rapide et fluide. Ils sont 63,6% à faire partie de ce même pattern entre T1 et T2, mais plus que 28,5% entre T2 et T3.

➤ *Site Diplomatie Française – Tâche 4, sécurité Andorre*

Enfin, pour la dernière tâche, nous identifions pour les trois sessions deux patterns de navigation.

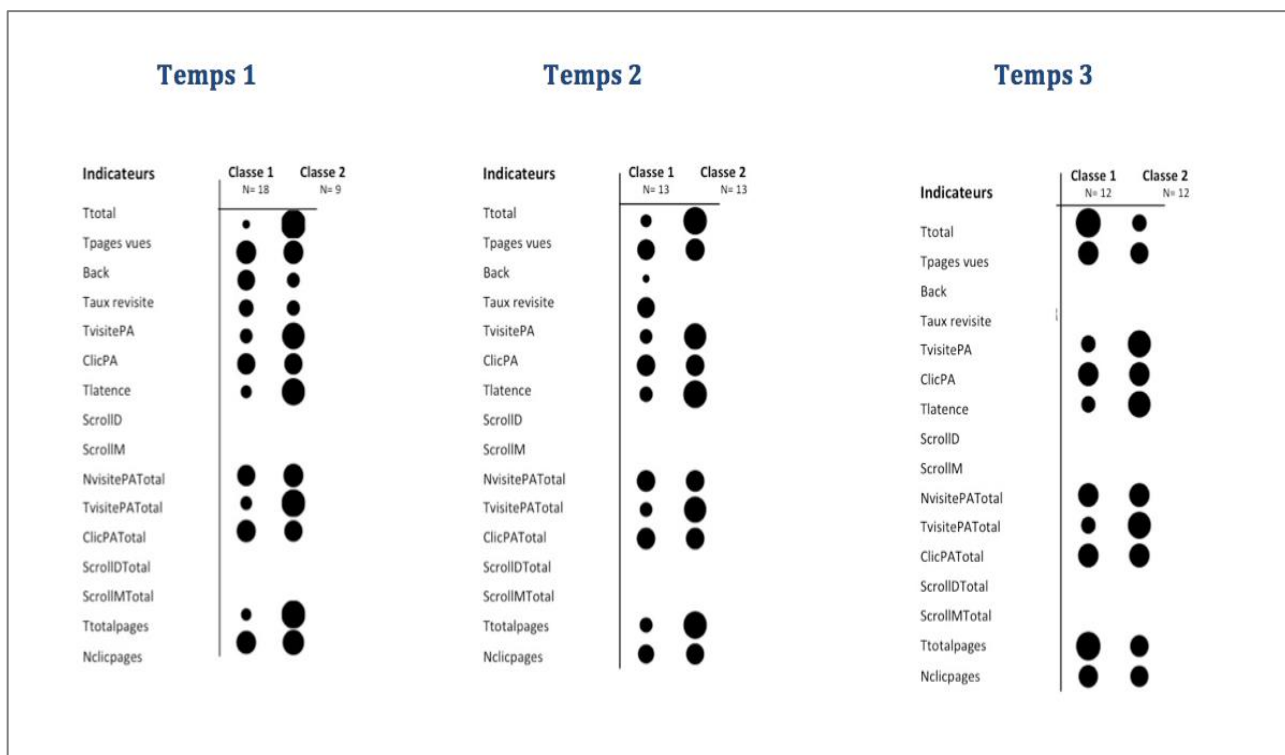


Figure 59 : Illustration des patterns pour la troisième Tâche à Temps 1, 2 et 3 – Diplomatie Française

*Description des patterns :*

À Temps 1, deux classes sont identifiées. Bien que tous les individus visitent et revisitent très sensiblement le même nombre de pages, on observe des différences d’exploration des pages visitées.

- Dans le pattern 2, les participants prennent deux fois plus de temps à explorer la page d’accueil, à réaliser un clic sur un hyperlien afin de la quitter et à naviguer sur les autres pages visitées. En l’occurrence, ils mettent également deux fois plus de temps à réaliser la tâche.
- Deux chemins étaient possibles pour accéder à la réponse ; via le lien « *Conseils aux voyageurs* » et l’autre via « *Pays Zone Euro* ». Ces deux liens étaient présentés côte à côte dans le menu principal horizontal. L’analyse des parcours identifie ces deux classes de chemins et montre une répartition quasi similaire entre les deux patterns.
- On retrouve donc ici **deux stratégies différentes** ; l’une qui amène les individus à naviguer de façon plus rapide, et l’autre qui conduit les participants à être plus exhaustifs et consciencieux sur le site, **mais qui au final conduisent à des performances sensiblement identiques où les individus éprouvent certaines difficultés pour effectuer la tâche**.



Tableau 25 : Répartition des parcours pour chaque Pattern Tâche 4-Temps 1 – Diplomatie Française

	Pattern 1	Pattern 2
Parcours classe 1	52,6% (n=10)	50% (n=4)
Parcours classe 2	47,4% (n=9)	50% (n=4)
Nombre total de participants	100% (n=19)	100% (n=8)

La seconde fois que les participants doivent réaliser cette tâche, on identifie également deux patterns présentant certaines similitudes avec ceux identifiés à la session précédente (dont un *outsider*) :

- Dans l'ensemble tous les participants ont optimisé la recherche d'information.
- La principale différence entre les deux patterns concerne **la vitesse d'exploration de la page d'accueil et des autres pages du site**.
- En effet, les individus du pattern 1 sont plus rapides pour explorer les pages, mais ils sont également les seuls à **se disperser au cours de leur navigation**.
- L'analyse des parcours met en exergue 3 classes de parcours différentes et on peut s'apercevoir qu'en fonction du pattern, la répartition est différente. Les individus du pattern 1 ont majoritairement emprunté les chemins passant par le lien « *Conseils aux voyageurs* », tandis que dans l'autre pattern ils ont autant emprunté ce parcours que celui impliquant de passer par « *Pays Zone Euro* ».

Tableau 26 : Répartition des parcours pour chaque Pattern Tâche 4-Temps 2 – Diplomatie Française

	Pattern 1	Pattern 2
Parcours classe 1	53,8% (n=7)	23,1% (n=3)
Parcours classe 2	23,1% (n=3)	38,4% (n=5)
Parcours classe 3	23,1% (n=3)	38,4% (n=5)
Nombre total de participants	100% (n=13)	100% (n=13)

Enfin, la dernière fois que les participants réalisent cette tâche, nous identifions également deux patterns comptant chacun 12 personnes. Trois individus ont été considérés comme *outsiders* :

- Par rapport à la précédente session, on constate que **tous ont pu davantage optimiser la recherche d'information**. En effet, plus aucun participant n'effectue un retour arrière et ne revisite une page.
- Comparativement **deux stratégies se distinguent** : l'une qui amène les individus à naviguer de façon plus rapide sur les pages (Pattern 1), et l'autre qui conduit les participants à être plus « posés » sur le site (Pattern 2). Ils prennent plus de temps sur la page d'accueil et avant de réaliser une action.
- L'analyse des parcours montre qu'en majorité dans les deux patterns, tous ont emprunté la même classe de parcours.

Tableau 27 : Répartition des parcours pour chaque Pattern Tâche 4-Temps 3 – Diplomatie Française

	Pattern 1	Pattern 2
Parcours classe 1	<b>66,6%</b> (n=8)	<b>75%</b> (n=9)
Parcours classe 2	<b>33,4%</b> (n=4)	<b>25%</b> (n=3)
Nombre total de participants	<b>100%</b> (n=12)	<b>100%</b> (n=12)

#### À retenir sur ces patterns :

Nous retrouvons ici des patterns que nous avons déjà identifiés dans les tâches précédentes :

- une **stratégie « consciencieuse, exhaustive » d'exploration** des pages mais impliquant une certaine **désorientation**,
- une **stratégie assez rapide** de navigation conduisant aussi les participants à se **perdre en chemin**,
- une façon « **consciencieuse, exhaustive** » de naviguer sur le site, mais conduisant les participants à accéder à la réponse de **façon linéaire**,
- Et enfin, une **stratégie de navigation rapide et directe** laissant percevoir une **bonne assimilation** de la structure du site et aucune difficultés.

### Evolution dans le temps:

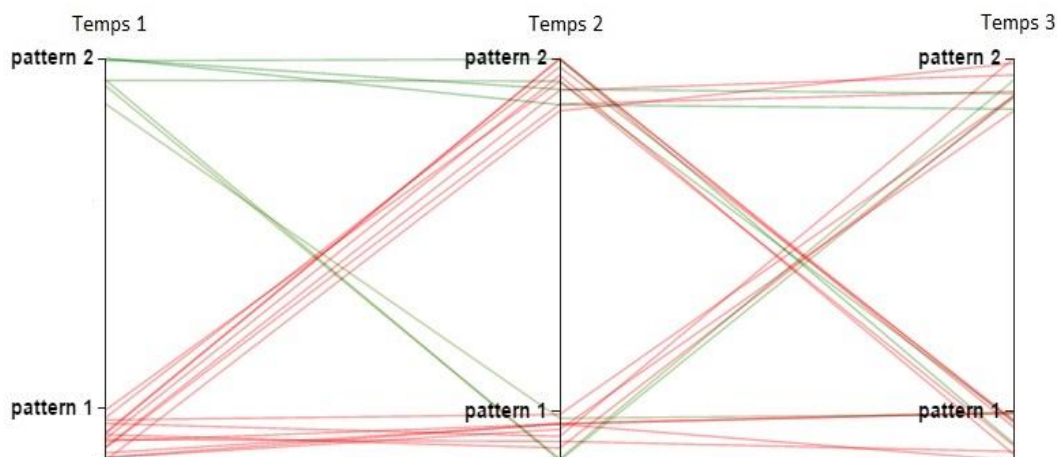


Figure 60 : Visualisation de l'évolution des participants à travers les répétitions de la dernière tâche sur le site de Diplomatie Française

À travers les patterns identifiés, on remarque que les performances des individus s'améliorent à chaque répétition de la tâche. Toutefois, en regardant si les gens conservent leur façon rapide ou plus exhaustive de naviguer, on s'aperçoit qu'entre T1 et T2, 50% des participants conservent cette stratégie rapide. Par ailleurs, même si à T3 cette rapidité n'est plus synonyme de désorientation, ils ne sont plus que 22,2% à être resté cohérent sur cette façon de naviguer entre la première et la dernière session.

#### 4.2.3.5

#### Réflexion intermédiaire

Plusieurs conclusions émanent de cette étude. La première concerne **l'homogénéité des patterns**. À travers les 4 tâches, nous en avons identifié quatre :

- une **stratégie dans laquelle les participants explorent assez finement les pages** (y compris de la page d'accueil) mais impliquant une certaine **désorientation dans la navigation**,
- une façon tout aussi « **exhaustive et consciencieuse** » de naviguer sur le site, mais conduisant les participants à accéder à la réponse sans grandes difficultés,
- une **stratégie plus rapide** de navigation sur les pages visitées dans laquelle les participants se **perdent en chemin**,
- Et enfin, une **stratégie de navigation rapide et directe** laissant percevoir une **bonne assimilation** de la structure du site et très peu de difficultés.

Toutefois, concernant la stabilité des patterns, nous avons pu remarquer que pour une même tâche, nous n'observons **pas toujours les mêmes patterns entre les 3 sessions**. En effet, à 72 heures pour la tâche 4 on ne retrouve plus de stratégies impliquant des difficultés pour accéder à l'information. À l'inverse pour la tâche 1, on observe toujours pour certains individus des difficultés quand ils doivent répéter la tâche pour la troisième fois. Dans ce cas en voulant rapidement naviguer sur les pages, ces individus empruntent les chemins les plus longs et reviennent parfois sur leurs traces.

Nous pourrions penser que le design du protocole et l'effet de « découverte » puisse expliquer ce résultat. En effet, il n'est pas commun de demander à trois reprises de faire une même tâche, certains individus ont peut-être pensé ne pas avoir trouvé la solution aux deux précédentes sessions. Cependant, nous retrouvons également ce résultat pour la tâche 3, session 3. **En fonction du contexte (tâche, site), certains individus parviennent plus difficilement à assimiler la structure du site et le chemin d'accès à la réponse.**

Par ailleurs, si l'on regarde pour la tâche 1 session 3 et tâche 3 session 3 quelle est **la stratégie d'exploration adoptée par ces individus éprouvant des difficultés, on n'observe que ce ne sont pas les mêmes.** Pour la tâche 1, il s'agit du pattern présentant une stratégie plus rapide et conduisant à une certaine désorientation alors que pour la tâche 3 il s'agit au contraire des individus montrant une certaine exhaustivité à naviguer et explorer le contenu mais qui finalement reflète une difficulté à s'orienter. Bien que nous avons qualifié le site de Diplomatie Française comme un peu moins organisé, il semblerait pertinent d'approfondir les travaux afin d'analyser si des caractéristiques du site peuvent influencer sur l'adoption d'un pattern plutôt qu'un autre. Est-ce qu'un pattern démontrant une navigation exhaustive mais désorientée peut s'expliquer par des pages avec trop de contenus ? Il serait pertinent d'approfondir et de tester plusieurs architectures et organisations de l'information dans ce sens.

Par ailleurs, dans le cas où l'on retrouve le même pattern pour les 3 sessions, on remarque, qu'en fonction de la tâche, les individus ne restent pas toujours cohérents dans leur façon de faire. Pour la tâche 3, entre la session 2 et 3 on observe que la majorité de ceux qui avaient une stratégie plutôt rapide deviennent plus « consciencieux » et inversement (stabilité comprise entre 11 et 28,5%). Au contraire, pour la tâche 2, concernant le pattern décrit comme rapide et direct, 62,5% des participants conservent cette même façon de faire pour toute la session. **Nous n'avons donc pas observé de tendance marquée à la stabilité ou à l'instabilité. Les individus adaptent leur stratégie en fonction du contexte.**

Un autre point important concerne la performance d'un pattern. Nous avons pris en compte les individus ayant échoué à la tâche 3. L'analyse réalisée met en exergue qu'un même pattern peut conduire à la réussite ou à l'échec. Il ne semble donc **pas possible de définir les caractéristiques propres à la réussite d'une tâche. En fonction des individus, une même façon de naviguer peut donc conduire à une performance différente.**

#### 4.2.4

### Discussion

Dans les deux expérimentations précédentes, nous avons souhaité voir comment évoluer les patterns dans le temps et si en faisant répéter une tâche à court terme et long terme on retrouverait de nouveaux patterns par rapport aux tâches initiales. Ainsi, à l'instar d'études descriptives comme celles de Tauscher et Greenberg (1997), Cockburn et McKenzie (2000) ou encore d'Obendorf & Weinreich (2007) ayant porté leur attention sur le comportement de revisite, nous avons souhaité étudier cet aspect de la navigation à travers la notion de patterns dans des protocoles définis et avec des tâches précises.

La répétition à court terme met en exergue trois patterns, tandis qu'à plus long terme on en observe quatre. Nous n'observons donc pas exactement les mêmes patterns pour une répétition à long terme et court terme. En effet, la stratégie dans laquelle les participants naviguent de façon rapide sur la première page du site ainsi que sur les autres pages visitées tout en les amenant à revenir sur leurs chemins n'a pas été identifiée. Ainsi à court terme lorsque les gens revisitent un site et qu'ils naviguent rapidement sur les pages, ils n'ont pas besoin de revenir sur leurs traces.

Par ailleurs, nous avons conclu que pour les répétitions à court terme, le pattern le plus efficace pour accéder à la réponse était toujours celui qui regroupait les gens naviguant rapidement sur les pages visitées, les explorant très peu, y compris la page d'accueil, et qui n'effectuaient que très peu, voir pas du tout de retour arrière et de revisites. À long terme, nous ne pouvons généraliser ce résultat. Nous avons identifié, par exemple, pour la tâche 1 à temps 3 ou tâche 4 à temps 2, que le pattern le plus efficient était celui dans lequel les gens exploraient plus posément le contenu des pages visitées. Ainsi, même si de façon générale, on observe que les gens optimisent leur façon de naviguer entre les sessions de répétitions (moins de temps pour réaliser la tâche, de pages visitées, de back, de temps passé sur ces pages, de scrolls, etc.) une répétition de tâche à court terme et long terme n'amène pas forcément les individus à être plus efficace pour accéder à l'information de la même façon. En se basant sur ces différences de navigation, il serait intéressant de poursuivre les travaux et de voir comment optimiser les pages et sites web en fonction du type de revisites. En effet, à notre connaissance la littérature propose surtout des recommandations établies pour la conception des navigateurs.

Pour ces deux expériences, nous avons sélectionné des participants ayant des profils homogènes : des étudiants d'environ une vingtaine d'année et ayant une bonne expérience sur le Web. Les résultats ont montré des similitudes lors de l'identification des patterns. Nous pouvons donc nous demander si l'homogénéité de nos échantillons influe sur nos 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.

## 4.3

# INFLUENCE DES CARACTERISTIQUES DES PARTICIPANTS SUR LES PATTERNS

---

### 4.3.1

## Problématique et contexte de l'étude

Cette étude repose sur plusieurs objectifs :

- 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. L'idée était donc de voir si avec des participants moins homogènes et avec un protocole différent impliquant un ordre aléatoire des tâches (sans remise) on retrouve des patterns.
- Nous avons remarqué dans la littérature que les patterns sont souvent expliqués à travers les styles cognitifs, mais certains chercheurs recommandent d'étendre les travaux et de s'intéresser à d'autres déterminants (Huang, 2003 ; Bousbia, Rebia, Labat & Balla 2010). Compte tenu de notre protocole et du fait que les gens devaient réaliser plusieurs tâches sur un même site, nous avons souhaité étudier si ces patterns pouvaient s'expliquer à travers le style d'apprentissage des participants.
- Enfin, si des individus ont échoué, est ce que l'on retrouve, comme pour l'étude précédente, les mêmes patterns conduisant à l'échec.

### 4.3.2

## Méthodologie expérimentale

#### 4.3.2.1

### Matériel et terrain

Comme pour l'ensemble de nos expérimentations, nous avons utilisé le système d'oculométrie déporté Tobii T120. Le logiciel Tobii Studi (v3.0) a permis de recueillir les fichiers logs pour chaque participant et chaque tâche.

La procédure expérimentale a eu lieu dans les locaux du Ministère de l'Égalité des Chances à Luxembourg pendant 25 jours. 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 contrôlé par l'expérimentateur pour lancer les phases de l'expérimentation.

#### 4.3.2.2

### Procédures

Comme pour les deux précédentes expérimentations, 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 n'était imposée.

Les tâches à réaliser, dans un ordre aléatoire sans remise, étaient les suivantes (depuis la page d'accueil du site [www.mega.public.lu](http://www.mega.public.lu)) :

1. « On fait souvent la différence entre sexe et genre. À 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, 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, trouvez la page du programme pour créatrices d'entreprises. »
4. « À partir de la page d'accueil du portail MEGA, 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, trouvez le lien permettant d'accéder aux statistiques sur l'écart salarial entre hommes et femmes au Luxembourg ».

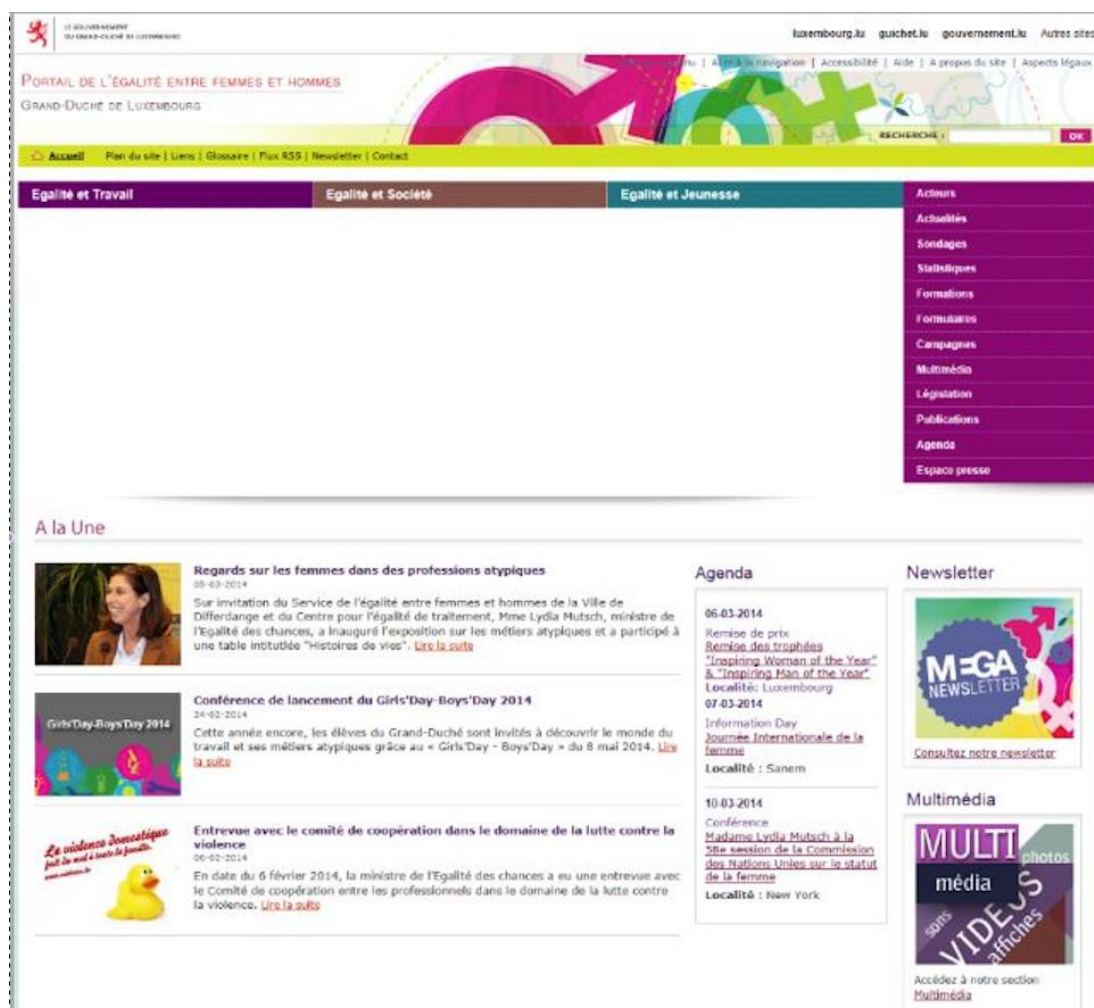


Figure 61 : Page d'accueil du site MEGA

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 :

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

Précisons toutefois que la tâche 3 n'a pas été analysée. En effet, à l'issue des passations, nous nous sommes aperçus que la notion de « créatrices » n'avait pas été correctement comprise amenant plus de 80% des utilisateurs à l'échec.

Enfin, un questionnaire post test concernant le style d'apprentissage était rempli par les participants. (Annexe C). Comme expliqué plus haut, il s'agit ici d'étudier s'il y a un lien entre les patterns de navigation et le style d'apprentissage.

Dans ce cadre, nous avons choisi d'utiliser le questionnaire ISALEM (Inventaire des Styles d'apprentissage du laboratoire d'Enseignement Multimédia) mis en point en 1997 par l'équipe pluridisciplinaire du Laboratoire d'Enseignement Multimédia de l'Université de Liège.

Basé sur la théorie de l'apprentissage expérientiel de Kolb (1984) et le learning-style inventory (LSI) du même auteur (1976), ce questionnaire présente douze situations rencontrées aussi bien dans le monde académique que dans la vie de tous les jours. Pour chaque situation, quatre réactions (correspondant chacune à un des styles d'apprentissage) sont proposées. Il convient de les classer dans l'ordre qui correspond le plus à sa préférence. ([http://www2.ulg.ac.be/lem/StyleApprent/StyleApprent\\_CG/page\\_05.htm](http://www2.ulg.ac.be/lem/StyleApprent/StyleApprent_CG/page_05.htm), consulté le 07 Aout 2016).

### **4.3.2.3**

#### **Echantillon**

Quarante-sept participants ont été recrutés, toutefois les données de 3 personnes n'ont pas fait l'objet d'une analyse en raison de leurs difficultés de compréhension de la langue française.

Les participants avaient 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.

57,5% des participants étaient des femmes et 42,5% des hommes. Tous les participants déclaraient utiliser Internet depuis plus de 6 ans et tous l'utilisaient dans le cadre de leur travail.

Par ailleurs, 52 % de l'échantillon jugeait ses compétences liées à Internet comme bonnes voir très bonnes, et 13 % comme excellentes. Les 10% restant s'estimait ne pas toujours être complètement à l'aise avec le Web.

Concernant les styles d'apprentissages de notre échantillon, nous avons identifiés :

- Plus de la moitié de notre échantillon (56,8%) comme faisant partie du style « Intuitif pragmatique » (IP) qui agit davantage par instinct qu'en fonction d'une analyse purement logique



- 18,1 % : « Méthodique pragmatique » (MP) qui est capable de résoudre des problèmes et de prendre des décisions de façon directe et sans tergiverser, tout en sélectionnant rapidement la solution optimale
- 16% : « Méthodique réflexif » (MR) qui excelle à synthétiser un vaste registre d'informations de manière logique et concise
- 9,1% : « Intuitif réflexif » (IR) qui observe énormément avant d'agir afin de considérer la situation sous des angles très variés.

### 4.3.3 Analyse des données

Concernant les modalités d'analyse des données, elles restent inchangées. Nous avons appliqué la procédure décrite précédemment en modifiant le paramétrage par défaut sur le logiciel R afin de permettre la détection de patterns à un niveau de granularité plus fin. Compte tenu du protocole de passation, et du fait que les tâches étaient réalisées dans un ordre de passation aléatoire, l'analyse de l'évolution d'une tâche à une autre n'avait pas de sens. Enfin, comme pour l'étude précédente, les individus ayant échoué ont été pris en compte dans les analyses.

Les données chiffrées figurent en Annexe C.

#### ➤ Tâche 1, différence entre sexe et genre

Trois patterns comportementaux différents ont été identifiés, et comptent respectivement 20, 16 et 8 participants.

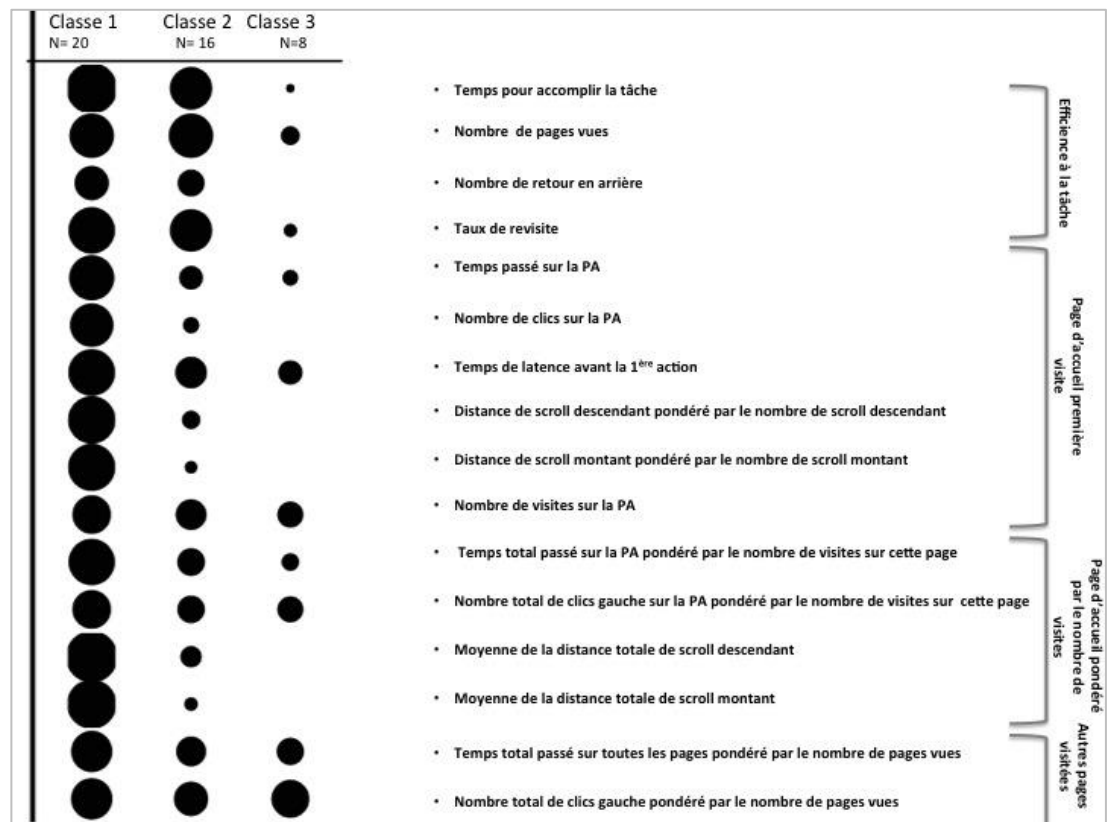


Figure 62 : Illustration des patterns pour la Tâche 1

Au sein du 1<sup>er</sup> pattern, la stratégie de navigation se caractérise surtout par une désorientation et une exploration approfondie des pages visitées:

- En effet, sur la page d'accueil ils passent en moyenne deux fois plus de temps à naviguer (47 secondes) et ils réalisent la distance moyenne de scroll la plus importante. De façon générale, **ils prennent plus de temps à explorer les pages vues**.
- Malgré cette exploration, leur taux de revisite et le nombre de retour en arrière laisse à penser qu'**ils se perdent en cherchant l'information**. Ils sont d'ailleurs ceux qui, en moyenne, reviennent 2 fois au point de départ sur la page d'accueil.
- Au sein de ce pattern la moitié des participants a échoué à la tâche.

Le pattern 2 présente également une désorientation mais ces participants sont plus rapides à naviguer sur l'ensemble des pages visitées et notamment la page d'accueil.

- Ces participants ont des difficultés à prélever la pertinence des informations car ils finissent par **revisiter des pages**. Ce constat est validé par l'étude des parcours, ils sont ceux **empruntant la classe de parcours la plus laborieuse**.
- Cette stratégie de navigation plus rapide conduit les trois quarts des individus à échouer à la tâche.

Enfin, le dernier pattern étudié se distingue par sa navigabilité rapide et directe.

- Aucun des participants n'est retourné sur une page précédemment visitée.
- Ces individus sélectionnent très rapidement le bon lien afin de quitter la page d'accueil pour accéder à la réponse. Ils n'éprouvent **aucune difficulté pour se repérer sur le site et pour se diriger**.
- 100% des individus ont réussi la tâche.

La classification des parcours a permis de classer les participants en trois groupes. Le tableau suivant montre la répartition de ces parcours au sein de chaque pattern et on perçoit facilement qu'en grande majorité ce sont les individus du Pattern 3 qui ont emprunté les chemins les plus optimaux (type 3) pour accéder à la réponse. La représentation des parcours est en annexe C.

Tableau 28 : Répartition des parcours pour chaque Pattern Tâche 1 – MEGA

	Pattern 1	Pattern 2	Pattern 3
Parcours type 1	35% (n=7)	56,25% (n=9)	0
Parcours type 2	45% (n=9)	31,25% (n=5)	0
Parcours type 3	20% (n=4)	12,5% (n=2)	100% (n=8)
Nombre total de participants	100% (n=20)	100% (n=16)	100% (n=8)

➤ *Tâche 2, mesures pour l'égalité des chances*

L'analyse de la tâche 2 met en exergue quatre patterns. L'un constitué de 19 participants, l'autre de 11 et les deux derniers de 7 individus chacun.

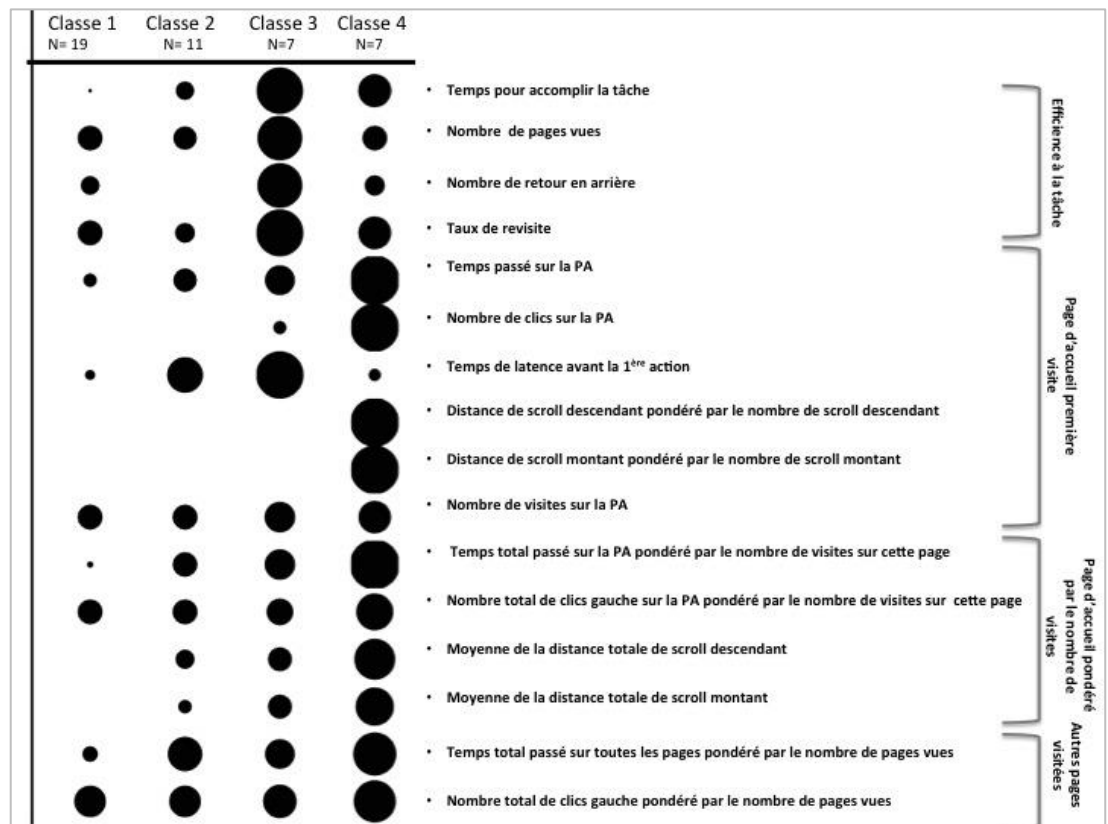


Figure 63 : Illustration des patterns pour la Tâche 2

Les participants du pattern 1 se caractérisent par la stratégie la plus rapide pour accéder à la réponse

- En effet, ils passent en moyenne **deux fois moins de temps que les autres participants sur la page d'accueil** (10,8 secondes), et presque 3 fois moins

de temps sur les autres pages visitées (8,4 secondes). Ils réalisent très rapidement un clic afin de quitter la page.

- Même si certains participants ont revisité certaines pages, l'analyse des parcours montre qu'ils empruntent en majorité les parcours optimaux.
- Cette rapidité dans la navigation est synonyme de réussite à 100%.

Les participants du pattern 2 présentent une stratégie de navigation moins rapide et qui semble aux premiers abords très efficaces.

- En effet, par rapport aux individus du pattern 1, ces participants passent en moyenne 7 secondes de plus sur la page d'accueil et avant de réaliser un clic sur un lien. De façon générale, ils réalisent la tâche assez rapidement (75secondes), tout en visitant le moins de pages. **Ces individus ne sembleraient donc pas se perdre en chemin.**
- Cependant, on observe ici un **haut taux d'échec** (81,8%).

Le pattern 3 regroupe les individus ayant littéralement « tourné en rond ».

- Avec en moyenne le taux de revisite le plus important (27%), le plus de retour arrière (2,4), le plus de pages visitées (16), et le temps le plus conséquent pour réaliser la tâche (265 secondes) **ces participants semblent être perdus.**
- Comparativement aux patterns 1, 2 ils prennent le **plus de temps pour explorer les nombreuses pages vues**. Malgré cette exploration, **71,5% des participants échouent.**

Enfin, le dernier pattern se caractérise aussi par une exploration approfondie des pages visitées et de la page d'accueil mais ces individus sont beaucoup moins désorientés,

- L'exploration se manifeste ici à la fois en terme de temps (57 secondes) c'est à dire 2 à 5 fois plus que les autres patterns, mais aussi par rapport au scroll.
- En passant en moyenne le plus de temps sur les pages visitées (31 secondes) et en y effectuant le plus de clics, **ces participants sont ceux qui « épiluchent » le plus les pages du site.**
- Leur taux de revisite de 13% laisse penser qu'ils ont éprouvé des difficultés pour s'orienter mais moins que pour le pattern 3.
- La majorité des participants a réussi la tâche.

La classification des parcours a permis de classer les participants en deux groupes. Avec une LD de 0,25 et un parcours minimal de 2 pages, il est intéressant de voir qu'en majorité ce sont les individus du pattern 1 qui ont emprunté les parcours optimaux de type 2.

Tableau 29 : Répartition des parcours pour chaque Pattern Tâche 2 – MEGA

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Parcours type 1	47,3% (n=9)	63,6% (n=7)	100% (n=7)	57,1% (n=4)
Parcours type 2	52,7% (n=10)	36,4% (n=4)	0	42,9% (n=3)
Nombre total de participants	100% (n=19)	100% (n=11)	100% (n=7)	100% (n=7)

*Tâche 4, demande de patronage et subvention*

À la suite de la classification, trois participants ont été considérés comme des *outsiders*. Les résultats mettent en exergue trois patterns comptant respectivement 20, 13 et 8 participants.

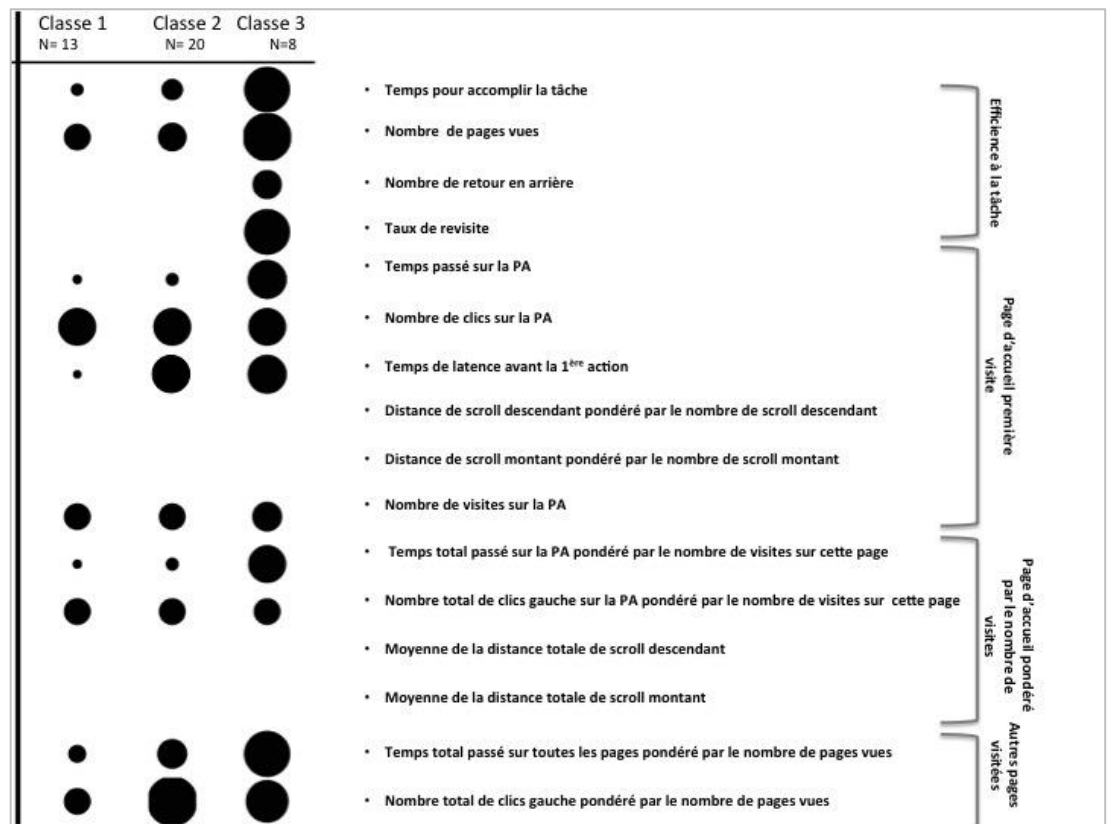


Figure 64 : Illustration des patterns pour la Tâche 4

Bien que les patterns 1 et 2 présentent de fortes similitudes et regroupent des participants ayant tous réussis la tâche, nous pouvons les distinguer notamment par la rapidité de navigation.

Les participants de P1 présentant la **navigation la plus rapide et directe**.

Au sein du pattern 2 les individus ont mis **deux fois plus de temps que ceux du pattern 1 à réaliser la tâche** (35 secondes) et ils naviguent de façon plus exhaustive :

- Ils mettent également **deux fois plus de temps pour explorer le contenu des pages visitées alors qu'ils en visitent en moyenne sensiblement le même nombre.**
- Comme pour dans le pattern 1, aucun individu ne revient sur ces traces.

Enfin le dernier pattern présente une désorientation et une exploration exhaustive des pages visitées :

- **Quatre à sept fois plus de temps pour explorer la page d'accueil.**
- D'importantes **difficultés pour s'orienter** avec 6,5 pages visitées (alors que la réponse était accessible une page après l'accueil) et un taux de revisite de 11%.
- Au sein de ce pattern, un seul individu a échoué à la tâche.

Tableau 30 : Répartition des parcours pour chaque Pattern Tâche 4 – MEGA

	Pattern 1	Pattern 2	Pattern 3
Parcours type 1	<b>85%</b> (n=17)	<b>92,3%</b> (n=12)	<b>50%</b> (n=4)
Parcours type 2	<b>15%</b> (n=3)	<b>7,4%</b> (n=1)	<b>50%</b> (n=4)
Nombre total de participants	<b>100%</b> (n=20)	<b>100%</b> (n=13)	<b>100%</b> (n=8)

### Tâche 5, statistiques écart salariales hommes femmes

L'analyse relève quatre patterns dont un participant exclu des analyses car considéré comme outsider.

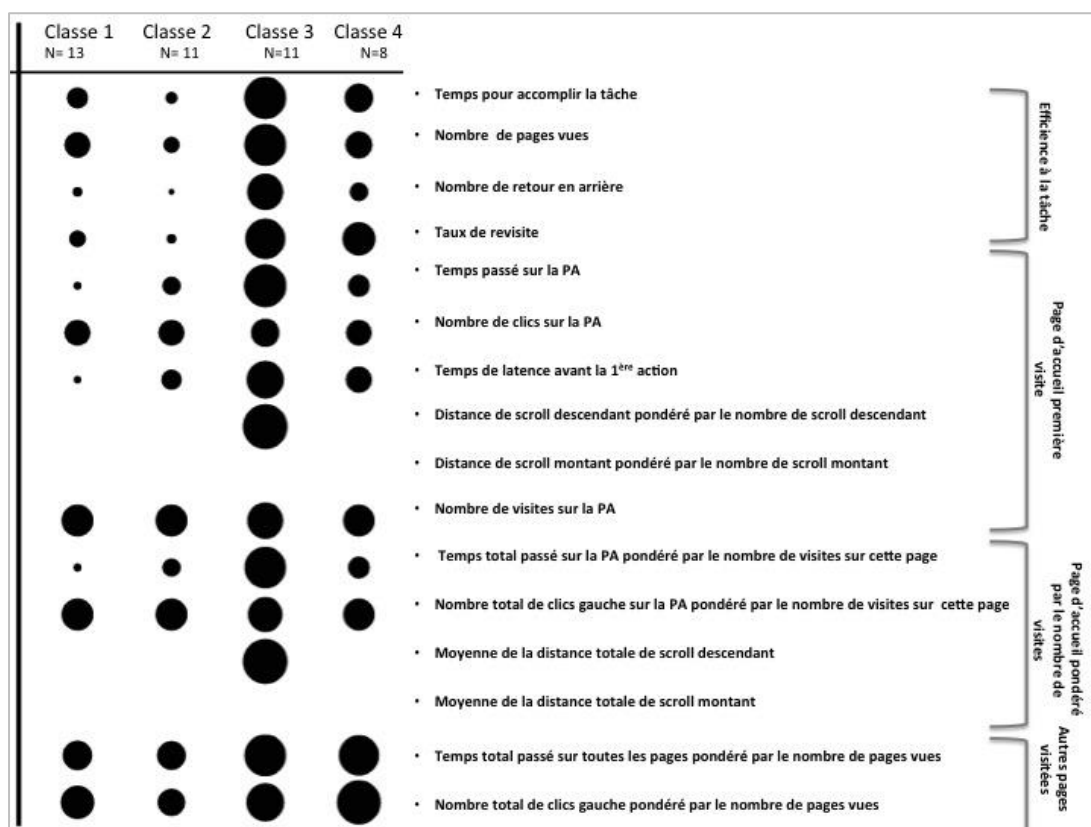


Figure 65 : Illustration des patterns pour la Tâche 5

Le pattern 1 rassemble les participants faisant parti des plus rapides à naviguer sur la page d'accueil :

- Même s'ils ont mis en moyenne **trois fois plus de temps (97secondes) que P2 à réaliser la tâche, ils sont beaucoup plus rapides que les participants du pattern 3 et 4.**
- Ils font également partis d'un des patterns ayant le moins recours au back et malgré un faible taux de revisite, tous les participants ont réussi la tâche.
- Cette rapidité de navigation amène cependant quelques individus à **rencontrer des difficultés d'orientation.**

Le second pattern présente une stratégie encore rapide et linéaire :

- Les participants vont **réaliser la tâche le plus plus rapidement.**
- Ils ont également le taux de revisite le plus faible, et avec 3 pages visitées en moyenne **ils sont ceux ayant la navigation la plus linéaire.** D'ailleurs, en moyenne avec le moins de clics effectués par pages visitées, ces individus sont **ceux ayant le moins exploré les contenus.**

- Toutefois, même si cette stratégie a déjà été identifiée et qu'elle était fructueuse pour les autres tâches, ici **tous les participants ont échoué à la tâche.**

Les individus du pattern 3 présentent aussi une stratégie déjà identifiée par ailleurs et se caractérisant par une importante désorientation et une exploration approfondie de la page d'accueil :

- En moyenne **20 pages visitées, plus de 3 back effectués et un taux de revisite de 32%**
- De façon générale, ces participants passent plus de temps que les autres à explorer le contenu des pages visitées.
- Leurs comportements laissent penser **qu'ils éprouvent des difficultés mais qu'ils persévèrent pour chercher l'information.** Cette stratégie s'avère efficace pour 36,7% d'entre eux qui accèdent à bonne la réponse.

Enfin, et dans une moindre mesure, le pattern 4 rassemble également des individus ayant éprouvés des difficultés d'orientation.

- Même si ces participants ont mis deux fois moins de temps que ceux de P3 a réalisé la tâche, ils ont aussi mis deux fois moins de temps sur la page d'accueil. Leur taux de revisite de 22% laisse percevoir qu'ils sont également désorientés.
- Avec cette stratégie de navigation, la moitié des participants a échoué à la tâche.

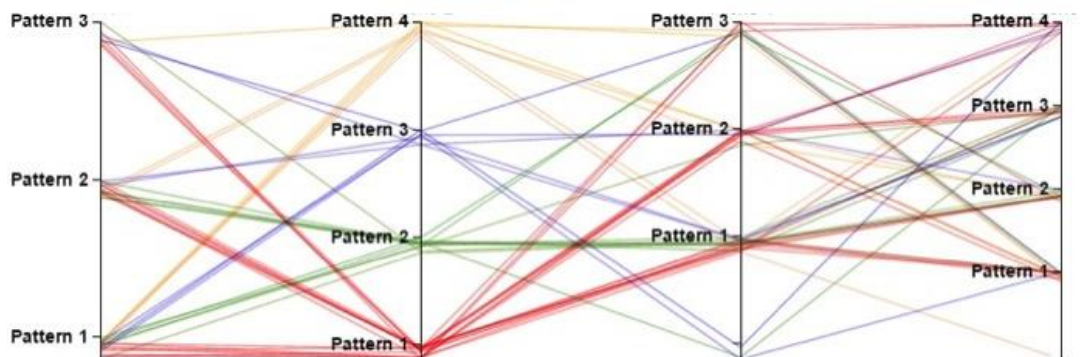
La classification des parcours a permis de classer les participants en deux groupes. La grande majorité des individus des patterns 3 et 4 ont emprunté les parcours de type 2 impliquant une partie de la navigation en dehors du site du Ministère.

Tableau 31 : Répartition des parcours pour chaque Pattern Tâche 5 – MEGA

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Parcours type 1	84,6% (n=11)	72,7% (n=8)	9% (n=1)	25% (n=2)
Parcours type 2	15,4% (n=2)	27,3% (n=3)	91% (n=10)	75%(n=6)
Nombre total de participants	100% (n=13)	100% (n=11)	100% (n=11)	100% (n=8)

Par ailleurs, les participants ayant réalisé les tâches dans un ordre aléatoire, nous ne présentons pas l'évolution de la stabilité d'une tâche à l'autre. Cependant, nous avons cherché à savoir si, de façon plus globale, pour l'ensemble de l'expérimentation, on retrouvait toujours les individus dans les mêmes patterns.





**Le nombre de patterns n'est pas stable d'une tâche à l'autre**, cependant nous identifions **deux patterns** qui ont toujours été récurrents :

- Parmi tous les participants ayant présenté une navigation désorientée impliquant une exploration approfondie des pages, aucun ne reste stable dans sa façon de naviguer pour les quatre tâches.
- Concernant le pattern décrit comme le plus rapide et linéaire, seul 2,6 % des participants restent cohérents sur cette façon de faire tout au long de l'étude.

Ce résultat invalide donc notre idée qui était celle de savoir si les patterns pouvaient s'expliquer à travers le style d'apprentissage de l'individu. En effet, les patterns n'étant pas toujours les mêmes d'une tâche à l'autre et les individus n'étant eux-mêmes pas stables dans leur façon de naviguer, **nous ne pouvons expliquer les patterns à travers un facteur interne tel que le style d'apprentissage. Les patterns ne sont donc pas propre aux individus.**

#### 4.3.4 Discussion

Dans cette étude, nous avons choisi un échantillon de participants plus hétérogènes afin de voir si nous identifions d'autres patterns.

Malgré des différences moins saillantes entre les patterns, notamment pour la tâche 2 et 5, nous retrouvons des fortes similitudes avec les 4 patterns identifiés précédemment :

- Une **stratégie de navigation rapide** avec laquelle les individus reviennent faiblement, voire pas du tout sur leurs traces et accèdent aisément à l'information. Leur **façon de naviguer est linéaire** et dans ce cas, il semblerait qu'ils sachent où aller.
- Une autre **stratégie de navigation rapide** amenant les individus à rapidement quitter la page d'accueil, et les pages visitées mais **impliquant une désorientation. Ces individus reviennent sur les pages précédemment vues.**

- Une stratégie de **navigation plus consciencieuse et exhaustive** impliquant une visite plus minutieuse sur la page d'accueil et les autres pages du site. Les individus reviennent très peu, voir pas du tout sur les pages vues précédemment.
- Une stratégie dans laquelle les individus passent aussi **beaucoup de temps sur les pages** mais avec une **navigation assez dispersée**, ils reviennent régulièrement sur leurs traces.

Concernant la stabilité, nous n'avons pas identifié le même nombre de patterns pour chaque tâche. Ce constat nous amène donc à penser que **sur un même site, en fonction de la tâche, certains individus ne naviguent pas toujours de la même façon**. Dans ce sens il serait pertinent de comprendre quels sont les facteurs inhérents à une tâche et qui peuvent influencer sur « l'adoption » d'un pattern ou non. Par exemple, il aurait été intéressant de faire passer un questionnaire afin d'évaluer la difficulté ressentie pour chaque scénarios. Peut-être que pour certaines tâches qui peuvent paraître plus facile, on ne retrouverait pas de stratégie de navigation exhaustive impliquant une visite minutieuse des pages visitées. C'est d'ailleurs ce qui pourrait expliquer l'absence de ce pattern pour la tâche 1. Ainsi, et comme nous l'avons évoqué plus haut, **les individus ne restent pas cohérents dans leur façon de naviguer**.

Par ailleurs, nous avons souhaité étudier le style d'apprentissage des individus afin de savoir si les patterns pouvaient s'expliquer par ce facteur. Compte tenu du fait que les gens ont tendance à ne pas garder la même stratégie de navigation, cette hypothèse est infirmée. Cependant, l'idée de prendre en compte les styles d'apprentissages nous avait paru judicieuse car nos participants réalisaient plusieurs tâches sur un même site, et cela supposait donc un apprentissage de la structure du site. Ce résultat confirme l'idée que les participants adoptent un style de navigation en fonction du contexte et non de leurs caractéristiques intrinsèques.

Un autre point faisant écho à notre précédente étude est également à souligner. En effet, nous avons vu qu'une même stratégie pouvait conduire des individus à l'échec ou à la réussite. Cependant, dans cette précédente expérimentation peu d'individus avaient échoué et cela ne portait que sur une tâche. Nos résultats montrent ici le même résultat. **Un style de navigation peut s'avérer efficace pour certains afin d'accéder à la réponse, alors qu'infructueux pour d'autres**. Même pour des stratégies qui au premier abord peuvent sembler efficiente et efficace car les individus ne semblent pas désorientés (navigation rapide, taux de revisite faible, peu de back et de pages visitées), on remarque des taux d'échecs élevés (Tâches 2 et 5). Dans ce cas, les participants pensaient donner la bonne réponse, mais il s'agissait d'une fausse réussite. Ce résultat va donc à l'encontre des travaux de Gwidzka et al.,(2007) expliquant que l'échec à la tâche peut être observé, notamment, à travers un style de navigation « inefficace » (beaucoup de pages visitées, de nombreuses revisites, et un temps important pour réaliser la tâche).

Par ailleurs, on aurait pu penser qu'une exploration plus minutieuse des pages du site et notamment de la page d'accueil à la première visite conduise plus souvent et facilement à une bonne performance, mais pour certains individus cela reflète plutôt une **difficulté à prélever l'information et à s'orienter**.

Néanmoins même si pour ces 4 tâches nous avons observé un pattern présentant 100% de réussite (pattern assez rapide pour lequel les gens reviennent faiblement voir pas du tout sur leurs traces), nous avons également pu voir que ce ne sont pas les mêmes personnes que l'on retrouve dedans tout au long de l'étude. Les individus adoptent donc une stratégie ou une autre en fonction des circonstances. Ce constat permet donc **difficilement d'imaginer d'émettre des recommandations pour l'apprentissage** en formant les individus à adopter cette stratégie plutôt qu'une autre car cela semble dépendre fortement du contexte.

# 5. DISCUSSION

---

## 5.1 SYNTHÈSE DE LA DEMARCHE

---

Cette thèse a eu pour objet d'étudier et de caractériser les comportements des internautes lors de leur navigation sur le Web. La littérature présente certaines études ayant déjà cherché à analyser ces interactions. En effet, à l'opposé d'études comme celle de Nielsen présentant un pattern d'exploration universel, certains chercheurs ont la volonté de prendre en compte la variabilité des comportements des internautes car aujourd'hui l'interface à taille unique ne peut plus satisfaire et le web ne peut rester « statique » en considérant tous les utilisateurs de la même manière. Ce type d'approche va à l'encontre de l'idée de design pour tous, impliquant que les caractéristiques du dispositif permettent qu'il soit utilisé par tous les types d'utilisateurs, indépendamment de leurs handicaps, de leurs compétences technologiques et de leurs expériences personnelles (Ruiz, Pajares, Utray, Moreno, 2011). Lorsque l'on sait que le nombre d'utilisateurs augmente chaque année, et que cela laisse place à une plus large variabilité en termes de compétences, besoins, et profils d'utilisateurs, le constat amène à s'intéresser à la prise en compte de la variabilité comportementale des individus.

La littérature met donc en exergue un certain nombre d'études qui ont cherché à découvrir la façon qu'ont les gens de se différencier lors de la navigation. Cependant, ces travaux sont souvent réalisés à un niveau macroscopique ne permettant pas vraiment de distinguer les gens entre eux et de décrire de façon précise la façon dont les usagers naviguent pour une tâche, un site et une page précise. Ce constat nous a amené à penser que le comportement des internautes devrait être analysé plus finement avant de formuler des recommandations. Le niveau d'analyse de la page permet justement une description fine des comportements, mais les études sont assez rares. De plus, la littérature présente souvent la page d'accueil comme une « clef de voute » déterminante pour expliquer la navigation sur les autres pages du site, ou l'abandon. L'un de nos objectifs a donc été d'étudier les comportements sur cette page au regard des styles de navigation.

D'un point de vue méthodologique, nous avons aussi constaté que la plupart de ces recherches sont réalisées dans le contexte naturel des personnes c'est-à-dire à leur domicile. Si dans cette approche de web mining, ce type d'environnement, familier aux individus, a l'avantage de ne pas les perturber, le contrôle de certains facteurs parasites devient dans ce cas plus difficile puisqu'on ne peut pas être certain que chaque sujet navigue sur le web de façon similaire. De cette façon, il semble compliqué de parler de standardisation des procédures. Ces recherches nous laissent également, parfois, avec un questionnement concernant les caractéristiques mêmes des

sites sur lesquels les individus sont amenés à naviguer. Pour cette raison, l'apport de la présente thèse est en partie méthodologique.

## **5.2 APPORTS ET LIMITES METHODOLOGIQUES**

---

Notre premier objectif a été d'identifier une méthode d'analyse permettant d'étudier la variabilité et de détecter des patterns comportementaux.

Une étude exploratoire a été réalisée afin de savoir, d'une part, si les indicateurs choisis permettaient de faire apparaître des styles ou patterns comportementaux lors de la consultation de la page d'accueil qui entraîneraient des niveaux de performance de navigation différents. Et, d'autre part, de découvrir la pertinence d'utiliser une méthode de classification. A l'issue de cette étude, trois patterns ont été découverts et nous avons développé de nouveaux outils d'analyse pour traiter les fichiers de données et affiner notre compréhension des comportements ayant lieu sur la page d'accueil.

Plus précisément, par rapport à l'étude exploratoire, les métriques d'efficacité (Temps total, nombre de pages vues, nombre de retour arrière et taux de revisite) n'ont pas changé. Concernant le comportement sur la page d'accueil à la première visite, l'indicateur temps de latence a été rajouté. En effet, il paraissait intéressant de pouvoir différencier les individus réalisant rapidement une première action sur le site, de ceux prenant plus de temps à explorer la page avant d'effectuer un clic ou un scroll. Dans le cadre de l'étude exploratoire, l'indicateur nombre de scroll montant et descendant a été pris en compte. Toutefois comme Buscher, White, Dumais et Huang (2012) l'ont fait sur la page de résultat de recherche Bing, la notion de distance est ici plus précise. Nous avons donc calculé un nouvel indicateur de scroll permettant de savoir quelle est la distance moyenne de scroll effectué par participant. Comme nous l'avons expliqué dans la méthodologie, cela a permis de différencier les individus réalisant de plus petites distances de scroll de ceux réalisant des distances moyennes de scroll beaucoup plus longue. Cet indicateur précisait le comportement de scroll des participants, et nous renseignait aussi si le participant avait scrollé sur une faible ou plus grande proportion de la page. Les indicateurs nombre de clic gauche et temps de visite sur la page à la première visite sont restés inchangés. De plus, afin de mieux décrire le comportement de revisite sur la page d'accueil, nous avons souhaité rajouter l'indicateur nombre de visites et distance totale de scroll montant et descendant. En mettant en évidence si les participants scrollaient plus qu'à leur première visite sur la page d'accueil, ces métriques nous ont permis de mieux expliquer les patterns identifiés. Enfin, deux autres indicateurs ont été rajouté afin d'avoir des données sur le comportement de navigation des participants sur les autres pages visitées : le temps total passé sur toutes les pages et le nombre total de clics gauches effectués en fonction du nombre de pages vues.

De plus, nous avons souhaité combiner une classification basée sur des métriques décrivant les comportements sur le web, et une classification reposant sur les distances de Levenstein inter-participants pour les parcours de navigation. À partir de ces résultats, il a été possible de mettre en perspective pour chacun des patterns, les parcours les plus empruntés et de distinguer les classes pour lesquelles les chemins de navigation semblaient les plus optimaux. Bien que dans la littérature Catledge et

Pitkow (1995), White et Drucker (2007) avaient déjà utilisé la LD pour définir des stratégies de navigation, les résultats n'avaient pas été croisés avec tant d'indicateurs comportementaux. A ce niveau, il aurait également pu être possible d'approfondir les analyses concernant le parcours de navigation. A l'instar d'étude comme celle de Herder et Juvina (2004), de futures recherches pourraient porter leur attention sur le nombre de liens suivis à partir d'une même page, ou encore la distance entre les pages.

Enfin, la méthode de classification en elle-même a été optimisée. En effet, dans l'étude exploratoire, la classification ascendante hiérarchique réalisée, demandait de découper le dendrogramme en se basant sur la chute des classes. Toutefois, cette méthode très répandue pour la classification ne permet pas de savoir si le nombre de classe identifié est optimal (Bittner, Meltzer et Trent, 1999) et cela constituait une des limites à cette première recherche. Une nouvelle méthode de classification a donc été choisie, la « Dynamic Hybrid Cut » permettant de détecter automatiquement le nombre de classes optimales. Par ailleurs, cette méthode ayant été testé au regard de 7 autres techniques de classification a été jugée comme la plus discriminante et robuste (Langfelder, Zhang et Horvath, 2008).

Au total, ce sont donc 16 indicateurs qui ont été analysées. Il aurait pu être envisagé, de réaliser une analyse factorielle sur ces métriques afin de réduire leur nombre et d'en dégager les facteurs principaux. On peut aisément supposer que le temps passé sur la page d'accueil soit corrélé avec la distance de scroll, par exemple. Toutefois, cette approche aurait été à l'encontre de notre objectif qui était celui d'être le plus précis possible dans la distinction des stratégies.

Par ailleurs, nous aurions pu penser à d'autres indicateurs qui auraient pu nous renseigner d'avantage sur le comportement d'interaction avec la première page vue. Par exemple, quelle est la première action réalisée sur la page (clic ou scroll), sur quel élément du site a lieu le clic, quel est le temps entre chaque action, ou encore quelle est la dynamique des comportements (première action ; clic, seconde action ; scroll, etc.). Nous nous sommes basés sur les études citées dans la littérature pour définir nos métriques, et nous avons identifié celles qui nous semblaient les plus pertinentes, mais finalement à quel degré de granularité faut-il s'arrêter ? Il serait important de poursuivre les travaux dans ce sens afin de savoir si certaines métriques sont plus pertinentes que d'autres. La diversité des métriques étudiées complexifie également la comparaison avec les autres stratégies étudiées dans la littérature. L'outil Tobii ne nous permettait pas de capturer les comportements avec la souris (mouvement, distance et vitesse), nous avons donc fait le choix de ne pas porter notre attention sur cette métrique. La question des données prises en compte en entrée de la classification et plus généralement les données prises en compte pour toute analyse de la navigation est donc à se poser.

Concernant la méthodologie, un autre point important concerne la caractérisation des classes. Nous avons choisi de présenter une visualisation d'indicateurs via l'application Bertifier, et cette technique nous a paru très efficace pour distinguer visuellement les patterns entre eux, et les expliquer. Toutefois, pour visualiser la dimension temporelle et la dynamique des interactions, il serait judicieux d'identifier de nouveaux outils et que des travaux dédiés à ce sujet soient menés.

Enfin, notre protocole était effectué dans un contexte de laboratoire avec des conditions strictement contrôlées. L'objectif était de standardiser les passations pour pouvoir comparer les gens entre eux. Mais, comme Byrne (1999) le précise, il est possible que les tâches étudiées ne puissent pas parfaitement refléter ce que les sujets font quand on leur laisse prendre des décisions propres. D'ailleurs, il est probable que dans notre dernière étude, certains individus ne se soient pas sentis libre d'abandonner la tâche quand il le souhaitait. Ce facteur n'est donc pas à négliger. Concernant les tâches et selon la classification de Broder (2002), nos protocoles se sont centrés sur des tâches « informationnelles ». Il serait intéressant de faire varier le type de tâches afin de voir si l'on retrouve les mêmes patterns.

Au final, notre méthode a été strictement appliquée sur plusieurs expérimentations ce qui nous a permis de reproduire les résultats. Mais malgré les précautions méthodologiques employées, et en l'absence de théorie sur le sujet, comment être certains que les classes obtenues ne soient pas de simples artefacts.

## **5.3 LES PATTERNS IDENTIFIES**

---

### **5.3.1 Description des patterns**

À partir de la méthodologie citée ci-dessus, notre second objectif était d'identifier des patterns comportementaux et d'étudier leur stabilité. Pour se faire, nous avons mené 3 expérimentations. Nous avons identifié pour chacune de ces études 3 ou 4 patterns distincts qui varient en fonction de la tâche et/ou du site.

Dans la première étude, 112 participants devaient réaliser 2 tâches sur 2 sites Web en les répétant immédiatement. Nous avons identifié au total 4 patterns distincts pour les tâches initiales, et 3 pour les tâches répétées. En effet, après la répétition d'une tâche à court terme, on observe une réduction du nombre de pattern. Ici, la répétition à tendance à faire varier à la baisse certains indicateurs (taux de revisite et back), et l'un des patterns décrit comme rapide mais désorienté n'est plus identifié.

Dans la deuxième expérimentation, la procédure expérimentale était similaire, mais 27 participants répétaient les tâches à 2 jours d'intervalle. Là aussi, nous avons identifié quatre patterns de structure proche qui varient en fonction du site ou de 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 avons souhaité reproduire nos analyses sur une population moins homogène et nous avons retrouvé quatre patterns ressemblant en grande partie à ceux identifiés par ailleurs.

Nous avons donc identifié sur l'ensemble des expérimentations 4 patterns distincts :

- Un pattern reflétant une façon **consciencieuse, exhaustive et linéaire** de naviguer.

Ces individus mettent généralement plus de temps avant de réaliser une première action sur la page d'accueil et pour en explorer le contenu (clic et

scroll quand nécessaire). On observe également cette même façon de faire sur les autres pages visitées. De façon générale, ces participants reviennent faiblement (ou pas du tout) sur leurs traces. Ils éprouvent très peu de difficultés pour s'orienter et empruntent souvent les chemins optimaux. Toutefois, il se peut que ces individus ne soient pas toujours les plus rapides pour finir la tâche.

- Le second pattern présente des similarités avec celui décrit juste au-dessus. Ces personnes prennent du temps pour naviguer sur les pages visitées (y compris la page d'accueil), **mais malgré cette exploration consciencieuse ils éprouvent ici des difficultés importantes pour s'orienter**. Leur taux de revisite, le nombre de retour en arrière et le chemin emprunté montrent qu'ils se dispersent en cherchant l'information. Ils sont souvent ceux qui mettent le plus de temps pour finir la tâche. Sur certains points, ce type de navigation ressemble aux « fragiles » de Herder et Juvina (2000) et à la navigation « inefficace » de Gwidkzka et Spence (2007).
- Un **pattern rapide et linéaire** dans lequel les individus passent peu de temps sur les pages visitées, y compris la page d'accueil, réalisent très vite une action sur cette page, revisitent très peu de pages et ce sont eux qui de façon générale passent le moins de temps sur les pages visitées. Ils explorent peu le contenu des pages (clics et scrolls). Dans ce cas, ces individus savent précisément où ils veulent aller.
- Un pattern dans lequel les **gens naviguent relativement vite** sur les pages visitées sans prendre le temps de les explorer en détail. Mais, ces individus reviennent sur leurs traces (y compris sur la page d'accueil), **ils paraissent se perdre et rencontrer des difficultés d'orientation**. Leur stratégie paraît infructueuse car ces personnes veulent rapidement effectuer la tâche, mais en explorant peu les pages elles sont obligées d'y revenir afin de collecter les informations manquantes. Ces personnes visitent et revisitent donc plus de pages.

Par ailleurs, cette volonté de naviguer rapidement et d'explorer faiblement les pages les amènent à emprunter le parcours le plus fastidieux. Il est possible que les « téléportés » décrits par Teevan et al. (2004) et les « économiques » décrits par Buscher et al. (2012) correspondent à ce pattern. À la différence de ces auteurs, notre expérimentation ne permettait pas aux participants de recourir à une navigation directe (l'utilisation du moteur de recherche était défendue). Dans ce cas, il semblerait qu'une page qui ne permet pas d'être accessible par un moteur de recherche performant (page Web non référencée, intranets, etc.) rendrait leur style de navigation inadapté. Par ailleurs, ce pattern présente d'autres points communs avec les « explorateurs » de White et Drucker (2007) décrits comme des utilisateurs faisant fréquemment des retours en arrière, visitant de nombreuses pages et ne parvenant pas directement à leurs objectifs.



### 5.3.2

## Performance des patterns

Au regard de ces patterns, plusieurs points concernant la notion de performance sont à préciser.

L'idée que nous avons et qui suggérait que l'interaction avec la page d'accueil influe sur les performances de navigation ne peut pas être affirmée. En effet, si l'on regarde la stratégie de navigation « d'exploration exhaustive » amenant les individus à prendre un certain temps sur les pages visitées, y compris la page d'accueil, il semblerait qu'en quittant cette première page certains usagers finissent par éprouver des difficultés à s'orienter, tandis que pour d'autres on observe une navigation plus linéaire et directe. Pour certains, ce temps d'exploration serait donc plutôt synonyme d'une difficulté à prélever des informations pertinentes les amenant à avoir une navigation dispersée.

Indépendamment de la tâche ou du site, il semblerait donc, qu'en fonction des individus, une stratégie de navigation impliquant une exploration exhaustive de la page d'accueil, et des autres pages du site, peut être fructueuse pour certains et synonyme de difficultés pour d'autres. A l'heure on l'on entend parler de plus en plus des systèmes de navigation adaptatifs qui se destinent à analyser le comportement de navigation sur une page afin d'adapter le reste du site en conséquence, il paraît important de mettre en garde les concepteurs et d'encourager d'autres travaux dans ce sens.

Par ailleurs, pour l'expérimentation 1 et 2, malgré certaines difficultés pour s'orienter, la quasi-totalité des participants ont toujours réussi à accéder à la bonne réponse. Au contraire, pour la dernière étude, nous observons que ces mêmes stratégies peuvent aussi être synonymes d'échecs. Une façon similaire de naviguer, et qui en apparence ne semble montrer aucune difficulté d'orientation, peut donc conduire à un échec comme à une réussite.

A travers ces résultats et à l'inverse de Gwidzka et Spence (2007), nous pensons qu'il n'est pas possible de définir les caractéristiques propres à la réussite ou l'échec d'une tâche. La réussite ou l'échec peuvent être atteint à travers des cheminements variables. En fonction des individus et du contexte, une même façon de naviguer peut conduire à une performance différente. Ce constat permet donc difficilement d'émettre des recommandations pour l'apprentissage à l'exploration des pages Web, ni pour une conception des pages Web qui favoriserait l'adoption d'un pattern plutôt qu'un autre.

### 5.3.3

## Stabilité des patterns

Pour chacune de nos expérimentations, nous avons souhaité évaluer la propension des individus à adopter le même pattern entre différentes sessions de navigation. La stabilité des patterns n'avait jamais, à notre connaissance, été étudiée. Un outil de visualisation sous forme de coordonnées parallèle a donc été développé.

Le fait que l'on ne retrouve pas toujours les mêmes patterns d'une tâche à l'autre, et qu'en fonction de la tâche le nombre varie, nous ne pouvons conclure à une stabilité des patterns.

Un même individu peut emprunter des patterns différents entre plusieurs tâches comme pour une même tâche, et ce, quelque soit la temporalité envisagée.

Ce résultat ne plaide donc pas en faveur d'une explication par des variables intrinsèques aux individus, tels que le style d'apprentissage, mais plutôt d'un contexte particulier favorisant l'adoption d'un pattern plutôt qu'un autre.

Afin de poursuivre ces travaux, il serait intéressant de réaliser des expériences qui font varier le contexte des participants, mais aussi les caractéristiques des tâches comme par exemple la difficulté ressentie.

## **5.4 PERSPECTIVES**

---

Nos travaux de recherches nous ont amené à pointer plusieurs limites méthodologiques citées plus haut et qui demanderaient à être approfondies dans le cadre de futurs travaux.

L'idée que nous avons et qui était de savoir s'il existe des patterns a été validée. À travers 3 études portant sur des échantillons différents, les individus se sont distingués selon 4 stratégies qui varient en fonction des tâches, des sites, et de la temporalité. Nous nous sommes aperçus que les individus ne se cantonnent pas à une seule et même façon de naviguer. Cette variabilité comportementale s'explique, notamment, par le contexte de la tâche. Toutefois, aujourd'hui, nous ne sommes pas en mesure d'en préciser les facteurs. Cette conclusion démontre donc toute la complexité qu'est celle de conceptualiser des recommandations pour un style de navigation en particulier. Il est donc difficilement envisageable de privilégier la conception d'un site ou d'une page Web en fonction d'un pattern et pas d'un autre. L'idée serait d'identifier des choix de conception « transversaux » qui permettraient d'aboutir à un bon compromis afin d'optimiser la navigation de ces 4 patterns.

Par ailleurs, depuis les années 2000 (Brusilovsky, 2001), la littérature montre un certain engouement pour les interfaces « adaptatives ». Ce type de système a pour objectif de s'adapter dynamiquement aux préférences de l'utilisateur en analysant sa navigation sur une ou plusieurs pages et, en conséquence, de réorganiser la structure des pages du site. Nos travaux nous amènent à croire que ce type d'environnement est à considérer avec précaution. Au regard de nos travaux, nous avons constaté, par exemple, que deux individus peuvent explorer et naviguer de la même façon sur la première page d'un site, mais cela ne veut pas dire pour autant qu'ils accéderont à l'information souhaitée avec le même cheminement.

Par ailleurs, les patterns ne semblant pas propres aux individus, ni aux tâches, il serait pertinent que de futurs travaux s'interrogent sur la façon dont il faut faire le lien entre les performances observées et les modèles de recherche d'information.

Enfin, il nous semblerait très intéressant de croiser ces résultats avec des données oculométriques. En effet, dans cette recherche, nous avons porté notre attention sur la page d'accueil car elle constituait le point de départ de nos tâches et elle est souvent citée comme la page centrale d'un site. Nous avons donc souvent parlé de l'exploration de cette page, mais quand est-il vraiment ? Sur quels éléments de la page les individus portent leur attention ? Il est possible que ces analyses complètent l'interprétation de nos patterns.

# BIBLIOGRAPHIE

---

## *A.*

- Abrams, D., Baecker, R., & Chignell, M. (1998). Information archiving with bookmarks: personal Web space construction and organization. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp.41–48). Los Angeles, USA: ACM
- Ackerman, M. S., and Halverson, C. (1998). Considering an organization's memory. *In Proceedings of the CSCW Conference* (pp.39-48). New York, USA: ACM
- Adar, E., Teevan, J., & Dumais, S. T. (2008). Large scale analysis of web revisitation patterns. *In Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems* (pp.1197-1206). Florence, Italy: ACM
- Adar, E., Teevan, J., & Dumais, S. T. (2009). Resonance on the web: web dynamics and revisitation patterns. *In Proceedings of the 27th international conference on Human factors in computing systems* (pp. 1381–1390). Boston, USA: ACM
- Albert, B., Tullis, T., & Tedesco, D. (2010). *Beyond the usability lab*. Amsterdam : Morgan Kaufmann (Eds), *L'ergonomie des documents électroniques*. Paris : PUF
- Anderson, J.R. (1990). The adaptative character of thought. Hillsdale, NJ: Erlbaum.
- Aula, A. (2003). Query formulation in information search. *In Proceedings of the IADIS International Conference WWW/Internet* (pp. 403–410). Lisboa, Portugal: IADIS Press
- Aula, A. (2004). Enhancing the readability of search result summaries. *In Proceedings of the Conference HCI: Design for Life* (pp. 1–4). Bristol, UK: Research Press International
- Aula, A., Jhaveri, N., & Käki, M. (2005). Information search and re-access strategies of experienced web users. *In Proceedings of the 14th International World Wide Web Conference*. (pp 583-592), Chiba, Japon : ACM
- Aula, A., & Norhaussen, K. (2006). Modeling Successful Performance in Web searching. *International Review of Research in Open and Distance Learning*, 57(12), 1678–1693

## *B.*

- Bastien, J. M. C., Drusch, G., & Dinet, J. (2010). Connaître les comportements des internautes : méthodes et outils. In L. Calderan, B. Hidoine & J. 168 Millet (Eds.), *L'usager numérique*. Séminaire INRIA (27 septembre – 1er octobre 2010 – Anglet) (pp. 39-62). Paris : ADBS
- Beauvisage, T. (2004). *Sémantique des parcours des utilisateurs sur le Web*. Thèse de doctorat, Université de Paris X, Nanterre
- Bevan, N. (2003). UsabilityNet methods for user centred design. *Human-Computer Interaction: Theory and Practice*, 1(1), 434–438

- Belk, M., Papatheocharous, E., Germanakos, P., & Samaras, G. (2013). Modeling users on the World Wide Web based on cognitive factors, navigation behavior and clustering techniques. *Journal of Systems and Software*, 86(12), 2995–3012
- Berendt, B., & Spiliopoulou, M. (2000). Analysis of navigation behaviour in web sites integrating multiple information systems. *The VLDB journal*, 9(1), 56–75
- Bernard, M., Baker, K. R., & Fernandez, M. (2002). Paging vs. scrolling: Looking for the best way to present search results. *Usability News*, 4(1)
- Bittner, M., Meltzer, P., & Trent, J. (1999). Data analysis and integration: of steps and arrows. *Nature Genetics*, 22(3), 213–215
- Blackmon, M.H., Mandalia, D.R., Polson, P.G., & Kitajima, M. (2007). Automating Usability Evaluation: Cognitive Walkthrough for the Web Puts LSA to Work on Real-World HCI Design Problems. In T.K. Landauer, D.S. McNamara, S. Dennis & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis* (pp. 345–375). Lawrence Erlbaum Associates
- Bousbia, N., Rebaï, I., Labat, J.-M., & Balla, A. (2010). Learners' navigation behavior identification based on trace analysis. *User Modeling and User-Adapted Interaction*, 20(5), 455–494
- Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novice: Analysis of a complex cognitive skill. *Computers in Human Behavior*, 21, 487–508.
- Brusilovsky, P., 2001. Adaptive hypermedia. *User Modeling and User-Adapted Interaction* 11 (1–2), 87–110
- Brusilovsky, P., Millán, E., 2007. User models for adaptive hypermedia and adaptive educational systems. In: Kobsa, A., Nejd, W. (Eds.), *the Adaptive Web*. LNCS, vol. 4321. Springer, Heidelberg, pp. 3–53
- Bucklin, R. E., & Sismeiro, C. (2003). A Model of Web Site Browsing Behavior Estimated on Clickstream Data. *Journal of Marketing Research*, 40(3), 249–267.
- Buscher, G., White, R. W., Dumais, S., & Huang, J. (2012). Large-scale analysis of individual and task differences in search result page examination strategies. In *Proceedings of the fifth ACM international conference on Web search and data mining*, (pp 373-382). Seattle, USA: ACM Press
- Broder, A. (2002). A taxonomy of web search. *ACM Sigir Forum*, 36(2), 3-10
- Byrne, M. D., John, B. E., Wehrle, N. S., and Crow, D. C. (Eds). (1999). The tangled web we wove: A taskonomy of WWW use. In *Proceedings of SIGCHI Human Factors in Computing Systems*, (pp.544-551). Pittsburg, USA: ACM Press
- Bidel, S., Lemoine, L., Piat, F., Artières, T., & Gallinari, P. (2003). Statistical machine learning for tracking hypermedia user behavior. In *Proceedings of the 2<sup>nd</sup> workshop on Machine Learning, Information Retrieval and User Modeling*, (pp. 56-65): ACM Press

## C.

- Calisir, F., & Gurel, Z. (2003). Influence of text structure and prior knowledge of the learner on reading comprehension, browsing and perceived control. *Computers in Human Behavior*, 19(2), 135–145
- Card, S. K., Pirolli, P., Wege, M. V. D., Morrison, J. B., Reeder, R. W., Schraedley, P. K., & Boshart, J. (2001). Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability. *In proceedings of SIGCHI* (pp. 498–505). Seattle, USA: ACM Press
- Catledge, L., & Pitkow, J. (1995). Characterizing browsing strategies in the World-Wide Web. *Computer Networks and ISDN systems*, 27(6), 1065–1073
- Chevalier, A., & Kicka, M. (2006). Web designers and web users: Influence of the ergonomic quality of the web site on the information search. *International Journal of Human Computer Studies*, 64(10), 1031–1048
- Chen, K., & Liu, L. (2003). Validating and Refining Clusters via Visual Rendering. *In Proceedings of the 3rd International Conference on Data Mining* (pp. 501–504). Melbourne, Florida: IEEE
- Chen, S. Y., & Macredie, R. (2010). Web-based interaction: A review of three important human factors. *International Journal of Information Management*, 30(5), 379–387
- Chi, E. H., Rosien, A., & Heer, J. (2003). LumberJack : Intelligent Discovery and Analysis of Web User Traffic Composition. *Mining Web Data for Discovering Usage Patterns and Profiles*, 27(03), 1–16.
- Chin, J., Fu, W., & Kannampallil, T. Adaptive information search: Age-dependent interactions between cognitive profiles and strategies. *In Proceedings of CHI'09* (pp. 1683-1692). Boston, USA: ACM Press
- Chin, J., & Fu, W.T. (2010). Interactive effects of age and interface differences on search strategies and performance. *In Proceedings of the 28th international conference on Human Factors in computing systems* (pp. 403-412). Atlanta, USA: ACM Press
- Cockburn, A., & McKenzie, B. (2001). What do web users do? An empirical analysis of web use. *International Journal of Human-Computer Studies*, 54(6), 903–922
- Cockburn, A., McKenzie, B., & JasonSmith, M. (2002). Pushing back: evaluating a new behaviour for the back and forward buttons in web browsers. *International Journal of Human-Computer Studies*, 57(5), 397–414
- Cohen, C. (1965) Maximum likelihood estimation in the weibull distribution based on complete and on censored samples. *Technometrics*, 7(4), 579–588, 1965
- Cooley, R., Mobasher, B. & Srivastava, J. Web Mining: Information and Pattern Discovery on the World Wide Web (1997). *In Proceedings of the International Conference on Tools in Artificial Intelligence*. (pp. 558-567)
- Cooley, R., Mobasher, B., & Srivastava J. (1999). Data preparation for mining World Wide Web browsing patterns. *Journal of Knowledge and Information Systems*, 1(1), 1-26

## **D.**

- Danielson, D. R. (2002). Web navigation and the behavioral effects of constantly visible site maps. *Interacting with Computers*, 14(5), 601–618
- Dinet, J., & Tricot, A. (2008). Recherche d'information dans les documents électroniques. In A. Chevalier & A. Tricot (Eds.), *L'ergonomie des documents électroniques*. Paris : PUF
- Dinet, J., Vivian, R., La Mantia, K., Kitajima, M., & Bertolo, D. (2013). How users explore homepages? The impact of culture on preference and visual exploration. In *Proceedings of the 7<sup>th</sup> International Conference on Digital Society*. (pp 81-86). Nice, France
- 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
- 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 *Proceedings of Ergo IHM*. Biarritz, France
- Drusch, G., Bastien, J.M.C, Becker, M., Rodriguez, P. (2012). Evalyzer : un outil d'évaluation à distance de l'expérience utilisateur. In *Proceedings of Ergo ErgoIHM'2012*. Biarritz, France
- Dubroy, P., & Balakrishnan, R. (2010). A study of tabbed browsing among mozilla firefox users. In *Proceedings of the 28th international conference on Human factors in computing systems CHI*. (pp.673-682). Atlanta, USA: ACM Press

## **E.**

- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215-251.

## **F.**

- Fang, X., Hu, P., Chau, M., Hu, H., Yang, Z., & Sheng, O. (2012). A data-driven approach to measure website navigability. *Manage Information System*, 29(2), 173–212
- Ford, N., Wood, F., & Walsh, C. (1994). Cognitive styles and searching. *Online and CDROM Review*, 18(2), 79-86
- Ford, N., & Chen, S. Y. (2000). Individual differences, hypermedia navigation and learning: an empirical study. *Journal of Educational Multimedia and Hypermedia*, 9(4), 281–312
- Ford, N. J., Wilson, T. D., Foster, A. E., Ellis, D., & Spink, A. (2002). Cognitive styles in information seeking analysis. *Journal of the American Society for Information Science and Technology*, 53(9), 728-735
- Foss, C.L. (1988). Effective browsing in hypertexts systems. In *Proceedings of RAIO Conference: User-oriented content based text and image handling*. Cambridge, UK

Fu, W., & Pirolli, P. (2007). SNIF-ACTA Cognitive Model of User Navigation on the World Wide Web. *Human-Computer Interaction*, 32, 198–208.

## G.

Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction*, 21(2), 220–231

Greenberg, S., & Cockburn, A. (1999). Getting back to back: Alternate behaviors for a Web browser's back button. In *Proceedings of the 5th Annual Human Factors and the Web Conference*. (pp. 1–7). Gaithersburg, USA

Guthrie, J.T. and Mosenthal, P. (1987). Literacy as multidimensional: locating information and reading comprehension. *Educational Psychologist*, 22, 279-297

Guthrie, J.T. (1988). Locating information in documents: examination of a cognitive model. *Reading Research Quarterly*, 23, 178-199

Guthrie, J.T., Briten, T., & Barker, K.G. (1991). Roles of document structure, cognitive strategy and awareness in searching for information. *Reading Research Quarterly*, 26, 300-321

Gwizdka, J., & Spence, I. (2007). Implicit measures of lostness and success in web navigation. *Interacting with Computers*, 19(3), 357–369

## H.

Habuchi, Y., Takeuchi, H., & Kitajima, M. (2007). Where Do Web Visitors Look and at What ? An Eye-Tracking Study. In *Proceedings of HCI International 2007*. (pp. 1084–1088)

Hawala, E., Murphy, E., Hawala, S., & Kathleen, A. (2010). Think-aloud protocols: a comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. In *Proceedings of CHI 2010*. (pp.2381–2390). Atlanta, USA: ACM Press

Hölscher, C., & Strube, G. (2000). Web search behavior of Internet experts and newbies. *Computer Networks*, 33(1), 337–346.

Hsieh-Yee, I. (2001). Research on Web search behavior. *Library & Information Science Research*, 23(2), 167–185.

Huang, A. H. (2003). Effects of multimedia on document browsing and navigation: an exploratory empirical investigation. *Information & Management*, 41(2), 189–198

Huang, J., & White, R. W. (2010). Parallel browsing behavior on the web. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia* (pp. 13–18). Toronto, Canada: ACM Press

Huang, J., & White, R. (2011). No clicks, no problem: using cursor movements to understand and improve search. In *proceedings of CHI 2011*. (pp.1225–1234). Vancouver, Canada: ACM Press



Huang, J., White, R. W., Buscher, G., & Wang, K. (2012). Improving searcher models using mouse cursor activity. *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '12*. (pp.195–204). Portland, USA

## *J.*

Jamet, E., Lemerrier, C., & Février, F. (2007). Les méthodes off-line. In A. Chevalier & A. Tricot (Eds.), *Ergonomie des documents électroniques*. Paris : Puf.

Jamet, E., & Février, F. (2008). Utilisabilité, utilité et acceptabilité des nouvelles technologies dans l'entreprise : Une approche de psychologie ergonomique. *Analyser les usages des systèmes d'information et des TIC*, 18–41

Jansen, B. J., Booth, D., and Spink, A. (2008) Determining the informational, navigational, and transactional intent of Web queries. *Information Processing & Management*. 44(3), 1251-1266

Jansen, B. J., Booth, D. L., & Spink, A. (2009). Patterns of query modification during Web searching. *Journal of the American Society for Information Science and Technology*. 60(7), 1358-1371.

Jiang, Q., Tan, C.-H., Phang, C. W., & Wei, K. K. (2012). Using Sequence Analysis to Classify Web Usage Patterns across Websites. *In proceedings of 45th Hawaii International Conference on System Sciences* (pp. 3600–3609). Hawaii, USA

## *K.*

Karavidas, M., Lim, N. K., & Katsikas, S. L. (2004). The effects of computers on older adult users. Poster session presented at the annual meeting of the American Psychological Association, Honolulu

Kellar, M., Watters, C., & Shepherd, M. (2007). A Goal-based Classification of Web Information Tasks. *Proceedings of the American Society for Information Science and Technology*, 43(1), 1–22

Kellar, M., Watters, C., & Shepherd, M. (2006). The Impact of Task on the Usage of Web Browser Navigation Mechanisms. *In Proceedings of Graphics Interface*. (pp. 235-242). Quebec City, Canada

Kim, K. S. and Allen, B. (2002). Cognitive and task influences on Web searching. *Journal of the American Society for Information Science Technology*, 52(2), 109-119

Kim, H., & Fesenmaier, D. R. (2008). Persuasive Design of Destination Web Sites: An Analysis of First Impression. *Journal of Travel Research OnlineFirst*, 47 (1), 3-13

Kitajima, M. (2003). Comprehension-Based Approach to HCI for Designing Interaction in Information Space. *In Proceedings of HCI International, Human-Centred Computing: Cognitive, Social and Ergonomic Aspects*. (pp. 1031-1035). Crete, Greece

- Kitajima, M., Blackmon, M.H. & Polson, P.G. (2005). Cognitive Architecture for Website Design and Usability Evaluation: Comprehension and Information Scent in Performing by Exploration. *In Proceedings of HCI International 2005*
- Kitajima, M. (2008). Vers un modèle de recherche d'information basé sur la compréhension. In J. Dinet (Ed.), *Usages, usagers et compétences informationnelles au 21e siècle* (pp.285-310). Paris: Hermes Science Publications
- Kaufman, L. & Rousseeuw, P. (2005). Finding groups in data: an introduction to cluster analysis. Wiley
- Kellar, M., C. Watters, & K. M. An exploration of Web-based monitoring: Implications for design. *In Proceedings of CHI'07*. (pp.377-286).
- Khoury, R., Dawborn, T., & Huang, W. (2011). Visualising web browsing data for user behaviour analysis. *In Proceedings of the 23rd Australian Computer-Human Interaction Conference on - OzCHI '11* (pp. 177–180). San José, USA: ACM Press
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. New Jersey: PrenticeHall
- Kolb, D. A. (1976). On management and the learning process. *California Management Review*, 18(3), 21-31

## *L.*

- Langfelder P., Zhang, B., & Horvath., S (2007). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics*, 24(5), 719-720
- Large, A., Beheshti, J., & Rahman, T. (2002). Gender differences in collaborative web searching behavior: an elementary school study. *Information Processing & Management*, 38, 427–443
- Lazonder, A.W., Biemans, H.J.A., & Worpeis, I.G.J.H. (2000). Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science*, 51, 576–581
- Lee, D., McCool, J., & Napieralski, L. (2000). Assessing adult learning preferences using the analytic hierarchy process. *International Journal of Life long Learning*, 19(3), 548–560
- Le Guel, F. (2004). *Analyse économique du comportement des internautes : Mesure, adoption et usages*, Thèse, Université de Rennes 1, Bretagne.
- Levensthein, V. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10(8), 707-710
- Lin, H., & Chen, L. (2008). Discovering learning pattern in different cognitive style of learners. *Convergence and Hybrid Information Technology*, 2, 268–273
- Ling, J., & van Schaik, P. (2006). The influence of font type and line length on visual search and information retrieval in web pages. *International Journal of Human Computer Studies*, 64(5), 395-404
- 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

- Liu, J., & Zhang, S. (2004). Characterizing Web usage regularities with information foraging agents. In *Proceedings of IEEE Transactions on Knowledge and Data Engineering*. (Pp.566-584)
- Liu, C., White, R.W., & Dumais, S. (2010). Understanding web browsing behaviors through Weibull analysis of dwell time. In *33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2010)*. (pp 379-386). Genève, Suisse
- Liu, W., Guo, F., Ye, G., & Liang, X. (2016). How homepage aesthetic design influences users' satisfaction: Evidence from China. *Displays*, 42, 25–35.
- Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F., & Pan, B. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*, 59(7), 1041–1052
- Lynch, P. J., & Horton, S. (2002). *Web style guide: Basic design principles for creating web sites*. New Haven, CT: Yale University Press

## **M.**

- Milic-Frayling, N., Jones, R., Rodden, K., Smyth, G., Blackwell, A., & Sommerer, R. (2004). Smartback: supporting users in back navigation. *Proceedings of the 13th international conference on World Wide Web* (pp. 63–71). New York, USA: ACM Press
- Minetou, C. G., Chen, S. Y., & Liu, X. (2008). Investigation of the use of navigation tools in web-based learning: A data mining approach. *International Journal of Human-Computer Interaction*, 24(1), 48–67
- Morrison, J.B., Pirolli, P., & Card, S.K. (2001). A taxonomic analysis of what World Wide Web activities significantly impact people's decisions and actions. In *Proceedings of the CHI'01 extended abstracts on human factors in Computing Systems*. (pp. 163-164). Seattle, USA
- Murtagh, F., & Legendre, P. (2011). Ward's hierarchical clustering method: clustering criterion and agglomerative algorithm. *Computing Research Repository*, 1–20

## **N.**

- Navarro-Pietro, R., Scaife, M., & Rogers, Y. (1999). Cognitive strategies in web searching. In *Proceedings of the 5<sup>th</sup> Conference on Human Factors & the Web* (pp. 43-56).
- Nielsen, J. (2000). *Designing Web Usability*. Indianapolis: New Riders Publishing
- Nielsen, J. (2003). *Homepage Real Estate Allocation*. Récupéré sur useit.com: <http://www.useit.com/alertbox/20030210.html>
- Nielsen, J., & Loranger, H. (2006). *Prioritizing Web Usability*. Berkeley, CA, USA: New Riders.
- Nielsen, J. (2006). F-Shaped Pattern For Reading Web Content.

Nielsen, J. (2011). *How Long Do Users Stay on Web Pages?* Récupéré sur useit.com: <http://www.useit.com/alertbox/page-abandonment-time.html>

## O.

Obendorf, H., Weinreich, H., Herder, E., & Mayer, M. (2007). Web page revisitation revisited: implications of a long-term click-stream study of browser usage. *In Proceedings of the SIGCHI conference on Human factors in computing systems.* (pp. 597–606). San José, USA: ACM Press

Odell, P., Korgen, K., Schumacher, P., & Delucchi, M. (2000). Internet use among female and male college students. *CyberPsychology & Behavior*, 855-862

## P.

Pak, R., & Price, M. M. (2008). Designing an information search interface for younger and older adults. *Human Factors*, 50(4), 614-628

Pallis, G., Angelis, L., & Vakali, A. (2007). Validation and interpretation of Web users' sessions clusters. *Information Processing and Management*, 43(5), 1348–1367

Palmquist, R.A., & Kim, K-S. (2000). Cognitive style and on-line database search experience as predictors of web search performance. *Journal of the American Society for Information Science*, 51, 558–566

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.* (pp. 147-154). ACM Press

Perin, C., Dragicevic, P., & Fekete, J. (2014). Revisiting Bertin's Matrices: New Interactions for Crafting Tabular Visualizations. *Visualization and Computer Graphics*, 20(12), 2082-2091

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, 106, 643-675

Pirolli, P., Fu, W.-T., Reeder, R., & Card, S. K. (2002). A user-tracing architecture for modeling interaction with the World Wide Web. *In Proceedings of the Working Conference on Advanced Visual Interfaces - AVI '02* (pp. 75–83). USA, New York

Pirolli, P., & Fu, W. (2003). SNIF-ACT : A Model of Information Foraging on the World Wide Web. *In Proceedings of the 9th international conference on User modeling* (pp. 45–54)

Potelle, H., & Rouet, J. (2003). Effects of content representation and readers? Prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, 58(3), 327-345.

Protopsaltis, A., & Bouki, V. (2008). Gender and information processing in electronic age. *In Proceedings of the 26th annual ACM international conference on Design of communication SIGDOC '08.* Lisbon, Portugal

## **R.**

- Riding, R., & Rayner, S. (1998). *Cognitive Styles and Learning Strategies*. David Fulton: Londres
- Riding, R. J. (1991). *Cognitive styles analysis*. Birmingham: Learning and Training Technology
- Rose, D. E., & Levinson, D. (2004). Understanding user goals in web search. *In The 13th international conference on World Wide Web*.(pp. 13–19). New York, USA: ACM Press
- Rouet, J.-F., Tricot, A. (1995). Recherche d'informations dans les systèmes hypertextes : des représentations de la tâche à un modèle de l'activité cognitive. *Sciences et Techniques Educatives*, 2(3), 307-331
- Rouet, J.-F., Tricot, A. (1998). Chercher de l'information dans un hypertexte : vers un modèle des processus cognitifs ”, in A. Tricot, J.-F. Rouet, (Eds.), *Les hypermédias, approches cognitives et ergonomiques*. Hermès, Paris
- Roy, M., Taylor, R., & Chi, M. T. H. (2003). Gender differences in patterns of searching the web. *Journal of Educational Computing Research*, 29, 335–348
- Ruiz, B., Pajares, J. L., Utray, F., & Moreno, L. (2011). Design for All in multimedia guides for museums. *Computers in Human Behavior*, 27(4), 1408–1415

## **S.**

- Scott, L. (2015). A Re-Examination of Fortune 500 Homepage Design Practices. *Transactions on professional communication*, 58(1), 20–44
- Sharit, J., Hernandez, M. A., Czaja, S. J., & Pirolli, P. L. (2008). Investigating the roles of knowledge and cognitive abilities in older adult information seeking on the Web. *ACM Transactions on Computer-Human Interaction*, 15(1), 3-28
- Shrestha, S. & Lenz, K. 2007. Eye gaze patterns while searching vs. browsing a Website. *Usability News* 9 (1)
- Sujatha, N., & Iyakutty, K. (2010). Refinement of Web usage Data Clustering from K-means with Genetic Algorithm. *European Journal of Scientific Research*, 42(3), 478–490
- Sujatha, V., & Punithavalli. (2012). Improved user navigation pattern prediction technique from web log data. *Procedia Engineering*, 30, 92–99
- Sullivan, T. (1997). Reading reader reaction: A proposal for inferential analysis of web server log files. *In Proceedings of the 3rd Conference Human Factors and the web*. USA, Colorado

## **T.**

- Tauscher, L. & Greenberg, S. (1997). Revisitation patterns in World Wide Web navigation. *Human Computer Interaction*, 399-406
- Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). The perfect search engine is not enough: a study of orienteering behavior in directed search. *In Proceedings of the SIGCHI conference on Human factors in computing systems*. (pp. 415-422). Vienna, Austria: ACM Press

- Teevan, J., Adar, E., Jones, R., & Potts, M. A. S. (2007). Information re-retrieval: repeat queries in Yahoo's logs. *In Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 151–158). Amsterdam, The Netherlands: ACM Press
- Tullis, B., & Albert, B. (2008). *Measuring the user experience*. Amsterdam: Morgan Kaufmann
- Thatcher, A. (2008). Web search strategies: The influence of web experience and task type. *Information Processing and Management*, 44 (3), 1308–1329
- Tricot, A., Pierre-demarcy, C., & Boussarghini, R. El. (1998). Un panorama des recherches sur l'activité mentale de l'utilisateur d'un hypermédia. *Sciences et Techniques Éducatives*, 5(4), 371–400
- Tricot, A., & Rouet, J.F. (2004). Activités de navigation dans les systèmes d'information. *Psychologie ergonomique : tendances actuelles*. PUF, Paris, 71-95.

## V.

- Vanderheiden G (2000) Fundamental principals and priority setting for universal usability. In: Scholtz J, Thomas J (eds) Proceedings of the ACM Universal Usability Conference. Washington, DC, pp 32–38
- Van Deursen, A., & Van Dijk, J. (2009). Using the Internet: Skill related problems in users' online behavior. *Interacting with Computers*, 21, 393-402
- Van Slyke, C., Comunale, C.L., & Belanger, F. (2002). Gender differences in perceptions of web-based shopping. *Communications of the ACM*, 45(8), 82-86

## W.

- Wang, P., Hawk, W. B., & Tenopir, C. (2000). Users' interaction with World Wide Web resources: An exploratory study using a holistic approach. *Information Processing & Management*, 36, 229–251
- Weber, I., & Jaimes, A. (2010) Demographic information flows. *In Proceedings on Information and knowledge management*. (pp. 1521–1524). Toronto, Canada: ACM Press
- Weber, I., & Jaimes, A. (2011). Who Uses Web Search for What ? And How ? In Proceedings on *WSDM'2011* (pp. 15–24). Hong-Kong, China: ACM
- Weinreich, H., Obendorf, H., Herder, E., & Mayer, M. (2008). Not quite the average : an empirical study of web usage. *ACM Transactions on the Web*, 2(1), 1–31
- Weiser, E. (2000). Gender differences in Internet use patterns and Internet application preferences: A two-sample comparison. *Cyberpsychology and behavior*, 67-178
- White, R. W., & Drucker, S. M. (2007), Investigating behavioral variability in web search. *In Proceedings of the 16th international conference on World Wide Web*. (pp. 21-30). Alberta, Canada: ACM Press
- White, R. W., Dumais, S., & Teevan, J. (2008). How medical expertise influences websearch interaction. *In Proceedings of the 31st annual international ACM SIGIR conference on research and development in information retrieval*. (pp. 781–792). Singapore, Singapore

- White, R. W., Dumais, S. T., & Teevan, J. (2009). Characterizing the influence of domain expertise on web search behavior. *In Proceedings of the Second ACM International Conference on Web Search and Data Mining WSDM 09.* (pp.132-142). Barcelona, Spain: ACM Press
- White, R. W., & Horvitz, E. (2012). Studies of the onset and persistence of medical concerns in search logs. *In Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (pp. 265–274). Portland, USA: ACM Press
- Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). Field-dependent and field independent cognitive styles and their educational implications. *Review of Educational Research*, 47(1), 1–64
- Won, S. S., Jin, J., & Hong, J. I. (2009). Contextual Web History : Using Visual and Contextual Cues to Improve Web Browser History. *Human-Computer Interaction*, 1457–1466

## Y.

- Yu, N., & Kong, J. (2016). User experience with web browsing on small screens: Experimental investigations of mobile-page interface design and homepage design for news websites. *Information Sciences*, 330, 427–443

## Z.

- Zhang, H., & Zhao, S. (2011). Measuring web page revisitation in tabbed browsing. *In Proceedings of the 2011 annual conference on Human factors in computing systems* (pp. 1831–1834). Vancouver, Canada: ACM Press
- Zhou, M. (2014). Gender difference in web search perceptions and behavior: Does it vary by task performance? *Computers and Education*, 78, 174–184





# TABLE DES FIGURES

---

Figure 1 : Schéma du modèle CoLiDeS (Kitajima et al., 2005) .....	18
Figure 2 : Schéma du modèle SNIF-ACT (Fu et Pirolli, 2007).....	19
Figure 3 : Schéma du modèle de recherche d'information par moteur de recherche ( Sharit et al., 2008).....	20
Figure 4 : Schématisation du processus de web mining .....	31
Figure 5 : Information pouvant être récupérées des fichiers d'évènements selon le World Wide Web Consortium .....	33
Figure 6 : le dispositif expérimental .....	57
Figure 7 : Schéma récapitulatif de la procédure expérimentale .....	58
Figure 8 : Page d'accueil du site Priceminister .....	59
Figure 9 : Page d'accueil du site Nature et Découverte .....	60
Figure 10 : Page d'accueil du site Service Public .....	61
Figure 11 : Page d'accueil du site Musée Carnavalet.....	62
Figure 12 : Principales étapes pour l'analyse des données.....	65
Figure 13 : Visualisation des classes avec la méthode Bertifier.....	68
Figure 14 : Illustration des classes en fonction du temps passé sur la PA (en ms) au début de la tâche par rapport au temps total (ms) pour réaliser la tâche.....	69
Figure 15 : Exemple de dendrogramme avec la méthode Dynamic Hybrid Cut.....	74
Figure 16 : Illustration des patterns pour la Tâche 3 – Service Public .....	75
Figure 17 : Chemin le plus représentatif pour le parcours n°1 Tâche 3 - Service Public (LD = 0,12) .....	77
Figure 18 : Chemin le plus représentatif pour le parcours n°2 Tâche 3 - Service Public (LD = 0,22) .....	77
Figure 19 : Chemin le plus représentatif pour le parcours n°3 Tâche 3 - Service Public LD = 0,65..	78
Figure 20 : Illustration des patterns pour la Tâche 4 - Service Public.....	79
Figure 21 : Chemin le plus représentatif pour le parcours n°1 Tâche 4 - Service Public LD = 0,06..	81
Figure 22 : Chemin le plus représentatif pour le parcours n°2 Tâche 4 - Service Public LD = 0,11..	81
Figure 23 : Chemin le plus représentatif pour le parcours n°3 Tâche 4 - Service Public LD = 0,55..	81
Figure 24 : Illustration des patterns pour la Tâche 5 – Musée Carnavalet.....	82
Figure 25 : Chemin le plus représentatif pour le parcours n°1 Tâche 5 – Musée Carnavalet LD = 0,12 .....	84
Figure 26 : Chemin le plus représentatif pour le parcours n°1 Tâche 5 – Musée Carnavalet LD = 0,38 .....	84
Figure 27 : Illustration des patterns pour la Tâche 6 – Musée Carnavalet.....	85
Figure 28 : Chemin le plus représentatif pour le parcours n°1 Tâche 6 – Musée Carnavalet LD = 0,2 .....	86
Figure 29 : Chemin le plus représentatif pour le parcours n°2 Tâche 6 – Musée Carnavalet LD = 0,12 .....	87
Figure 30 : Chemin le plus représentatif pour le parcours n°3 Tâche 6 – Musée Carnavalet LD = 0,52 .....	87
Figure 31 : Visualisation de l'évolution des participants à travers les différentes tâches et patterns .....	88

Figure 32 : Évolution des participants caractérisés comme ayant une « navigation consciencieuse » de la première tâche et jusqu'à la fin de l'étude .....	90
Figure 33 : Illustration des patterns pour la Tâche 3R - Service Public .....	96
Figure 34 : Chemin le plus représentatif pour le parcours n°1 Tâche 3R - Service Public (LD = 0,03) .....	97
Figure 35 : Chemin le plus représentatif pour le parcours n°2 Tâche 3R- Service Public (LD = 0,13) .....	97
Figure 36 : Visualisation de l'évolution des participants à travers la répétition de la première tâche sur le site du Service Public .....	98
Figure 37 : Illustration des patterns pour la Tâche 4R – Service Public .....	99
Figure 38 : Chemin le plus représentatif pour le parcours n°1 Tâche 4R – Service Public (LD = 0,07) .....	100
Figure 39 : Chemin le plus représentatif pour le parcours n°2 Tâche 4R – Service Public (LD = 0,02) .....	100
Figure 40 : Visualisation de l'évolution des participants à travers la répétition de la deuxième tâche sur le site du Service Public .....	101
Figure 41 : Illustration des patterns pour la Tâche 5R – Musée Carnavalet .....	102
Figure 42 : Chemin le plus représentatif pour le parcours n°1 Tâche 5R - Musée Carnavalet (LD = 0,01) .....	103
Figure 43 : Chemin le plus représentatif pour le parcours n° 2 Tâche 5R - Musée Carnavalet (LD = 0,16) .....	103
Figure 44 : Visualisation de l'évolution des participants à travers la répétition de la première tâche sur le site du Musée Carnavalet .....	104
Figure 45 : Illustration des patterns pour la Tâche 6R - Musée Carnavalet .....	105
Figure 46 : Chemin le plus représentatif pour le parcours n°1 Tâche 6R - Musée Carnavalet (LD = 0,05) .....	106
Figure 47 : Chemin le plus représentatif pour le parcours n°2 Tâche 6R - Musée Carnavalet .....	106
Figure 48 : Visualisation de l'évolution des participants à travers la répétition de la deuxième tâche sur le site du Musée Carnavalet .....	107
Figure 49 : Représentation Bertifier des sessions 3R, 4R, 5R et 6R .....	108
Figure 50 : Schéma récapitulatif de la procédure expérimentale .....	110
Figure 51 : Page d'accueil du site Musée Carnavalet .....	111
Figure 52 : Page d'accueil du site Diplomatie Française .....	112
Figure 53 : Illustration des patterns pour la première Tâche à Temps 1, 2 et 3 – Musée Carnavalet .....	115
Figure 54 : Visualisation de l'évolution des participants à travers les répétitions de la première tâche sur le site du Musée Carnavalet .....	117
Figure 55 : Illustration des patterns pour la deuxième Tâche à Temps 1, 2 et 3 – Musée Carnavalet .....	118
Figure 56 : Visualisation de l'évolution des participants à travers les répétitions de la deuxième tâche sur le site du Musée Carnavalet .....	121
Figure 57 : Illustration des patterns pour la troisième Tâche à Temps 1, 2 et 3 – Diplomatie Française .....	122
Figure 58 : Visualisation de l'évolution des participants à travers les répétitions de la troisième tâche sur le site de Diplomatie Française .....	125
Figure 59 : Illustration des patterns pour la troisième Tâche à Temps 1, 2 et 3 – Diplomatie Française .....	126

<i>Figure 60 : Visualisation de l'évolution des participants à travers les répétitions de la dernière tâche sur le site de Diplomatie Française.....</i>	<i>129</i>
<i>Figure 61 : Page d'accueil du site MEGA.....</i>	<i>133</i>
<i>Figure 62 : Illustration des patterns pour la Tâche 1.....</i>	<i>135</i>
<i>Figure 63 : Illustration des patterns pour la Tâche 2.....</i>	<i>137</i>
<i>Figure 64 : Illustration des patterns pour la Tâche 4.....</i>	<i>139</i>
<i>Figure 65 : Illustration des patterns pour la Tâche 5.....</i>	<i>141</i>

# TABLE DES TABLEAUX

---

Tableau 1 : Principales méthodes pour le recueil de données sur la navigation .....	28
Tableau 2 : Effectifs pour chaque tâche, étude 1 .....	64
Tableau 3 : Caractéristiques de l'échantillon, étude 1 .....	64
Tableau 4 : Moyenne des indicateurs relevées pour chaque pattern.....	67
Tableau 5 Répartition des parcours pour chaque pattern pour la Tâche 3 - Service Public.....	77
Tableau 6 : Répartition des parcours pour chaque Pattern pour la Tâche 4 - Service Public.....	80
Tableau 7 : Répartition des parcours pour chaque Pattern pour la Tâche 5 - Musée Carnavalet....	83
Tableau 8 : Répartition des parcours pour chaque Pattern pour la Tâche 6 - Musée Carnavalet....	86
Tableau 9 : Effectifs pour chaque tâche.....	95
Tableau 10 : Répartition des parcours pour chaque pattern pour la Tâche 3R - Service Public .....	97
Tableau 11 : Répartition des parcours pour chaque Pattern pour la Tâche 4R – Service Public ...	100
Tableau 12 : Répartition des parcours pour chaque Pattern pour la Tâche 5R - Musée Carnavalet .....	103
Tableau 13 : Répartition des parcours pour chaque Pattern pour la Tâche 6R - Musée Carnavalet .....	106
Tableau 14 : Effectifs pour chaque tâche, étude 2 .....	114
Tableau 15 : Caractéristiques de l'échantillon, étude 2 .....	114
Tableau 16 : Répartition des parcours pour chaque Pattern Tâche 1-Temps 1 - Musée Carnavalet .....	116
Tableau 17 : Répartition des parcours pour chaque Pattern Tâche 1-Temps 2 - Musée Carnavalet .....	116
Tableau 18 : Répartition des parcours pour chaque Pattern Tâche 1-Temps 3 - Musée Carnavalet .....	117
Tableau 19 : Répartition des parcours pour chaque Pattern Tâche 2-Temps 1 - Musée Carnavalet .....	119
Tableau 20 : Répartition des parcours pour chaque Pattern Tâche 2-Temps 2 - Musée Carnavalet .....	120
Tableau 21 : Répartition des parcours pour chaque Pattern Tâche 2-Temps 3 - Musée Carnavalet .....	120
Tableau 22 : Répartition des parcours pour chaque Pattern Tâche 3-Temps 1 – Diplomatie Française .....	123
Tableau 23 : Répartition des parcours pour chaque Pattern Tâche 3-Temps 2 – Diplomatie Française .....	123
Tableau 24 : Répartition des parcours pour chaque Pattern Tâche 3-Temps 3 – Diplomatie Française .....	124
Tableau 25 : Répartition des parcours pour chaque Pattern Tâche 4-Temps 1 – Diplomatie Française .....	127
Tableau 26 : Répartition des parcours pour chaque Pattern Tâche 4-Temps 2 – Diplomatie Française .....	127
Tableau 27 : Répartition des parcours pour chaque Pattern Tâche 4-Temps 3 – Diplomatie Française .....	128
Tableau 28 : Répartition des parcours pour chaque Pattern Tâche 1 – MEGA.....	137

<i>Tableau 29 : Répartition des parcours pour chaque Pattern Tâche 2 – MEGA.....</i>	<i>139</i>
<i>Tableau 30 : Répartition des parcours pour chaque Pattern Tâche 4 – MEGA.....</i>	<i>140</i>
<i>Tableau 31 : Répartition des parcours pour chaque Pattern Tâche 5 – MEGA.....</i>	<i>142</i>
<i>Tableau 32 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3 - Service Public.....</i>	<i>180</i>
<i>Tableau 33 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Service Public.....</i>	<i>181</i>
<i>Tableau 34 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 5 – Musée Carnavalet.....</i>	<i>182</i>
<i>Tableau 35 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 6 – Musée Carnavalet.....</i>	<i>183</i>
<i>Tableau 36 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3R - Service Public.....</i>	<i>184</i>
<i>Tableau 37 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4R - Service Public.....</i>	<i>186</i>
<i>Tableau 38 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 5R - Musée Carnavalet.....</i>	<i>187</i>
<i>Tableau 39 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 6R - Musée Carnavalet.....</i>	<i>188</i>
<i>Tableau 40 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 1 – Temps 1 - Musée Carnavalet.....</i>	<i>189</i>
<i>Tableau 41 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 1 – Temps 2 - Musée Carnavalet.....</i>	<i>191</i>
<i>Tableau 42 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 1 – Temps 3 - Musée Carnavalet.....</i>	<i>193</i>
<i>Tableau 43 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 2 – Temps 1 - Musée Carnavalet.....</i>	<i>195</i>
<i>Tableau 44 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 2 – Temps 2 - Musée Carnavalet.....</i>	<i>197</i>
<i>Tableau 45 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 2 – Temps 3 - Musée Carnavalet.....</i>	<i>199</i>
<i>Tableau 46 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3 – Temps 1 - Diplomatie Française.....</i>	<i>201</i>
<i>Tableau 47 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3 – Temps 2 - Diplomatie Française.....</i>	<i>203</i>
<i>Tableau 48 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3 – Temps 3 - Diplomatie Française.....</i>	<i>205</i>
<i>Tableau 49 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Temps 1 - Diplomatie Française.....</i>	<i>207</i>
<i>Tableau 50 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Temps 2 - Diplomatie Française.....</i>	<i>209</i>
<i>Tableau 51 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Temps 3 - Diplomatie Française.....</i>	<i>211</i>
<i>Tableau 52 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 1 – Différence entre sexe et genre.....</i>	<i>216</i>
<i>Tableau 53 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 2 – Mesures pour l'égalité des chances.....</i>	<i>218</i>

<i>Tableau 54 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Demande de patronage et subvention .....</i>	<i>220</i>
<i>Tableau 55 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 5 – Mesures pour l'égalité des chances .....</i>	<i>222</i>

# SOMMAIRE DES ANNEXES

---

<b>ANNEXES A .....</b>	<b>175</b>
A.1 ANNEXES DE L'ETUDE « METHODES ET ANALYSES POUR L'IDENTIFICATION DE PATTERNS COMPORTEMENTAUX SUR DES SITES WEB ».....	175
A.1.1 <i>Questionnaire Post test</i> .....	175
A.1.2 <i>Résultats des analyses</i> .....	180
<b>ANNEXES B.....</b>	<b>184</b>
B.1 ANNEXES DE L'ETUDE « ÉVOLUTION TEMPORELLE DES PATTERNS COMPORTEMENTAUX ». 184	
B.1.1 <i>Résultats des analyses pour l'étude des patterns comportementaux à court terme.....</i>	184
B.1.2 <i>Résultats des analyses pour l'étude des patterns comportementaux à long terme.....</i>	189
<b>ANNEXES C .....</b>	<b>213</b>
C.1 ANNEXES DE L'ETUDE « INFLUENCE DES CARACTERISTIQUES DES PARTICIPANTS SUR LES PATTERNS » .....	213
C.1.1 <i>Questionnaire ISALEM Style d'apprentissage.....</i>	213
C.1.2 <i>Résultats des analyses.....</i>	216





# ANNEXES A

---

## A.1 ANNEXES DE L'ETUDE « METHODES ET ANALYSES POUR L'IDENTIFICATION DE PATTERNS COMPORTEMENTAUX SUR DES SITES WEB »

---

### A.1.1 Questionnaire Post test

#### PARTIE 1 : QUESTIONS CONCERNANT L'ETUDE

1) Avez-vous déjà participé à une étude comme celle à laquelle vous venez de participer ?

Oui  Non

2) Si oui, pouvez-vous décrire (le plus précisément possible) les tâches que vous aviez effectué ?

- ...

3) Si, les tâches se rapportaient à internet, quels sont le ou les sites sur lesquels vous avez navigué ?

- ...

4) Selon vous, quel était l'objectif de notre étude ?

- ...

5) Aviez-vous déjà surfé sur le site « [priceminister.com](http://priceminister.com) » (site 1) ?

Oui  Non

Si oui,

- à quelle fréquence ?

Très rarement

Parfois

Souvent

Tous les jours

- Quel est votre degré de connaissance de ce site selon vous (entourez le chiffre qui correspond) ?

Je ne le connais pas du tout 1 - 2 - 3 - 4 - 5 Je le connais parfaitement

6) Aviez-vous déjà surfé sur le site « nature et découverte » (site 2) ?

Oui  Non

**Si oui,**

- à quelle fréquence ?

- Très rarement
- Parfois
- Souvent
- Tous les jours

- Quel est votre degré de connaissance de ce site selon vous (entourez le chiffre qui correspond) ?

Je ne le connais pas du tout    1 – 2 – 3 – 4 – 5    Je le connais parfaitement

**7) Avez-vous déjà surfé sur le site « [www.service-public.fr](http://www.service-public.fr) » ?**

- Oui
- Non

**Si oui,**

- à quelle fréquence ?

- Très rarement
- Parfois
- Souvent
- Tous les jours

- Quel est votre degré de connaissance de ce site selon vous (entourez le chiffre qui correspond) ?

Je ne le connais pas du tout    1 – 2 – 3 – 4 – 5    Je le connais parfaitement

**8) Avez-vous déjà surfé sur le site du musée Carnavalet ?**

- Oui
- Non

**Si oui,**

- à quelle fréquence ?

- Très rarement
- Parfois
- Souvent
- Tous les jours

- Quel est votre degré de connaissance de ce site selon vous (entourez le chiffre qui correspond) ?

Je ne le connais pas du tout    1 – 2 – 3 – 4 – 5    Je le connais parfaitement

## PARTIE 2 : FICHE SIGNALÉTIQUE

1) Age :.....ans

2) Sexe :

Masculin  Féminin

3) Dernier diplôme obtenu :

→ Si vous n'êtes plus étudiant, passez directement à la question 7).

4) Diplôme préparé actuellement ?

- ...

5) Vous êtes en :

Licence 1

Licence 2

Licence 3

Master 1

Master 2

Doctorat

6) Filière d'étude :

- ...

7) Quelle est votre profession ?

- ...

8) Depuis combien de temps utilisez-vous Internet ?

Moins de 6 mois

De 6 mois à 1 an

De 1 an à 2 ans

De 2 ans à 3 ans

3 ans et plus

**9) Combien d'heures naviguez-vous ?**

- Par jour (ou).....heures
- Par semaine (ou).....heures
- Par mois.....heures

**10) Quels sont les sites que vous visitez le plus souvent ?**

~...

**11) Quel navigateur Web utilisez-vous ? (Plusieurs réponses possibles)**

- Firefox
- Safari
- Microsoft Internet Explorer
- Autres (Précisez) .....

**PARTIE 3 : VOTRE CONNAISSANCE D'INTERNET**

(Entourez la réponse qui vous correspond pour chaque question)

**1) Les ordinateurs permettent d'accéder à des informations sur le Web :**

- je n'ai jamais entendu parler du Web
- j'en ai entendu parler mais je ne sais pas utiliser le Web
- je sais comment le utiliser le Web
- j'ai une connaissance experte de la façon dont on utilise le Web
- j'en ai une connaissance tellement experte que je peux enseigner aux autres comment utiliser le Web

**2) <http://univ-paris12.fr/www/labos/lmp/watzky/L/Murphy.html> est une adresse URL:**

- je n'ai jamais entendu parler des URL
- j'en ai entendu parler mais je ne sais pas comment utiliser les URL
- je sais comment utiliser les URL
- j'ai une connaissance experte de la façon dont on utilise les URL
- j'en ai une connaissance tellement experte que je peux enseigner aux autres comment utiliser les URL

**3) Des documents électroniques ou des fichiers peuvent être téléchargés depuis le Web :**

- je n'ai jamais entendu dire que l'on pouvait télécharger des fichiers sur le Web
- j'en ai entendu parler mais je ne sais pas comment faire
- je sais comment télécharger des fichiers
- j'ai une connaissance experte de la façon dont on télécharge les fichiers
- j'en ai une connaissance tellement experte que je peux enseigner aux autres comment télécharger des fichiers

**4) Le Web peut permettre d'effectuer une recherche en utilisant des mots-clefs :**

- je n'ai jamais entendu parler de mot-clef pour rechercher sur le Web
- j'en ai entendu parler mais je ne sais pas comment utiliser les mots-clefs pour rechercher sur le Web
- je sais comment utiliser les mots-clefs pour rechercher sur le Web
- j'ai une connaissance experte de la façon dont on utilise les mots-clefs pour rechercher sur le Web
- j'en ai une connaissance tellement experte que je peux enseigner aux autres comment utiliser les mots-clefs pour rechercher sur le Web

**5) Internet peut être utilisé pour envoyer et recevoir des courriers électroniques :**

- je n'ai jamais entendu parler des courriers électroniques
- j'en ai entendu parler mais je ne sais pas comment utiliser les courriers électroniques
- je sais comment utiliser les courriers électroniques
- j'ai une connaissance experte de la façon dont on utilise les courriers électroniques
- j'en ai une connaissance tellement experte que je peux enseigner aux autres comment utiliser les courriers électroniques

**6) Des documents ou des fichiers électroniques peuvent être joints à des courriers électroniques :**

- je n'ai jamais entendu parler de pièces jointes concernant les courriers électroniques
- j'en ai entendu parler mais je ne sais pas comment joindre une pièce à un courrier électronique
- je sais comment utiliser les pièces jointes dans les courriers électroniques
- j'ai une connaissance experte de la façon dont on joint une pièce dans un courrier électronique
- j'en ai une connaissance tellement experte que je peux enseigner aux autres comment joindre une pièce dans un courrier électronique

**7) Les fichiers Portable Document Format (.pdf ou PDF) sont communément trouvés sur internet :**

- je n'ai jamais entendu parler de fichiers PDF
- j'en ai entendu parler mais je ne sais pas comment lire un fichier PDF
- je sais comment lire un fichier PDF
- j'ai une connaissance experte de la façon dont on lit un fichier PDF
- j'en ai une connaissance tellement experte que je peux enseigner aux autres comment lire un fichier PDF

*Questionnaire inspiré de 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*

## A.1.2

## Résultats des analyses

Tableau 32 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3 - Service Public

Moyennes pour les métriques analysées	Pattern 1 (n=63)	Pattern 2 (n=30)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	219 944	135 455
Nombre total de pages visitées pour réaliser la tâche	13,26	7,56
Nombre de retours en arrière	1,8	1,7
Taux de revisite (en pourcentage)	19%	10%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	34 348	38 565
Nombre de clics gauche réalisés	1,22	1
Temps de latence avant la première action (en millisecondes)	11 187	33 076
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	115,13	12,66
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	106,88	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	2,87	1,46
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	31 607	33 128
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1,30	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	125,05	16,66
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	121,98	5,03
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	16 681	17 420
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,92	0,88

Tableau 33 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Service Public

Moyennes pour les métriques analysées	Pattern 1 (n=47)	Pattern 2 (n=24)	Pattern 3 (n=24)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	71 226	38 050	127 531
Nombre total de pages visitées pour réaliser la tâche	5,82	5,87	12,5
Nombre de retours en arrière	0,08	0	2,58
Taux de revisite (en pourcentage)	3%	0,7%	29%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	18 642	6 971	10 970
Nombre de clics gauche réalisés	1	1	1
Temps de latence avant la première action (en millisecondes)	17 628	6540	10 655
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	14,25	23,08	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>			
Nombre de visites sur cette page	1	1	2,83
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	17 628	6971	12 109
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	14,25	23,08	30,04
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0	0
<b>Comportements sur les autres pages visitées</b>			
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	12 252	6 226	10 381
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,86	0,83	0,69

Tableau 34 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 5 – Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=62)	Pattern 2 (n=37)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	65 531	114 273
Nombre total de pages visitées pour réaliser la tâche	5,03	4,3
Nombre de retours en arrière	0,40	0,21
Taux de revisite (en pourcentage)	6%	2%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	15 024	31 570
Nombre de clics gauche réalisés	2	2,3
Temps de latence avant la première action (en millisecondes)	8 543	9 299
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	10,02	365,50
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	5,82	199,39
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1,1	1,1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	15 102	29 581
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	2,1	2,2
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	22,36	365,50
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	11,65	199,39
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	15 638	27 425
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1,45	1,65



Tableau 35 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 6 – Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=36)	Pattern 2 (n=30)	Pattern 3 (n=20)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	151 483	130 047	126 490
Nombre total de pages visitées pour réaliser la tâche	11,25	4,16	5,25
Nombre de retours en arrière	1,80	0,56	1,05
Taux de revisite (en pourcentage)	33%	13%	23%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	9841	38 398	18 335
Nombre de clics gauche réalisés	1,3	2,03	1,4
Temps de latence avant la première action (en millisecondes)	7605	13 099	9908
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	11,11	282,57	60,16
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	11,11	247,57	58,12
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>			
Nombre de visites sur cette page	2,83	1,6	2,1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	23 187	39 154	31 054
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1,35	2,03	1,625
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	276,61	289,71	173,36
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	222,43	248,83	118,19
<b>Comportements sur les autres pages visitées</b>			
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	15 480	31 143	23 669
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1	1,73	1,34

# ANNEXES B

## B.1 ANNEXES DE L'ETUDE « ÉVOLUTION TEMPORELLE DES PATTERNS COMPORTEMENTAUX »

### B.1.1 Résultats des analyses pour l'étude des patterns comportementaux à court terme

Tableau 36 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3R - Service Public

Moyennes pour les métriques analysées	Pattern 1 (n=41)	Pattern 2 (n=41)	Pattern 3 (n=9)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	18 030	28 699	39 367
Nombre total de pages visitées pour réaliser la tâche	5	5	7,3
Nombre de retours en arrière	0	0	1,1
Taux de revisite (en pourcentage)	0%	0%	19%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	3 458	6 432	8 177
Nombre de clics gauche réalisés	1	1	1
Temps de latence avant la première action (en millisecondes)	2 934	5 612	7 788
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	11,65	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	6,34	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>			
Nombre de visites sur cette page	1	1	1,4
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	3 458	6 432	7 652
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	11,65	0

Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	6,34	0
<b>Comportements sur les autres pages visitées</b>			
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	3 586	5 625	5 440
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,80	0,85	0,78

Tableau 37 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4R - Service Public

Moyennes pour les métriques analysées	Pattern 1 (n=59)	Pattern 2 (n=41)
<b>Efficiencie à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	22 748	16 635
Nombre total de pages visitées pour réaliser la tâche	4,9	4,2
Nombre de retours en arrière	0	0
Taux de revisite (en pourcentage)	0%	0%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	4 969	1 972
Nombre de clics gauche réalisés	1	1
Temps de latence avant la première action (en millisecondes)	4 375	1 745
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	4 969	1 972
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	4 864	3 747
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,8	0,8

Tableau 38 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 5R - Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=21)	Pattern 2 (n=80)
<b>Efficiencie à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	40 591	18 245
Nombre total de pages visitées pour réaliser la tâche	3,3	3,2
Nombre de retours en arrière	0	0
Taux de revisite (en pourcentage)	0%	0%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	16 722	5 591
Nombre de clics gauche réalisés	3,9	2
Temps de latence avant la première action (en millisecondes)	5 739	2 982
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	16 722	5 591
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	3,9	2
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	12 266	5 624
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	2,5	1,2

Tableau 39 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 6R - Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=34)	Pattern 2 (n=28)	Pattern 3 (n=24)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	12 116	11 568	33 885
Nombre total de pages visitées pour réaliser la tâche	2,26	2	3,9
Nombre de retours en arrière	0	0	0,8
Taux de revisite (en pourcentage)	0%	0%	22%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	5 060	5 510	6 999
Nombre de clics gauche réalisés	1	1,2	1,3
Temps de latence avant la première action (en millisecondes)	3 186	4 299	3 319
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	108	487,64	32,27
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>			
Nombre de visites sur cette page	1	1	2
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	5 060	5 510	7 775
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	1,2	1,5
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	108	487,64	391,47
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0	0
<b>Comportements sur les autres pages visitées</b>			
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	5309	5 433	8 522
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,57	0,70	1,27

## B.1.2

# Résultats des analyses pour l'étude des patterns comportementaux à long terme

Tableau 40 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 1 – Temps 1 - Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=15)	Pattern 2 (n=8)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	128 673	176 567
Nombre total de pages visitées pour réaliser la tâche	3,8	6,4
Nombre de retours en arrière	0,13	1,3
Taux de revisite (en pourcentage)	2%	27%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	27 275	31 625
Nombre de clics gauche réalisés	2	2,7
Temps de latence avant la première action (en millisecondes)	11 494	11 985
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	210	357
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	153	548
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	27 275	31 625
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	2	2,7
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	210	357
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	153	548
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	35 075	29 228
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1,14	1,03

## Représentation des parcours types pour la Tâche 1 à Temps 1



Chemin le plus représentatif pour le parcours n°1 Tâche 1 Temps 1 (LD = 0,30)



Chemin le plus représentatif pour le parcours n°2 Tâche 1 Temps 1 (LD = 0,56)



Chemin le plus représentatif pour le parcours n°3 Tâche 1 Temps 1 (LD = 0,23)



*Tableau 41 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 1 – Temps 2 - Musée Carnavalet*

Moyennes pour les métriques analysées	Pattern 1 (n=10)	Pattern 2 (n=9)	Pattern 3 (n=8)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	46 945	162 538	125 012
Nombre total de pages visitées pour réaliser la tâche	2,8	8,8	5,6
Nombre de retours en arrière	0	1	0,3
Taux de revisite (en pourcentage)	0%	27%	11%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	8 513	9 846	28 565
Nombre de clics gauche réalisés	1,2	1	2,5
Temps de latence avant la première action (en millisecondes)	7 995	9 440	11 453
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0	332
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0	567
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>			
Nombre de visites sur cette page	1	1	1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	8 513	9 846	28 565
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1,2	1	2,5
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0	332
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0	567
<b>Comportements sur les autres pages visitées</b>			
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	17 253	17 399	22 300
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,7	0,8	1,5

## Représentation des parcours types pour la Tâche 1 à Temps 2



Chemin le plus représentatif pour le parcours n°1 Tâche 1 Temps 2 (LD = 0,46)



Chemin le plus représentatif pour le parcours n°2 Tâche 1 Temps 2 (LD = 0,42)

Tableau 42 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 1 – Temps 3 - Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=19)	Pattern 2 (n=6)
<b>Efficiencie à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	44 961	52 071
Nombre total de pages visitées pour réaliser la tâche	4,3	3,6
Nombre de retours en arrière	0,3	0
Taux de revisite (en pourcentage)	4,6%	1,3%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	6 891	14 350
Nombre de clics gauche réalisés	1	2,16
Temps de latence avant la première action (en millisecondes)	6 514	8 014
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	6 891	14 350
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	2,16
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	10 733	13 383
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,7	1,15

## Représentation des parcours types pour la Tâche 1 à Temps 3



Chemin le plus représentatif pour le parcours n°1 Tâche 1 Temps 3 (LD = 0,41)



Chemin le plus représentatif pour le parcours n°2 Tâche 1 Temps 3 (LD = 0,24)

Tableau 43 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 2 – Temps 1- Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=16)	Pattern 2 (n=8)
<b>Efficiencie à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	38 290	104 619
Nombre total de pages visitées pour réaliser la tâche	4	7,5
Nombre de retours en arrière	0	1
Taux de revisite (en pourcentage)	0%	22%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	7 175	9 986
Nombre de clics gauche réalisés	1	1
Temps de latence avant la première action (en millisecondes)	6 764	9 388
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	7 175	9 986
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	9 540	13 681
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,7	0,8

## Représentation des parcours types pour la Tâche 2 à Temps 1



Chemin le plus représentatif pour le parcours n°1 Tâche 2 Temps 1 (LD = 0,43)



Chemin le plus représentatif pour le parcours n°2 Tâche 2 Temps 1 (LD = 0,06)



Chemin le plus représentatif pour le parcours n°2 Tâche 2 Temps 1 (LD = 0,3)

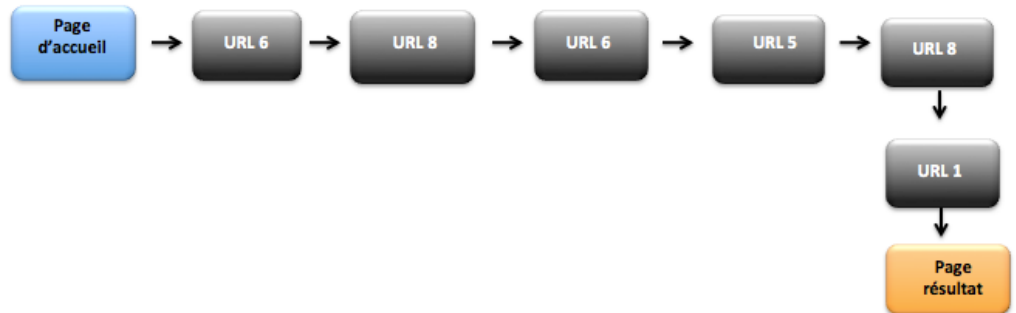
Tableau 44 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 2 – Temps 2 - Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=10)	Pattern 2 (n=9)	Pattern 3 (n=7)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	16 337	27 443	107 801
Nombre total de pages visitées pour réaliser la tâche	3,1	3,6	10,7
Nombre de retours en arrière	0	0	0,8
Taux de revisite (en pourcentage)	0%	0%	26%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	4 442	8 440	6 528
Nombre de clics gauche réalisés	1	1	1
Temps de latence avant la première action (en millisecondes)	3 835	7 019	7 528
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>			
Nombre de visites sur cette page	1	1	1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	4 442	8 440	6 528
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0	0
<b>Comportements sur les autres pages visitées</b>			
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	5 284	7 361	11 191
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,7	0,7	0,9

## Représentation des parcours types pour la Tâche 2 à Temps 2



Chemin le plus représentatif pour le parcours n°1 Tâche 2 Temps 2 (LD = 0,09)



Chemin le plus représentatif pour le parcours n°2 Tâche 2 Temps 2 (LD = 0,44)



Tableau 45 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 2 – Temps 3 - Musée Carnavalet

Moyennes pour les métriques analysées	Pattern 1 (n=18)	Pattern 2 (n=8)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	14 440	31 466
Nombre total de pages visitées pour réaliser la tâche	3	4
Nombre de retours en arrière	0	0
Taux de revisite (en pourcentage)	0%	0%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	4 683	8 363
Nombre de clics gauche réalisés	1	1,25
Temps de latence avant la première action (en millisecondes)	4 112	7 350
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	4 683	8 363
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	4 722	7 869
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,6	0,9

## Représentation des parcours types pour la Tâche 2 à Temps 3



Chemin le plus représentatif pour le parcours n°1 Tâche 2 Temps 3 (LD = 0,01)



Chemin le plus représentatif pour le parcours n°2 Tâche 2 Temps 3 (LD = 0,06)

Tableau 46 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3 – Temps 1 - Diplomatie Française

Moyennes pour les métriques analysées	Pattern 1 (n=16)	Pattern 2 (n=11)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	265 428	115 736
Nombre total de pages visitées pour réaliser la tâche	9	7,3
Nombre de retours en arrière	0,7	0,2
Taux de revisite (en pourcentage)	13%	4%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	61 174	14 859
Nombre de clics gauche réalisés	2,8	1
Temps de latence avant la première action (en millisecondes)	12 953	18 014
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	347	18
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	481	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1,5	1,2
Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	57 993	17 251
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	2,6	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	356	39
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	496	42
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	28 455	16 180
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1,19	0,9

## Représentation des parcours types pour la Tâche 3 à Temps 1



Chemin le plus représentatif pour le parcours n°1 Tâche 3 Temps 1 (LD = 0,7)



Chemin le plus représentatif pour le parcours n°2 Tâche 3 Temps 1 (LD = 0,2)

Tableau 47 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3 – Temps 2 - Diplomatie Française

Moyennes pour les métriques analysées	Pattern 1 (n=13)	Pattern 2 (n=13)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	42 878	213 877
Nombre total de pages visitées pour réaliser la tâche	4,6	10,6
Nombre de retours en arrière	0	0,8
Taux de revisite (en pourcentage)	0%	11%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	7 808	22 354
Nombre de clics gauche réalisés	1,15	1,38
Temps de latence avant la première action (en millisecondes)	5 885	11 473
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	24
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	24
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	7 808	22 354
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	1,15	1,38
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	24
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	24
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	8 172	20 537
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,9	1

## Représentation des parcours types pour la Tâche 3 à Temps 2



Chemin le plus représentatif pour le parcours n°1 Tâche 3 Temps 2 (LD = 0,64)



Chemin le plus représentatif pour le parcours n°2 Tâche 3 Temps 2 (LD = 0,09)

Tableau 48 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 3 – Temps 3 - Diplomatie Française

Moyennes pour les métriques analysées	Pattern 1 (n=15)	Pattern 2 (n=9)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	112 176	15 867
Nombre total de pages visitées pour réaliser la tâche	9,2	3,4
Nombre de retours en arrière	0,8	0
Taux de revisite (en pourcentage)	8%	0%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	9 697	5 469
Nombre de clics gauche réalisés	1	1
Temps de latence avant la première action (en millisecondes)	8 397	3 926
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	9 697	5 469
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	12 330	4 694
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1	0,8

## Représentation des parcours types pour la Tâche 3 à Temps 3



Chemin le plus représentatif pour le parcours n°1 Tâche 3 Temps 3 (LD = 0,57)



Chemin le plus représentatif pour le parcours n°2 Tâche 3 Temps 3 (LD = 0,12)



Tableau 49 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Temps 1 - Diplomatie Française

Moyennes pour les métriques analysées	Pattern 1 (n=19)	Pattern 2 (n=8)
<b>Efficiencie à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	70 646	169 364
Nombre total de pages visitées pour réaliser la tâche	6	6
Nombre de retours en arrière	0,47	0,25
Taux de revisite (en pourcentage)	4%	5%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	8 236	15 039
Nombre de clics gauche réalisés	1	1
Temps de latence avant la première action (en millisecondes)	8 236	15 039
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1,2
Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	8 236	17 660
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	11 963	27 629
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1,19	0,95

## Représentation des parcours types pour la Tâche 4 à Temps 1



Chemin le plus représentatif pour le parcours n°1 Tâche 4 Temps 1 (LD = 0,07)



Chemin le plus représentatif pour le parcours n°2 Tâche 4 Temps 1 (LD = 0,55)

Tableau 50 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Temps 2 - Diplomatie Française

Moyennes pour les métriques analysées	Pattern 1 (n=13)	Pattern 2 (n=13)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	27 353	52 651
Nombre total de pages visitées pour réaliser la tâche	5,3	4,6
Nombre de retours en arrière	0,15	0
Taux de revisite (en pourcentage)	2,3%	0%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	4 667	8 078
Nombre de clics gauche réalisés	1	1
Temps de latence avant la première action (en millisecondes)	3 586	6 913
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	4 667	8 078
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	5 765	9 615
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,77	0,95

## Représentation des parcours types pour la Tâche 4 à Temps 2



Chemin le plus représentatif pour le parcours n°1 Tâche 4 Temps 2 (LD = 0,08)



Chemin le plus représentatif pour le parcours n°2 Tâche 4 Temps 2 (LD = 0)



Chemin le plus représentatif pour le parcours n°3 Tâche 4 Temps 2 (LD = 0,24)

Tableau 51 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 – Temps 3 - Diplomatie Française

Moyennes pour les métriques analysées	Pattern 1 (n=12)	Pattern 2 (n=12)
<b>Efficienc e à la tâche</b>		
Temps total pour accomplir la tâche (en millisecondes)	24 726	21 695
Nombre total de pages visitées pour réaliser la tâche	5	4
Nombre de retours en arrière	0	0
Taux de revisite (en pourcentage)	0%	0%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>		
Temps passé sur la page (en millisecondes)	4 305	7 208
Nombre de clics gauche réalisés	1	1
Temps de latence avant la première action (en millisecondes)	3 229	6 047
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0
<b>Comportements sur la page d'accueil en fonction du nombre de visites sur cette page</b>		
Nombre de visites sur cette page	1	1
Temps total passé sur la PA pondéré par le nombre de visites sur cette page (en millisecondes)	4 305	7 208
Nombre total de clics gauche sur la PA pondéré par le nombre de visites sur cette page	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0
<b>Comportements sur les autres pages visitées</b>		
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	4 794	5 672
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,9	0,8

## Représentation des parcours types pour la Tâche 4 à Temps 3



Chemin le plus représentatif pour le parcours n°1 Tâche 4 Temps 3 (LD = 0,12)



Chemin le plus représentatif pour le parcours n°2 Tâche 4 Temps 3 (LD = 0,8)

# ANNEXES C

---

## C.1 ANNEXES DE L'ETUDE « INFLUENCE DES CARACTERISTIQUES DES PARTICIPANTS SUR LES PATTERNS »

---

### C.1.1 Questionnaire ISALEM Style d'apprentissage

#### CONSIGNE

Assis au côté du participant, nous étions disponible s'il avait des questions. Nous lui expliquions la consigne suivante.

Il n'y a pas de bonnes, mauvaises réponses, ni une réponse « plus intelligente » que les autres.

Pour chacune des 12 situations, il faut :

- Mettre un chiffre dans les 4 cases (de 1 à 4°)
- Utiliser chaque fois un chiffre différent
- Utiliser les 4 chiffres

Signification des chiffres :

- 1 correspond à tout à fait moi
- 2 correspond à souvent moi
- 3 correspond à souvent moi
- 4 correspond à rarement moi

Il faut répondre le plus spontanément possible et répondre à toutes les questions

#### 1) Quand j'utilise un nouvel appareil

- a) j'analyse soigneusement le mode d'emploi et j'essaie de bien comprendre le fonctionnement
- b) je procède par essais et erreurs, je tâtonne
- c) je me fie à mes intuitions ou je demande à un copain de m'aider
- d) j'écoute et j'observe attentivement les explications de celui qui s'y connaît

#### 2) En général, face à un problème

- a) je prends tout mon temps et j'observe
- b) j'analyse rationnellement le problème, j'essaie de rester logique
- c) je n'hésite pas, je fonce et j'agis
- d) je réagis plutôt instinctivement ; je me fie à mes impressions ou à mes sentiments

### **3) Pour m'orienter dans une ville inconnue lorsque je n'ai pas mon GPS**

- a) je me fie à mon intuition, je « sens » la direction générale. Si cela ne va pas, je demande à quelqu'un
- b) j'observe calmement et attentivement, j'essaie de trouver des points de repère
- c) je me repère rationnellement, je consulte le plan
- d) l'important pour moi c'est de réagir rapidement ; parfois je demande, parfois j'essaie un itinéraire, quitte à revenir sur mes pas

### **4) Si je dois étudier un cours ou une formation**

- a) j'essaie surtout de faire des exercices et de découvrir des applications pratiques
- b) je décortique le document, j'analyse et je raisonne
- c) je prends mon temps, je lis et relis attentivement le document
- d) j'aime travailler en groupe et je m'attache à ce qui me paraît important

### **5) Quand je dois faire un achat important, pour choisir,**

- a) j'observe, j'écoute les avis et contre avis, je prends tout mon temps
- b) je fais confiance à mon intuition
- c) j'essaie de calculer le meilleur rapport qualité prix
- d) ce qui m'intéresse, c'est d'abord de faire un essai

### **6) Le formateur qui me convient le mieux est quelqu'un**

- a) qui expose la discipline avec rigueur, logique et précision
- b) qui fait agir aussi souvent que possible
- c) qui, avant tout, fait appel à l'expérience antérieure
- d) qui a le souci de faire observer et réfléchir avant d'agir

### **7) Pour apprendre une langue étrangère, je préfère**

- a) lire et écouter pour bien m'imprégner de la langue
- b) étudier les bases avant de me lancer
- c) me plonger dans la pratique et exercer le plus tôt possible
- d) improviser : tout dépend de la langue, de mes rencontres et de mon état d'esprit

### **8) Pour préparer une présentation**

- a) je le construis en fonction de mon public. S'il faut, j'improvise sur place
- b) je répète seul ou en comité
- c) je m'inspire d'exemples que j'ai pu observer et apprécier
- d) je construis une structure logique, une analyse et une synthèse



**9) Pour passer de bonnes vacances**

- a) je me décide rapidement, je prépare mon matériel et je fonce
- b) je rassemble une solide documentation, je pèse le pour et contre et je choisis en connaissance de cause
- c) j'aime voir sur place et risquer un peu d'imprévu
- d) j'hésite à me décider, j'ai besoin d'avis

**10) Si je dois lire un livre compliqué**

- a) j'analyse la table des matières, j'essaie d'assimiler chaque élément avant de passer au suivant
- b) je commence par le parcourir pour en avoir un aperçu et pour voir si cela vaut la peine d'insister
- c) je recherche surtout les exemples, les aspects concrets et les applications
- d) je ne me presse pas, je demande parfois des avis, des appréciations

**11) Si je dois préparer un plat**

- a) je m'adresse à quelqu'un qui s'y connaît et je l'observe
- b) j'analyse la recette ; il faut de la rigueur et de la précision
- c) je me fie plutôt à mon expérience et à mon coup d'œil
- d) je me lance, je tâtonne, je goûte... Je mets tout de suite la main à la pâte

**12) Pour choisir une profession**

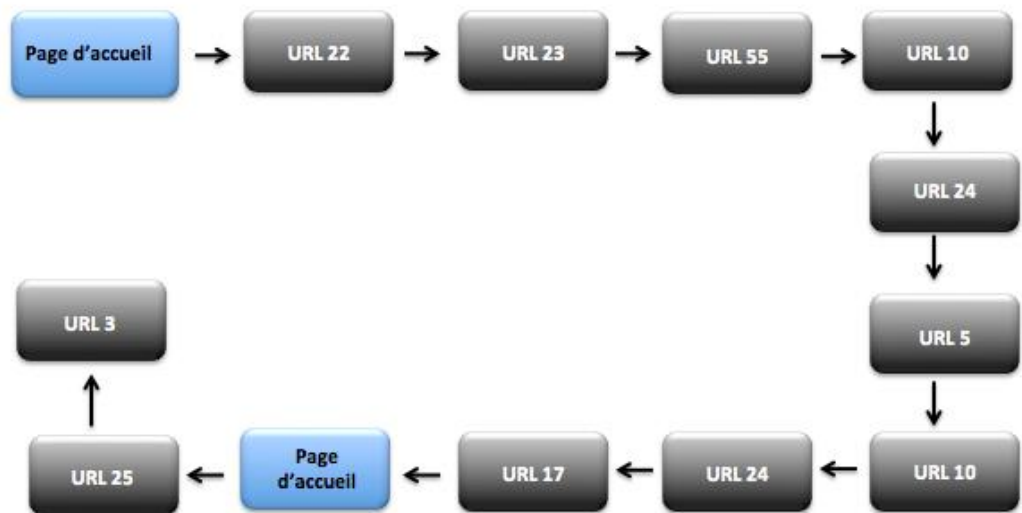
- a) le mieux c'est d'essayer en faisant un stage
- b) pour moi, le plus important est de se fier à ses intuitions et à ses relations
- c) l'idéal est d'aller sur le terrain observer et solliciter des témoignages
- d) l'essentiel est d'analyser méthodologiquement tous les éléments ; aptitudes, débouchés, salaires

## C.1.2 Résultats des analyses

Tableau 52 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 1 –  
Différence entre sexe et genre

Moyennes pour les métriques analysées	Pattern 1 (n=20)	Pattern 2 (n=16)	Pattern 3 (n=8)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	276 417	216 732	56 610
Nombre total de pages visitées pour réaliser la tâche	11	11,6	2,7
Nombre de retours en arrière	1,3	0,8	0
Taux de revisite (en pourcentage)	24%	20%	0%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	47 766	20 618	16 636
Nombre de clics gauche réalisés	2,4	1,1	1
Temps de latence avant la première action (en millisecondes)	26 193	17 357	14 636
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	125,7	19,5	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	124,5	9,7	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>			
Nombre de visites sur cette page	2,2	1,4	1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	41 886	21 360	14 636
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	2,2	1,1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	186	34,2	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	175,7	14,2	0
<b>Comportements sur les autres pages visitées</b>			
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	24 931	19 874	19 245
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1,2	0,8	1,06

## Représentation des parcours types pour la Tâche 1



Chemin le plus représentatif pour le parcours n°1 Tâche 1 (LD = 0,72)



Chemin le plus représentatif pour le parcours n°2 Tâche 1 (LD = 0,66)



Chemin le plus représentatif pour le parcours n°3 Tâche 1 (LD = 0,26)

Tableau 53 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 2 –  
Mesures pour l'égalité des chances

Moyennes pour les métriques analysées	Pattern 1 (n=19)	Pattern 2 (n=11)	Pattern 3 (n=7)	Pattern 4 (n=7)
<b>Efficienc e à la tâche</b>				
Temps total pour accomplir la tâche (en millisecondes)	41 403	75 686	365 079	154 509
Nombre total de pages visitées pour réaliser la tâche	5,1	4,4	16,1	5,1
Nombre de retours en arrière	0,42	0	2,42	0,5
Taux de revisite (en pourcentage)	7,5%	4,8%	26,6%	13,4%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>				
Temps passé sur la page (en millisecondes)	10 848	18 662	35 966	57 379
Nombre de clics gauche réalisés	1	1	1,1	2,8
Temps de latence avant la première action (en millisecondes)	10 664	18 386	34 789	11 005
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0	0	178,1
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0	0	167,5
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>				
Nombre de visites sur cette page	1,05	1,09	1,5	1,7
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	10 585	19 746	44 826	48 457
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	1	1,1	2,1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	22,5	35,4	166,3
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	11,8	35,4	144,4
<b>Comportements sur les autres pages visitées</b>				
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	8 393	18 601	22 347	31 201
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	0,80	0,87	0,93	1,43

## Représentation des parcours types pour la Tâche 2



Chemin le plus représentatif pour le parcours n°1 Tâche 2 (LD = 0,67)



Chemin le plus représentatif pour le parcours n°2 Tâche 2 (LD = 0,25)

Tableau 54 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 4 –  
Demande de patronage et subvention

Moyennes pour les métriques analysées	Pattern 1 (n=13)	Pattern 2 (n=20)	Pattern 3 (n=8)
<b>Efficienc e à la tâche</b>			
Temps total pour accomplir la tâche (en millisecondes)	18 921	35 236	124 002
Nombre total de pages visitées pour réaliser la tâche	2,07	2,35	6,5
Nombre de retours en arrière	0	0	0,37
Taux de revisite (en pourcentage)	0%	0%	12%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>			
Temps passé sur la page (en millisecondes)	6 671	8 370	34 476
Nombre de clics gauche réalisés	1	1	1
Temps de latence avant la première action (en millisecondes)	3 835	7 019	7 528
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>			
Nombre de visites sur cette page	1	1	1,25
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	6 671	8 370	32 501
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	1	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0	0
<b>Comportements sur les autres pages visitées</b>			
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	9 000	15 937	21 075
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1	1,8	1,3

## Représentation des parcours types pour la Tâche 4



Chemin le plus représentatif pour le parcours n°1 Tâche 4 (LD = 0)



Chemin le plus représentatif pour le parcours n°2 Tâche 4 (LD = 0,27)

Tableau 55 : Moyennes relevées pour chaque métriques concernant les patterns de la Tâche 5 –  
Mesures pour l'égalité des chances

Moyennes pour les métriques analysées	Pattern 1 (n=13)	Pattern 2 (n=11)	Pattern 3 (n=11)	Pattern 4 (n=8)
<b>Efficienc e à la tâche</b>				
Temps total pour accomplir la tâche (en millisecondes)	96 622	38 219	353 330	169 741
Nombre total de pages visitées pour réaliser la tâche	7,6	3	20,9	8,5
Nombre de retours en arrière	0,23	0,09	3,27	0,87
Taux de revisite (en pourcentage)	5,5%	1,8%	32,3%	22,7%
<b>Comportements sur la page d'accueil (PA) lors de la première visite (PV)</b>				
Temps passé sur la page (en millisecondes)	5 469	7 492	18 503	8 647
Nombre de clics gauche réalisés	1	1	1,27	1
Temps de latence avant la première action (en millisecondes)	5 284	7 041	11 942	8 463
Distance de scroll descendant pondéré par le nombre de scroll descendant (en pixels)	0	0	4	0
Distance de scroll montant pondéré par le nombre de scroll montant (en pixels)	0	0	0	0
<b>Comportements sur la PA en fonction du nombre de visites sur cette page</b>				
Nombre de visites sur cette page	1	1	1,3	1
Temps moyen passé sur la PA en fonction du nombre de visites sur cette page (en millisecondes)	5 469	7 492	17 884	8 647
Nombre moyen de clics gauche sur la PA en fonction du nombre de visites sur cette page	1	1	1,2	1
Distance totale de scroll descendant en fonction du nombre de scroll descendant sur la PA (en pixels)	0	0	4	0
Distance totale de scroll montant en fonction du nombre de scroll montant sur la PA (en pixels)	0	0	0	0
<b>Comportements sur les autres pages visitées</b>				
Temps moyen passé sur toutes les pages en fonction du nombre de pages visitées (en millisecondes)	12 268	12 315	18 530	19 562
Nombre moyen de clics gauche effectués pendant la tâche en fonction du nombre de pages vues	1,16	0,75	1,43	1,96



## Représentation des parcours types pour la Tâche 5



Chemin le plus représentatif pour le parcours n°1 Tâche 5 (LD = 0,42)



Chemin le plus représentatif pour le parcours n°2 Tâche 5 (LD = 0,58)