



**HAL**  
open science

# Effets de la complexité de l'information sur les intentions et les comportements de mobilité urbaine: Construction d'un outil d'aide au changement de comportement

Thomas Chambon

## ► To cite this version:

Thomas Chambon. Effets de la complexité de l'information sur les intentions et les comportements de mobilité urbaine: Construction d'un outil d'aide au changement de comportement. Architecture, aménagement de l'espace. Université de La Rochelle, 2023. Français. NNT : 2023LAROS017 . tel-04446437

**HAL Id: tel-04446437**

**<https://theses.hal.science/tel-04446437v1>**

Submitted on 8 Feb 2024

**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.



LA ROCHELLE UNIVERSITÉ

*ÉCOLE DOCTORALE EUCLIDE*

LABORATOIRE L3i

**THÈSE** présentée par :

**Thomas Chambon**

soutenue le : **mercredi 30 août 2023**

pour obtenir le grade de : **Docteur en Informatique**

Discipline : **Informatique et Applications**

---

**Effets de la complexité de l'information sur les intentions et les comportements de mobilité urbaine : Construction d'un outil d'aide au changement de comportement.**

---

<b>RAPPORTEURS</b>	Jean-Paul DELAHAYE Nicolas ROUGIER	Professeur des universités Directeur de recherche	Université de Lille INRIA
<b>EXAMINATEURS</b>	Karell BERTET Gilles SÉRÉ DE LANAUZE	Professeure des universités Maître de conférences	La Rochelle Université Université de Montpellier
<b>DIRECTION</b>	Jeanne LALLEMENT Jean-Loup GUILLAUME	Professeure des universités Professeur des universités	La Rochelle Université La Rochelle Université



Communauté  
d'**Agglomération** de  
**La Rochelle**



**THÈSE RÉALISÉE AU** Laboratoire L3i  
La Rochelle Université - Institut LUDI  
Avenue Michel Crépeau  
17042 La Rochelle cedex 01

Tel : +33 5 46 45 82 62

Web : <http://l3i.univ-larochelle.fr>

**SOUS LA DIRECTION DE** Jeanne LALLEMENT Professeure des universités  
Jean-Loup GUILLAUME Professeur des universités  
Arnaud REVEL Professeur des universités

**FINANCEMENT** Allocation de recherche de la Communauté d'Agglomération de La Rochelle



## Résumé

---

En réponse au défi que représente le changement climatique, il est impératif de prendre des mesures rapides visant à réduire notre impact et par extension nos émissions de CO<sub>2</sub>. La mobilité, qui représente près d'un quart des émissions mondiales, est un facteur central à prendre en compte pour atteindre cet objectif. Il est alors essentiel que chacun de nous s'engage individuellement à modifier son comportement en adoptant de nouvelles habitudes de déplacement plus respectueuses de l'environnement. Ce changement pouvant être perçu comme long et fastidieux, il doit être accompagné et expliqué avec une grande régularité. Cet accompagnement intensif, difficilement réalisable il y a encore quelques années, est maintenant rendu possible grâce à la démocratisation de moyens de communication rapides et performants tels que les applications mobiles sur téléphones portables ou les réseaux sociaux en ligne. Grâce à cette évolution, l'information est alors devenue une composante influente dans nos actions et nos prises de décision. Devant la multitude d'informations reçues et transmises quotidiennement, la personnalisation et l'adaptation de l'information sont obligatoires pour devenir un potentiel déclencheur d'un changement de comportement durable. La question se pose alors de savoir quelle est la nature et la forme de l'information à transmettre pour maximiser l'adoption d'un comportement de mobilité durable.

Les travaux de recherche menés dans le cadre de cette thèse ont pour objectif de contribuer à trois problématiques liées à la nature de l'information, à sa perception et à sa transmission à l'aide d'une application mobile comme outil d'aide au changement. Nos contributions sur la nature de l'information et plus particulièrement sa complexité sont destinées à faciliter la compréhension de l'information. Nous avons proposé une nouvelle méthode de classement basée sur la théorie algorithmique de l'information (complexité de Kolmogorov). Ces travaux nous ont permis de mieux comprendre l'impact de la complexité de l'information chez les individus. Nous avons également porté une attention particulière à la perception de l'information en analysant les caractéristiques exerçant une influence sur le processus de décision chez l'humain. À partir de cette analyse, nous avons proposé une étude exploratoire d'oculométrie permettant de mettre en lumière les effets de la complexité et de la position de l'information sur le choix d'un individu dans une tâche de mobilité urbaine. Notre recherche sur la création d'un outil d'aide au changement de comportement, nous a mené à la conception et au développement d'une application mobile de suivi de sa propre consommation de CO<sub>2</sub>. L'objectif de cette application est de comprendre les facteurs d'acceptation de ces outils et l'influence de la complexité de l'information sur ces facteurs. Cette contribution constitue également une expérimentation en conditions réelles de l'ensemble de nos travaux, dans un contexte de recherche encore peu exploité, la mobilité urbaine.

L'originalité de nos travaux réside également dans l'approche pluridisciplinaire que nous avons appliquée à la majorité des problématiques présentées. Cette thèse est le résultat de méthodologies et de recherches issues de trois domaines scientifiques différents : l'informatique, les sciences de gestion et les neurosciences computationnelles.

**Mots-clés :** mobilité urbaine, comportement éco-responsable, théorie algorithmique de l'information, complexité de Kolmogorov, oculométrie, application mobile d'autosuiwi, systèmes d'information durables.



---

**Effects of information complexity on urban mobility intentions and behaviors : Construction of a tool to help change behavior.**

---



## Abstract

---

In response to the challenge of climate change, it is imperative to act promptly to reduce our impact and by extension our CO<sub>2</sub> emissions. Mobility, which accounts for nearly a quarter of global emissions, is a central factor in achieving this goal. It is therefore essential that all of us individually commit to changing our behavior by adopting new, more environmentally responsible transportation habits. This change can be perceived as long and tedious and must be accompanied and explained with high regularity. This intensive support, which was difficult to achieve a few years ago, is now possible due to the democratization of fast and efficient communication means such as cell phone applications or online social networks. With this evolution, information has become an influential component in our actions and decision making. Given the multitude of information received and transmitted on a daily basis, the personalization and adaptation of information are mandatory to become a potential trigger for a sustainable change in behavior. The question then arises as to the nature and form of the information to be transmitted in order to maximize the adoption of a sustainable mobility behavior.

The research work conducted in the framework of this thesis aims to contribute to three issues related to the nature of information, its perception and its transmission using a mobile application (tool for change). Our contributions on the nature of information and more particularly its complexity are intended to facilitate the understanding of information. Based on the algorithmic theory of information and more particularly the complexity of information (Kolmogorov complexity), a new classification method has been proposed. This work has allowed us to better understand the impact of information complexity on individuals. We also paid particular attention to the perception of information by analyzing the characteristics that influence the decision process in humans. From this analysis, an exploratory eye-tracking study was carried out to highlight the effects of information complexity and position on an individual's choice in an urban mobility task. Our research on the creation of a behavior change tool led us to the design and development of a mobile application to monitor one's own CO<sub>2</sub> consumption. The objective of this application is to understand the factors of acceptance of these tools and the influence of the complexity of the information on these factors. This contribution is also an experimentation in real conditions of all our work, in a research context still little exploited, the urban mobility.

The originality of our research also resides in our multidisciplinary approach to the set of issues presented. This thesis is the result of methodologies and research from three different scientific fields : computer science, management science and computational neuroscience.

**Keywords :** urban mobility, eco-responsible behavior, algorithmic information theory, Kolmogorov complexity, eye tracking, mobile self-tracking application, green information systems.



## Remerciements

---

Je souhaite exprimer ma gratitude envers Jean-Paul Delahaye et Nicolas P. Rougier pour avoir accepté d'être les rapporteurs de ma thèse. Je tiens également à remercier chaleureusement Gilles Séré de Lanauze et Karell Bertet d'avoir accepté de faire partie de mon jury en tant qu'examineurs et pour leur suivi tout au long de mes années de thèse.

Je tiens à exprimer ma reconnaissance envers mes parents, sans lesquels je ne serais pas arrivé là où je suis aujourd'hui, ainsi que pour leur soutien indéfectible durant ces quatre années.

Je souhaite exprimer ma profonde gratitude envers Jeanne Lallement, Arnaud Revel et Jean-Loup Guillaume pour leur confiance et pour m'avoir guidé tout au long de ce difficile parcours qu'est le doctorat. Je souhaite remercier mon compère, Ulysse Soulat, dans cette aventure du projet Ecomob avec qui j'ai grandement apprécié travailler.

Je souhaite remercier tout particulièrement Marie-Neige pour sa capacité à me supporter au quotidien et à toujours voir le positif dans chaque situation.

Je tiens à remercier Guillaume Bernard pour la qualité de nos échanges scientifiques et techniques au cours de toutes ces années, ainsi que pour ses précieux conseils.

Je remercie Damien pour ses conseils précieux durant ma courte carrière d'enseignant et Julien pour les moments partagés lors de nos pauses.

Je remercie Mélanie pour sa disponibilité à nous écouter (et à supporter nos plaintes) ainsi que pour son soutien au quotidien au laboratoire.

Enfin, je souhaite remercier toutes les personnes que j'ai eues l'occasion de côtoyer au sein du laboratoire ou dans mon « deuxième bureau ».



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Comportements de mobilité urbaine : un changement urgent ?	9
1.2	Approche pluridisciplinaire : le projet Ecomob	11
1.3	Problématiques scientifiques	12
1.4	Contributions	14
1.5	Organisation de la thèse	15
<b>2</b>	<b>Complexité de l'information</b>	<b>17</b>
2.1	Introduction	19
2.2	État de l'art	20
2.2.1	L'information du point de vue de la transmission	20
2.2.2	L'information d'un point de vue génératif	24
2.2.3	Notion de similarité et distances informationnelles	27
2.2.4	Structure de l'information : la profondeur logique de Bennett	30
2.3	Information Complexity Ranking (ICR)	31
2.3.1	Preuve	32
2.4	Expérimentations	33
2.4.1	Benchmark	34
2.4.1.1	Protocole	34
2.4.1.2	Résultats	35
2.4.2	Dataset SAVOIAS	37
2.4.2.1	Description	37
2.4.2.2	Protocole	37
2.4.2.3	Résultats	40
2.4.3	Expérimentation avec la profondeur logique de Bennett	42
2.4.3.1	Protocole	43
2.4.3.2	Résultats	43
2.5	Conclusion	44
<b>3</b>	<b>Traitement de l'information</b>	<b>47</b>
3.1	Introduction	49
3.2	État de l'art	50
3.2.1	Attention visuelle du point de vue psychologique	50

3.2.2	Attention visuelle du point de vue computationnel . . . . .	54
3.2.3	Perception de la complexité visuelle . . . . .	57
3.2.4	Processus de décision : Le choix . . . . .	59
3.2.5	Position de l'information et comparaison par paire . . . . .	61
3.3	Étude d'oculométrie - protocole et hypothèses . . . . .	62
3.3.1	Protocole expérimental . . . . .	62
3.3.2	Hypothèses . . . . .	66
3.4	Résultats . . . . .	68
3.4.1	Effets de la complexité de l'information . . . . .	68
3.4.2	Effets sur le choix des individus . . . . .	69
3.4.3	Effets de la position de l'information . . . . .	70
3.4.4	Effets sur le processus de décision . . . . .	71
3.5	Conclusion . . . . .	73
<b>4</b>	<b>Acceptation de l'information</b>	<b>77</b>
4.1	Introduction . . . . .	79
4.2	État de l'art . . . . .	81
4.2.1	Technologies persuasives ou d'autosuiivi . . . . .	81
4.2.2	Les systèmes d'information durables . . . . .	83
4.2.3	Changement de comportement vers une consommation plus durable . . . . .	84
4.3	Contribution : Application mobile (Eco-Mobilité) . . . . .	87
4.3.1	Contexte . . . . .	87
4.3.2	Développement et fonctionnement de l'application . . . . .	88
4.3.2.1	Suivi des utilisateurs . . . . .	89
4.3.2.2	Interface utilisateur . . . . .	90
4.3.2.3	Déploiement de l'application . . . . .	91
4.3.3	Protocole expérimental . . . . .	91
4.3.3.1	Méthodologie de l'étude préliminaire . . . . .	92
4.3.3.2	Méthodologie de l'étude en conditions réelles . . . . .	93
4.3.4	Hypothèses . . . . .	93
4.3.4.1	Étude préliminaire . . . . .	94
4.3.4.2	Expérience en conditions réelles . . . . .	95
4.3.5	Résultats . . . . .	95
4.3.5.1	Étude préliminaire . . . . .	95
4.3.5.2	Étude en conditions réelles . . . . .	97
4.4	Conclusion . . . . .	100
<b>5</b>	<b>Conclusion</b>	<b>103</b>
5.1	Problématiques et contributions . . . . .	104
5.2	Limites . . . . .	105
5.3	Perspectives . . . . .	106
5.3.1	Amélioration de notre méthode de la mesure de la complexité . . . . .	106
5.3.2	Nouvelles études d'oculométrie . . . . .	107
5.3.3	Nouvelles versions de l'application Ecomobilité . . . . .	107

5.3.4 Vers une personnalisation de l'information . . . . . 109



# Table des figures

1.1	Total des émissions de CO <sub>2</sub> mondiales issues du transport en 2019. Les émissions imputables à l’aviation et au transport maritime ne sont pas pris en compte [Ritchie et al., 2022]. . . . .	10
1.2	Exemples de correspondances possibles pour quantifier le CO <sub>2</sub> émis. . . . .	12
1.3	Positionnement des trois thèses du projet Ecomob. . . . .	13
1.4	Chapitres principaux de cette thèse. . . . .	16
2.1	Schéma du modèle de Shannon et Weaver (1949) . . . . .	21
2.2	Codage de Huffman pour la chaîne de caractères $s_1$ . . . . .	22
2.3	Codage de Huffman pour la chaîne de caractères $s_2$ . . . . .	22
2.4	Exemple d’information de type texte générée . . . . .	35
2.5	Exemple de nombre (4 chiffres) généré . . . . .	35
2.6	Exemple de pastille de couleur générée . . . . .	35
2.7	Exemples d’images du dataset SAVOIAS pour chaque catégorie. . . . .	38
2.8	Exemple de deux images du dataset (à gauche : IMAGE A et à droite : IMAGE B) pour notre méthode Information Complexity Ranking (ICR). . . . .	39
2.9	Exemples d’images dans la catégorie <i>Suprematism</i> et <i>Visualizations</i> . . . . .	42
3.1	Représentation simplifiée des deux étapes du modèle de la Feature Integration Theory. . . . .	51
3.2	Représentation théorique de la FIT selon [Kristjansson and Egeth, 2020] . . . . .	52
3.3	Taxonomie du domaine de la modélisation de l’attention visuelle issue de [Borji and Itti, 2010] . . . . .	54
3.4	Architecture générale du modèle d’Itti, Koch et Niebur issue de [Itti et al., 1998] . . . . .	56
3.5	Courbe de Wundt . . . . .	58
3.6	Instructions et contexte de l’étude . . . . .	63
3.7	Stimulus avec un niveau de complexité faible . . . . .	63
3.8	Stimulus avec un niveau de complexité modéré . . . . .	63
3.9	Stimulus avec un niveau de complexité élevé . . . . .	63
3.10	Exemple d’une série d’écrans d’accueil . . . . .	65
3.11	Modèle d’hypothèses de l’étude d’oculométrie. . . . .	67
3.12	Valeur moyenne des fixations pour chaque série . . . . .	74

3.13	Valeur moyenne des revisites pour chaque série . . . . .	74
3.14	Valeur moyenne du temps de visionnage pour chaque série . . . . .	74
4.1	Adaptation du modèle UTAUT2 [Venkatesh et al., 2012] à notre contexte de recherche. L'adaptation (en rouge) correspond à l'ajout de la complexité de l'information comme variable explicative de l'intention comportementale et du comportement d'utilisation. . . . .	83
4.2	Écrans d'accueil avec les trois niveaux de complexité (simple, modéré et élevé, de gauche à droite). . . . .	88
4.3	Temporalité du déploiement de l'étude préliminaire et de celle en conditions réelles. . . . .	91
4.4	Modèle d'hypothèses des différentes études. . . . .	94

# Liste des tableaux

2.1	Performance des algorithmes de compression en calculant le taux de compression et le respect pour chaque propriété requise pour satisfaire le concept d'un compresseur normal. . . . .	36
2.2	Nombre d'images générées pour tester ICR. Le facteur 2 appliqué dans la colonne <i>Nombre de comparaisons</i> représente la concaténation horizontale et verticale de chaque image. . . . .	39
2.3	Comparaison entre ICR et NCD par rapport à la perception humaine. Les valeurs de NCD et d'ICR présentées sont la moyenne des valeurs obtenues avec les algorithmes PNG, PPM et LZ77. . . . .	41
2.4	Comparaison entre ICR et différentes mesures de la complexité visuelle dans la littérature, par rapport à la perception humaine. Les valeurs des cinq premières colonnes sont reprises de la Table 3 dans l'article [Saracee et al., 2018]. . . . .	41
2.5	Comparaison des performances d'ICR et de la profondeur logique de Bennett par rapport à la perception humaine. Les valeurs de la profondeur logique de Bennett et d'ICR présentées sont la moyenne des valeurs obtenues avec les algorithmes PNG, PPM et LZ77, en utilisant le temps de décompression. . . . .	44
3.1	Valeurs des mesures d'oculométrie dans l'étude. Les lignes correspondent aux valeurs moyennes pour chaque série et chaque participant. . . . .	68
3.2	Valeurs des métriques d'oculométrie des participants en fonction de la position de l'image. . . . .	70
3.3	Choix de l'écran d'accueil pour chaque série. . . . .	73
4.1	Effet de la complexité de l'information sur l'acceptabilité de l'application. La première ligne correspond à la moyenne des valeurs issues de l'échelle de Likert associée. Il est à noter que les participants ont vu un seul et unique niveau de complexité. La deuxième ligne correspond à l'estimation de l'erreur type de chaque groupe. . . . .	96

- 4.2 Effet des différents niveaux de complexité de l'information les uns par rapports aux autres sur l'acceptabilité de l'application. La première colonne présente la différence entre la valeur moyenne d'un niveau de complexité en comparaison à un autre niveau. La deuxième colonne correspond à l'estimation de l'erreur-type de chaque groupe. La troisième colonne (P-valeur) indique le niveau de signification de nos données. La valeur  $F$  correspond à la valeur du test de Fisher qui correspond au ratio de la variance inter-classes sur la variance intra-classes. . . . . 97
- 4.3 Effet de l'acceptabilité de l'application sur l'intention d'utiliser des moyens de transport qui émettent moins de CO<sub>2</sub> que la voiture. Les valeurs  $R$  et  $R^2$  représentent le coefficient de corrélation linéaire et le coefficient de détermination. La valeur  $F$  correspond à la valeur du test de Fisher qui représente au ratio de la variance inter-classes sur la variance intra-classes. La valeur  $B$  représente l'ordonnée à l'origine de l'équation du modèle. La valeur  $T$  ou test de Student représente la comparaison des moyennes des groupes entre complexité et acceptabilité. . . . . 97
- 4.4 Résultats de l'expérimentation en conditions réelles. Statistiques descriptives de l'effet de la complexité sur le comportement. La première ligne correspond à la moyenne des valeurs issues de l'échelle de Likert associée. Il est à noter que les participants ont vu un seul niveau de complexité. La deuxième ligne correspond à l'estimation de l'écart type de chaque groupe. 98
- 4.5 Résultats de l'expérimentation en conditions réelles. Effets de la complexité visuelle sur le comportement de mobilité urbaine. La première colonne présente la différence entre la valeur moyenne d'un niveau de complexité en comparaison à un autre niveau. La deuxième colonne correspond à l'estimation de l'erreur-type de chaque groupe. La troisième colonne P-valeur indique le niveau de signification de nos données. . . . . 99
- 4.6 Résultats de l'expérimentation en conditions réelles. Effet de l'intention comportementale sur le comportement de mobilité urbaine. Les valeurs  $R$  et  $R^2$  représentent le coefficient de corrélation linéaire et le coefficient de détermination. La valeur  $F$  correspond à la valeur du test de Fisher qui correspond au ratio de la variance inter-classes sur la variance intra-classes. La valeur  $B$  correspond à l'ordonnée à l'origine de l'équation du modèle. La valeur  $T$  ou test de Student correspond à une comparaison des moyennes des groupes entre complexité et l'intention comportementale. . . 99

# Chapitre 1

## Introduction

---

### Sommaire

---

<b>1.1</b>	<b>Comportements de mobilité urbaine : un changement urgent ?</b>	<b>9</b>
<b>1.2</b>	<b>Approche pluridisciplinaire : le projet Ecomob . . . . .</b>	<b>11</b>
<b>1.3</b>	<b>Problématiques scientifiques . . . . .</b>	<b>12</b>
<b>1.4</b>	<b>Contributions . . . . .</b>	<b>14</b>
<b>1.5</b>	<b>Organisation de la thèse . . . . .</b>	<b>15</b>

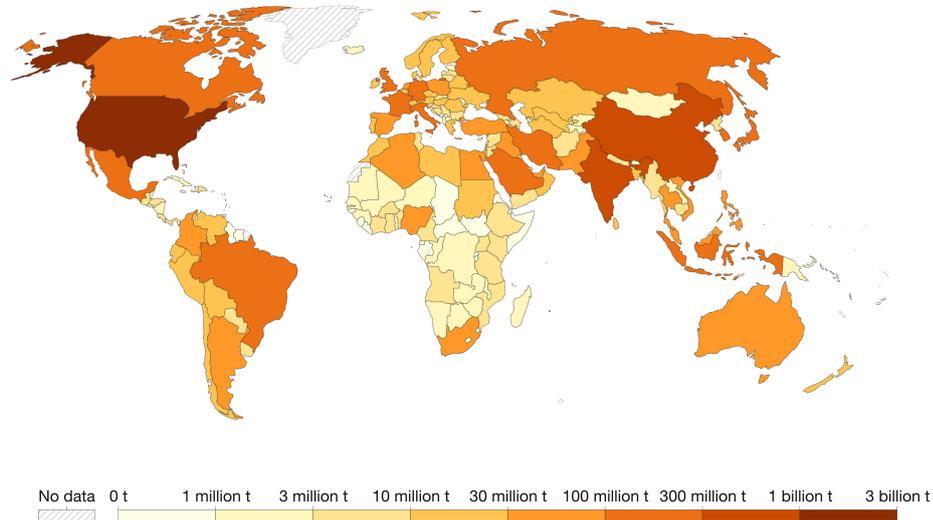
---

### 1.1 Comportements de mobilité urbaine : un changement urgent ?

Le changement climatique est l'un des principaux enjeux du siècle : les scénarios les plus optimistes prédisent une augmentation des températures de plusieurs degrés en quelques décennies, avec des conséquences catastrophiques. Pour endiguer ce fléau, chacun d'entre nous doit adapter son comportement et réduire ses émissions de gaz à effet de serre. Parmi les facteurs contribuant à ces émissions, la mobilité représente environ 14% du total des émissions et joue donc un rôle central [et L.A. Meyer, 2014]. On constate également qu'à l'échelle mondiale les émissions de CO<sub>2</sub> dues au transport sont très disparates et se concentrent en majorité dans les pays développés (voir Figure 1.1). Au delà de l'aspect environnemental, les émissions de CO<sub>2</sub> provoquent d'autres problèmes impactant plus directement les personnes. Selon l'OMS le changement climatique est responsable d'au moins 150000 décès par an, chiffre qui devrait doubler d'ici 2030 [Organization, 2023]. Toujours selon l'OMS, en 2019, 99% de la population mondiale vivaient dans des zones ne respectant pas les seuils préconisés relatifs à la qualité de l'air [Organization, 2021]. Cependant des solutions existent telles que la promotion des transports publics ou les déplacements actifs qui pourrait à la fois réduire les émissions de carbone et alléger le fardeau de la pollution domestique et ambiante. Cette réduction permettrait d'éviter

sept millions de décès prématurés par an [Organization, 2021].

CO<sub>2</sub> emissions from transport, 2019



Source: Our World in Data based on Climate Analysis Indicators Tool (CAIT).  
Note: International aviation and shipping emissions are not included.

OurWorldInData.org/transport • CC BY

FIGURE 1.1 – Total des émissions de CO<sub>2</sub> mondiales issues du transport en 2019. Les émissions imputables à l’aviation et au transport maritime ne sont pas pris en compte [Ritchie et al., 2022].

Même s’il existe de nombreuses sources d’émissions de CO<sub>2</sub>, celles associées au transport individuel restent importantes, mais peuvent être réduites. Il convient de souligner qu’une majorité d’usagers de ce type de transport veulent modifier leur comportement. Toutefois, cette volonté de changement exige une sensibilisation et une connaissance précise de l’impact CO<sub>2</sub> de nos propres choix de mobilité. Il est donc primordial de trouver des mécanismes pour inciter les personnes à adopter un comportement plus responsable. Cette incitation au changement ne repose en aucun cas sur la contrainte ou la culpabilisation, mais sur un changement « en douceur », en accompagnement de certaines politiques publiques qui peuvent être incitatives ou contraignantes. On peut citer la théorie du paternalisme libéral ou nudge (« coup de pouce », en français) [Thaler and Sunstein, 2008] comme un exemple d’incitation au changement « en douceur » afin de maximiser l’adoption de ce comportement sur le long terme. Le but de cette théorie est de mettre en évidence les choix « les plus pertinents » face à un objectif à atteindre (par exemple l’amélioration de la sécurité routière, l’incitation à faire du sport, etc.). Avec la mise en oeuvre de tels outils, le choix est suggéré et n’est jamais imposé, ce qui influence directement la motivation des individus. Cependant, il est important de préciser que cette théorie repose sur un mécanisme non conscient, ce qui peut poser des problèmes éthiques.

À titre de comparaison, notre approche de recherche est fondée sur un mécanisme de sensibilisation aux actions individuelles, de faire prendre conscience aux gens de l'impact de leur propre comportement. Il sera ainsi possible, grâce aux informations communiquées à l'utilisateur, de l'inciter à modifier son comportement de mobilité.

Pour encourager l'évolution de ces comportements, les chercheurs en management et en sciences du comportement ont souligné l'importance de l'information fournie [Lehto and Oinas-Kukkonen, 2015, Bruck et al., 2012]. Le rôle positif de l'information a été démontré dans différents cas, comme dans celui de l'alimentation pour prévenir l'obésité infantile par exemple [Langarizadeh et al., 2021]. Cependant, l'impact de cette information peut varier en fonction de plusieurs critères propres à chaque individu, comme la nature de l'objectif à atteindre, la motivation, le gain perçu ou encore la culture et les expériences vécues. Au vu du nombre élevé de variables qui peuvent influencer ce changement de comportement, il est important de trouver des solutions qui permettent d'orienter ce choix dès les premières étapes cognitives, comme l'attention visuelle guidant l'information vers le cerveau pour la traiter mais également tout au long du processus de décision.

Au delà de la nécessité d'informer les gens sur leurs comportements, une difficulté propre au contexte de la mobilité apparaît, celle de la connaissance de l'information et sa représentation mentale. En effet, qu'est ce qu'un kilogramme de CO<sub>2</sub> ? Le simple fait d'émettre du CO<sub>2</sub> étant une action impalpable et indirecte, cela ne facilite pas sa compréhension. Tandis que la représentation d'un kilogramme lorsque l'on se pèse, est une action plus visible et directe. Pour pallier cette difficulté, un travail supplémentaire devra être effectué sur l'information transmise. Cette information pourrait intégrer des éléments de comparaison simples comme un indicateur possédant un code couleur (par exemple de type Nutriscore [Julia and Herberg, 2017]) (voir Figure 1.2a). Ce type d'indicateur étant utilisé au quotidien par un grand nombre de personnes, cela facilite la représentation de l'information qu'il contient. D'autres équivalences facilement compréhensibles peuvent être mise en place [de la transition écologique , ADEME] (voir Figure 1.2b).

## 1.2 Approche pluridisciplinaire : le projet Ecomob

L'originalité de nos travaux réside en partie sur l'approche pluridisciplinaire que nous adoptons. Notre recherche est regroupée au sein d'un projet nommé Ecomob [Projet Ecomob]. Ce projet débuté en 2019, est l'association de trois disciplines, **les sciences de gestion, les neurosciences computationnelles et l'informatique**. Au total, ces travaux mobilisent dix chercheurs répartis sur deux sites, La Rochelle et Bordeaux. Le projet porte également trois thèses réparties dans chaque domaine. Chaque discipline a des objectifs qui lui sont propres, cependant ils s'articulent autour d'une démarche globale d'exploration de l'impact de l'information sur le comportement de mobilité individuelle. Du point de vue des sciences de gestion, la recherche s'inscrit dans un contexte de communication sociale, avec l'étude du cadrage de l'information et de l'effet de sa complexité sur les intentions et les comportements de mobilité urbaine. Dans le domaine des neurosciences computationnelles, la modélisation informatique de certains mécanismes

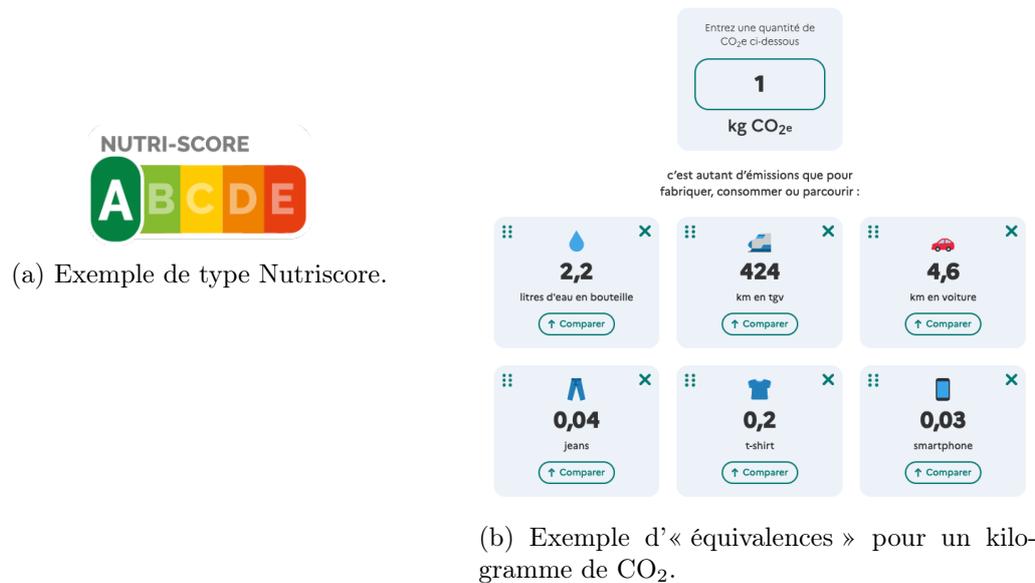


FIGURE 1.2 – Exemples de correspondances possibles pour quantifier le CO<sub>2</sub> émis.

cérébraux présents lors de la prise de décision permet de mieux comprendre la représentation de l'information dans le cerveau dans le but d'acquiescer un comportement spécifique par rapport à un contexte particulier (voir Figure 1.3).

Du point de vue de cette thèse, notre recherche a pour objectif d'inciter les individus à changer de comportement de mobilité à l'aide d'une application mobile. Pour cela nous avons concentré nos efforts sur l'information transmise à l'utilisateur et plus particulièrement sur sa complexité. Nous avons défini trois étapes principales dans nos travaux : la nature de l'information (voir Section 2), la perception de l'information (voir Section 3) et la transmission de l'information (voir Section 4).

### 1.3 Problématiques scientifiques

Compte tenu de la nécessité d'informer les gens au sujet de leur mobilité, il est primordial de mieux comprendre l'impact de l'information sur le changement de comportement. Il est aussi important de concentrer les travaux de recherche sur l'acceptation de l'information pour maximiser son effet sur le changement de comportement. Tout au long de ce manuscrit, le terme complexité couvrira deux notions intimement liées : la complexité visuelle, qui traduit la difficulté de la perception visuelle, comme le nombre d'éléments, leur densité, leur diversité, etc. et la complexité algorithmique qui est une mesure mathématique de la complexité visuelle.

La première étape consiste à mieux comprendre l'impact de l'information, pour cela il faut trouver des méthodes de mesure de la compréhension humaine de l'information. Une mesure possible de cette compréhension est la notion de complexité algorithmique.

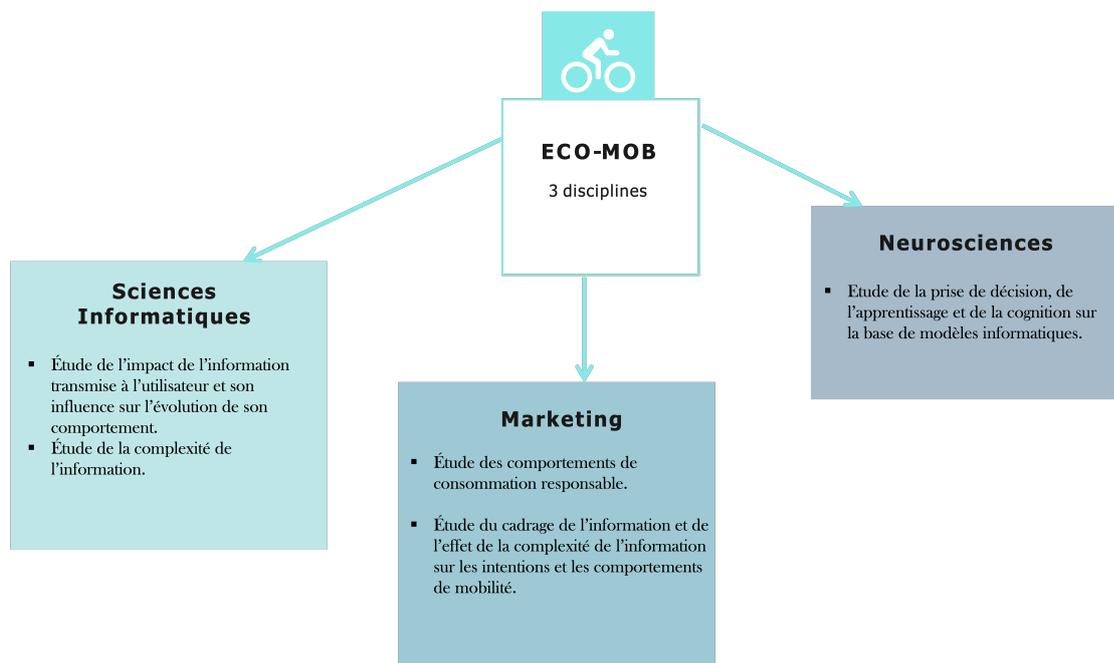


FIGURE 1.3 – Positionnement des trois thèses du projet Ecomob.

Plusieurs travaux prenant place dans des domaines différents apportent un éclairage sur l'effet de la complexité visuelle sur la compréhension chez l'individu. Le contexte d'incitation au changement de comportement nous oblige à transmettre plusieurs informations en même temps pour informer de la manière la plus exhaustive possible l'individu de son comportement actuel. Cependant, les méthodes de bases proposées dans le domaine de la mesure de la complexité algorithmique de l'information ne prennent en compte qu'un seul objet à la fois, ce qui ne correspond qu'en partie à notre contexte de recherche. L'objectif de cette première étape de recherche est donc de classer plusieurs informations en fonction de leur complexité algorithmique afin de sélectionner celles qui maximisent la compréhension. Les informations transmises étant centrées sur le même sujet, à savoir les émissions de CO<sub>2</sub>, elles devront prendre différentes formes pour attirer l'individu. Il est par conséquent important d'inclure la notion de similitude dans cette classification afin de savoir si la forme de l'information a une influence. Pour cela l'utilisation d'une mesure combinant ces deux aspects est nécessaire.

Une fois la mesure de la complexité de l'information acquise, il est pertinent de comprendre comment elle est perçue afin de maximiser son acceptation au sein du processus de décision. Il est donc primordial de développer des applications d'aide au changement basées, dès leur conception, sur la notion de complexité visuelle de l'information transmise. De plus, cette étape d'intégration conditionne sur le long terme l'adoption ou le rejet d'un comportement. La mise au point de ce type d'outils entraînant parfois des coûts importants, il serait intéressant de pouvoir intégrer des informations qui sont comprises et traitées par l'utilisateur. L'objectif serait de pouvoir disposer de niveaux de complexité

visuelle déjà testés auprès d'un échantillon de personnes avant leur intégration finale.

La dernière étape de ce processus d'information pour inciter au changement de comportement consiste à trouver un vecteur de communication adapté. À l'heure actuelle, l'application sur smartphone reste le vecteur le plus performant, en termes de temps d'utilisation et de taux d'utilisation, 77% de la population française détenant un smartphone en 2022 [Stéphane Legleye, 2022]. Néanmoins, il est à noter que l'application doit être acceptée sur le long terme par l'utilisateur pour qu'elle provoque un quelconque bénéfice. En parcourant la littérature consacrée aux applications d'autosuiivi (self-tracking) et aux systèmes d'information durables (SI verts), on constate que la majorité de ces dispositifs sont centrés autour d'un intérêt individuel de type éducatif ou lié à la santé. Les résultats de ces applications dans un contexte environnemental sont mitigés voire faibles. On peut alors se demander si l'information transmise ne serait pas en cause dans ces résultats. Pour tester cette hypothèse, il faudrait intégrer la notion de complexité de l'information dans un outil d'aide au changement de comportement. À terme, cette intégration permettra de savoir si la complexité est une variable incitatrice au changement de comportement, ou pas. De plus, dans un contexte de mobilité urbaine, beaucoup de facteurs externes influencent le choix du moyen de transport. Il peut d'agir, par exemple, des conditions météorologiques, des habitudes ou de la possession d'un moyen de transport personnel. La complexité de l'information pourrait-elle être un atténuateur de ces facteurs ?

## 1.4 Contributions

Trois contributions majeures ont émergé de notre recherche améliorant la compréhension de l'impact de la complexité de l'information mais également sa perception auprès des individus. Nos travaux ont également produits des méthodes et des outils qui facilitent la transmission de l'information auprès d'un large public. Nos contributions se déclinent comme suit :

1. **Une méthode générique de classement de l'information par sa complexité**, dans le but résoudre le problème qui se pose concerne la mesure et le classement de plusieurs informations. Aucune des solutions existantes ne permettaient de retranscrire à la fois la complexité intrinsèque d'une information et son taux de similarité par rapport à une autre information. Nous avons construit une nouvelle méthode de classement se basant sur la notion de similarité entre deux objets, à l'aide de la distance de compression normalisée (Normalized Compression Distance ou NCD) et la complexité de Kolmogorov.
2. **Une étude d'oculométrie**, menée auprès de 24 personnes, dans le but de présélectionner le niveau de complexité des stimuli à intégrer dans un outil d'incitation au changement. La mise en oeuvre de cette étude vient du fait que les concepteurs et les développeurs d'applications d'aide au suivi portent peu d'attention à la complexité du message qu'ils transmettent. Cette étude est donc une approche exploratoire afin de mieux comprendre les mécanismes de l'attention visuelle, de la perception de la complexité et du choix.

3. **Une application mobile d'aide au changement de comportement**, permettant de tester en conditions réelles l'influence de la complexité de l'information sur les comportements de mobilité au quotidien. Cette application proposée à plus de 1000 étudiants et utilisée par 186 étudiants de La Rochelle Université durant trois mois nous a permis de mieux comprendre l'impact de l'information sur le choix du moyen de transport. Les données et les résultats issus de cette application pourront servir de guide pour les concepteurs d'outils d'aide au changement durable.

## 1.5 Organisation de la thèse

Ce manuscrit de thèse est organisé en cinq chapitres. Le présent chapitre introduit notre contexte de recherche ainsi que le contexte pluridisciplinaire dans lequel nos travaux sont ancrés. Nous faisons ensuite le point sur les problématiques scientifiques soulevées et enfin nous listons nos contributions. Le chapitre 2 présente nos travaux sur l'étude de la nature de l'information et sa quantification grâce à la notion de complexité. Différentes méthodes de la mesure de la complexité de l'information et la notion de similarité entre deux objets (au sens algorithmique) sont détaillées. Dans le même chapitre, une nouvelle méthode de mesure de la complexité de plusieurs pièces d'information est également proposée. Le chapitre 3 est consacré à l'étude de la perception humaine et de l'attention visuelle en rapport avec la complexité d'un ou plusieurs stimulus. Les notions de position de l'information et de phases d'attention sont aussi traitées pour établir une méthodologie de mesure fiable pour construire le stimuli le plus incitatif au changement de comportement. Dans le chapitre 4, nous proposons une expérimentation en conditions réelles reposant sur l'utilisation d'une application mobile, appelée Ecomobilité, que nous avons développée et déployée dans le cadre de notre projet. Cette application basée sur les concepts de systèmes d'information durable et d'applications d'autosuiwi permet de mesurer l'influence de la complexité de l'information sur les comportements de mobilité des individus. Le dernier chapitre de ce manuscrit conclut cette thèse en présentant les liens entre les problématiques et les contributions de nos travaux. Il présente également certaines limites de nos études dans un objectif d'amélioration continue. Pour finir, les perspectives de nos travaux à plus ou moins long terme sont énoncées.

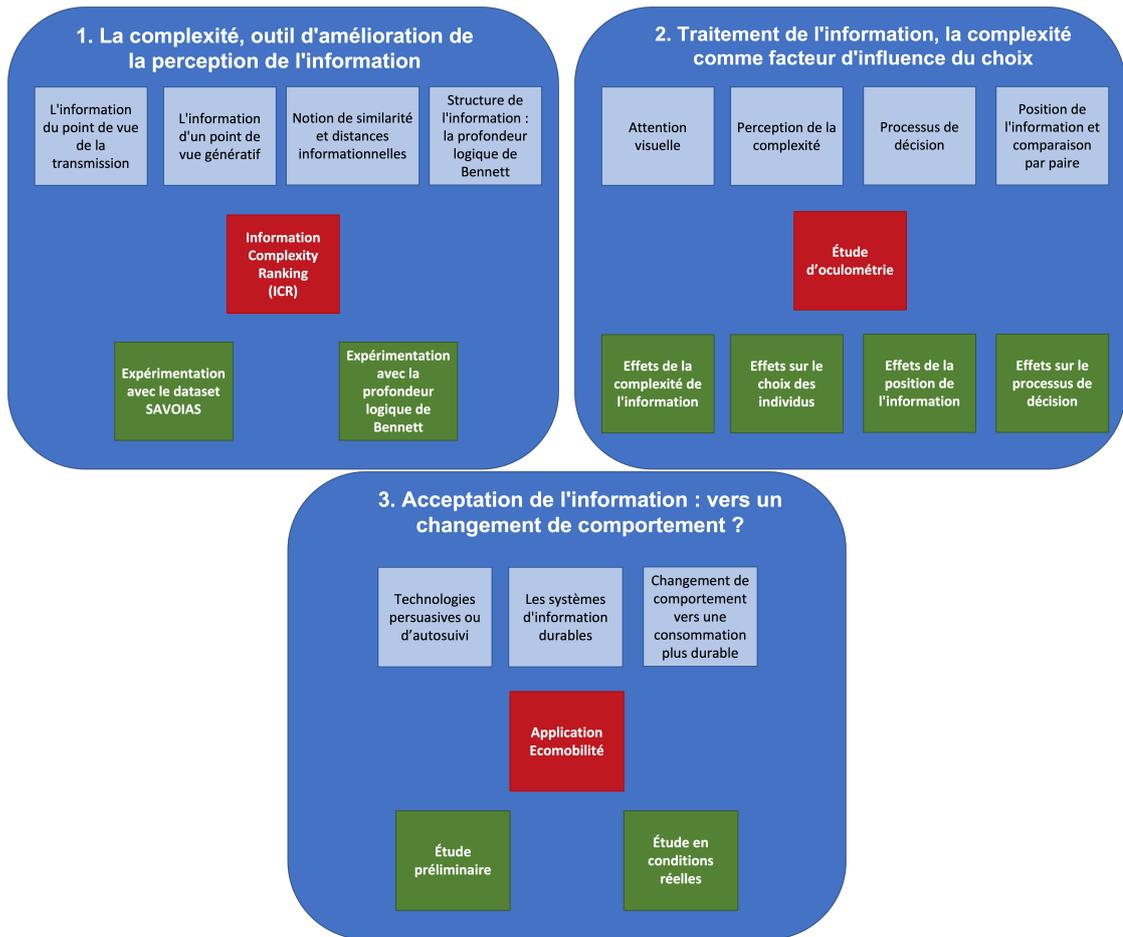


FIGURE 1.4 – Chapitres principaux de cette thèse.

## Chapitre 2

# La complexité, outil d'amélioration de la perception de l'information

---

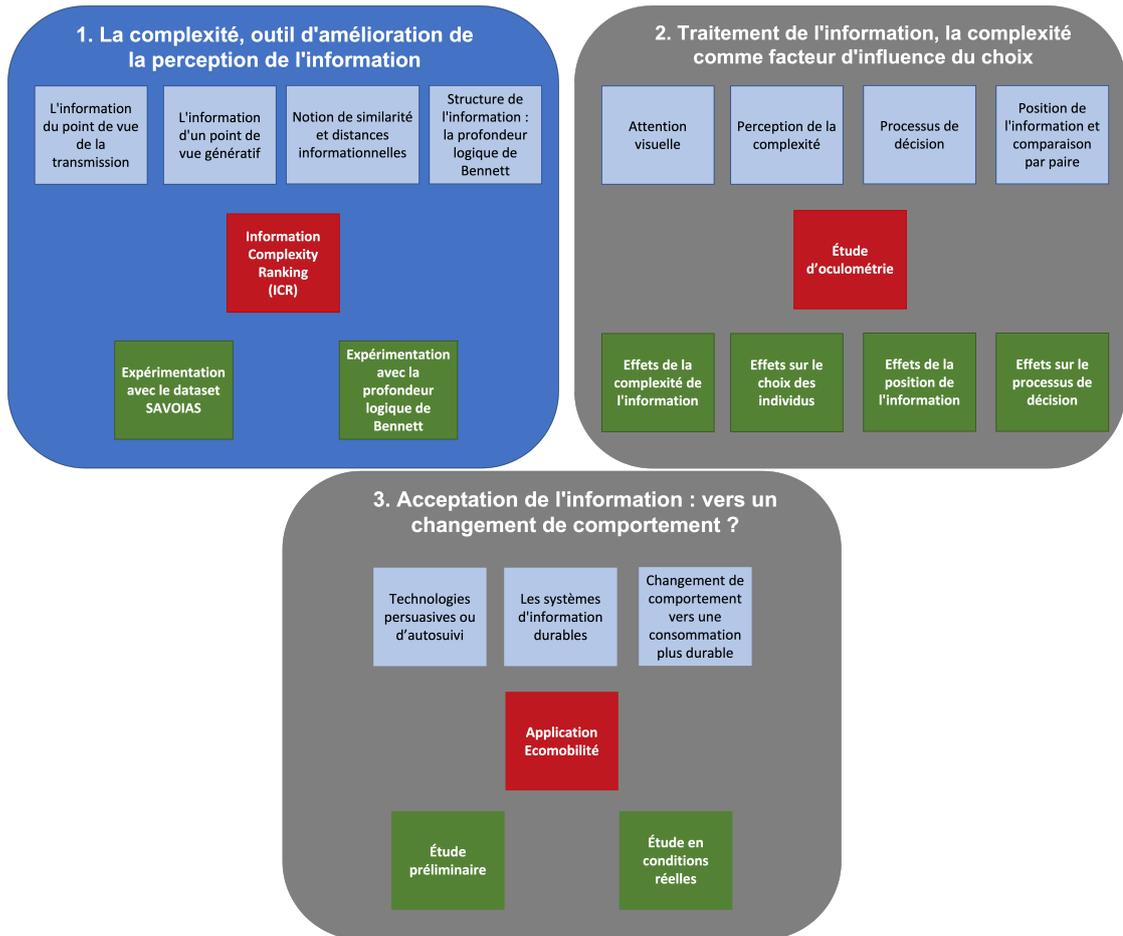
### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>19</b>
<b>2.2</b>	<b>État de l'art</b>	<b>20</b>
2.2.1	L'information du point de vue de la transmission	20
2.2.2	L'information d'un point de vue génératif	24
2.2.3	Notion de similarité et distances informationnelles	27
2.2.4	Structure de l'information : la profondeur logique de Bennett	30
<b>2.3</b>	<b>Information Complexity Ranking (ICR)</b>	<b>31</b>
2.3.1	Preuve	32
<b>2.4</b>	<b>Expérimentations</b>	<b>33</b>
2.4.1	Benchmark	34
2.4.2	Dataset SAVOIAS	37
2.4.3	Expérimentation avec la profondeur logique de Bennett	42
<b>2.5</b>	<b>Conclusion</b>	<b>44</b>

---

Les résultats présentés dans ce chapitre sont basés sur l'article « Information Complexity Ranking : A New Method of Ranking Images by Algorithmic Complexity. » publié dans la revue Entropy.



## 2.1 Introduction

Dans un contexte de mobilité éco-responsable, adopter aussi souvent que possible des moyens de transport plus doux s'avère nécessaire. Le point de départ de cette prise de conscience est souvent un événement particulier ou une information pertinente transmise à un moment propice. Cependant, pour entrevoir de potentiels résultats, les individus doivent être accompagnés dans cette démarche au quotidien. L'information et sa récurrence ont donc un rôle de premier plan à jouer dans cette transition. Il est alors important de rendre cette information compréhensible pour un large public. Prenons l'exemple de la conceptualisation de la quantité de CO<sub>2</sub> émise lors d'un trajet domicile-travail. Cette tâche cognitive s'avère particulièrement complexe du fait des multiples définitions à mobiliser mais également par le côté impalpable de l'activité en elle-même. Cet exemple classique de la mobilité urbaine prouve la nécessité d'améliorer la représentation mentale de l'information chez les individus. Un des leviers, pour d'atteindre cet objectif est de travailler sur la complexité de l'information en elle-même.

Dans son ouvrage « Complexité aléatoire et complexité organisée », Jean-Paul Delahaye propose deux définitions précises du mot complexe [Delahaye, 2009]. Ce terme traduirait d'un côté un aspect aléatoire ou chaotique avec une absence d'ordre, appelé « complexité aléatoire » et d'un autre côté une organisation très fortement structurée qu'il convient de nommer « complexité organisée ». Il précise également que la distinction entre complexité aléatoire et complexité organisée est difficile à caractériser.

Pour quantifier la complexité de l'information, deux méthodes sont généralement mises en avant, l'entropie de Shannon [Shannon, 1948] et la complexité de Kolmogorov [Kolmogorov, 1965]. Claude Elwood Shannon a été le premier à donner une définition de l'information basée sur la notion d'entropie. Cependant, cette méthode souffre d'une limitation importante lorsqu'on veut évaluer la complexité d'une image. Le point de vue observationnel de Shannon ne permet pas de comparer plusieurs images, car l'entropie résulte d'une sommation des composantes de l'histogramme et ne tient donc pas compte de l'ordre de ces symboles, c'est-à-dire de la sémantique du message.

Andreï Nikolaïevitch Kolmogorov, inspiré par la théorie de Shannon, a tenté de donner une autre définition de la complexité d'un objet, basée sur les notions algorithmiques de programme et de langage universel. Cette approche peut être résumée par la phrase suivante : ce qui est simple peut être décrit brièvement. Cependant la complexité de Kolmogorov étant une mesure de la complexité d'un seul objet, elle ne convient pas totalement, à notre contexte de recherche qui nécessite de comparer plusieurs objets et donc de pouvoir calculer leur complexité simultanément. De plus, la notion de structure n'est pas totalement prise en compte avec cette approche.

Pour pallier ces limitations, la notion de similarité entre deux objets peut aider à mettre en évidence de potentielles structures communes. Pour calculer la similarité entre deux objets, la distance de compression normalisée (NCD) [Cilibrasi and Vitányi, 2005] est appropriée dans le contexte algorithmique. En effet, cette distance universelle permet de détecter toutes les similarités entre deux objets de n'importe quel type (par exemple des chaînes de caractères, des images, des vidéos, etc.). Le degré de similitude entre deux

objets est important mais insuffisant pour classer les informations en fonction de leur complexité.

Par conséquent, en nous basant sur la classification de la complexité et la notion de similarité entre deux objets, nous proposons de définir une mesure de la complexité indépendante du type d'objet et qui retranscrit le plus fidèlement possible la complexité perçue chez les individus. Le but de notre proposition est à la fois de mesurer la complexité de chaque objet et de pouvoir les comparer pour établir un classement.

Expérimentalement, nous nous intéresserons particulièrement à la mesure de l'impact de la complexité d'une image transmise à une personne. Les images étant un vecteur d'information, il est pertinent de pouvoir les classer afin de mesurer plus précisément l'effet provoqué par ces dernières. Pour cela, nous testons les performances de notre méthode appelée Information Complexity Ranking (ICR) en la comparant à la perception humaine de la complexité.

Nos travaux s'inscrivent dans plusieurs cadres théoriques et applicatifs qui sont détaillés dans la section 2.2. Dans la section 2.3, nous présentons une nouvelle méthode de classement des informations en fonction de leur complexité en nous basant sur les fondements théoriques de la théorie algorithmique de l'information. Enfin, dans la section 2.4, nous présentons le protocole ainsi que les résultats des deux expérimentations associées à notre proposition scientifique.

## 2.2 État de l'art

### 2.2.1 L'information du point de vue de la transmission

La qualification de l'information a évolué au cours du temps. Cette évolution s'est accompagnée de nouvelles définitions qui ont apporté un regard nouveau sur la notion d'information. L'un des pionniers à contribuer à la compréhension de la nature même de l'information a été Claude Shannon en 1948. Dans son article, « A mathematical theory of communication », Shannon tente de résoudre des problèmes de transmission d'un signal analogique entre une source et une destination. Plus spécifiquement, il souhaitait maximiser la transmission de signaux sur un canal, mais également améliorer la correction d'un signal bruité émis par une source. Comme on peut le constater dans la Figure 2.1, un message est transmis à partir d'une source qui va être transformé en un signal par un émetteur, pour être envoyé vers un récepteur qui transformera à nouveau ce signal en message pour le transmettre à destination. La notion de bruit, ou d'ajout d'information perturbatrice, est représentée par une source de bruit se trouvant au milieu du canal de communication.

Pour résoudre ce problème, il proposa une fonction mathématique similaire à l'entropie de Boltzmann et lui donnera également le même nom. Au premier abord, il n'est pas évident de faire un parallèle entre la thermodynamique et la théorie de l'information. Mais si l'on considère un système thermodynamique équivalent à une source qui n'envoie aucun message vers un récepteur (désordre minimal) ou inversement (désordre maximal), la filiation entre les deux entropies est claire. En s'inspirant de la théorie des probabilités,

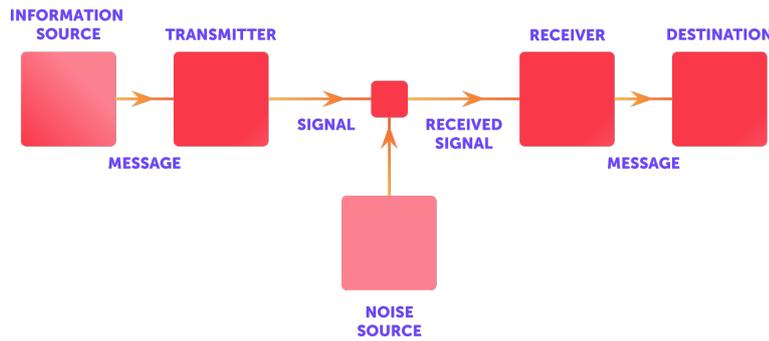


FIGURE 2.1 – Schéma du modèle de Shannon et Weaver (1949)

des statistiques et de la thermodynamique, l'idée de l'entropie de Shannon était ainsi de mesurer l'incertitude d'un message transmis. Autrement dit de lier la complexité d'un message à la quantité de bits nécessaires pour le transmettre et pour qu'il soit reçu sans ambiguïté sur son contenu. L'entropie est alors interprétée comme étant une mesure du nombre de bits nécessaires pour coder l'information à envoyer et définie comme suit :

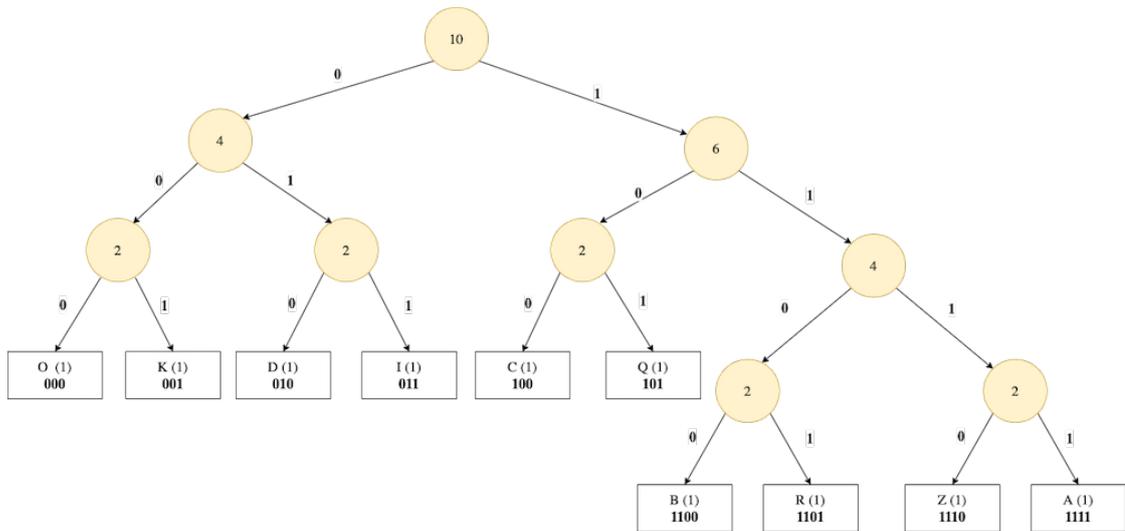
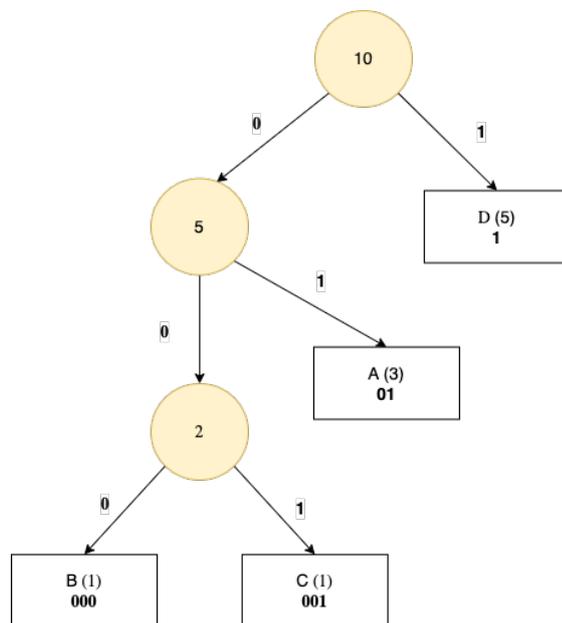
$$H(m) = - \sum_{i=1}^n P_i \log_2 P_i \quad (2.1)$$

où  $m$  est le message à transmettre,  $n$  est le nombre de symboles et  $P_i$  ( $P_i = n_i/n$ ) est la fréquence d'utilisation ou d'apparition de chaque symbole dans  $m$ . Pour faciliter la compréhension, dans la suite de ce chapitre, nous utiliserons la notation  $H_n(m)$  pour décrire la longueur totale du message à la place de  $H(m) * n$  du fait que  $H(m)$  peut être considéré comme la longueur moyenne de l'encodage de chaque symbole du message.

Pour mieux comprendre ce principe, prenons l'exemple de deux chaînes de caractères  $s_1 = \text{« cqzabrdiok »}$  et  $s_2 = \text{« aaabccddddd »}$ , pour lesquelles on obtient une entropie respective de  $H_{10}(s_1) = \mathbf{33.1}$  et  $H_{10}(s_2) = \mathbf{16.9}$ . Cette différence d'entropie s'explique par l'absence de répétitions dans  $s_1$ , tandis que  $s_2$  contient trois répétitions pour le premier caractère et cinq pour le quatrième caractère. On constate alors grâce à l'entropie qu'il est moins coûteux (en termes de bits) d'encoder un message possédant de multiples répétitions.

Cette asymétrie entre les occurrences d'un caractère dans un message et la difficulté à l'encoder est à l'origine du codage de Huffman [Huffman, 1952]. Proposé en 1952, ce codage consiste à générer une table de codage de longueur variable, dans laquelle chaque caractère du message à transmettre est codé avec un nombre de bits inversement proportionnel à la fréquence d'apparition de chaque caractère dans le message. Pour mieux comprendre l'application directe de l'entropie de Shannon dans le codage de Huffman reprenons l'exemple précédent.

Le codage de la première chaîne de caractères  $s_1$ , représenté par la figure 2.2, attribue trois bits pour encoder six des dix caractères et quatre bits pour encoder les quatre restant. Dans le cas du codage de la chaîne de caractères,  $s_2$ , représenté par la figure 2.3,

FIGURE 2.2 – Codage de Huffman pour la chaîne de caractères  $s_1$ FIGURE 2.3 – Codage de Huffman pour la chaîne de caractères  $s_2$

elle requiert un seul bit pour le caractère le plus fréquent, deux bits pour le second et les deux derniers caractères sont encodés à l'aide de trois bits.

La chaîne  $s_1$  ne possédant pas de répétitions, le nombre d'occurrences de chaque caractère est égal à un. Le nombre de bits requis pour encoder la chaîne est alors égal à la somme des bits pour encoder chaque caractère individuellement, **34 bits**  $\simeq H_{10}(s_1) = 33.1$ . Pour la chaîne  $s_2$ , le nombre de bits requis pour l'encodage est la longueur du code de chaque caractère multiplié par le nombre d'occurrences de chaque caractère, soit  $(2 \times 3) + (3 \times 1) + (3 \times 1) + (1 \times 5) = \mathbf{17 \text{ bits}}$   $\simeq H_{10}(s_2) = 16.9$ .

En adoptant un point de vue probabiliste, Shannon ne traite pas de l'ordre dans lequel les éléments d'information sont transmis. Il concentre sa théorie sur la qualité de la transmission d'un message, avec pour objectif de minimiser le nombre de bits pour encoder l'information. De manière symétrique, il assure le décodage de cette même information sans erreur. La fréquence d'occurrence de chaque symbole dans la mesure de l'entropie ne permet en aucun cas de définir leur ordre dans un texte ou encore leur position en deux dimensions (2D) dans une image (les symboles de l'image étant les valeurs de chaque pixel qui la composent). Cependant, la complexité d'un texte ou d'une image dépend certainement de cet ordre ou de ces positions [Delahaye, 2013]. De manière plus générale la théorie de Shannon ne traite pas les aspects sémantiques, c'est-à-dire la signification, le sens du message transmis. Pour citer Shannon lui-même :



Ces aspects sémantiques de la communication ne sont pas pertinents pour le problème d'ingénierie.

*Claude Shannon, A mathematical theory of communication, 1948, citation traduite de l'anglais.*

En outre, l'entropie de Shannon attribue une quantité d'information à un ensemble de messages finis possibles. Les messages de cet ensemble sont considérés comme étant équiprobable, cette quantité est donc le nombre de bits nécessaires pour compter toutes les possibilités. Par conséquent, chaque message de l'ensemble peut être transmis en utilisant ce même nombre de bits. Il nous est donc impossible de calculer le nombre requis de bits pour envoyer un objet de l'ensemble individuellement. Pour illustrer notre propos, nous utilisons l'exemple de Ming Li et Paul Vitanyi [Li and Vitányi, 2019] :



Considérons l'ensemble constitué de toutes les chaînes binaires de longueur 9999999999999999. Selon la mesure de Shannon, nous avons besoin de 9999999999999999 bits en moyenne pour coder une chaîne dans un tel ensemble. Cependant, la chaîne constituée de 9999999999999999 peut être codée en 55 bits environ en exprimant 9999999999999999 en binaire et en ajoutant le motif répété « 1 ».

*Ming Li et Paul Vitanyi, An Introduction to Kolmogorov Complexity and Its Applications, 2019, citation traduite de l'anglais*

La vision observationnelle de Shannon est longtemps restée la norme. Elle fut fon-

datrice dans l'émergence d'un nouveau concept, qui peut se résumer ainsi : « un objet complexe ne peut être décrit brièvement ». En effet, le lien entre la richesse de l'information contenue dans un objet et sa description semble intuitif. Plus un objet est complexe (c'est-à-dire riche en informations), plus son explication sera longue. La longueur de la description non ambiguë d'un objet permet donc de quantifier son contenu en information, et donc sa complexité. En parallèle du développement de la théorie de la calculabilité, la relation entre la notion de simplicité d'un objet et le petit nombre de symboles pour le décrire, a été démontrée.

D'un point de vue purement applicatif, l'information que nous souhaitons transmettre aux individus participant au cas d'application de nos travaux n'est pas assez diversifiée pour légitimer l'utilisation d'une méthode statistique comme celle de Shannon. Nous utilisons des images dans lesquelles l'information est contrôlée (non aléatoire) pour évaluer son impact sur le comportement. De plus, la non prise en compte de l'aspect sémantique de l'information et son ordre par l'outil de Shannon ne correspond pas au traitement et à la perception de la complexité chez l'individu.



**À retenir :** La théorie de Shannon fut la première tentative de mesure de la complexité, avec pour objectif de résoudre des problèmes de transmission de l'information. Bien que fondatrice dans son domaine, l'utilisation de cette théorie statistique ne correspond pas à notre contexte de recherche, celui de la perception humaine de la complexité.

### 2.2.2 L'information d'un point de vue génératif

L'évolution de cette idée vers des définitions mathématiques formelles a permis le développement du concept de la théorie algorithmique de l'information (Algorithmic Information Theory ou AIT). Cette nouvelle vision fut initiée par plusieurs travaux de Gregory Chaitin et Ray Solomonoff, mais la contribution principale dans ce domaine est due à Andreï Kolmogorov dans son article fondateur de 1965, « Three Approaches to Information » [Kolmogorov, 1965]. L'objectif principal de cette théorie algorithmique est de fournir des outils pour mesurer la complexité des chaînes de caractères, mais aussi des autres structures de données. Elle est définie par un de ses principaux contributeur de la manière suivante :



La théorie algorithmique de l'information (AIT) est le résultat de la mise en commun de la théorie de l'information de Shannon et de la théorie de la calculabilité de Turing dans un shaker et d'une agitation vigoureuse. L'idée de base est de mesurer la complexité d'un objet par la taille en bits du plus petit programme permettant de le calculer.

*G. J. Chaitin, citation traduite de l'anglais.*

L'objectif principal de cette théorie algorithmique est ainsi de fournir des outils pour mesurer la complexité d'un ensemble de données au sens large (chaînes de caractères, mais aussi des structures de données plus complexes). Se basant sur la théorie des probabilités,

de la théorie de l'information de Shannon et des notions d'aléatoire [Li and Vitányi, 2019], une implémentation de cette nouvelle approche qui se concentre sur la manière dont est générée l'information a été proposée. Kolmogorov a émis l'idée de considérer la taille du programme implémentant l'algorithme qui génère l'information dans son calcul de complexité. La longueur d'un programme (sa description) permet de mesurer la complexité de l'objet et donc le contenu de l'information. Au vu de cette contribution majeure au domaine, cette nouvelle forme d'évaluation de la complexité est couramment appelée par le nom de son auteur, la complexité de Kolmogorov.

La complexité de Kolmogorov d'un objet  $x$ , appelée  $K\varphi(x)$ , est définie comme étant la longueur du plus court programme permettant de calculer  $x$ . Ce programme générateur doit être obligatoirement écrit avec un langage de description universelle Turing complet (par exemple C, C++, Python, JAVA).

$$K\varphi(x) = \min\{|p| : \varphi(p) = x\} \quad (2.2)$$

où  $p$  est le programme générateur (sans entrée),  $\varphi$  exécute le programme et  $\varphi(p)$  est la sortie de l'exécution du programme  $p$ . Ainsi pour  $x \in \mathcal{O} \rightarrow \{0,1\}^*$  (toute séquence finie composée de 0 et de 1),  $K\varphi(x)$  est la longueur du plus court programme  $p$  avec lequel  $\varphi$  calcule  $x$  [Ferbus-Zanda and Grigorieff, 2010].

D'après les travaux d'Alan Turing sur l'indécidabilité de la terminaison des programmes [Turing, 1937], ou encore la publication plus récente « A formal theory of inductive inference » [Solomonoff, 1964] de Solomonoff,  $K\varphi(x)$  n'est pas effectivement calculable au sens de Turing, car non récursive [Grunwald and Vitanyi, 2004]. Néanmoins,  $K\varphi(x)$  est approchable. Plus précisément, Kolmogorov a montré dans son théorème d'invariance [Kolmogorov, 1965] que tout couple de programmes (U,V), implémentant le même algorithme dans deux langages différents, ont la même longueur à une constante additive près.

$$|K_U(x) - K_V(x)| < C_{UV} \quad (2.3)$$

où  $K_U(x)$  (respectivement  $K_V(x)$ ) représente la complexité de Kolmogorov avec une machine de Turing universelle  $U$  (respectivement  $V$ ), et  $C_{UV}$  une constante.

Ce théorème nous permet d'affirmer (à une constante près) que le changement de machine de Turing universelle (c'est-à-dire le langage de programmation) n'affecte pas la complexité de Kolmogorov de la séquence  $x$ . Dans ce contexte, le programme le plus court peut être considéré comme une description minimale de l'objet et comme une approximation satisfaisante de la complexité de Kolmogorov.

Cette méthode algorithmique permet de considérer la description comme une version compressée de l'objet. Le programme calcule l'objet avec un nombre de symboles inférieur au nombre de symboles qui composent l'objet ( $K\varphi(x) \leq |x|$ ), par opposition à la notion d'incompressibilité qui traduit l'absence de description plus courte que la description littérale de l'objet  $x$  [Kolmogorov, 1965, Li and Vitányi, 2019]. Ces outils théoriques nous offrent une nouvelle méthode de quantification de l'information en introduisant la notion de description comme mesure de la complexité et par extension comme mesure du contenu

de l'information. En effet, la description de l'objet (le programme générateur) et la version compressée sont basées sur un principe commun, la recherche de régularités (répétitions). Ce cadre théorique s'applique parfaitement à différents problèmes comme trouver le plus court programme générateur ou réduire la taille d'un fichier le plus possible. Le but commun est d'exploiter l'information répétée contenue dans l'objet.

L'application directe de l'exploitation des répétitions est le ratio de compression  $C(o)$  d'un objet  $o$  défini comme :

$$C(o) = \frac{|o|}{|C(o)|} \quad (2.4)$$

où  $|o|$  est la taille du fichier de l'objet original, et  $|C(o)|$  celle de sa version compressée. Ce ratio défini par  $C(o)$  est une approximation de la complexité de Kolmogorov, car il fournit une indication sur la taille de la description d'un objet et par extension sur sa complexité.

Afin de mieux comprendre l'utilité de ce concept théorique, reprenons l'exemple de nos deux chaînes de caractères  $s_1 = \text{« cqzabrдиok »}$  et  $s_2 = \text{« aaabccccdd »}$  précédemment utilisées. Ces deux objets ont intuitivement des complexités différentes, respectivement élevée et faible au sens de la complexité de Kolmogorov. Afin de déterminer la complexité de ces deux chaînes de caractères, un algorithme de compression sans perte<sup>1</sup> peut être utilisé.

À des fins d'illustration uniquement, car non optimal du point de vue de la compression, nous utilisons le compresseur Run-Length Encoding (RLE) [Tsukiyama et al., 1986]. Après la compression avec RLE, l'encodage de la première chaîne est « 1c1q1z1a1b1r1d1i1o1k » avec un total de vingt caractères et celui de la seconde chaîne est « 3a1b1c5d » avec un total de huit caractères. Ces deux sorties reflètent parfaitement la notion de répétition et son influence sur la complexité.

La complexité de Kolmogorov est considérée comme un bon estimateur de la complexité d'un objet représenté par une séquence binaire. Cette qualité est utile pour définir la notion absolue d'une séquence aléatoire [Martin-Löf, 1966], mais ne permet pas de capturer la notion de complexité structurée [Delahaye and Vidal, 2016]. En effet, elle attribue une complexité élevée aux objets aléatoires et hautement organisés, les plaçant au même niveau [Zenil et al., 2012]. La complexité structurée serait davantage en adéquation avec l'information que nous transmettons aux individus dans nos travaux. Comme cité précédemment pour l'entropie de Shannon, l'information est contrôlée et non aléatoire. Dans le but d'améliorer la mesure et le classement de notre propre méthode, nous avons cherché des alternatives à la complexité de Kolmogorov (voir Sections 2.2.3 et 2.2.4).

Le point de vue génératif de l'information nous offre un cadre théorique et pratique plus en adéquation avec certaines de nos problématiques de recherche en particulier sur l'aspect sémantique, mais elle ne résout pas une question centrale, celle du choix de l'individu entre différentes informations. En effet, dans nos travaux pour mesurer efficacement l'impact de l'information sur le comportement, nous ne pouvons pas nous limiter

---

1. Un compresseur est sans pertes s'il permet de reconstruire sans perte d'information la donnée originale.

à présenter une seule et unique information à la fois. De plus, le processus de décision peut être influencé par différentes informations au fur à mesure de leur perception. La complexité de Kolmogorov étant une mesure de la complexité d'un seul objet, il nous est impossible de connaître la complexité de plusieurs images simultanément par exemple. Quel serait alors l'impact de chaque information dans le processus de décision ?



**À retenir :** La complexité de Kolmogorov, issue des travaux sur la théorie algorithmique de l'information, est un bon estimateur d'une forme de complexité aléatoire. L'idée fondatrice met en exergue le lien entre longueur de la description et complexité d'un objet, ce qui est utile pour la prise en compte d'un aspect sémantique dans notre recherche. De plus, cette notion est approximable facilement à l'aide d'algorithmes de compression sans perte (compresseurs normaux).

### 2.2.3 Notion de similarité et distances informationnelles

Pour répondre à la question de l'influence de chaque information dans le processus de décision, nous nous sommes intéressés à la notion de similarité entre deux objets. En effet, il est important d'avoir une vision plus globale du message transmis (ensemble d'informations) à l'individu tout au long de ce processus. Si nous considérons la complexité comme une mesure de l'information, nous pouvons faire l'hypothèse que la comparaison de la complexité entre différents objets peut permettre de classer l'information qu'ils représentent.

En intégrant ce nouvel aspect de comparaison entre les objets, la complexité de Kolmogorov n'en devient pas pour autant obsolète, grâce à la notion de description de l'information. En effet, deux objets sont similaires du point de vue de leur complexité si l'un peut être décrit par la description de l'autre [Cilibrasi and Vitányi, 2005].

Pour détecter la similarité entre plusieurs objets, l'idée de base est de calculer l'information commune contenue dans la description de chacun d'eux afin de pouvoir les classer [Ferbus-Zanda, 2010]. Une des premières tentatives d'exploitation de l'information descriptive a été proposée par Bennett et al. En 1998, ils définissent la notion de *distance informationnelle* entre deux mots  $x$  et  $y$  comme étant la taille du plus court programme pouvant faire correspondre  $x$  et  $y$  et inversement (notion de calcul réversible [Ferbus-Zanda, 2010]) [Bennett et al., 1998]. Nous présentons deux versions de cette distance une définition formelle, désignée  $ID'$  et une définition alternative, désignée  $ID$  [Ferbus-Zanda, 2010].

$$ID'(x, y) = \min\{|p| : U(p, x) = y \text{ et } U(p, y) = x\} \quad (2.5)$$

où  $U$  est une machine universelle de Turing optimale pour calculer la complexité conditionnelle  $K(\cdot)$  [Ferbus-Zanda and Grigorieff, 2010]. La complexité conditionnelle de Kolmogorov représente la quantification des informations nécessaires pour retrouver  $x$  à partir de  $y$  [Li and Vitányi, 2019].

Une définition alternative de cette distance  $ID$  est couramment utilisée :

$$ID(x, y) = \max\{K(x|y), K(y|x)\} \quad (2.6)$$

Cependant cette distance informationnelle ne permet pas de définir formellement une classification [Ferbus-Zanda, 2010]. Pour résoudre ce problème, une normalisation est nécessaire. Cette distance est donc appelée distance informationnelle normalisée (Normalized Information Distance - NID). NID est une distance universelle théorique basée sur la complexité de Kolmogorov [Li et al., 2004]. Une définition formelle de cette dernière est la suivante [Ferbus-Zanda, 2010] :

$$NID(x, y) = \frac{\max\{K(x|y), K(y|x)\}}{\max\{K(x), K(y)\}} \quad (2.7)$$

où  $K(x|y)$  (respectivement  $K(y|x)$ ) représente la complexité conditionnelle de Kolmogorov de  $x$  (respectivement  $y$ ) par rapport à  $y$  (respectivement  $x$ ). Ces définitions sont pertinentes lorsque l'on s'intéresse à la notion de similarité, car le plus court programme possible qui calcule l'objet  $x$  depuis l'objet  $y$  et inversement prend en compte toutes les similarités entre  $x$  et  $y$  [Ferbus-Zanda, 2010].

Malheureusement, la distance informationnelle normalisée se basant sur la complexité de Kolmogorov, il a été prouvé que la distance est elle-même incalculable [Ferbus-Zanda, 2010], mais elle est approximable [Bennett et al., 1998].

Quelques années plus tard, Cilibrasi et Vitányi ont proposé une approximation calculable de la distance informationnelle normalisée, appelée distance de compression normalisée (Normalized Compression Distance - NCD) [Cilibrasi and Vitányi, 2005] dans laquelle les compresseurs normaux sont considérés comme une approximation de la complexité de Kolmogorov. NCD se définit comme suit :

$$NCD(x, y) = \frac{C(xy) - \min\{C(x), C(y)\}}{\max\{C(x), C(y)\}} \quad (2.8)$$

où  $C(x)$  et  $C(y)$  représentent respectivement la taille compressée de l'objet  $x$  et de l'objet  $y$ .  $C(xy)$  représente la taille compressée de la concaténation (opération qui joint les deux objets) des objets  $x$  et  $y$ . La similarité maximale entre deux objets est atteinte quand la valeur de la distance est égale à 0. La dissimilarité maximale est obtenue quand à elle lorsque la valeur de la distance atteint  $1 + \epsilon$  (le  $\epsilon$  est dû aux imperfections des techniques de compression) [Cilibrasi and Vitányi, 2005].

NCD est une métrique de similarité quasi universelle si et seulement si le compresseur utilisé dans le calcul de la distance est *normal* [Cilibrasi and Vitányi, 2005]. Un compresseur normal est un compresseur capable de détecter les répétitions dans les données à compresser. En détectant ces répétitions, il est alors possible de supprimer les données redondantes dans la description d'un objet et ainsi atteindre une approximation satisfaisante de la distance de similarité entre deux objets. Le compresseur doit aussi respecter certaines propriétés telles que : l'idempotence, la symétrie, la monotonie et la distributivité [Cilibrasi and Vitányi, 2005].

**Propriété d'idempotence**

$$C(xx) = C(x) \quad (2.9)$$

**Propriété de symétrie**

$$C(xy) \geq C(x) \quad (2.10)$$

**Propriété de monotonie**

$$C(xy) = C(yx) \quad (2.11)$$

**Propriété de distributivité**

$$C(xy) + C(z) \leq C(xz) + C(yz) \quad (2.12)$$

Il a également été prouvé que la NCD est une distance et qu'elle respecte donc les propriétés suivantes :

**Propriété de séparation**

$$NCD(x, y) = 0 \Leftrightarrow x = y \quad (2.13)$$

**Propriété de symétrie**

$$NCD(x, y) = NCD(y, x) \quad (2.14)$$

**Propriété d'inégalité triangulaire**

$$NCD(x, y) \leq NCD(x, z) + NCD(z, y) \quad (2.15)$$

L'optimalité de cette distance est intéressante dans nos travaux, dans le but de proposer une méthode de classement générique, indépendante de la nature de l'information à classer.

La notion de similarité nous donne une indication de la ressemblance entre deux objets, ce qui est important, mais pas suffisant pour évaluer l'impact des informations sur le comportement. Dans notre contexte de recherche sur le choix du moyen de transport, il est nécessaire de présenter des informations de nature et de forme différentes. Cette diversité d'informations permet d'analyser leur impact sur l'évolution possible du choix de chaque individu. Avant de pouvoir mesurer cette évolution, il est important de hiérarchiser les informations transmises en fonction de leur complexité. Nous proposons donc une mesure alternative de la complexité de l'information basée sur la complexité algorithmique et la notion de similarité entre deux objets. Cette mesure, appelée Information Complexity Ranking (ICR) sera développée dans la Section 2.3.



**À retenir :** La distance de compression normalisée (Normalized Compression Distance - NCD) proposée par Rudi Cilibrasi et Paul Vitanyi, est une mesure de similarité universelle entre deux objets. Cette proposition issue de travaux sur la distance informationnelle à l'avantage d'être approximable à l'aide de compresseurs *normaux*. Dans le cadre de notre recherche, la prise en compte de la similarité entre différentes informations est utile pour analyser l'impact d'un changement progressif dans sa nature ou dans sa forme sur le comportement de mobilité des individus. L'optimalité et l'universalité de cette distance sont des points à retenir dans l'optique d'une méthode de classement indépendante de la nature de l'objet à classer.

#### 2.2.4 Structure de l'information : la profondeur logique de Bennett

Dans les sections précédentes, nous avons mis en avant des méthodes reposant sur la complexité de Kolmogorov, qui traduit la complexité aléatoire d'un objet en calculant la longueur de son programme générateur. Pour quantifier convenablement une information, il est également important d'en comprendre toutes ses composantes et de les prendre en compte. De fait, si la notion d'aléatoire dans l'information existe, la notion de structure existe également. Pour prendre cette dernière en compte dans la complexité d'un objet, nous avons donc opté pour la notion de profondeur logique. La principale différence avec la profondeur logique de Bennett, est la prise en compte du temps comme représentant de la complexité. Concrètement, la profondeur logique ne prend pas en compte le nombre d'instructions du programme générateur comme mesure de la complexité algorithmique, mais le temps d'exécution de ce même programme. L'équation suivante présente ce principe [Zenil et al., 2012] :

$$D_s(x) = \min\{T(p) : (|p| - |p^*| < s) \wedge (U(p) = x)\} \quad (2.16)$$

soit  $D_s(x)$  la profondeur d'une chaîne  $x$  au niveau de signification  $s$ ,  $T(p)$  le temps le plus court requis pour calculer la chaîne  $x$  par un programme  $s$ -incompressible  $p$  avec une machine de Turing universelle  $U$ .

L'idée fondatrice de cette nouvelle mesure de la valeur réelle de l'information contenue dans une séquence réside sur l'utilisation de la notion de programme minimal. Pour mieux comprendre l'intérêt de cette notion, utilisons l'analogie formelle entre les programmes universels, leurs sorties et la relation entre la théorie et l'observation en science (c'est-à-dire l'explication de la théorie à partir de données récoltées). Dans ce contexte, le programme minimal représente l'explication la plus économique de sa sortie et donc a priori la plus plausible. Il convient tout de même de préciser que cette analogie est basée sur la capacité des machines universelles à exécuter tous les processus déductifs formels et sur leur capacité présumée à simuler tous les processus de causalité physique. En l'acceptant, on peut logiquement dire qu'un objet complexe ou profond a pour origine la plus probable un long calcul par un processus efficace (programme minimal) [Bennett, 1988b].

La raison principale de notre choix repose sur le fait que les travaux de Bennett com-

plètent ceux de Kolmogorov. Toutefois, contrairement à la complexité algorithmique, la profondeur attribue une faible complexité aux objets aléatoires et triviaux [Zenil et al., 2012] et permet également de prendre en compte uniquement l'information pertinente [Delahaye and Vidal, 2016].

Il est cependant intéressant de souligner l'interdépendance de la complexité algorithmique et de la profondeur logique, car si on ne trouve pas le plus court programme générateur, il est impossible de calculer son temps d'exécution. Ce lien nous permet également d'exploiter la notion de compression pour l'approximer. En effet, si la compression sans perte nous donne une mesure satisfaisante de la complexité algorithmique, il suffit d'inverser le processus en utilisant le temps d'exécution de l'algorithme de décompression et ainsi calculer sa profondeur logique [Zenil et al., 2012]. Une expérimentation pour évaluer ses performances par rapport à la perception humaine sera présentée dans la Section 2.4.3.



**À retenir :** La profondeur logique de Bennett, mesurant le temps d'exécution du programme minimal générant un objet, permet de mesurer plus fidèlement la complexité structurée de l'information. Tout comme la complexité de Kolmogorov, elle est facilement approximable à l'aide d'algorithmes de compression sans perte. Dans le cadre de nos recherches, la profondeur logique de Bennett est un estimateur à prendre en compte au regard de l'information que nous transmettons à l'individu (information non-aléatoire).

## 2.3 Information Complexity Ranking (ICR), une nouvelle mesure de la complexité

Le principal objectif de nos travaux est de mesurer l'impact de la complexité de l'information sur le comportement des individus. Nous devons donc définir le contenu informationnel de chaque message que l'on transmet, et nous utilisons pour cela la théorie algorithmique de l'information et la profondeur logique (Section 2.2). Une fois cette quantification effectuée, nous devons comparer les différentes informations transmises aux individus au cours du temps. Cette étape de comparaison multiple nécessite une nouvelle méthode de classement prenant en compte la similarité entre les objets. Afin de répondre à ce problème, nous proposons une nouvelle mesure qui permet d'ordonner la complexité de plusieurs objets du plus simple au plus complexe et ainsi, de les comparer.

L'idée fondamentale de la méthode ICR est de fournir une mesure plus fine de la complexité. D'un point de vue purement algorithmique, deux images peuvent être très similaires, phénomène que NCD retranscrit fidèlement, mais l'utilisation de NCD ne donne aucune information sur la complexité de ces deux images. C'est cette lacune que nous tentons de combler en utilisant la complexité intrinsèque de chaque image.

Nous introduisons d'abord le concept de classement de la complexité de l'information (ICR) basé sur la mesure de la similarité entre deux objets (NCD) pondérée par la complexité intrinsèque de chaque objet. Nous utilisons le taux de compression,  $C(x)$ , comme approximation de la complexité de Kolmogorov et incorporons la complexité des

deux éléments d'information en effectuant un produit avec NCD comme suit :

$$ICR(x, y) = NCD(x, y) \times \frac{C(y)}{C(x)} \quad (2.17)$$

Nous affirmons alors que la relation binaire  $R(x, y)$  définie par :

$$R(x, y) = xRy \Leftrightarrow ICR(x, y) \geq ICR(y, x) \quad (2.18)$$

est un préordre, c'est-à-dire une relation binaire respectant les propriétés de réflexivité et de transitivité. Il aurait été pertinent de proposer un ordre pour pouvoir comparer directement des objets mais on ne peut donc pas prouver que la relation est anti-symétrique. En effet, si  $R(x, y)$  et  $R(y, x)$  cela signifie seulement que  $x$  et  $y$  ont le même taux de compression (ou la même complexité de Kolmogorov) mais pas qu'ils sont identiques. Cependant, en faisant le quotient avec la relation d'équivalence sur les objets qui ont le même taux de compression on obtient un ordre ce qui nous permet de comparer les objets ayant un taux de compression différent.

### 2.3.1 Preuve

Les objets traités ici sont génériques, il peut s'agir de toute séquence de symboles binaires formant un ensemble de bits. Comme mentionné ci-dessus, un préordre est une relation binaire satisfaisant les propriétés suivantes :

$$\begin{cases} R(x, x) & \text{Réflexivité} \\ R(x, y) \text{ ET } R(y, z) \Rightarrow R(x, z) & \text{Transitivité} \end{cases} \quad (2.19)$$

#### Propriété de réflexivité

Regardons si  $R(x, x)$  est vérifié. En le remplaçant par sa définition, nous devons vérifier si  $ICR(x, x) \geq ICR(x, x)$  ce qui peut être réécrit :

$$NCD(x, x) \times \frac{C(x)}{C(x)} \geq NCD(x, x) \times \frac{C(x)}{C(x)}$$

ou encore

$$NCD(x, x) \geq NCD(x, x)$$

En raison de la propriété de séparation (eq. 2.13) de NCD,  $NCD(x, x) = 0$  qui nous permet de conclure que  $ICR(x, x) \geq ICR(x, x)$  et donc que  $R(x, y)$  est **réflexive**.

#### Propriété de transitivité

Supposons maintenant que  $ICR(x, y) \geq ICR(y, x)$  et  $ICR(y, z) \geq ICR(z, y)$

Nous pouvons l'écrire comme suit :

$$\begin{cases} NCD(x, y) \times \frac{C(y)}{C(x)} \geq NCD(y, x) \times \frac{C(x)}{C(y)} \\ NCD(y, z) \times \frac{C(z)}{C(y)} \geq NCD(z, y) \times \frac{C(y)}{C(z)} \end{cases} \quad (2.20)$$

On suppose que  $x$ ,  $y$  et  $z$  sont distincts, la transitivité est triviale dans le cas contraire. Cela nous permet d'assurer que tous les termes NCD sont non nuls du fait de la propriété de séparation (eq. 2.13).

Par symétrie de NCD, on a donc

$$\begin{cases} \frac{C(y)}{C(x)} \geq \frac{C(x)}{C(y)} \\ \frac{C(z)}{C(y)} \geq \frac{C(y)}{C(z)} \end{cases} \quad (2.21)$$

On utilise la propriété  $a/b \geq b/a \Leftrightarrow a \geq b$  si  $a$  et  $b$  sont deux nombres strictement positifs. En effet si  $a \geq b$  alors  $a/b \geq 1$  et  $1 \geq b/a$ . Inversement si  $a/b \geq b/a$  alors  $a^2 \geq b^2$  et donc  $a \geq b$ .

On a ainsi  $C(y) \geq C(x)$  et  $C(z) \geq C(y)$ , donc  $C(z) \geq C(x)$  et  $\frac{C(z)}{C(x)} \geq \frac{C(x)}{C(z)}$ . A nouveau par symétrie de NCD on obtient :

$$NCD(x, z) \times \frac{C(z)}{C(x)} \geq NCD(z, x) \times \frac{C(x)}{C(z)}$$

Ce qui prouve la transitivité de notre relation.

Pour conclure, comme  $R(x, y)$  respecte les propriétés de réflexivité et de transitivité, cela nous permet d'affirmer qu'il s'agit d'un préordre.  $\square$

Notre méthode est une mesure abstraite reposant sur des hypothèses fortes concernant les propriétés que devraient avoir le compresseur et l'opérateur de concaténation, et doit être validée sur des applications concrètes. Pour cela, nous avons validé nos hypothèses sur une expérience de benchmark, puis nous avons utilisé une base de données de référence dans le domaine de la complexité visuelle.



**À retenir :** Le but de notre méthode ICR est de fournir un outil de classement de l'information en fonction de sa complexité algorithmique, indépendante de la nature même de l'information. Cette proposition repose sur deux notions principales, la similarité entre deux objets, représentée par la distance de compression normalisée (NCD) et la complexité intrinsèque d'un objet, représentée par l'utilisation de la complexité de Kolmogorov. Cette méthode nous permettra de quantifier le contenu de l'information transmise à un individu et ainsi de mesurer la réponse associée à la complexité de cette même information.

## 2.4 Expérimentations

Pour tester la validité de notre méthode ICR, nous avons mis en oeuvre deux expérimentations, la première de type benchmark pour sélectionner les compresseurs normaux, condition indispensable à l'approximation de la complexité de Kolmogorov. La deuxième expérimentation repose sur un jeu de données d'images évaluant la complexité visuelle. L'avantage de ce jeu de données est double pour notre recherche, de par sa diversité

d'images et sa vérité terrain qui offre des informations sur la perception humaine de la complexité, avec plus d'un millier de contributeurs interrogés. Nous avons également mené une troisième expérimentation avec le même jeu de données pour comparer notre méthode avec la profondeur logique de Bennett pour étudier si la mesure d'une complexité structurée possède une corrélation avec la perception humaine de la complexité.

## 2.4.1 Benchmark

### 2.4.1.1 Protocole

Notre méthode ICR utilise deux éléments principaux : le taux de compression d'un compresseur normal et la capacité à concaténer deux éléments (NCD). Dans cette expérimentation, nous allons tester huit compresseurs sans perte couramment utilisés pour vérifier leur capacité à respecter certaines propriétés. Pour être qualifié de *normal*, un compresseur doit respecter les propriétés d'idempotence, de symétrie, de monotonie et de distributivité. Pour avoir une mesure de la complexité la plus fidèle possible, nous allons tester leur performance en termes de compression, à l'aide du taux de compression.

Pour que notre méthode soit cohérente, l'opération de concaténation doit respecter les propriétés précédemment citées. La complexité de Kolmogorov prenant en compte l'ordre de l'information, il est important également de vérifier l'influence du sens de la concaténation de chaque paire d'images. A cette fin, les images testées sont concaténées par paire, soit verticalement, soit horizontalement. La qualité du compresseur utilisé est importante concernant l'applicabilité de notre méthode. A cette fin, nous avons également étudié les paramètres de ces compresseurs et adopté le paramètre maximisant le taux de compression.

Concernant les images utilisées pour sélectionner les compresseurs, nous avons opté pour des images réalistes dans notre contexte de recherche avec des images de type texte, des nombres et des pastilles de couleur. En effet, ce format est cohérent avec les étiquettes visuelles que nous voulons utiliser dans notre expérience avec une application mobile (chapitre 4) afin de tester l'impact de la complexité de l'information sur le processus de décision (comme les étiquettes nutritionnelles [Grunert et al., 2012]).

De plus, ces trois types d'éléments peuvent être facilement synthétisés. En pratique, l'image générée est composée d'une chaîne de 20 caractères (Figure 2.4), concaténée avec un nombre possédant de 1 à 4 chiffres (Figure 2.5) et une pastille de couleur sélectionnée aléatoirement (Figure 2.6). Comme on peut le remarquer, ce modèle n'est pas complètement aléatoire pour deux raisons : premièrement, nous voulons pouvoir comparer des structures similaires avec de légères différences afin de tester la sensibilité de notre méthode. Pour calculer la complexité d'une image, le format original de l'image utilisée doit être sans compression. Nous avons donc opté pour le format Bitmap (BMP), car il respecte ce pré requis et peut être édité et manipulé sans perte de qualité au niveau de l'image. Au final, 7200 images ont été créées (3600 images concaténées horizontalement et 3600 images concaténées verticalement).

ijlzxqbtthpsjdzxjkzx

FIGURE 2.4 – Exemple d’information de type texte générée

3170

FIGURE 2.5 – Exemple de nombre (4 chiffres) généré

### 2.4.1.2 Résultats

Pour le choix de l’algorithme de compression, nous avons testé dans quelle mesure ils respectent les propriétés requises pour la NCD dans le but de rendre notre méthode généralisable (voir Table 2.1). Nous testons dans un premier temps les huit algorithmes de compression par rapport au taux de compression moyen, dans lequel les valeurs les plus élevées seront retenues (première ligne de la Table 2.1). Par la suite nous testons les différentes propriétés d’un compresseur normal dans les deux sens de concaténation. Pour les propriétés d’idempotence, de symétrie et de monotonie (de la deuxième à la sixième ligne de la Table 2.1) les valeurs présentées correspondent à la valeur moyenne obtenue pour chaque algorithme testé. Du point de vue de l’implémentation de ce test, nous avons concaténé différentes images entre elles pour vérifier le respect des différentes propriétés. De part leur nature les images ne possédant pas les mêmes dimensions, nous avons opté pour l’ajout de marge de couleur blanche pour ajuster la taille de la plus petite image à celle de la plus grande. a consisté pour la propriété d’idempotence, a concaténer deux fois la même image pour ensuite comparer le ratio de compression obtenu avec le ratio de compression issu de l’image seule. Pour les propriétés de symétrie et de monotonie, la concaténation s’est effectuée sur deux images différentes, mais de type identique. Ces opérations ont été menés sur l’ensemble des images générées. Compte tenu du grand nombre de possibilités pour tester la propriété de distributivité, les valeurs présentées sont obtenues à partir d’un échantillon de 30 inégalités entre types d’images (dernière ligne de la Table 2.1). Pour les valeurs concernant les propriétés du compresseur, seules les valeurs les plus faibles sont retenues.

Sur la base de ces résultats, trois compresseurs ont été retenus pour des tests supplémentaires, à savoir PNG, PPM et LZ77. Ces trois algorithmes donnent les meilleurs résultats concernant la propriété d’idempotence. Cette propriété est importante pour nous, car elle illustre la détection des répétitions de la part des algorithmes de compression, ce qui nous permettra de quantifier le plus précisément possible le contenu informationnel. En ce qui concerne la propriété de symétrie, les algorithmes PNG et LZ77 présentent l’un des meilleurs résultats en matière de concaténation horizontale et verticale (quatrième et cinquième lignes de la Table 2.1). Néanmoins, ils n’ont pas les meilleurs résultats concernant les propriétés de monotonie et de distributivité (sixième et septième lignes de



FIGURE 2.6 – Exemple de pastille de couleur générée

<b>Algorithm</b>	<b>PNG</b>	<b>PPM</b>	<b>LZ77</b>	<b>ALZ</b>	<b>LZP</b>	<b>LZW</b>	<b>AC</b>	<b>Huffman</b>
Taux de compression	3.174	3.676	3.509	3.745	3.300	3.875	2.028	0.760
Idempotence horizontale	0.088	0.091	0.098	0.196	0.209	0.555	0.533	0.216
Idempotence verticale	0.043	0.047	0.058	0.130	0.133	0.631	0.591	0.237
Symétrie horizontale	0.004	0.014	0.004	0.007	0.007	0.004	0.001	0.000
Symétrie verticale	0.006	0.027	0.009	0.031	0.005	0.007	0.009	0.000
Monotonicité horizontale	0.318	0.202	0.285	0.197	0.334	0.323	0.731	1.423
Monotonicité verticale	0.344	0.204	0.298	0.199	0.329	0.343	0.909	1.588
Distributivité horizontale	0.313	0.197	0.273	0.203	0.327	0.314	0.736	1.414
Distributivité verticale	0.329	0.202	0.287	0.203	0.313	0.333	0.894	1.549

TABLE 2.1 – Performance des algorithmes de compression en calculant le taux de compression et le respect pour chaque propriété requise pour satisfaire le concept d'un compresseur normal.

la Table 2.1). En ce qui concerne la performance du taux de compression, l'algorithme PPM présente le troisième meilleur résultat. En ce qui concerne les algorithmes PNG et LZ77, leurs résultats sont proches de ceux de PPM (première ligne de la Table 2.1 - avec neuf comme paramètre de compression, le niveau maximum). Pour l'algorithme PPM, la taille maximale du dictionnaire a été utilisée (32768 Ko) pour obtenir le meilleur taux de compression.

## 2.4.2 Dataset SAVOIAS

### 2.4.2.1 Description

Pour aller plus loin et tester les performances de notre approche dans un contexte plus réaliste, nous avons expérimenté sa robustesse sur le jeu de données SAVOIAS [Sarae et al., 2018]. L'objectif de l'équipe de l'Université de Boston qui a fourni ces données était de créer un jeu de données de complexité visuelle de plus de **1400 images**. L'un de ses avantages est la pertinence du choix des images afin d'aborder le problème de la complexité visuelle. La diversité des images est également intéressante avec diverses caractéristiques de bas et de haut niveau. Les images sont ainsi regroupées en sept catégories, à savoir : Scènes, Publicités, Visualisation et infographie, Objets, Décoration intérieure, Art et Suprématisme (voir Figure 2.7). Les auteurs comparent leur vérité terrain à cinq algorithmes (c'est-à-dire la densité des contours, le nombre de régions, l'encombrement des caractéristiques, le ratio de compression et l'entropie des sous-bandes) pour mesurer leurs performances.

La vérité terrain du dataset a été obtenue en faisant comparer par 1 600 contributeurs un sous-ensemble d'images par paires dans un ensemble de plus de 37 000 images.

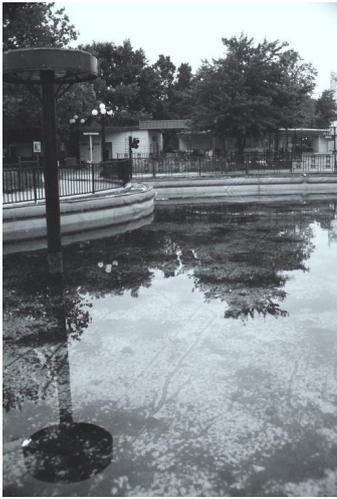
Pour illustrer, prenons l'exemple de deux images que l'on peut trouver dans l'ensemble de données (voir Figure 2.8). La sortie de l'ICR pour ces deux images est alors  $ICR(A, B)=4,032$  et  $ICR(B, A)=0.229$ , nous classons donc l'image A comme étant plus complexe que l'image B ce qui semble intuitivement cohérent.

Pour aller au-delà de la première impression « intuitive », nous avons proposé un protocole expérimental rigoureux.

### 2.4.2.2 Protocole

Nous avons utilisé la vérité terrain de l'expérience SAVOIAS pour faire une comparaison entre l'état de l'art et le classement de la complexité de l'information (ICR). La Table 2.2 donne le détail du nombre d'images par catégorie et du nombre de comparaisons résultantes.

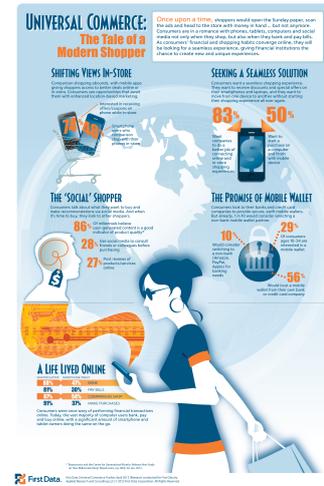
A la fin du processus, nous obtenons quatre valeurs pour le calcul de l'ICR. Les deux premières valeurs correspondent au calcul de  $ICR(A, B)$  pour lequel l'image A est concaténée avec l'image B dans la direction horizontale (respectivement verticale) tandis que les deux dernières valeurs correspondent à  $ICR(B, A)$ , pour lequel l'image B est concaténée avec l'image A dans les deux directions. Pour comparer les valeurs obtenues et la vérité terrain, nous classons les images en fonction de leur complexité. Pour cela,



(a) Exemple d'image de la catégorie Scènes



(b) Exemple d'image de la catégorie Publicités



(c) Exemple d'image de la catégorie Visualisation



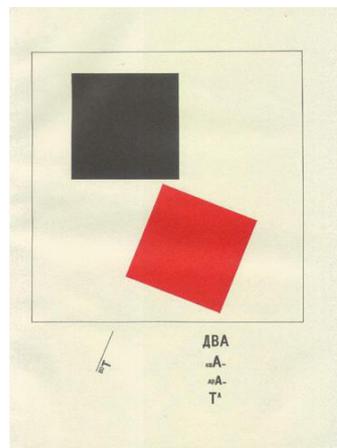
(d) Exemple d'image de la catégorie Objets



(e) Exemple d'image de la catégorie Décoration intérieure



(f) Exemple d'image de la catégorie Art



(g) Exemple d'image de la catégorie Suprématisme

FIGURE 2.7 – Exemples d'images du dataset SAVOIAS pour chaque catégorie.

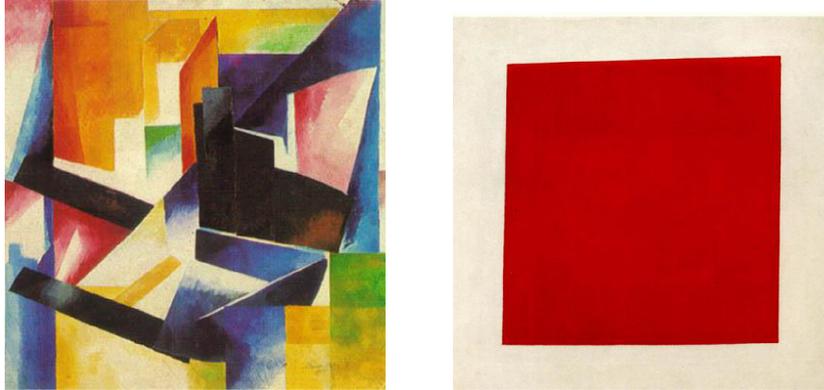


FIGURE 2.8 – Exemple de deux images du dataset (à gauche : IMAGE A et à droite : IMAGE B) pour notre méthode Information Complexity Ranking (ICR).

Catégorie	Nombre d'images	Nombre de comparaisons	Nombre total d'images traitées
Scènes	200	$(200 \times 200) \times 2$	80 000
Publicités	200	$(200 \times 200) \times 2$	80 000
Visualisations	200	$(200 \times 200) \times 2$	80 000
Objets	200	$(200 \times 200) \times 2$	80 000
Décoration intérieure	100	$(100 \times 100) \times 2$	20 000
Art	420	$(420 \times 420) \times 2$	352 800
Suprématisme	100	$(100 \times 100) \times 2$	20 000

TABLE 2.2 – Nombre d'images générées pour tester ICR. Le facteur 2 appliqué dans la colonne *Nombre de comparaisons* représente la concaténation horizontale et verticale de chaque image.

nous utilisons la méthode de Condorcet [Young, 1988]. Ce système de vote, établi en 1785 par Nicolas de Condorcet, permet de désigner le vainqueur d'une élection, en contrastant le nombre de voix dans une comparaison directe. En application de cette méthode, la valeur maximale de chaque comparaison par paire était considérée comme un « vote ». Tous les votes sont ensuite additionnés pour chaque image, ce qui donne finalement un classement de toutes les images. Enfin, afin de comparer le rang de complexité donné par la vérité terrain et le rang donné par notre méthode, nous calculons la corrélation de Spearman, qui est un estimateur robuste, ce qui permet de limiter la potentielle influence de données aberrantes venant d'un ou plusieurs contributeurs du jeu de données. La valeur du coefficient de Spearman est comprise entre -1 et 1, une valeur nulle indique que les deux rangs ne sont pas corrélés, contrairement à une valeur tendant vers 1 indiquant une forte corrélation entre les deux rangs. À l'opposé, si la valeur du coefficient tend vers -1, une forte corrélation existe, mais avec une variation inverse.

La corrélation de Spearman est définie comme suit :

$$r_S = \frac{\text{cov}(rg_X, rg_Y)}{\sigma_{rg_X} \sigma_{rg_Y}} \quad (2.22)$$

où  $\text{cov}(rg_X, rg_Y)$  est la covariance des variables de rang  $rgX_i$  et  $rgY_i$  et  $\sigma_{rg_X}$  et  $\sigma_{rg_Y}$  sont les écarts-types des variables de rang.

Pour démontrer la robustesse de notre proposition, nous l'avons également comparée à cinq autres mesures de complexité visuelle de pointe utilisées dans l'article du jeu de données SAVOIAS et présentées précédemment :

- **la densité des contours** est le pourcentage de pixels qui sont des pixels de contours dans une image ;
- **l'encombrement des caractéristiques** représente l'encombrement visuel et la difficulté d'ajouter un nouvel élément susceptible d'attirer l'attention sur un affichage ou une scène [Rosenholtz et al., 2007] ;
- **l'entropie des sous-bandes** repose sur l'idée que l'encombrement est lié au nombre de bits nécessaires pour coder l'image en sous-bandes (ondelettes) [Rosenholtz et al., 2007] ;
- **le nombre de régions** équivaut à compter les régions les plus denses dans l'espace des caractéristiques d'une image [Comaniciu and Meer, 2002] ;
- **le taux de compression** correspond au bppR (bit-per-pixel ratio), qui est le rapport entre le bpp d'une image déformée (par un sous échantillonnage chromatique) et le bpp d'une version compressée sans perte de l'image originale (en JPEG) [Corchs et al., 2014].

À des fins de généralisation, il est important de rappeler qu'ICR peut être appliquée à tous les types de données, contrairement à ces méthodes qui ne s'appliquent qu'aux images.

### 2.4.2.3 Résultats

Notre proposition est basée sur la combinaison de NCD et du taux de compression ( $ICR(x, y) = NCD(x, y) \times C(y)/C(x)$ ), il est donc important de démontrer ses perfor-

Catégorie	NCD	ICR
Scènes	0.23	0.40
Publicités	0.16	0.58
Visualisation	0.13	0.31
Objets	0.05	0.16
Décoration intérieure	0.52	0.67
Art	0.28	0.58
Suprématisme	0.47	0.73

TABLE 2.3 – Comparaison entre ICR et NCD par rapport à la perception humaine. Les valeurs de NCD et d’ICR présentées sont la moyenne des valeurs obtenues avec les algorithmes PNG, PPM et LZ77.

mances par rapport aux éléments individuels qui la composent. Comme présenté dans la Table 2.3, notre méthode ICR présente de meilleures corrélations avec la perception humaine dans toutes les catégories du jeu de données par rapport à l’utilisation de NCD seule. En comparant les résultats de la Table 2.3 et le taux de compression seul de la Table 2.4 (sixième colonne), nous constatons un apport de la NCD dans les catégories *Scènes*, *Publicités*, *Art* et *Suprématisme*. Dans les catégories *Objets* et *Décoration intérieure*, les résultats de notre méthode sont similaires au taux de compression seul et supérieurs au NCD seul. Le coefficient sur la catégorie *Visualisations* est moins performant, car les images de cette catégorie sont très hétérogènes, donc la contribution de NCD est par conséquent presque inexistante.

Catégorie	Densité contours	Nombre régions	Congestion des caractéristiques	Entropie des sous-bandes	Ratio de Compression	ICR
Scènes	0.16	<b>0.57</b>	0.42	0.16	0.30	0.40
Publicités	0.54	0.41	0.56	0.54	0.56	<b>0.58</b>
Visualisations	0.57	0.38	0.52	<b>0.61</b>	0.55	0.31
Objets	0.28	0.29	<b>0.30</b>	0.10	0.16	0.16
Décoration intérieure	0.61	0.67	0.58	0.31	<b>0.68</b>	0.67
Art	0.48	<b>0.65</b>	0.22	0.33	0.51	0.58
Suprématisme	0.18	<b>0.84</b>	0.48	0.39	0.60	0.73

TABLE 2.4 – Comparaison entre ICR et différentes mesures de la complexité visuelle dans la littérature, par rapport à la perception humaine. Les valeurs des cinq premières colonnes sont reprises de la Table 3 dans l’article [Saracee et al., 2018].

Comme présenté dans la Table 2.4, nous pouvons remarquer que notre méthode présente les meilleurs résultats pour les catégories *Publicités* et *Design intérieur*. Les résultats sur les catégories *Art*, *Suprématisme* sont meilleurs que les mesures de densité des bords, de congestion des caractéristiques, d’entropie des sous-bandes et de taux de compression. Dans la catégorie *Visualisations*, les résultats sont inférieurs à ceux des autres méthodes. En ce qui concerne les images représentant des objets, les performances sont inférieures aux autres mesures, sauf pour l’entropie de sous-bande.

Cette différence pour la catégorie *Visualisations* (voir Figure 2.9) pourrait s’expliquer

par la grande diversité des images, ce qui implique une faible similarité. Cette notion étant centrale dans notre proposition, les résultats produits sont moins bons que dans les autres catégories. Le style de peinture Suprematism, quant à lui, est basé sur des formes géométriques telles que des cercles, des carrés ou des rectangles. Il se caractérise également par une palette de couleurs réduite. Cela renforce la capacité de notre mesure à mieux classer la complexité en détectant les nombreuses répétitions contenues dans ces images. Les améliorations possibles seront développées dans la section suivante. Notons que, quel que soit le sens de concaténation (horizontale ou verticale), le coefficient de corrélation de notre approche est identique, ce qui traduit l'absence d'influence du sens de concaténation sur la complexité de l'image.

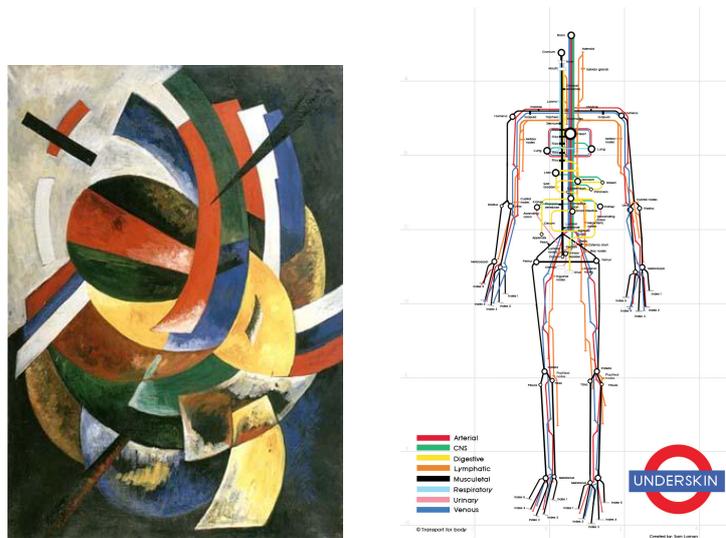


FIGURE 2.9 – Exemples d'images dans la catégorie *Suprematism* et *Visualizations*.

### 2.4.3 Expérimentation avec la profondeur logique de Bennett

Au regard de la littérature dans le domaine, présentée dans la section 2.2 et de la nature de l'information transmise, principalement des images, nous avons testé les performances de la profondeur logique de Bennett. L'objectif de cette expérimentation est de trouver une meilleure corrélation entre la perception humaine de la complexité et une mesure algorithmique de cette dernière. Afin de mesurer la profondeur logique de Bennett le plus fidèlement possible, nous nous basons sur les travaux sur la décompression pour la classification d'image d'Hector Zenil [Zenil, 2011, Zenil et al., 2012]. La mise en œuvre de cette mesure passe donc par le calcul du temps de décompression de chaque image.

### 2.4.3.1 Protocole

Comme pour l'expérimentation précédente, nous avons repris la même base de comparaison, à savoir les algorithmes de compression sans perte, PNG, PPM et LZ77, ainsi que les images du dataset SAVOIAS [Saracee et al., 2018]. Pour implémenter cette mesure du temps, nous avons compressé et décompressé à plusieurs reprises les différentes images à l'aide de trois bibliothèques Python de référence, **gzip** [Foundation, 2023a] pour l'algorithme LZ77, **zlib** [Foundation, 2023c] pour l'algorithme PNG et **py7zr** [Miura, 2023] pour l'algorithme PPM.

Pour obtenir une mesure la plus précise possible, nous utilisons le profileur **cProfile** [Rosen and Czotter, 2023] intégré à Python. L'avantage de cette bibliothèque est son relevé d'information détaillé comprenant par exemple le temps d'exécution total du code profilé, le temps d'exécution de chaque fonction, mais aussi le temps cumulé d'une fonction particulière ainsi que celui des sous-fonctions associées. C'est cette dernière information qui nous intéresse particulièrement. Pour formater les résultats ainsi obtenus, nous utilisons la bibliothèque **pstats** [Foundation, 2023b] qui permet d'imprimer les valeurs souhaitées. Dans cette expérimentation, nous utilisons uniquement le temps cumulé de chaque fonction de décompression.

La machine sur laquelle ont été menés les calculs n'exécutait que les programmes nécessaires au bon déroulement de la compression et de la décompression des images. À savoir le système d'exploitation (Windows 10), l'interpréteur Python et l'environnement de développement intégré (Visual Studio Code). Tous les autres programmes ont été arrêtés lors des mesures. Toutes les opérations ont été menées sur le CPU de la machine<sup>2</sup>. Du fait de la possible instabilité des temps de calcul de la machine, nous avons pris la moyenne des trois algorithmes de compression comme mesure de référence. Nous avons réutilisé la vérité terrain de l'expérience SAVOIAS pour faire une comparaison entre la perception humaine et le classement des images établi grâce à la profondeur logique. Pour terminer, afin de comparer le rang de complexité donné par la vérité terrain et le rang donné par le temps de décompression, nous calculons la corrélation de Spearman.

### 2.4.3.2 Résultats

Comme présenté dans la Table 2.5, nous pouvons constater une très faible corrélation dans les catégories *Publicités*, *Visualisations*, *Objets et Art*. Compte tenu de la difficulté de mesurer le temps d'exécution et l'instabilité possible de l'outil de mesure, nous pouvons conclure à une absence de corrélation entre la profondeur logique de Bennett et la perception humaine de la complexité dans ces catégories. Concernant les trois autres catégories *Scènes*, *Décoration intérieure et Suprématisme*, nous pouvons constater un coefficient de corrélation bien inférieur à celui obtenu grâce à notre méthode ICR mais équivalent à certains algorithmes de mesure de la complexité présentés dans la section 2.4.2.

---

2. Intel Core I7-8665U 1,90Ghz / 32 Go de RAM

Catégorie	Bennett	ICR
Scènes	0.15	<b>0.40</b>
Publicités	0.02	<b>0.58</b>
Visualisations	-0.06	<b>0.31</b>
Objets	-0.04	<b>0.16</b>
Décoration intérieure	0.23	<b>0.67</b>
Art	-0.02	<b>0.58</b>
Suprématisme	0.34	<b>0.73</b>

TABLE 2.5 – Comparaison des performances d'ICR et de la profondeur logique de Bennett par rapport à la perception humaine. Les valeurs de la profondeur logique de Bennett et d'ICR présentées sont la moyenne des valeurs obtenues avec les algorithmes PNG, PPM et LZ77, en utilisant le temps de décompression.

## 2.5 Conclusion

Comprendre l'impact de l'information sur le comportement des gens est un domaine important dans les systèmes adaptatifs ou dans l'aide à la décision. Il est également un indicateur explicatif important du comportement des consommateurs et notamment de leur traitement de l'information. Pour améliorer cette compréhension, nous proposons une nouvelle méthode pour hiérarchiser les informations en fonction de leur complexité. Ce classement est accompagné d'une nouvelle proposition de comparaison d'objets. Elle est basée sur la similarité entre deux objets (distance de compression normalisée, NCD) et sur la complexité intrinsèque au sens algorithmique du terme. La pondération de la similarité de plusieurs objets par leur complexité intrinsèque permet de repérer l'objet le plus complexe, et donc de les classer pour évaluer leur implication potentielle dans le processus de décision.

Au regard des résultats présentés dans ce chapitre, on peut légitimement dire que la mesure algorithmique de la complexité est un estimateur correct, mais incomplet dans le cadre de la perception humaine de la complexité. Il est difficile de rendre compte de la perception humaine de la complexité, que ce soit comme une complexité aléatoire mesurable à l'aide de la complexité de Kolmogorov, soit comme une complexité structurée mesurable à l'aide de la profondeur logique de Bennett. À la suite de nos expériences, nous pouvons dire qu'un outil composé d'une distance de similarité et d'une mesure de la complexité aléatoire offre une meilleure corrélation avec la perception humaine qu'une mesure prenant en compte la structure de l'information. Dès lors, nous pouvons émettre l'hypothèse que la structure informationnelle d'un objet est un acteur secondaire de la perception de la complexité chez l'être humain. Il est alors pertinent de s'interroger sur ce qui est réellement perçu par l'individu et si d'autres variables internes ou externes jouent un rôle particulier.

C'est dans ce cadre que nous avons décidé d'approfondir la perception de l'information chez l'être humain. Pour se faire, nous allons nous intéresser au traitement de l'information chez les individus. Ce processus cognitif nous paraît primordial pour comprendre quelle information peut engendrer un comportement spécifique. Ces aspects seront dé-

veloppés dans le chapitre 3. Au-delà d'une meilleure compréhension de l'impact de la complexité, ces premiers travaux, nous permettent également d'établir les bases d'un système de transmission de l'information optimisé (application mobile) pour faciliter son adoption et son utilisation sur le long terme. Ces aspects seront développés dans le chapitre 4.

Même si notre nouvelle méthode de classement basée sur la complexité offre des résultats encourageants, de futurs travaux sur la quantification de l'information pourront être menés afin d'améliorer la corrélation entre la méthode algorithmique et la perception humaine de la complexité, ces pistes de recherche seront développées dans le chapitre 5.3. Une application d'ICR est également possible dans d'autres cas d'utilisation. Certains domaines comme le traitement du langage naturel (NLP), la cybersécurité avec la détection de logiciels malveillants [Deng et al., 2011] ou même en génétique et phylogénie se basent sur des expériences utilisant NCD [Cilibrasi and Vitányi, 2005] ou la complexité de Kolmogorov.



#### Contributions de ce chapitre :

- Nous proposons une nouvelle méthode de classement de l'information au regard de sa complexité.
- Cette méthode, appelée Information Complexity Ranking (ICR), repose sur les notions de similarité entre deux objets et de complexité intrinsèque de chaque objet. La notion de similarité est représentée par la distance de compression normalisée (NCD) [Cilibrasi and Vitányi, 2005].
- Nous avons comparé cette méthode ICR à la perception de la complexité chez l'humain à l'aide d'un jeu de données venant de l'attention visuelle [Saraee et al., 2018].
- Nous avons étudié les performances de notre nouvelle méthode (ICR) et celles de la profondeur logique de Bennett [Bennett, 1988a] par rapport à la complexité perçue par 1600 personnes [Saraee et al., 2018].



## Chapitre 3

# Traitement de l'information, la complexité comme facteur d'influence du choix

---

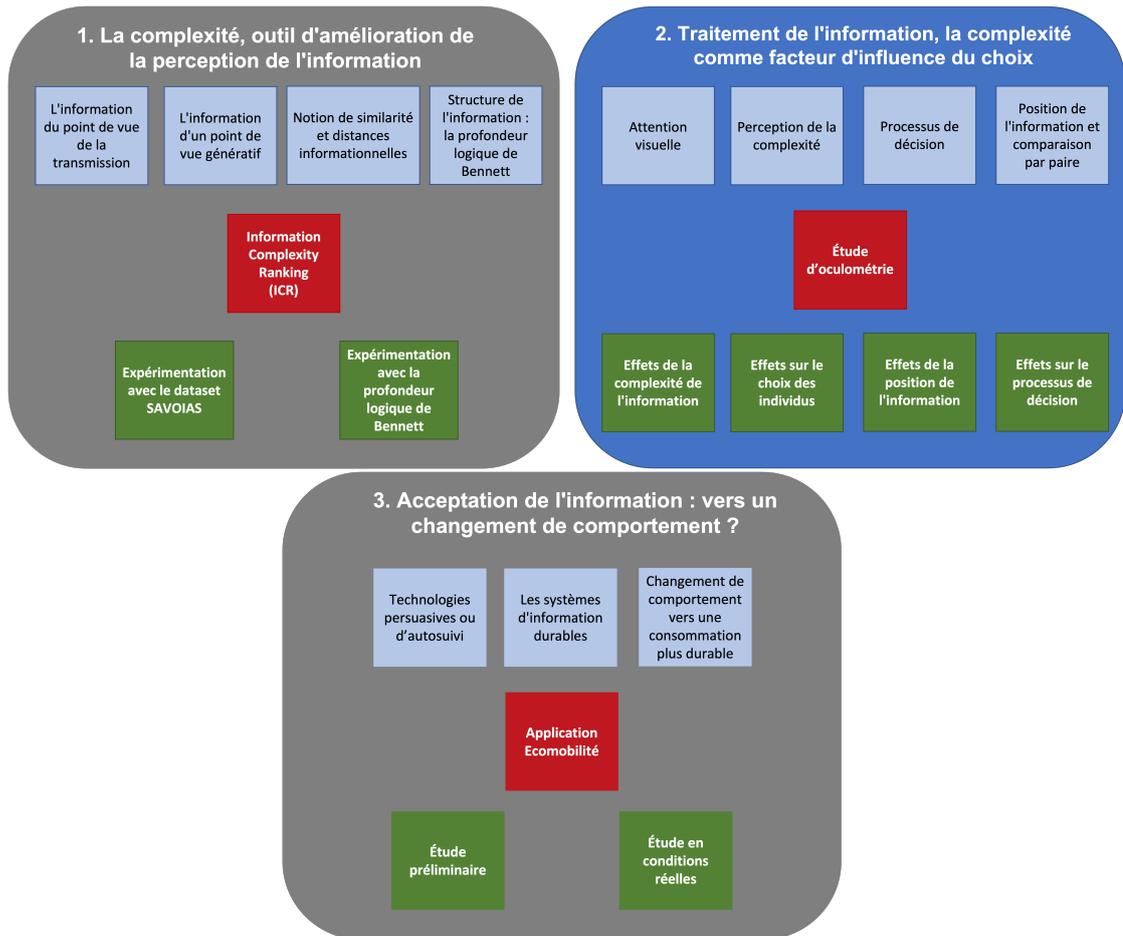
### Sommaire

---

<b>3.1</b>	<b>Introduction</b>	<b>49</b>
<b>3.2</b>	<b>État de l'art</b>	<b>50</b>
3.2.1	Attention visuelle du point de vue psychologique	50
3.2.2	Attention visuelle du point de vue computationnel	54
3.2.3	Perception de la complexité visuelle	57
3.2.4	Processus de décision : Le choix	59
3.2.5	Position de l'information et comparaison par paire	61
<b>3.3</b>	<b>Étude d'oculométrie - protocole et hypothèses</b>	<b>62</b>
3.3.1	Protocole expérimental	62
3.3.2	Hypothèses	66
<b>3.4</b>	<b>Résultats</b>	<b>68</b>
3.4.1	Effets de la complexité de l'information	68
3.4.2	Effets sur le choix des individus	69
3.4.3	Effets de la position de l'information	70
3.4.4	Effets sur le processus de décision	71
<b>3.5</b>	<b>Conclusion</b>	<b>73</b>

---

Les résultats présentés dans ce chapitre sont basés sur l'article « Effects of information complexity on user decision making : an eye-tracking study to build a mobile application for carbon footprint estimation. » actuellement soumis.



## 3.1 Introduction

La réduction des émissions de CO<sub>2</sub> dans nos trajets quotidiens présente des avantages à la fois individuels et collectifs. L'amélioration de la qualité de l'air permettrait à chaque habitant d'être en meilleure santé et de profiter d'une qualité de vie urbaine supérieure. Individuellement, les personnes ont conscience des bienfaits et de la nécessité de diminuer rapidement les émissions de CO<sub>2</sub> mais, même si les alternatives de transport moins émissives existent, elles ne sont que trop peu utilisées comme le rappelle l'Insee [Brutel Chantal, 2021] :



Pour des distances inférieures à 5 kilomètres, la voiture représente encore 60% des déplacements domicile-travail.

*INSEE 2021*

Une question se pose alors : comment inciter les personnes à adopter durablement des comportements moins émetteurs de CO<sub>2</sub> ? De manière plus générale, l'idée de construire des outils d'aide au changement a alors émergée et des chercheurs de multiples champs disciplinaires, notamment en sciences du comportement, travaillent sur de tels outils [Bruck et al., 2012, Lehto and Oinas-Kukkonen, 2015]. Malgré quelques succès [Dubois et al., 2021, Nabec et al., 2022], ces tentatives de changement durable des comportements sont au mieux limitées à de petits groupes de personnes sensibilisées [Froehlich et al., 2009, Tulusan et al., 2012].

Pour mieux comprendre ces résultats mitigés, l'un des axes de recherche est de travailler sur les causes potentielles du fossé entre intention comportementale et comportement réel. En particulier, sur la manière dont la complexité visuelle est perçue et influence une attitude et forge un comportement. Cet aspect est absent des études précédemment citées bien qu'elle joue un rôle majeur dans le processus d'incitation au changement : si une information est peu complexe ou au contraire trop complexe, son traitement cognitif après sa visualisation sera ignoré ou aura un coût cognitif prohibitif. Dans les deux cas, elle ne déclenchera pas de motivation durable au changement.

Pour comprendre comment s'effectue la sélection d'une information quelle que soit sa complexité, l'étude des mécanismes d'attention visuelle est un préalable indispensable. L'attention jouant un rôle tout au long du traitement perceptif de l'information [Kristjansson, 2006, Noudoost et al., 2010], il est nécessaire de maîtriser les caractéristiques attractives d'un stimulus et de comprendre comment elles peuvent être améliorées [Itti and Koch, 2000, Itti and Baldi, 2009, Koch and Ullman, 1985, Treisman and Gelade, 1980].

Dans un souci de compréhension globale du processus de décision, nous explorons dans un premier temps, la notion d'attention visuelle qui participe activement à la construction cognitive de la préférence et au processus de sélection de l'information [Shimojo et al., 2003]. Dans un second temps nous intégrons la complexité dans les réponses cognitives des individus et présentons la notion psychologique de la perception de la complexité [Berlyne, 1960]. Enfin, nous mobilisons des notions concernant les phases d'attention associées au

choix [Russo and Leclerc, 1994, Wedell and Senter, 1997, Clement, 2007].

Pour prendre en compte ces trois éléments et mieux comprendre comment l'information transmise est prise en compte dans le processus de décision, nous proposons une étude d'oculométrie. Nous proposons pour cela plusieurs écrans d'accueil de complexité variable d'une application fictive de suivi de la consommation quotidienne de CO<sub>2</sub> et étudions les réponses des participants. Cette étude nous permettra d'établir des points de repères sur le processus de décision et les réponses associées à des informations de complexité différente.

Dans ce chapitre nous présentons nos travaux exploratoires sur le choix d'une information en fonction de sa complexité par le biais d'une expérience d'oculométrie. Plus précisément nous décrivons les aspects théoriques de l'attention visuelle ainsi que de la préférence visuelle et le processus de décision, dans la Section 3.2. Le protocole de l'expérience et les hypothèses sont présentés dans la Section 3.3 et les résultats associés dans la Section 3.4.

## 3.2 État de l'art

### 3.2.1 Attention visuelle du point de vue psychologique

Pour pouvoir mesurer l'éventuel impact d'une information sur un individu, il est obligatoire de commencer par comprendre ce qui attire son attention. Néanmoins, il convient de souligner l'importance de cette première étape dans notre démarche, car l'attention visuelle peut être accompagnée d'un phénomène appelé cécité d'inattention ou cécité perceptive [Mack and Rock, 1998]. Cet effet de cécité temporaire survient lorsqu'un individu ne perçoit pas un stimulus inattendu mais parfaitement visible, par exemple une nouvelle information. L'occurrence de ce phénomène est due à l'attention de l'observateur qui est déjà portée sur d'autres stimuli. De nombreuses études [Most et al., 2001, Hyman et al., 2009, Memmert, 2006] ont démontré l'existence d'un tel phénomène. La plus célèbre d'entre elles, est celle de Simons et Chabris [Simons and Chabris, 1999]. Dans cette étude, il est demandé aux participants de compter le nombre de passes effectuées par deux équipes composées de trois joueurs de basket-ball. À la fin de la vidéo, en plus de répondre à la tâche principale, les participants sont invités à répondre à une série de questions additionnelles, portant sur la perception d'un événement inattendu lors de la vidéo. « L'événement surprenant » est soit une femme portant un parapluie traversant la scène ou une femme portant un costume de gorille. Sur 192 observateurs retenus, seulement 54% a perçu l'événement. Compte tenu de ce phénomène, une complexité de l'information trop élevée pouvant entraîner une surcharge cognitive, il est crucial de minimiser un besoin d'attention supplémentaire, afin de ne pas exclure d'informations du processus de décision.

C'est dans le domaine de l'attention que la théorie de l'intégration des caractéristiques (Feature Integration Theory ou FIT) de Treisman et Gelade définit le mécanisme de la perception de l'information [Treisman and Gelade, 1980]. Dans leur modèle, les caractéristiques qui peuvent être considérées comme attractives sont la couleur, l'orien-

tation, la fréquence spatiale, la luminosité et la direction du mouvement. En 2004, Wolfe et Horowitz, ont défini la couleur et le mouvement comme étant deux caractéristiques attractives [Wolfe and Horowitz, 2004]. Au-delà de définir des « fixateurs » d'attention, les travaux de Treisman et Gelade proposent une compréhension du mécanisme de la perception de ces éléments particuliers au sein d'une scène. Les caractéristiques sont perçues automatiquement de manière précoce et les objets en eux-mêmes sont identifiés lors d'une étape ultérieure, avec le concours d'une attention spécifique (voir Figure 3.1). En d'autres termes, l'individu perçoit en premier les caractéristiques séparables (attracteurs) d'un objet, puis perçoit l'objet dans son entièreté par intégration de ces dernières. Cette intégration repose sur la combinaison des caractéristiques d'une même fixation centrale, l'attention focale est alors considérée comme la « colle » qui permet de percevoir les objets (voir Figure 3.2).

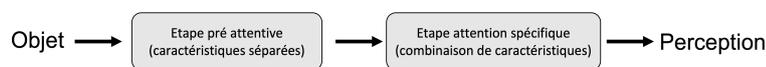


FIGURE 3.1 – Représentation simplifiée des deux étapes du modèle de la Feature Integration Theory.

Cette théorie explique également le rôle précoce de l'attention, qui est opérationnelle dès les premiers stades de la vision. L'intégration de toutes les combinaisons possibles de toutes les caractéristiques d'une scène est une tâche particulièrement lourde d'un point de vue cognitif. Afin d'éviter l'explosion combinatoire qui en résulte, le système d'attention ne fournit que la combinaison de caractéristiques simples à des endroits spécifiques de la scène [Hoffman, 1986]. Les entrées visuelles sont décomposées en dimensions distinctes par un ensemble de canaux parallèles. Cette décomposition renforce la cohérence de la FIT au regard des preuves physiologiques et psychophysiques [Regan, 1982].

La localisation de l'attention est aussi une notion importante à prendre en considération. Des travaux dans ce domaine ont prouvé l'existence de la notion de localisation de l'attention visuelle qui agit comme un focalisateur de l'attention, permettant de meilleures réponses aux stimuli de la part de l'individu. Prenons l'exemple de la perception des lettres et des mots qui sont des formes familières. La détection de ce type d'éléments est « automatique » chez un individu, cependant elle nécessite tout de même une attention particulière. Cette attention spatiale joue également un rôle fonctionnel dans la détection de la présence d'un élément cible. De plus, les réponses associées à ces formes sont réduites si elles ne se trouvent pas au centre de l'attention [Francolini and Egeth, 1980, Hoffman et al., 1983, Kahneman and Treisman, 1984].

Hillyard et Munte ont prouvé que la localisation joue un rôle spécial dans l'attention visuelle [Hillyard and Münte, 1984]. Dans cette étude, ils ont enregistré le potentiel cérébral (Event-Related potential en anglais) d'individus en réponse à des tâches de discrimination visuelles basées sur la localisation spatiale, la couleur et la longueur de barres verticales. Les résultats montrent que l'analyse de la caractéristique de la couleur du stimulus dépendait d'un test de localisation au préalable. Ce rôle particulier est renforcé par le fait que la sélection sur cette caractéristique est complètement différente de la

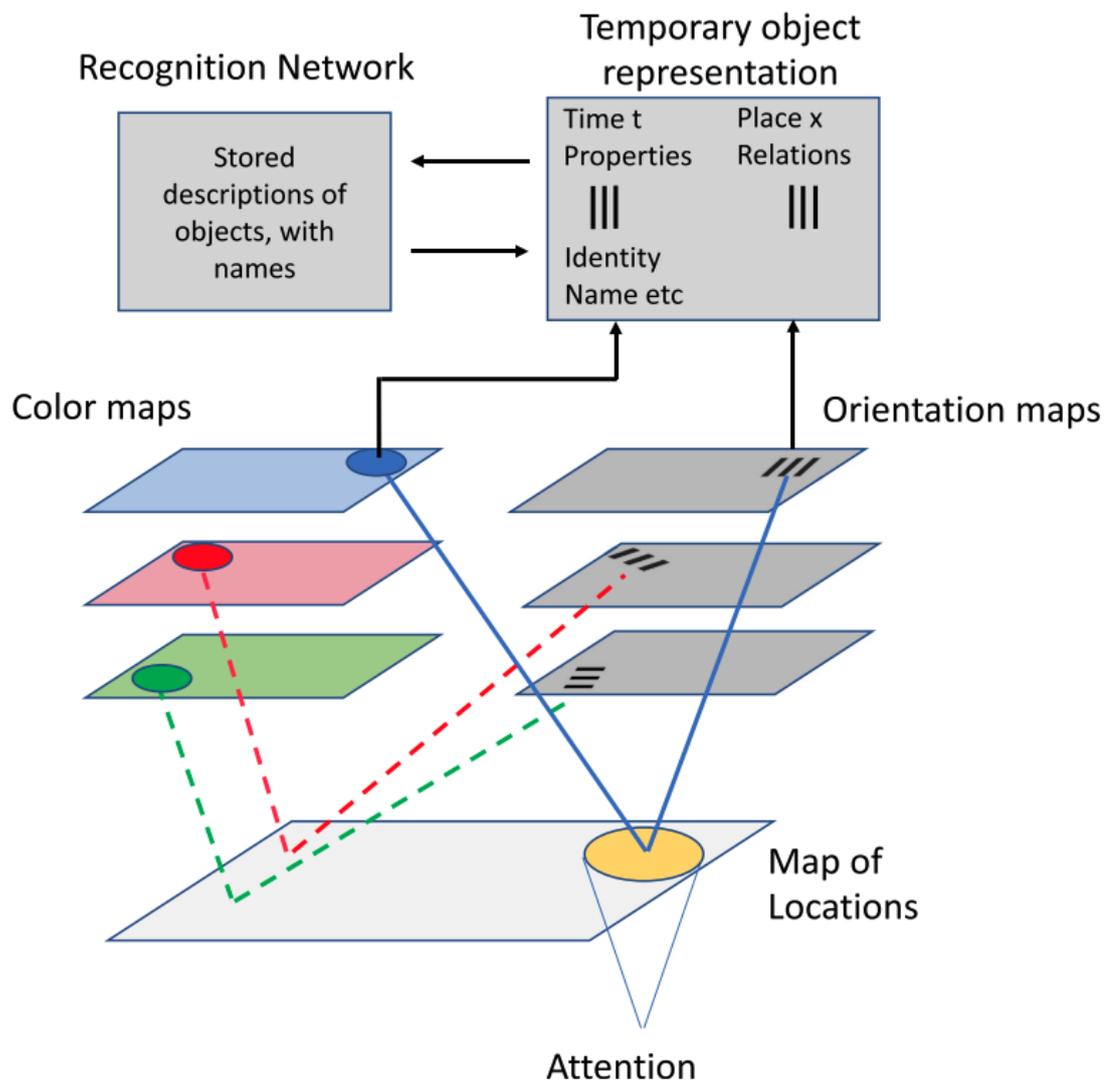


FIGURE 3.2 – Représentation théorique de la FIT selon [Kristjansson and Egeth, 2020]

sélection sur d'autres caractéristiques comme la couleur ou l'orientation [Nissen, 1985].

La principale preuve de la robustesse de la FIT provient de la distinction désormais classique entre la recherche visuelle basée sur la conjonction et celle basée sur les caractéristiques. Lorsque les participants doivent rechercher un élément cible donné, leur efficacité est meilleure lorsque la cible est définie par une différence dans l'une de ses caractéristiques par rapport aux distracteurs. Selon cette théorie, les cibles caractéristiques ressortent dans la recherche parce qu'elles peuvent être détectées sur la base de l'activation non ambiguë de la carte des détecteurs répondant à cette caractéristique [Humphreys, 2016]. La variation de la recherche de caractéristiques a été expliquée en recourant aux notions d'activité des caractéristiques sur les cartes associées et aux processus d'établissement de la moyenne de cette activité pour produire une réponse [Treisman and Souther, 1985]. Quelques années plus tard, Anne Treisman proposa une deuxième version de la FIT [Treisman, 1988, 1991, 1992] afin de prendre en compte l'évolution de la compréhension du balayage attentionnel. Cette nouvelle version a également remis en cause la distinction entre les deux étapes initiales, le stage préattentif et le traitement attentif.

Nous pouvons donc résumer la FIT en cinq points importants [Quinlan, 2003] :

- Les caractéristiques sont modélisées par une valeur spécifique sur une carte de caractéristiques particulières.
- Le partage de caractéristiques entre les éléments cibles et les éléments non-cibles, comme dans la tâche de recherche de conjonction, peuvent engendrer un rétrécissement du centre de l'attention et produire une fonction de recherche sérielle.
- La sélection attentionnelle est produite au niveau d'une carte des emplacements supérieure. L'accès à une localisation particulière de cette carte entraîne une récupération immédiate de tous les attributs provenant de cartes de caractéristiques correspondantes.
- La sélection attentionnelle est relative à un filtre sélectif, appelé fenêtre attentionnelle. Ce filtre parcourt la carte supérieure des emplacements. L'attention est répartie dans un premier temps sur l'ensemble de la scène. À la suite de la détection d'un élément cible, le filtre attentionnel est réduit sur la carte à l'emplacement correspondant (approche descendante).
- La notion d'inhibition sélective d'emplacements particuliers sur la carte supérieure à l'aide des détecteurs de caractéristiques permet une accélération de la recherche. De ce fait, les éléments non-cibles représentés sur la carte supérieure sont inhibés, ce qui permet de concentrer l'attention sur les éléments correspondants à la tâche de recherche.

Malgré de nombreuses interrogations sur la pertinence de la FIT, les travaux d'Anne Treisman ont posé les bases d'un domaine entier, celui de la recherche visuelle et des mécanismes associés, et a mis en avant une notion centrale, celle de l'attention [Wolfe, 2020]. Cette théorie fut également la base de développement de modèles d'attention visuelle computationnels permettant de mieux comprendre la perception humaine.

### 3.2.2 Attention visuelle du point de vue computationnel

Le développement de modèles reproduisant les mécanismes de l'attention visuelle directement ou indirectement inspirés par des concepts cognitifs a permis de mieux simuler et comprendre l'attention visuelle chez l'humain. Le domaine de la modélisation de l'attention visuelle est très actif et compte plusieurs approches réparties en trois principaux domaines : la psychophysique (relation entre un stimulus et sa perception), la modélisation computationnelle et la neurophysiologie. Si l'on se concentre sur la modélisation computationnelle, les deux grandes familles de modèles sont les modèles à filtres et les modèles connexionnistes, plus couramment appelés réseaux de neurones. Dans cette section nous nous intéressons uniquement à l'apport de la première catégorie de modèles (voir Figure 3.3). Nous nous concentrerons sur les modèles à filtres, car ils permettent de mieux comprendre les fondements biologiques de l'attention visuelle, tout en ayant une corrélation avec les mouvements oculaires humains dans des tâches de type visionnage libre qui correspondent à notre contexte de recherche. En effet, pour mesurer l'impact de la complexité de l'information sur les intentions et les comportements, il est essentiel de ne pas donner d'instructions de visionnage aux individus.

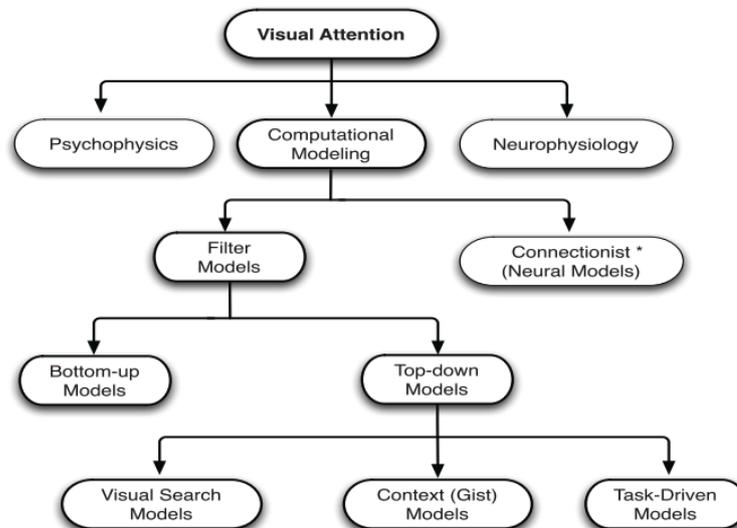


FIGURE 3.3 – Taxonomie du domaine de la modélisation de l'attention visuelle issue de [Borji and Itti, 2010]

De ces premières recherches basées sur la FIT ont émergé la notion de saillance et de combinaisons de caractéristiques [Koch and Ullman, 1985]. La notion de saillance est centrale dans la modélisation de l'attention visuelle, car elle permet de caractériser intuitivement certaines parties de la scène qui seront « mises en avant » par rapport aux autres [Borji and Itti, 2010]. Le principe de ces modèles biologiquement inspirés repose sur le calcul d'une carte de saillance qui permet de déterminer le niveau d'activité de chaque élément composant une scène. Cette carte générale de saillance est elle-même construite

à partir de différentes cartes spécifiques propres à chaque caractéristique élémentaire se trouvant dans la scène [Koch and Ullman, 1985]. Le but de cette carte topographique est de prédire la conspécuité d'un endroit précis et par extension de prédire où l'attention visuelle va se concentrer. La première implémentation complète et testée du modèle de Christof Koch et Shimon Ullman a été proposée, une dizaine d'années plus tard par Laurent Itti, Koch lui-même et Ernst Niebur [Itti et al., 1998].

Pour mieux comprendre quelles sont les caractéristiques à privilégier, il est pertinent de décrire son fonctionnement (voir Figure 3.4), qui inspirera d'autres modèles et deviendra une référence pour la comparaison. Ce modèle, plus généralement appelé modèle d'Itti repose sur l'utilisation de trois canaux de caractéristiques : l'intensité, la couleur et l'orientation [Itti et al., 1998]. Il est composé d'une carte de saillance dite spécialisée pour traiter la couleur, et de deux autres cartes pour traiter l'intensité et l'orientation. La première étape de traitement consiste en la décomposition de l'image d'entrée en suivant les trois caractéristiques principales, et également en fonction de la spatialité de chaque élément. Cette étape donne lieu à la création de cartes de caractéristiques. La combinaison ascendante de ces dernières produit une carte de saillance principale (master, en anglais). Cette carte principale, code topographiquement la perception de chaque caractéristique sur l'ensemble de la scène [Itti et al., 1998]. Pour prédire la prochaine position de l'attention, le modèle propose un réseau de neurones winner-take-all en tant que dernière étape de traitement en sélectionnant le point le plus saillant de la carte principale. Ce réseau est accompagné d'un mécanisme d'inhibition permettant le changement de la position de l'attention, en sélectionnant le prochain point le plus saillant, pour éviter la fixation indéfinie d'une même position. Plusieurs évolutions de ce modèle ont été proposées [Itti and Koch, 2000], dont une avec l'intégration d'une quatrième caractéristique, le mouvement [Itti, 2004].

La variation de ces propriétés peut être d'une grande aide pour créer des attracteurs ou des éléments dissimilaires au regard de leurs proches voisins, ce qui permet d'attirer l'attention. Ces éléments différenciés peuvent prendre place dans deux dimensions différentes, l'espace et le temps. Pour les décrire dans l'espace on utilisera la notion de saillance et celle de nouveauté pour les décrire dans le temps [Itti and Baldi, 2009].

L'évolution de ces modèles de simulation et leur corrélation avec l'attention visuelle humaine ont permis l'émergence d'une multitude d'applications. Celles-ci s'inscrivent principalement dans les domaines de la vision assistée par ordinateur, l'infographie et la robotique [Borji and Itti, 2010]. Un autre champ disciplinaire, proche de notre contexte de recherche est celui de la conception d'interfaces utilisateur et de la visualisation d'informations. En se basant sur les prédictions de ces mêmes modèles, une étude a été menée auprès d'équipes de conception d'interfaces pour savoir si les prédictions résultantes étaient un bon outil d'aide à la conception. Les résultats montrent qu'elles sont un facilitateur de communication entre concepteurs et un améliorateur de la perception de la qualité de conception. Cependant l'intégration des modèles devrait être utilisée avec parcimonie pour éviter des interférences dans le processus de conception [Rosenholtz et al., 2011]. Une autre étude a montré l'efficacité de ces modèles dans le domaine de la publicité vidéo avec des systèmes d'insertion de contenu virtuel [Liu et al., 2008].

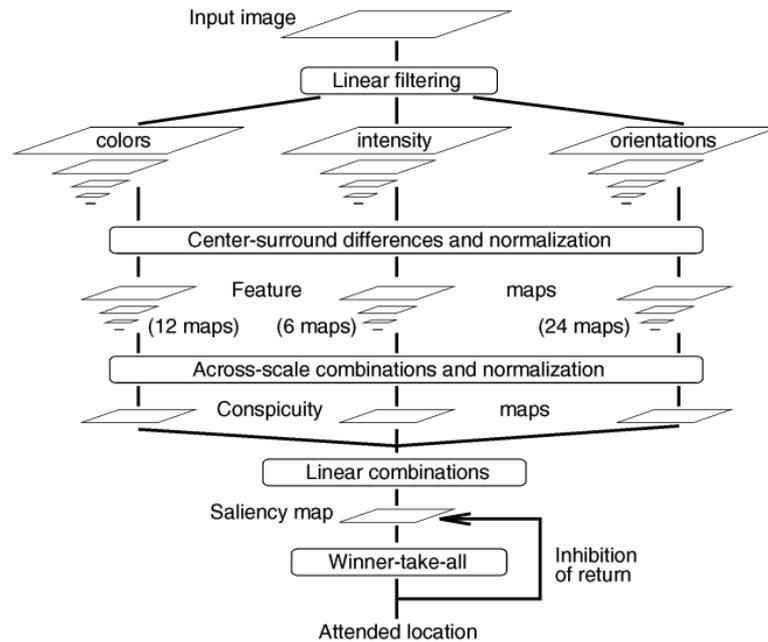


FIGURE 3.4 – Architecture générale du modèle d'Itti, Koch et Niebur issue de [Itti et al., 1998]

Le domaine de l'attention visuelle s'intéresse à la compréhension et la sélection des propriétés discriminantes pour attirer l'attention et donc maximiser le traitement de l'information, ce qui pourrait éventuellement conduire à une intention comportementale. Nous devons donc axer notre travail sur la différenciation des caractéristiques d'un élément pour le rendre plus attractif, telles que l'orientation, la couleur ou encore l'intensité. La modélisation de l'attention visuelle a permis également de tester la cohérence et la prépondérance de certaines caractéristiques jouant le rôle d'attracteur de l'attention visuelle.



**À retenir :** Les travaux de Treisman et Gelade et plus particulièrement leur *Feature Integration Theory* ont posé les bases d'une nouvelle compréhension de l'attention visuelle chez l'humain. Une des principales contributions, outre le traitement en deux étapes, est la définition des caractéristiques qui focalisent l'attention (couleur, orientation, fréquence spatiale, luminosité et direction du mouvement), ce qui est confirmé par les modèles d'attention visuelle de type cognitifs, par exemple avec les modèles de Koch et Ullman, ou celui d'Itti.

Les recherches associées au domaine de l'attention visuelle nous permettent de présélectionner les caractéristiques les plus influentes, afin d'étudier le rôle de l'information dans le processus de décision.

### 3.2.3 Perception de la complexité visuelle

Afin de savoir ce qui encourage une personne à choisir un stimulus visuel par rapport à un autre, il est important de comprendre les travaux de l'un des pionniers dans le domaine de la perception et du comportement en réponse à un stimulus. Daniel Ellis Berlyne, avec son livre intitulé « Conflict, Arousal, and Curiosity » [Berlyne, 1960], définit quatre concepts à prendre en compte dans la sélection d'un stimulus : la nouveauté, l'incertitude, le conflit et la complexité. Dans les travaux présentés dans cette thèse, nous nous intéressons uniquement à la dernière de ces propriétés, la complexité.



La complexité est « le concept le plus impalpable des quatre concepts insaisissables ».

*Daniel Ellis Berlyne, Conflict, Arousal, and Curiosity, 1960, citation traduite de l'anglais.*

Berlyne définit certaines propriétés qui influent sur le concept de complexité :

- **Le nombre d'éléments** permet d'augmenter la complexité d'un stimulus.
- **La nature des éléments** permet de faire baisser la complexité avec des éléments possédant des caractéristiques similaires et inversement avec les caractéristiques d'éléments dissimilaires. La complexité peut être réduite en plaçant des éléments identiques à différents endroits (par exemple avec une répétition cyclique ou une symétrie de translation).
- **L'orientation des éléments** provoquera automatiquement une augmentation de la complexité perçue d'un stimulus s'ils sont orientés différemment.
- **La perception des éléments** en tant qu'unité est importante. La complexité diminue si les éléments sont perçus comme un ensemble (par exemple la proximité des lettres formant un mot).

Tout comme les travaux d'Anne Treisman évoqués précédemment, Berlyne se concentre sur la localisation et plus précisément sur la proximité entre les éléments dans le traitement des stimuli par les individus. Outre ces éléments, la variation de la complexité dépend également de facteurs psychologiques. La perception des propriétés intrinsèques des stimuli sera équivalente pour tous les individus, mais son effet psychologique fluctuera pour chaque individu.

Une autre contribution majeure des travaux de Berlyne est la notion d'efficacité [Berlyne, 1960]. L'efficacité permet d'analyser la relation entre la complexité d'un stimulus et sa sélection ou non par un individu. La théorie proposée par Berlyne prédit une relation inverse et curviligne entre la complexité et l'efficacité d'un stimulus [Berlyne, 1970]. Comme on peut le constater dans la Figure 3.5, la notion de plaisir représentée par l'axe des ordonnées (Hedonic value) augmente progressivement jusqu'à atteindre un pic avec une complexité modérée, représentée par l'axe des abscisses (Arousal potential). Avec cette courbe, on peut également noter le côté aversif d'un stimulus très complexe, provoquant un fort effet négatif sur le plaisir perçu. Cette théorie a été validée par d'autres études [Day, 1967, Morrison and Dainoff, 1972].

Même si le concept de complexité reste abstrait, il n'en reste pas moins modulable,

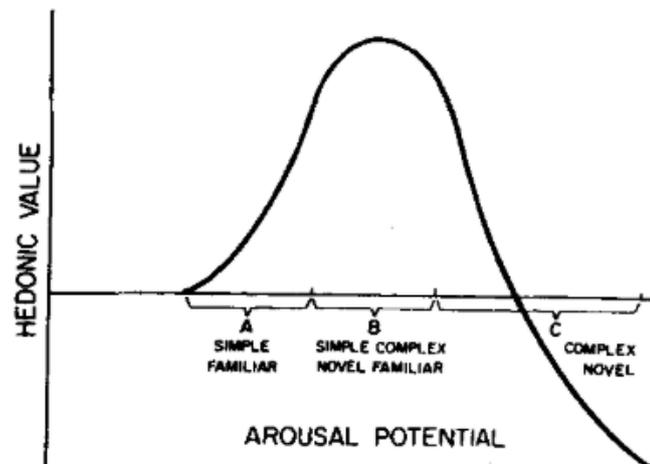


FIGURE 3.5 – Courbe de Wundt

justifiant l'intérêt de poursuivre les recherches sur ses potentiels effets. Il est à noter qu'il n'y a ce jour pas de consensus sur un unique niveau de complexité bénéfique pour l'individu. Certains chercheurs mettent en avant les effets positifs d'une complexité modérée et les effets négatifs d'une complexité élevée. Geissler et al. [Geissler et al., 2006] ont démontré que les individus répondent plus favorablement à des pages web d'accueil de complexité modérée. Précédemment à cette étude, deux autres études ont prouvé l'influence de la complexité des pages web sur les attitudes et les intentions de consommation associées [Bruner and Kumar, 2000, Stevenson et al., 2000]. Un lien a également été établi entre les réactions émotionnelles (c'est-à-dire les propriétés psychologiques) d'individus et une complexité visuelle élevée. Une complexité visuelle trop élevée provoque ainsi des réactions émotionnelles négatives [Deng and Poole, 2012]. Il est à noter qu'un ensemble important de travaux, basés sur la théorie de Berlyne, concluent également qu'une complexité modérée a un effet plus significatif sur les attitudes et les intentions [Wang and Lin, 2019, Deng and Poole, 2012]. La fluidité du traitement de l'information montre que les personnes ayant une capacité d'acquisition limitée sont plus réceptives aux informations peu complexes [Mosteller et al., 2014]. Les effets négatifs d'une complexité trop importante peut également influencer sur les préférences esthétiques [Wang and Lin, 2019], mais aussi sur les intentions comportementales [Wu et al., 2016], ce qui est un résultat exploitable dans le cadre de nos travaux. Une faible complexité visuelle engendre des niveaux de confiance plus élevés et des jugements plus positifs sur l'esthétique d'un site web [Wu et al., 2016]. On peut également le constater sur les intentions d'achat [Chen, 2018].

À l'inverse, d'autres chercheurs ont mis en avance une préférence des individus pour une complexité élevée. Certaines études ont montré que la satisfaction des individus est plus élevée lorsque la complexité d'une page web est élevée [Palmer, 2002]. Dans un contexte de consommation, Pieters et al. ont constaté que les publicités à forte complexité visuelle augmentent l'attention des consommateurs et génèrent une attitude positive [Pie-

ters et al., 2010]. Une étude plus ancienne, explore les attitudes des consommateurs envers les produits visuellement complexes (par opposition aux produits simples) et montre que les consommateurs préfèrent les produits complexes en cas d'exposition répétée [Cox and Cox, 2002]. Il est à noter qu'un ensemble important de travaux, basé sur la théorie de Berlyne, conclut également qu'une complexité modérée a un effet plus significatif sur les attitudes et les intentions [Wang and Lin, 2019, Deng and Poole, 2012].

Dans un premier temps, ces travaux fondateurs et les études qui en découlent nous permettent de définir une « échelle de valeurs » afin de mesurer l'impact de la complexité de l'information sur la perception humaine. Dans un second temps, il est utile d'investiguer plus largement sur la construction du choix d'un stimulus aux dépens d'un autre, et ainsi mieux comprendre comment la complexité influence le processus de décision.



**À retenir :** La théorie de Berlyne relie la courbe de Wundt à la motivation, à la recherche de nouveauté, mais surtout fait le lien entre effet de plaisir et perception de la complexité. Cette théorie met en lumière les effets négatifs voir aversifs d'un stimulus trop complexe, contrairement à un stimulus de complexité modérée qui provoque un effet positif (agréable) sur les individus. Cela nous permet donc de définir une échelle de mesure de la complexité avec trois stimulus représentant une complexité simple, modérée et élevée.

### 3.2.4 Processus de décision : Le choix

L'attention est généralement définie comme l'étape de sélection dans le mécanisme de perception. Elle possède un rôle direct dans le processus de décision en déterminant ce qui est perçu ou non. De plus, si l'information (c'est-à-dire le stimulus) ne reçoit aucune fixation de la part de l'individu et est en dehors de la portée perceptuelle de la fixation la plus proche, elle risque de ne pas être identifiée. Cette non-identification a pour résultat l'absence de l'information dans le processus de décision [Orquin and Loose, 2013].

À partir de ce constat, il est primordial d'analyser le niveau de complexité de l'information disponible pour qu'il puisse être intégré au processus décisionnel et de voir si ces niveaux de complexité doivent être adaptés dans le temps. Dans ce contexte, nous nous appuyons sur les différentes stratégies de prise de décision, en analysant les étapes qui les composent. Ces étapes font référence à des processus ou à des activités qualitativement différents les uns des autres [Russo and Leclerc, 1994]. La majorité des modèles dans la littérature possèdent deux étapes. La première, appelée filtrage (screening en anglais), permet d'encoder rapidement les éléments du stimulus visuel pour obtenir une « vue globale » des alternatives proposées. La deuxième appelée évaluation a pour objectif d'obtenir des informations supplémentaires concernant des alternatives restantes pour aboutir à une décision. En sciences de gestion, le découpage du processus de décision est différent, l'attention visuelle représente la première étape du processus conduisant à une étape de traitement de l'information. À la suite de cette étape de traitement, l'individu développe une attitude qui provoque le choix (ou la décision).

L'analyse des processus de décision ayant évolué avec le temps, le nombre d'étapes dans ce même processus a également évolué. Dans leurs travaux sur l'analyse des pro-

cessus de choix par fixation du regard, Russo et Leclerc [Russo and Leclerc, 1994] ont simulé en laboratoire un environnement de supermarché composé de différents produits de consommation et ont enregistré les fixations oculaires des participants. Ils sont arrivés à une conclusion contraire au processus généralement admis de deux étapes et ont donc défini un nouveau modèle composé de trois étapes, appelées orientation, évaluation et vérification. La première étape consiste en une vue d'ensemble de chaque produit accompagné d'un premier tri des alternatives proposées. La seconde étape, la plus longue du processus, est une succession de comparaisons directes entre deux ou trois alternatives issues du premier tri. La dernière étape, celle de vérification, consiste quant à elle à analyser de nouveau d'autres alternatives ayant reçu peu ou pas de fixations lors des deux premières étapes. Cette étape finale permet au participant de conforter son choix.

Le cadre fixé par Russo et Leclerc a été repris par plusieurs autres études, démontrant l'existence de phases d'attention. En 1997, Wedell et Senter [Wedell and Senter, 1997] dans leur étude sur l'importance du regard dans le jugement et le choix ont évalué la pertinence d'un processus en trois étapes pour ce type de tâche, basé sur celui de Russo et Leclerc. Selon eux, le processus de décision ne repose que sur deux étapes qui correspondent aux étapes d'orientation et d'évaluation du cadre de Russo et Leclerc. Cependant, pour les tâches de choix, les résultats ne montrent pas si le comportement observé correspond à une étape d'évaluation ou de vérification. Cette non-détermination de la seconde étape est due aux expériences menées qui ne comportaient que deux alternatives ne provoquant pas le même comportement que celle de Russo et Leclerc qui en possédait seize [Wedell and Senter, 1997]. En effet, le nombre réduit d'alternatives de cette étude ne peut pas mettre en lumière la phase de vérification qui se traduit par la tendance à rechercher rapidement les autres alternatives de l'ensemble, afin de confirmer son choix.

Dans un contexte similaire d'action d'achat en magasin, Jesper Clement a cherché à déterminer l'influence visuelle du design de l'emballage d'un produit sur l'action d'achat en magasin. Il a demandé aux participants de son étude d'acheter ce qu'ils voulaient sans limite de temps. L'étude s'est déroulée dans des conditions réelles de consommation (supermarché). Grâce aux données d'oculométrie (*eye-tracking* en anglais) recueillies, il conclut à l'existence de cinq phases dans le processus de décision [Clement, 2007]. Son modèle se compose d'une phase de préattention, dans laquelle les produits attirent l'attention du client. La phase suivante appelée « d'attention réussie » est le moment où le visuel de l'emballage se développe dans l'esprit du consommateur. La phase suivante, celle où le client prend le produit en main est un moment de basculement au cours duquel il entre dans une phase d'action physique. Cette phase particulière entraîne l'action d'achat, quatrième phase de ce modèle. La cinquième et dernière phase dans laquelle l'emballage fait office de distracteur visuel, comme lors des premières phases est appelée phase de post-achat (post-sélection).

Au vu de la littérature sur le sujet, nous pouvons légitimement dire que le nombre d'étapes dépend en premier lieu de la nature de la tâche, du nombre d'alternatives proposées et du temps disponible pour effectuer la tâche.

Il est à préciser que dans notre étude, en raison d'un biais de position de l'image de

complexité modérée, nous ne pouvons pas valider d'hypothèses concernant la construction du choix. Nous reviendrons sur ce point dans la Section 3.4.4.



**À retenir :** En s'intéressant au processus de décision, Russo et Leclerc ont mis en avant l'existence de différentes étapes, appelées phases d'attention. Le modèle en trois étapes (orientation, évaluation et vérification) avancé ne fait cependant pas consensus. À la suite de différentes études, le nombre d'étapes au cœur de ce même processus a évolué selon la nature de la tâche. La notion de phases d'attention qui nous intéresse est cependant communément admise, ce qui nous permettra dans le futur d'évaluer l'évolution de la perception de la complexité chez les individus.

### 3.2.5 Position de l'information et comparaison par paire

L'attention visuelle étant guidée en partie par la localisation du stimulus (voir Section 3.2.1 et Section 3.2.2), il est pertinent de se pencher sur la question de la position physique de l'information. Pour tenter de comprendre l'influence de la position, la notion de position optimale de visualisation (Optimal Viewing Position ou OVP) est intéressante. Elle montre que l'identification d'un mot est meilleure lorsque les yeux se fixent d'abord près du centre du mot. Cette position centrale minimise le temps de reconnaissance de ce dernier [Rayner, 2009]. On peut donc se demander si cette position optimale a une quelconque influence sur l'évaluation des choix. Il a été prouvé, dans un contexte de consommation et donc de choix, que les marques de produits en position centrale sur un présentoir reçoivent plus d'attention et sont plus susceptibles d'être choisies que les marques situées au bas du présentoir [Chandon et al., 2009]. L'effet de la position centrale a également été constatée dans les décisions fondées sur la valeur à l'aide d'un modèle de dérive-diffusion (drift-diffusion en anglais) pour observer le choix et le temps de réponse associé à la prise de décision [Krajbich and Rangel, 2011]. De plus, l'effet de la position centrale est robuste à la pression temporelle dans un contexte de choix. Les premières fixations du regard dans des ensembles de différentes tailles se portent sur l'objet se trouvant au centre de ce dernier [Reutskaja et al., 2011]. Cette position optimale peut également provoquer certaines réponses en fonction de la proximité spatiale de certains éléments.

Lors de l'étape d'évaluation dans le processus de décision (voir Section 3.2.4), la comparaison directe se porte sur deux ou trois alternatives, justifiant notre choix d'investiguer la notion de comparaison par paire. Le lien entre la proximité spatiale des alternatives, autrement dit de la position et l'utilisation de la comparaison par paire, a été prouvé. Ce mécanisme sert à minimiser les coûts d'attention en utilisant des saccades plus courtes [Russo and Rosen, 1975]. Certaines études définissent les comparaisons de séquences comme X-Y-X ou X-Y-X-Y, où X et Y sont des alternatives différentes [Russo and Rosen, 1975, Russo, 1978, Russo and Leclerc, 1994]. La stratégie de comparaison par paire minimise également la mémoire de travail à l'aide de refixations (c'est-à-dire de revisites d'une même alternative) [Glaholt and Reingold, 2011].

Avec l'ensemble de la littérature présentée dans cette section concernant les compa-

raisons par paire ou la position optimale, nous supposons que l'information modérément complexe servira de référence dans une comparaison par paires. C'est-à-dire que la comparaison se fera entre l'information modérée et l'information complexe ou entre l'information modérée et l'information simple. Nous avons également émis une hypothèse sur l'attention générée par une information placée au centre et son impact sur le choix des participants. Nous développons ces hypothèses dans la Section 3.3.



**À retenir :** La position centrale d'une information est perçue d'une meilleure manière par les individus. Plusieurs études ont montré une influence sur le choix. Dans notre cas de recherche, la notion de comparaison par paires est utile, car nous présentons une multitude d'informations, au sein d'un même visuel. Cela nous permet de mettre en avant certaines informations spécifiques afin de valoriser l'adoption d'un comportement.

### 3.3 Étude d'oculométrie - protocole et hypothèses

Dans le cadre de nos recherches nous essayons d'inciter les individus à avoir un comportement de mobilité plus vertueux sur le long terme à l'aide d'informations pertinentes et factuelles. L'objectif principal de cette étude d'oculométrie est d'analyser les mécanismes de choix dans le processus de décision en lien avec l'attention visuelle. Le but est d'étudier si l'attention visuelle et la complexité de l'information présentée exercent ou non une influence sur le choix de l'individu. Au cours de cette étude, trois niveaux de complexité seront présentés aux participants dans le but de mieux cerner cette influence potentielle. Les résultats de cette étude nous permettront également d'alimenter notre expérimentation de transmission de l'information via une application mobile (voir Section 4.3).

#### 3.3.1 Protocole expérimental

Au total, 24 sujets ont été recrutés pour cette expérience, âgés de 19 à 45 ans (moyenne = 25, écart-type = 6,413), dont 20 hommes et 4 femmes. La population est composée de chercheurs, d'étudiants et de personnes ayant un rôle de soutien administratif dans le domaine de l'informatique.

Avant de présenter les différents stimuli aux participants, nous réalisons une mise en situation en lien avec la mobilité urbaine. Chaque participant est un utilisateur quotidien d'une application qui calcule sa consommation de CO<sub>2</sub> en fonction de ses déplacements. Nous affinons cette situation, en lui précisant qu'il vient d'effectuer un trajet en bus de 10 kilomètres. Les instructions de posture et d'utilisation de la souris pour répondre à chaque question sont affichées en même temps (Figure 3.6). Nous lui présentons alors trois stimuli (Figures 3.7, 3.8 et 3.9) représentant différents écrans d'accueil (récapitulatif de consommation de CO<sub>2</sub> quotidien) susceptibles d'être intégrés à une application de ce type.

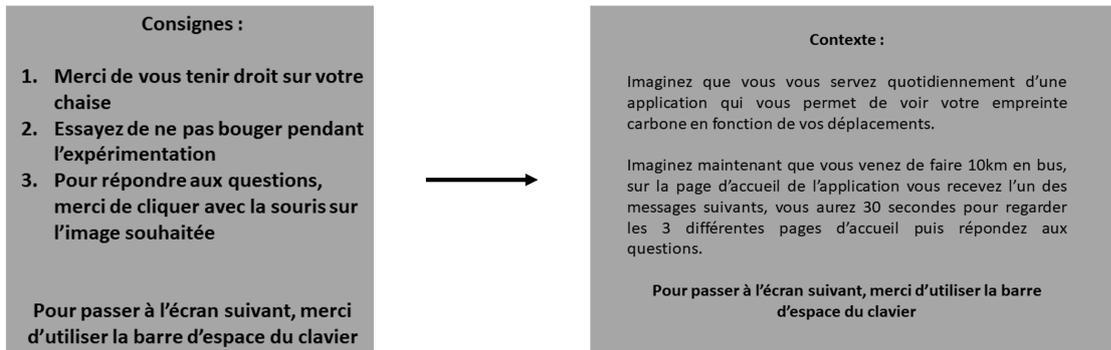


FIGURE 3.6 – Instructions et contexte de l'étude



FIGURE 3.7 – Stimulus avec un niveau de complexité faible



FIGURE 3.8 – Stimulus avec un niveau de complexité modéré



FIGURE 3.9 – Stimulus avec un niveau de complexité élevé

Pour mesurer et isoler l'impact de la complexité de l'information transmise, il est important de construire nos différents stimuli en partant d'éléments communs, puis d'ajouter ou de supprimer au fur et à mesure des informations pour faire augmenter ou diminuer la complexité. Cette variation de la complexité nous permet d'analyser de potentiels changements dans le choix des individus et, sur le plus long terme, dans leurs comportements. Pour faire varier la complexité nous utilisons deux propriétés : le nombre d'éléments et leur position. L'information ainsi ajoutée pour augmenter la complexité nous permet également de faire un retour informationnel plus précis auprès des individus, ce qui peut potentiellement influencer sur un comportement de mobilité plus vertueux sur le long terme.

Pour cette étude, nous avons conçu trois prototypes différents de page d'accueil d'application mobile. Ces trois images représentent trois niveaux de complexité différents appelés *faible*, *modéré* et *élevé* et présentés dans les figures 3.7, 3.8 et 3.9.

L'architecture de chaque stimulus peut être divisée en deux parties distinctes : la partie supérieure contient un ou plusieurs indicateurs de la consommation journalière de CO<sub>2</sub> en fonction du niveau de complexité, tandis que la partie inférieure propose un récapitulatif plus ou moins détaillé, toujours en fonction du niveau de complexité, de la consommation hebdomadaire.

En se basant sur la littérature présentée précédemment, il serait judicieux de pouvoir faire varier un ensemble de caractéristiques pour chaque élément présenté sur les différents écrans. Cependant, le choix du smartphone avec une application mobile comme vecteur de changement de comportement nous impose certaines contraintes : la taille, parfois très réduite de l'écran du téléphone et son hétérogénéité entre tous les modèles nous ont obligés à condenser les informations présentées pour plus d'attractivité et de compréhension (tous les éléments sont sur l'écran sans besoin de les faire défiler). Nous avons donc fait varier la présentation des informations et non leur nature, en affichant des informations identiques sous différentes formes.

Tous les stimuli possèdent des éléments communs tels que les titres, les bannières de présentation (respectivement en haut et en bas de l'image) et l'image suggérant le transport et l'écomobilité. Le code couleur est également identique pour les trois niveaux. Le niveau de complexité *faible* propose un compteur quotidien de CO<sub>2</sub> consommé sous deux formes différentes couramment affichées dans les applications du même type. La première forme est la simple valeur numérique affichée et la deuxième forme est une échelle avec différents niveaux (de A à E) correspondant à la consommation de CO<sub>2</sub>. Le niveau de complexité *modérée* comprend tous les éléments du niveau simple avec en plus, dans la partie inférieure, la valeur moyenne de la semaine et le titre associé « données hebdomadaires » de la partie. Enfin, le niveau de complexité *élevée* comprend tous les éléments des niveaux précédents ainsi que, dans la partie supérieure, une troisième forme d'affichage de la valeur quotidienne (la lettre unique à droite, C sur la figure 3.9), les valeurs minimale et maximale dans la partie inférieure et les icônes de nuage. En outre, la position des éléments est modifiée pour provoquer une perturbation de l'attention visuelle du participant et ainsi délimiter ostensiblement le plus haut niveau de complexité.

Pour l'affichage de nos stimuli, nous nous sommes basés sur les notions du Paradigme de Préférence Visuelle (PPV), appliqué par Robert Fantz en 1958 pour identifier les

éléments qui attirent l'attention du nourrisson [Fantz, 1958]. Dans notre contexte d'application mobile, afin de correspondre aux conditions réelles, le PPV est utile car il n'est pas nécessaire d'avoir un retour verbal des participants. Plus tard, ce paradigme a été mis en œuvre avec des adultes [Shimojo et al., 2003], où l'on a découvert que la préférence peut être confondue avec un biais d'habituation. Pour remédier à ce problème, il suffit d'afficher les stimuli de manière aléatoire. Chaque image est donc présentée dans toutes les combinaisons possibles de positions (c'est-à-dire six fois) par ensemble de trois images représentant les trois niveaux de complexité, que l'on nommera série (voir figure 3.10) pour limiter le biais d'accoutumance. Toujours pour limiter ce biais, il est important de « ré-initialiser le regard » du participant. Pour cela, un écran comportant une croix au centre est inséré entre chaque série. Les sujets doivent fixer cette croix pendant 1000 ms.

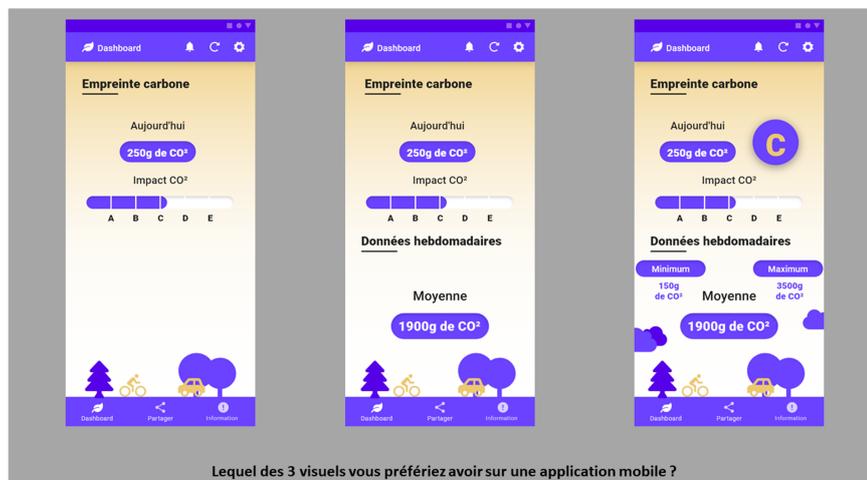


FIGURE 3.10 – Exemple d'une série d'écrans d'accueil

Une fois les stimuli présentés, le choix des participants en analysant leur regard est parfois difficile voir impossible à discriminer, cette situation est appelée l'effet Midas [Jacob, 1991]. Pour contrer cet effet, nous avons opté pour un contrôle manuel, avec un périphérique de sélection externe, à savoir une souris. La mise en pratique de ce contrôle est la suivante : en dessous de chaque écran présenté, la question « *Lequel de ces trois visuels préférez-vous avoir dans une application mobile ?* » est posée. Pour y répondre, les sujets de l'étude doivent cliquer sur l'image de leur choix, tout en étant limité à une image par série. Le sujet est bien entendu libre de changer d'image entre chaque série. Chaque série d'images est présentée à l'écran pendant 30 secondes. Le temps restant pour répondre à la question n'est pas indiqué au participant afin de ne pas engendrer une pollution visuelle ainsi qu'une pression temporelle trop importante qui pourrait engendrer un biais.

Afin d'analyser précisément et de détecter à chaque instant où le regard du participant se concentre, des zones d'intérêts (AOI) ont été prédéfinies. Ces zones sont positionnées aux coordonnées exactes des trois stimuli présentés sur l'écran. Ces zones seront par la suite activées, reconnues et analysées par le logiciel de suivi de regard que nous utilisons

dans le cadre de cette expérience. Pour que ces mesures soient correctes, il est demandé aux sujets d'ajuster leur position par rapport à l'écran et d'adopter une position assise confortable avant la tenue de l'expérience. La distance entre l'écran et la chaise est d'environ 70 cm. La distance entre l'oculomètre et la chaise quant à elle est d'environ 50 cm. Ces distances peuvent légèrement varier en fonction des sujets. Une procédure de calibration de l'oculomètre est dans tous les cas réalisée pour chaque sujet avant le début de l'étude, dans le but de contrôler la position du participant et la qualité de l'enregistrement de son regard.

Cette étude a été conduite en laboratoire, dans une salle isolée de tous distracteurs externes, dans des conditions propices à la concentration et avec un éclairage suffisant. Chaque sujet a été informé des objectifs et des procédures de l'expérience et a signé un formulaire de consentement éclairé avant l'expérience formelle. Les sujets ont réalisé l'expérience indépendamment les uns des autres. Après l'expérience, les données brutes ont été exportées et un retour d'informations est donné aux sujets sur leur participation.

### 3.3.2 Hypothèses

Dans le cadre de cette étude, nous émettons plusieurs séries d'hypothèses (voir Figure 3.11) qui correspondent aux questions suivantes :

- L'attention accordée à un stimulus dépend-elle de sa complexité ?
- La complexité d'un stimulus influence-t-elle le choix des individus ?
- La position d'un stimulus influence-t-elle le choix des individus indépendamment de sa complexité ?
- Un stimulus de complexité modérée est-il une référence servant à la comparaison dans le processus de sélection d'un individu ?

Selon la littérature [Lohse and Johnson, 1996, Reutskaja et al., 2011, Russo and Doshier, 1983], le nombre de fixations devrait augmenter avec la complexité de l'information. Ce postulat constitue notre première hypothèse : **H1 : L'écran d'accueil de complexité élevée aura un nombre de fixations plus important que les écrans d'accueil de complexité faible ou modérée.**

Le travail cognitif supplémentaire induit par une complexité plus élevée, notamment lié au plus grand nombre d'éléments composant le stimulus, peut se traduire au niveau visuel par un temps de visionnage et un nombre de revisites plus importants. Deux hypothèses découlent de ce postulat, **H2 : L'écran d'accueil avec une complexité élevée reçoit un temps de visionnage plus important que les écrans d'accueil de complexité faible ou modérée** et **H3 : L'écran d'accueil de complexité élevée reçoit un plus grand nombre de revisites que les écrans d'accueil de complexité faible ou modérée.**

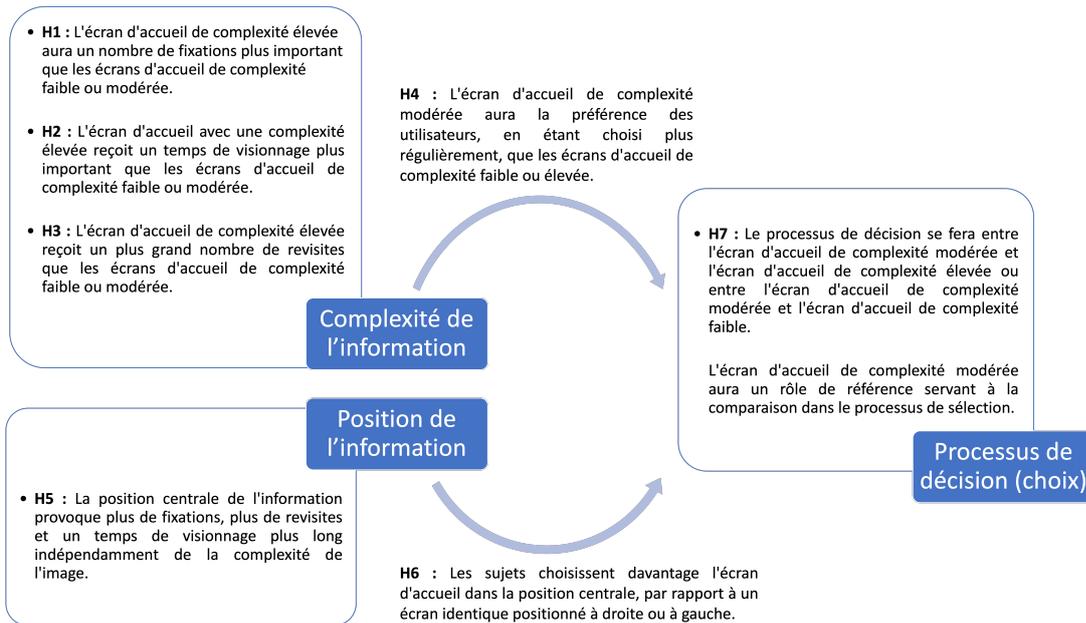


FIGURE 3.11 – Modèle d'hypothèses de l'étude d'oculométrie.

Pour répondre à la question du choix, nous émettons l'hypothèse suivante :

**H4 : L'écran d'accueil de complexité modérée aura la préférence des utilisateurs, en étant choisi plus régulièrement, que les écrans d'accueil de complexité faible ou élevée.** Cette hypothèse est fondée sur les travaux issus des articles [Geissler et al., 2006, Bruner and Kumar, 2000, Stevenson et al., 2000].

En se basant sur la littérature concernant l'effet de la position de l'information sur l'attention et le choix [Chandon et al., 2009, Krajbich and Rangel, 2011, Reutskaja et al., 2011], nous émettons l'hypothèse suivante :

**H5 : La position centrale de l'information provoque plus de fixations, plus de revisites et un temps de visionnage plus long indépendamment de la complexité de l'image.** Nous émettons également l'hypothèse que la position centrale de l'information provoque une adoption (choix) plus importante indépendamment de sa complexité. **H6 : Les sujets choisissent davantage l'écran d'accueil dans la position centrale, par rapport à un écran identique positionné à droite ou à gauche.**

Enfin, dans la lignée de nos premières hypothèses et des travaux sur la comparaison par paires [Russo and Rosen, 1975, Russo, 1978, Russo and Leclerc, 1994], nous proposons l'hypothèse suivante qui, si elle est vérifiée, nous permettra de faire une analyse plus approfondie de l'impact de la complexité. **H7 : Le processus de décision se fera entre l'écran d'accueil de complexité modérée et l'écran d'accueil de complexité élevée ou entre l'écran d'accueil de complexité modérée et l'écran d'accueil de complexité faible. L'écran d'accueil de complexité modérée aura un rôle**

de référence servant à la comparaison dans le processus de sélection.

## 3.4 Résultats

### 3.4.1 Effets de la complexité de l'information

Nous présentons ici les résultats des mesures de suivi oculaire (fixations, revisites et temps de visionnage) en fonction du niveau de complexité des écrans d'accueil. Comme le montre la Table 3.1, l'écran d'accueil de complexité élevée est celui qui a reçu le plus grand nombre de fixations, avec une moyenne supérieure à 30. L'écran de complexité modérée a reçu en moyenne 25 fixations par participant, et l'écran le moins complexe a reçu en moyenne 14 fixations. Ces valeurs nous permettent d'affirmer que le nombre absolu de fixations augmente avec la complexité. De même, le temps de visionnage est plus important en moyenne pour les écrans de complexité élevée (9,32s), par rapport au temps de visionnage pour les écrans de complexité modéré (7,67s) et faible (4,82s). Ce résultat peut s'expliquer par le nombre d'attributs entre les trois écrans d'accueil. Par conséquent, nos hypothèses **H1** (L'écran d'accueil de complexité élevée aura un plus grand nombre de fixations que les écrans d'accueil de complexité faible ou modérée) et **H2** (L'écran d'accueil avec une complexité élevée reçoit un temps de visionnage plus important que les écrans d'accueil de complexité faible ou modérée) sont confirmées.

	Complexité faible	Complexité modérée	Complexité élevée
Nombre moyen de fixations	14.71	25.34	30.51
Nombre moyen de revisites	4.98	7.43	7.41
Temps moyen de visionnage (en sec.)	4.82 s	7.67 s	9.32 s
Choix des sujets	3	13	7

TABLE 3.1 – Valeurs des mesures d'oculométrie dans l'étude. Les lignes correspondent aux valeurs moyennes pour chaque série et chaque participant.

D'autre part, le nombre moyen de revisites est similaire pour l'écran d'accueil modéré (7,43) et l'écran complexe (7,41) (voir Table 3.1). L'écran peu complexe a un nombre plus faible de revisites (4,98). Notre hypothèse **H3** (L'écran d'accueil de complexité élevée reçoit un plus grand nombre de revisites que les écrans d'accueil de complexité faible ou modérée.) est donc rejetée. Une explication possible est qu'une complexité accrue peut modifier le modèle de recherche, ce qui peut conduire les individus à ignorer certaines informations dans le processus de décision ou à traiter certaines alternatives de manière plus approfondie [Glaholt and Reingold, 2011]. Un phénomène équivalent a été démontré en sciences de gestion avec l'influence de la réputation sur le biais de confirmation menant à l'évitement de certaines informations [Lallement et al., 2020]. Avec le nombre de revisites, on constate que l'écran le moins complexe est traité de manière superficielle, contrairement aux deux autres écrans qui sont traités avec plus d'attention. Ce traitement prolongé des alternatives modérée et élevée, peut être corrélé

avec le processus de décision des participants : si leur choix préliminaire est fait tôt dans le processus, les revisites serviraient comme un « outil de contrôle » pour assurer leur décision. Ce point sera développé dans la section 3.4.4.

Ces résultats indiquent également une attention visuelle accrue sur deux des trois alternatives présentées. D'une part avec un plus grand nombre de fixations et leur durée sur l'écran de haute complexité et d'autre part avec un nombre similaire de revisites sur les écrans de complexité modérée et élevée. Le choix se fera donc entre ces alternatives. Ce point sera développé dans la section suivante.



**À retenir :** L'écran d'accueil complexe attire plus de fixations et un temps de visionnage plus élevé de la part des individus que les écrans simple et modéré. Toutefois l'écran d'accueil complexe a un nombre de revisites égal à celui de l'écran de complexité moyenne.

### 3.4.2 Effets sur le choix des individus

Dans cette section, nous étudions les résultats des mesures oculométriques (fixations, revisites et temps de visionnage) en fonction de l'image choisie par les sujets. Nous considérons qu'un choix est effectué si un participant sélectionne la même image au minimum dans les trois dernières séries. Sur les 24 sujets, 20 ont fait le même choix sur les six séries, 3 ont fait le même choix sur les trois dernières séries et 1 a fait plusieurs changements (modéré – élevé – élevé – modéré – modéré – élevé) et a donc été considéré comme indécis. L'image de niveau de complexité modérée est choisie par 13 sujets tandis que les images de niveau de complexité élevée et faible sont respectivement choisies par sept et trois sujets.

Comme le montre la Table 3.1, l'image de haute complexité a le plus grand nombre de fixations, cependant elle est la deuxième image la plus choisie. Le nombre moyen de revisites est similaire pour les écrans d'accueil de complexité modérée et élevée alors qu'il y a une différence significative dans le choix. Enfin, l'écran d'accueil de complexité élevée est le plus regardé, suivi de l'écran d'accueil de complexité modérée. Avec ces résultats, la métrique de fixation montre qu'il n'y a pas de corrélation avec le choix des sujets. Les revisites montrent quant à elles une évaluation avant que le choix final ne soit fait entre une complexité modérée et une complexité élevée. De même, la durée de visualisation d'un écran d'accueil n'influence pas le choix final des sujets.

Sur la base de tous les résultats présentés, nous pouvons affirmer que la majorité des participants (13 sur 23) choisissent l'écran d'accueil modérément complexe pour être informés de leur comportement en matière d'émissions de CO<sub>2</sub>. Nous pouvons donc conclure que le nombre d'attributs a un impact sur le choix, un nombre trop faible ou trop élevé d'attributs n'encourageant pas le choix. Notre hypothèse **H4** est donc confirmée.



**À retenir :** L'écran d'accueil de complexité modérée a la préférence des participants. Le nombre d'attributs a un impact sur le choix : un nombre trop faible d'attributs n'encourage pas le choix, un nombre trop élevé d'attributs poussent les individus à ignorer certaines informations.

### 3.4.3 Effets de la position de l'information

Dans cette section, nous présentons les résultats des métriques du nombre de fixations, du nombre de revisites et de temps de visionnage en rapport avec la position de l'image. Comme le montre la Table 3.2 (en haut à gauche), nous observons au cours des six séries d'images que la position centrale entraîne un plus grand nombre de sélections de l'image dans cette position, quel que soit son niveau de complexité. Concernant les images de complexité faible et élevée, le nombre de sélections reste stable, quelle que soit la position. Le classement intracatégorie reste identique au choix global des participants (1. modéré, 2. élevé, 3. faible).

Comp.	Gauche	Centrale	Droite
Faible	5	<b>6</b>	5
Modérée	26	<b>32</b>	28
Élevée	14	<b>15</b>	13

**Nombre de sélections**

Comp.	Gauche	Centrale	Droite
Faible	25	<b>49</b>	24
Modérée	41	<b>69</b>	42
Élevée	50	<b>78</b>	46

**Nombre moyen de fixations**

Comp.	Gauche	Centrale	Droite
Faible	7	<b>18</b>	8
Modérée	10	<b>24</b>	10
Elevée	14	<b>21</b>	11

**Nombre moyen de revisites**

Comp.	Gauche	Modérée	Droite
Faible	7.79	<b>14.6</b>	6.54
Modérée	13.02	<b>19.78</b>	13.22
Élevée	15.88	<b>25.35</b>	14.7

**Temps moyen de visionnage**

TABLE 3.2 – Valeurs des métriques d'oculométrie des participants en fonction de la position de l'image.

Le nombre moyen de fixations dans la Table 3.2 (en haut à droite) indique que la position centrale sur l'écran reçoit plus de fixations que toute autre position, indépendamment de la complexité des trois écrans d'accueil. Ce résultat peut s'expliquer en partie par « l'écran de réinitialisation » inséré entre chaque série d'images. La croix se trouvant au centre de l'écran, cela assure une fixation initiale à l'image positionnée au centre dans la série suivante. Comme pour les résultats concernant le choix des sujets, le nombre de fixations est similaire pour les images en position gauche et droite sur l'écran.

Comme présenté dans la Table 3.2 (en bas à gauche), le nombre moyen de revisites est également plus important en position centrale. Dans cette même position, c'est l'image de complexité modérée qui obtient le plus de revisites. Il est à noter que l'image en position droite sur l'écran, reçoit un nombre de revisites stable quel que soit le niveau de complexité des images. En position gauche, c'est l'image la plus complexe qui en reçoit le plus.

Concernant le temps de visualisation dans la Table 3.2 (en bas à droite), la position centrale bénéficie d'un temps de visionnage supérieur pour chaque image au contraire des deux autres positions. Les positions gauche et droite ont des valeurs similaires pour les trois images. L'image la plus regardée dans les trois positions est l'image de complexité élevée. Comme pour la métrique des fixations, la croix de « réinitialisation de l'écran »

influence légèrement la position centrale.

Les images de complexité modérée et élevée, lorsqu'elles sont placées au centre, attirent plus de fixations, de revisites et de temps de visionnage que n'importe quelle image placée à gauche ou à droite. Même l'image de faible complexité attire plus de revisites que n'importe quelle autre image lorsqu'elle est placée au centre. Cependant, cette même image placée au centre obtient moins de fixations et de temps de visionnage (même si la différence est faible) que l'image complexe placée sur le côté gauche. Compte tenu de ces résultats, notre hypothèse **H5** peut être partiellement confirmée : la position centrale de l'image induit plus de fixations, plus de revisites et un temps de visionnage plus long, mais cela n'est pas totalement indépendant du niveau de complexité de chaque image. Enfin, notre hypothèse **H6** peut être validée : un écran d'accueil donné sera plus sélectionné s'il est au centre que s'il est sur un côté.



**À retenir :** La position de l'image a une influence sur le choix, mais le niveau de complexité prédomine : une image peu ou trop complexe au centre de l'écran ne sera pas sélectionnée, même si elle bénéficie d'un temps de visionnage plus long, d'un plus grand nombre de revisites et de fixations.

#### 3.4.4 Effets sur le processus de décision

Pour des raisons de facilité de mise en oeuvre de l'étude, l'ordre des stimuli présentés aux participants n'est pas aléatoire. Concrètement, l'écran d'accueil de complexité modérée a été présenté en position centrale lors des deux premières séries (S1 et S2). Cet ordre prédéfini constitue un biais dans notre étude. Il conduit par conséquent à l'impossibilité d'établir des conclusions fortes par rapport à la construction du choix. Il convient dès lors d'examiner avec une certaine mesure les résultats de la présente section.

Selon la Figure 3.12, le nombre moyen de fixations pour l'écran d'accueil de complexité modérée est stable pour les 2 premières séries (S1 = 35 et S2 = 34), puis diminue pour se situer autour de 20 fixations (S3 = 21, S4 = 22, S5 = 20 et S6 = 20). Cette baisse s'explique par la position identique du stimulus lors des deux premières séries qui entraîne un nombre de fixations plus élevées (voir Section 3.4.3). L'écran d'accueil de complexité faible obtient un nombre moyen de fixations stable pour les 3 premières séries (S1 = 15, S2 = 12 et S3 = 12). On observe ensuite une fluctuation pour les 3 dernières séries, avec une valeur de 25 pour la quatrième série, une baisse (S5 = 10) et enfin une nouvelle augmentation à 24 fixations en moyenne pour la dernière série. En ce qui concerne la dernière image, la plus complexe, on observe davantage de variations dans le nombre moyen de fixations. L'écran d'accueil commence avec une valeur de 21 fixations, puis augmente légèrement (S2 = 28). Pour la troisième série, la valeur moyenne augmente fortement pour atteindre 39 fixations. Elle diminue à nouveau pour la quatrième série, à 22 fixations. Le nombre augmente à nouveau pour atteindre 39 fixations en moyenne. Pour la dernière série, le nombre diminue (S6 = 25). Il est à noter qu'à position identique, on ne constate pas de variations, ce qui traduit une prédominance de la position de l'image.

L'écran d'accueil de complexité faible obtient le même nombre moyen de revisites (Figure 3.13) pour les trois premières séries (c'est-à-dire 4 revisites). Le nombre de revisites

augmente pour la quatrième série ( $S_4 = 10$ ). C'est au cours de la cinquième série que le nombre de revisites atteint son minimum ( $S_5 = 3$ ). Pour la dernière série, l'écran compte huit revisites. L'écran d'accueil modéré a la même valeur moyenne de douze revisites pour les deux premières séries. Ensuite, il perd la moitié de ses revisites ( $S_3 = 6$ ). Pour les quatrième et cinquième séries, la valeur moyenne est de cinq revisites. Elle reste stable pour la dernière série avec quatre revisites. L'écran d'accueil de complexité élevée a une valeur moyenne de six revisites pour la première série. Elle augmente ensuite légèrement pour atteindre huit revisites. L'écran atteint onze revisites lors de la troisième série affichée. Le nombre de revisites diminue jusqu'à six revisites, puis augmente ( $S_5 = 10$ ) et se termine à la valeur moyenne de cinq. On constate également que la valeur moyenne des revisites diminue au cours des séries.

En se basant sur ces résultats, nous pourrions émettre l'éventualité de deux phases concernant les revisites, avec une première phase, qu'on pourrait qualifier de « balayage » dans laquelle les participants regardent l'ensemble des alternatives, ce qui provoque plus de revisites. Puis une seconde phase apparaît, dans laquelle le nombre de revisites diminue, ce qui pourrait traduire une attention accrue portée à des alternatives spécifiques afin de se préparer à choisir. Cette possibilité devrait être testée avec une nouvelle méthodologie excluant le biais de position évoqué précédemment.

Comme le montre la Figure 3.14, nous pouvons observer que l'écran d'accueil le plus regardé est l'écran modéré pendant les deux premières séries avec une valeur de temps de visionnage d'environ 10 secondes ( $S_1 = 10,0$  et  $S_2 = 9,79$ ). Au cours de la troisième série, la durée moyenne de visualisation chute à 6,58 secondes, puis remonte à 7,11 secondes pour la série suivante. Enfin, la valeur du temps de visionnage se stabilise autour de 6 secondes ( $S_5 = 6,11$  et  $S_6 = 6,43$ ). Au cours des deux premières séries, l'écran à haute complexité est visualisé en moyenne en 6,58 secondes et 8,50 secondes. La valeur augmente de manière significative dans la troisième série ( $S_3 = 12,28$  secondes). Elle diminue ensuite au cours de la quatrième série ( $S_4 = 7,38$  secondes) pour augmenter à nouveau au cours de la cinquième série ( $S_5 = 13,06$ ) et atteindre la durée moyenne la plus élevée des trois écrans d'accueil. Lors de la dernière série, il obtient un temps de visionnage moyen de 8,12 secondes. Quant à l'écran à complexité simple, il a la durée moyenne la plus faible de tous les écrans pour les cinq premières séries ( $S_1 = 4,58$ ,  $S_2 = 3,26$ ,  $S_3 = 3,28$ ,  $S_4 = 6,93$  et  $S_5 = 3,21$ ). Pour la dernière série, la valeur augmente et dépasse celle de l'écran modéré, atteignant 7,67 secondes.

Avec le nombre de choix par série dans la Table 3.3, on peut observer que l'image de complexité modérée est plus sélectionnée dans les séries  $S_1$  et  $S_2$  et se stabilise ensuite entre 13 et 14 sélections. Cette baisse correspond à un transfert du choix vers l'image de complexité élevée qui obtient entre deux et trois sélections supplémentaires. Compte tenu du biais détaillé au début de cette section, nous pensons que ce transfert est potentiellement lié à la position centrale du stimulus, cependant un processus de choix n'est pas exclu et devrait également être testé.

L'écran de complexité faible commence avec le temps de visualisation moyen le plus bas. Il reçoit ensuite un temps de visionnage similaire pour les séries 2, 3 et 5. Ces valeurs, qui sont les plus basses de l'ensemble, reflètent le faible nombre d'attributs présentés

	Complexité faible	Complexité modérée	Complexité élevée
Série 1	2	<b>17</b>	5
Série 2	2	<b>15</b>	7
Série 3	3	<b>13</b>	8
Série 4	3	<b>14</b>	7
Série 5	3	<b>14</b>	7
Série 6	3	<b>13</b>	8

TABLE 3.3 – Choix de l'écran d'accueil pour chaque série.

sur cet écran et donc le faible besoin de cognition pour le comprendre. Enfin, lors de la dernière série, comme pour les mesures des revisites et des fixations, les temps de visionnage sont similaires. Cette similarité dans les résultats de cette série pourrait être due à l'apparition d'une vérification du choix. Cette dernière phrase est cependant une intuition et devra être confirmée par une nouvelle étude.

Il est intéressant d'observer que les valeurs de fixations et de temps de visionnage entre tous les écrans d'accueil de la quatrième série sont similaires, ce qui peut indiquer la nécessité d'une comparaison globale à un certain stade du processus.

En ce qui concerne la question de la comparaison par paire, **notre hypothèse H7 est confirmée**. Au-delà du biais de la position centrale de l'image modérée sur les séries S1 et S2, l'écran de complexité modérée joue son rôle de référence. Les valeurs de toutes les métriques sont stables pendant la séquence S3 - S4 - S5. Si nous examinons les liaisons, nous constatons que la comparaison par paire entre les écrans modérés et élevés est plus forte que la paire des écrans faibles et modérés, ce qui reflète une comparaison plus détaillée entre les images dont le niveau de complexité est supérieur.



**À retenir :** L'écran moyennement complexe joue un rôle de référence en matière de comparaison dans le processus de sélection des personnes.

### 3.5 Conclusion

La caractérisation d'un niveau de complexité de l'information mentalement acceptable est primordiale dans l'objectif de générer un changement de comportement chez l'individu. Pour choisir le niveau de complexité approprié, nous proposons une étude d'oculométrie afin d'évaluer différentes informations pour encourager l'adoption d'un comportement plus vertueux. Pour cela, nous mobilisons les domaines de l'attention visuelle, de la perception de la complexité, du processus de décision chez l'individu et de l'impact de la position de l'information sur le choix. En nous basant sur cette littérature, nous émettons des hypothèses sur le niveau de complexité des informations à transmettre à l'utilisateur, mais aussi sur la présentation de ces informations afin de construire une application d'autosuiivi. Les résultats de notre étude montrent que le niveau de complexité le plus apprécié est le niveau modéré. Le niveau de complexité élevé recueille le

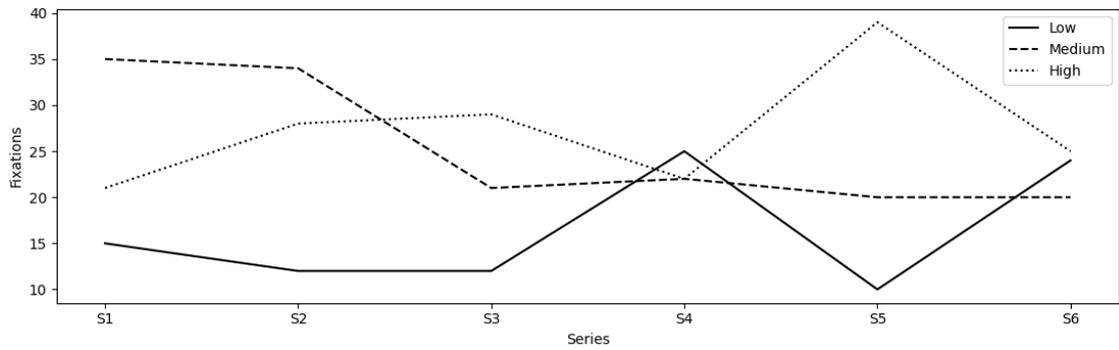


FIGURE 3.12 – Valeur moyenne des fixations pour chaque série

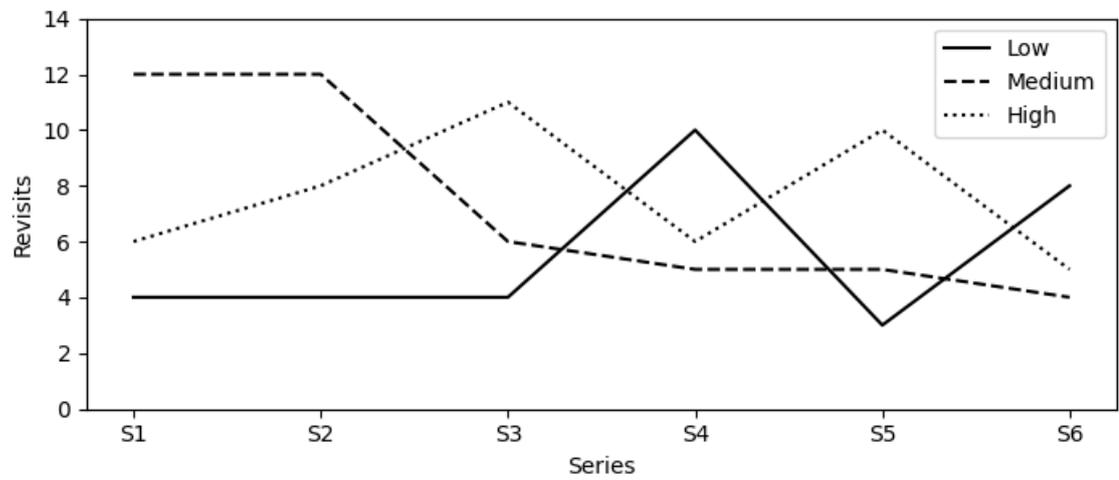


FIGURE 3.13 – Valeur moyenne des revisites pour chaque série

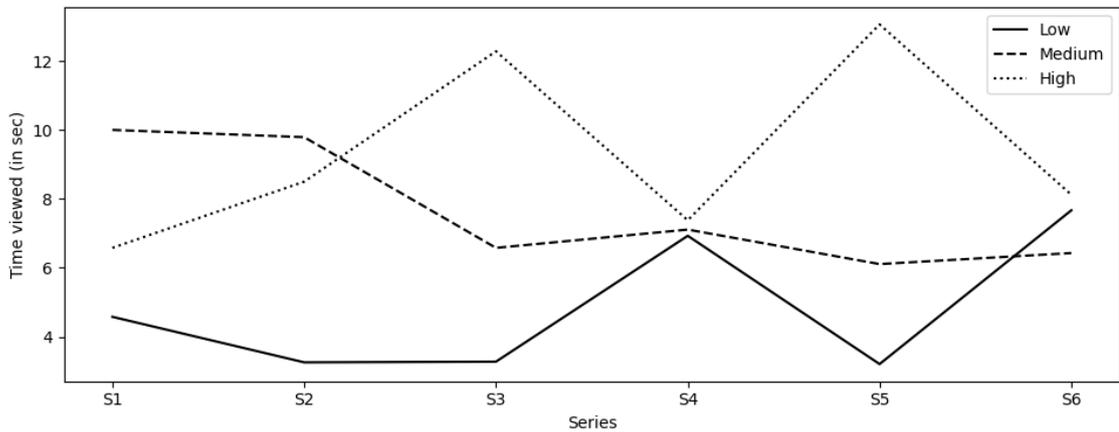


FIGURE 3.14 – Valeur moyenne du temps de visionnage pour chaque série

plus grand nombre de fixations et la plus grande durée de fixation. Cependant, le nombre de revisites est similaire entre le niveau modéré et le niveau complexe. La position centrale de l'information a une influence sur le choix mais doit être fortement corrélée avec le niveau de complexité de l'écran d'accueil. Par ces résultats, nous voyons clairement que la complexité de l'information transmise joue un rôle dans l'attention et donc dans la prise de décision des individus. Cette notion doit donc être prise en compte dès la conception d'une application mobile d'autosuiwi pour permettre de faciliter son acceptation et ainsi maximiser l'avènement d'un changement de comportement.

À partir de ces résultats, nous avons intégré les notions associées dans une application mobile. Cette nouvelle expérience en conditions réelles auprès d'un large public d'étudiants nous permettra de valider les résultats de l'impact de la complexité informationnelle dans une tâche cognitive plus large. Les participants à cette nouvelle expérience seront divisés en trois groupes (faible, modéré et complexe), chacun ayant un écran d'accueil avec un niveau de complexité correspondant. Elle nous permettra également de démontrer une éventuelle corrélation entre le niveau de complexité de l'information transmise et le comportement réel.



#### Contributions de ce chapitre :

- Nous avons mené une expérience de suivi oculaire en laboratoire avec 24 participants pour tester l'impact potentiel de la complexité de l'information sur les processus de choix et d'attention.
- Nous avons observé que les informations d'une grande complexité provoquent plus de fixations et un temps de visualisation plus long que les informations d'une complexité moyenne ou faible. Elles entraînent également un nombre de revisites équivalent aux informations de complexité modérée et un nombre supérieur comparé à celles de complexité faible.
- Nous avons confirmé qu'il existe une relation entre la position de l'information et l'augmentation de l'attention, qui n'est pas corrélée avec la complexité de l'information présentée. La position centrale de l'information provoque également plus de sélection que toute autre position pour la même information à complexité égale.



## Chapitre 4

# Acceptation de l'information : vers un changement de comportement ?

---

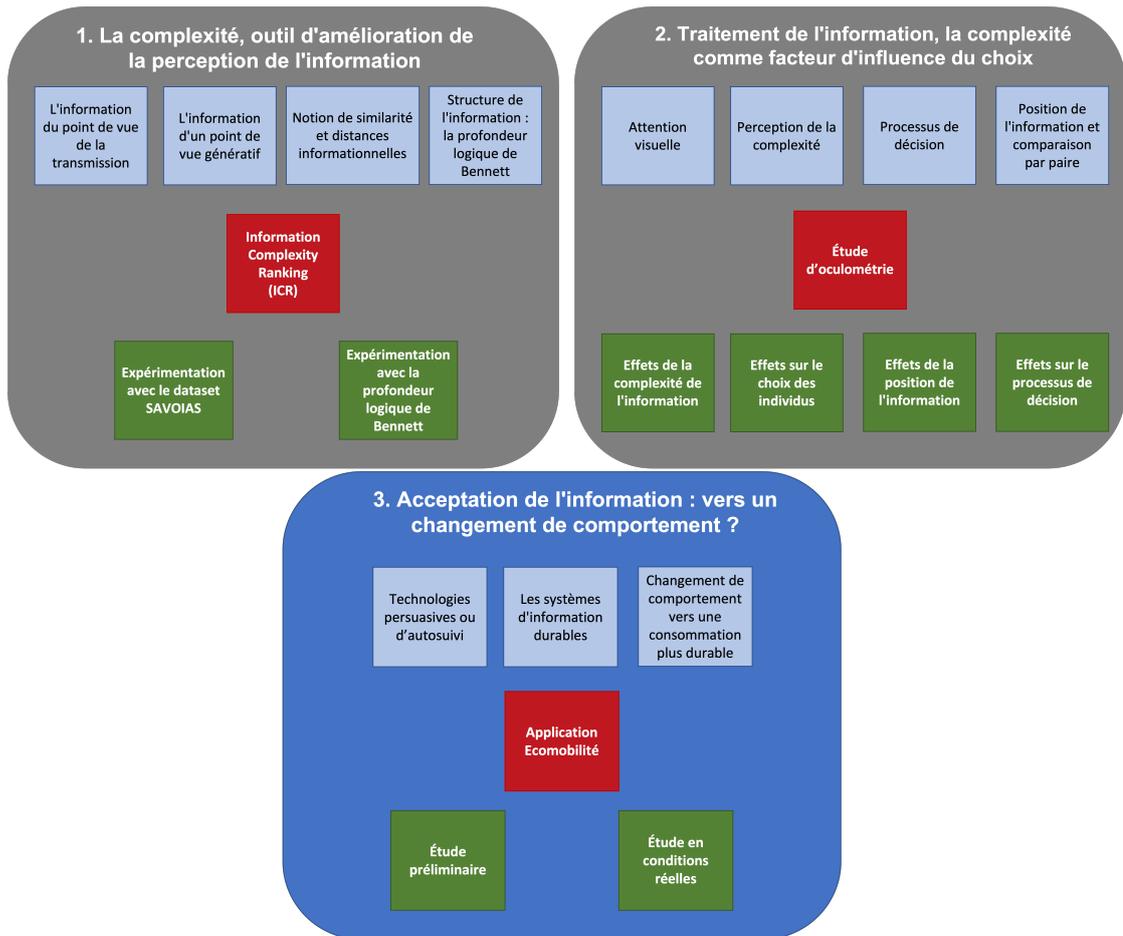
### Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>79</b>
<b>4.2</b>	<b>État de l'art</b>	<b>81</b>
4.2.1	Technologies persuasives ou d'autosuiivi	81
4.2.2	Les systèmes d'information durables	83
4.2.3	Changement de comportement vers une consommation plus durable	84
<b>4.3</b>	<b>Contribution : Application mobile (Eco-Mobilité)</b>	<b>87</b>
4.3.1	Contexte	87
4.3.2	Développement et fonctionnement de l'application	88
4.3.3	Protocole expérimental	91
4.3.4	Hypothèses	93
4.3.5	Résultats	95
<b>4.4</b>	<b>Conclusion</b>	<b>100</b>

---

Les résultats présentés dans ce chapitre sont basés sur l'article « The effect of visual information complexity on urban mobility intention and behavior. » publié dans la conférence RCIS 2023.



## 4.1 Introduction

Dans les pratiques de mobilité quotidienne, le choix du moyen de transport est souvent influencé par l'habitude. L'enjeu est donc d'inciter les usagers à moins utiliser leur voiture pour faire baisser les émissions de gaz à effet de serre. Mais pour y parvenir, une prise de conscience est nécessaire, qui peut être favorisée par une information comprise et acceptée par les individus. Il est aussi essentiel de choisir judicieusement l'outil de communication approprié pour renforcer l'impact de cette information. À l'heure actuelle, le smartphone est l'appareil de communication le plus utilisé au quotidien avec un fort taux de pénétration dans la population globale [Stéphane Legleye, 2022, Croutte Patricia, 2021] :



En France, plus de 99% de la population âgée de 15 ans ou plus est équipée d'un téléphone, fixe ou mobile. L'équipement mobile est la norme : 95% de la population possède un téléphone mobile et 77% détient plus particulièrement un smartphone.

*INSEE 2022*



La durée moyenne de temps passé sur un écran s'élève à 32 heures par semaine.

*CREDOC 2021*

Cependant, il reste une question fondamentale : quelle est la perception de l'information reçue ? La question est d'autant plus importante dans un contexte d'écomobilité, au vu de la difficulté cognitive de se représenter l'impact de ce choix. L'information permet alors d'orienter le choix et ainsi de réduire la charge mentale. Les chercheurs ont ainsi étudié les liens entre la personnalisation de l'information et l'impact de cette dernière pour apporter une aide au changement de comportement. Au-delà de la personnalisation, la possibilité de suivre ses propres données peut être un autre levier invitant les individus à réfléchir à leur propre comportement (démarche de réflexivité). Dans un cadre purement individuel ce suivi semble fonctionner et encourage, par exemple, les personnes à prendre soin de leur santé [Jin et al., 2022]. Si la dimension personnelle est un levier efficace au changement de comportement, qu'en est-il pour une dimension collective, comme la réduction de la consommation de CO<sub>2</sub> lors de nos déplacements ? Autrement dit, les outils efficaces pour un bénéfice « personnel », le sont-ils toujours pour un bénéfice « commun » ?

Pour savoir si ce changement de perception des avantages d'un outil d'autosuiivi (self-tracking, en anglais) au bénéfice d'un collectif n'invalide pas notre application, nous nous intéressons aux effets des systèmes d'information durables ou SI verts (Green Information Systems ou green IS). Le concept de SI vert vise à utiliser les systèmes d'information (SI) pour atteindre et favoriser les objectifs liés à la durabilité [Brauer et al., 2016]. Ce domaine de recherche étant encore jeune et polarisé sur des thèmes individuels (par exemple, santé

ou éducation), nous souhaitons l'appliquer au domaine de la mobilité urbaine.

Il convient de rappeler que plusieurs variables situationnelles (par exemple, conditions météorologiques, infrastructures de transports, etc.) ou sociales (par exemple, style de vie, pression sociale, etc.) influencent ce choix de mobilité quotidiennement. Pour aider à comprendre le poids de ces variables externes, les chercheurs dans le domaine de la consommation durable proposent des définitions et des modèles qui ont pour objectif d'aider à l'adoption de comportement plus vertueux sur le long terme [Trudel, 2019, Allcott, 2011, Schultz et al., 2007].

Pour atteindre notre objectif, nous avons opté pour une démarche scientifique pluridisciplinaire qui mêle à la fois un point de vue informatique et un point de vue sciences de gestion. L'intégration d'un point de vue non informatique est un avantage pour prendre en compte tous les aspects de notre question de recherche. À notre connaissance, aucune recherche n'a étudié l'effet de la complexité avec un appareil mobile sur une dimension environnementale. Notre contribution se porte donc sur l'analyse de l'intégration de la notion de complexité dans des applications de suivi personnel et durables (green SI). Dans un premier temps, nos travaux permettent de connaître l'effet de la complexité visuelle avec un outil numérique (application mobile) sur une intention et un comportement de mobilité durable. Dans un second temps, nous analysons l'influence de l'acceptation de la technologie sur les intentions et les comportements durables. Pour approfondir notre analyse du rôle de la complexité, nous nous basons sur la définition de variables incitatives à l'adoption d'un comportement durable.

Pour tester l'efficacité de la complexité et les effets de l'acceptation de la technologie, nous avons proposé deux études. En premier lieu, nous avons mis en place une étude préliminaire permettant de tester la complexité d'un visuel sur les intentions d'utilisation d'une application d'aide au changement de comportement. À partir de ce visuel, nous avons également testé les intentions de changement de comportements des individus. Les résultats de cette première étude, nous ont permis de mieux comprendre les éventuels effets de la complexité d'un visuel sur des individus avant leur intégration dans notre application mobile. Dans un second temps, nous avons déployé notre application mobile expérimentale de suivi de sa propre consommation de CO<sub>2</sub>. Dans cette étude en conditions réelles, nous comparons les comportements suivant différents niveaux de complexité de l'information transmise. Cette dernière est divisée en trois niveaux différents de complexité (simple-moderé-complexe). Chaque participant est soumis à un unique niveau de complexité afin d'analyser plus précisément son potentiel impact sur un changement de comportement. Ce test auprès d'utilisateurs finaux, nous a aidé à mieux comprendre l'impact de l'information mais également la nécessité de combiner cette information à d'autres variables afin de maximiser l'adoption de nouveaux comportements durables.

Les différentes notions utiles à nos travaux sont présentées dans la Section 4.2. Dans la Section 4.3.1, le contexte de notre étude avec une application mobile expérimentale auprès d'étudiants est présenté. Dans la Section 4.3.2, nous présentons le développement et les caractéristiques de celle-ci. Dans la Section 4.3.3, nous présentons le protocole de nos deux études sur l'impact de la complexité de l'information dans une application mobile durable. Les hypothèses et les résultats associés à ces études sont finalement présentés

dans les Sections 4.3.4 et 4.3.5.

## 4.2 État de l'art

### 4.2.1 Technologies persuasives ou d'autosuivi

Les technologies d'autosuivi permettent aux gens d'accéder à leurs données personnelles, de les surveiller et de les analyser [Constantiou et al., 2022]. Ce suivi individuel est désormais une pratique courante dans de nombreux domaines : pour évaluer ses performances, son nombre de pas, son alimentation, son sommeil... [Jin et al., 2022]. Bien que ces solutions numériques se développent pour amener à un comportement responsable, peu d'études ont été menées sur ce sujet. Par exemple, [Gabrielli et al., 2013] a exploré l'impact d'une application mobile proposant différentes options de transport sur les décisions des individus. Cette étude montre une légère augmentation des choix de transport durable sur une période d'un mois.

La recherche sur l'adoption de la technologie d'autosuivi est basée sur la théorie unifiée de l'acceptation de la technologie (Unified Theory of Acceptance and Use of Technology ou UTAUT, en anglais) [Venkatesh et al., 2012]. UTAUT1 propose un modèle explicatif des intentions d'utilisation d'une technologie [Venkatesh et al., 2003]. Cette acceptation repose sur quatre piliers : **l'espérance de performance, la durée de l'effort à fournir, la pression sociale et des conditions propices d'utilisation**. L'espérance de performance est définie comme le degré de bénéfice apporté par l'utilisation de la technologie. La durée de l'effort recouvre quant à elle la facilité avec laquelle l'utilisateur utilise la technologie. Le jugement d'autres personnes comme la famille, les amis, les collègues, etc. sur l'utilisation ou non de la technologie est intégrée au modèle avec l'élément de la pression sociale. Le dernier élément composant UTAUT, à savoir les conditions propices d'utilisation, décrit la perception des utilisateurs des ressources et du support à leur disposition pour accomplir le comportement voulu. Les trois premiers, sont des déterminants directs de l'intention et du comportement de l'utilisateur et sont modérés par une combinaison de facteurs (l'âge, le genre et l'expérience). Le dernier pilier est un déterminant direct du comportement de l'utilisateur et est modéré par l'âge et l'expérience d'utilisation [Venkatesh et al., 2012]. Une deuxième version de ce modèle présentant d'autres variables explicatives comme **la motivation hédoniste** (c'est-à-dire la volonté de se faire plaisir), **la valeur de prix de la technologie** et **les habitudes** a été proposée (voir Figure 4.1). Cette deuxième version a été testée dans un contexte de consommation via un questionnaire en ligne en deux étapes auprès de 1512 consommateurs. Les résultats obtenus permettent d'expliquer la variance dans l'intention comportementale à hauteur de 74% et dans l'utilisation de la technologie à hauteur de 52%. Comparé à la base de référence du modèle qui expliquait 56% de la variance dans l'intention comportementale et 40% dans l'utilisation de la technologie, UTAUT2 présente des résultats significativement meilleurs. Dans le cadre de notre recherche, nous adaptons ce modèle, en intégrant une variable explicative supplémentaire, la complexité de l'information (voir Figure 4.1).

Différentes applications et études ont prouvé l'efficacité de ce modèle dans différents contextes :

- Koivumäki et al. se sont intéressés à l'influence de la durée d'utilisation de services mobiles sur leur acceptation auprès de 243 personnes dans la ville Oulu en Finlande. L'étude prouve qu'un temps d'utilisation supérieur n'a pas d'effet sur les perceptions des consommateurs à l'égard de ces services contrairement à la familiarité de l'appareil et aux compétences des utilisateurs [Koivumäki et al., 2008].
- Eckhardt et al. ont étudié l'impact de l'influence sociale sur l'acceptation de la technologie auprès de 152 entreprises de tailles diverses (entre 500 et 200 000 employés) en Allemagne. Leur étude a montré un impact significatif de l'influence sociale des collègues et des supérieurs hiérarchiques sur l'adoption de la technologie de bases de données de *curriculum vitae* [Eckhardt et al., 2009].
- Curtis et al. ont lancé en 2009 une enquête auprès de 409 personnes travaillant dans les relations publiques pour des ONG aux États-Unis. Les auteurs ont montré que les ONG dotées de services de relations publiques adoptent et utilisent plus facilement les réseaux sociaux pour atteindre leurs objectifs [Curtis et al., 2010].
- Verhoeven et al. ont utilisé le modèle UTAUT en Belgique auprès de 714 étudiants de première année d'université, pour mettre en exergue l'évolution des compétences et d'utilisation en matière de SI entre l'école secondaire et l'université [Verhoeven et al., 2010].

Cependant certaines limitations subsistent dans l'utilisation de ce modèle concernant l'âge moyen de l'échantillon participant à l'étude. Cela peut en effet influencer l'acceptation et l'utilisation de la technologie, comme le précisent Venkatesh et al. [Venkatesh et al., 2012]. Une autre influence peut également venir du taux de déploiement de la technologie en question dans le pays concerné par l'étude. Dans notre étude ces variables seront prises en compte. L'intégration du modèle UTAUT permettrait de voir si la complexité de l'information joue un rôle facilitateur ou non sur l'intention et l'utilisation de notre vecteur informationnel, l'application mobile.



**À retenir :** Le modèle de la théorie unifiée de l'acceptation de la technologie (UTAUT) permet de comprendre les leviers d'acceptation de la technologie. Cette théorie repose sur la performance du dispositif, de l'effort à fournir par l'utilisateur, de ses compétences à l'utiliser et de l'injonction sociale à l'utiliser. Accepter la technologie, dans notre cas l'application mobile, est une condition *sine qua non* à la construction d'un futur comportement. En effet, si le public visé par l'information n'utilise pas ou peu l'application par laquelle l'information est transmise, les effets seront inexistantes. Cette théorie nous permet également d'optimiser le développement de vecteurs d'information, par exemple une application mobile.

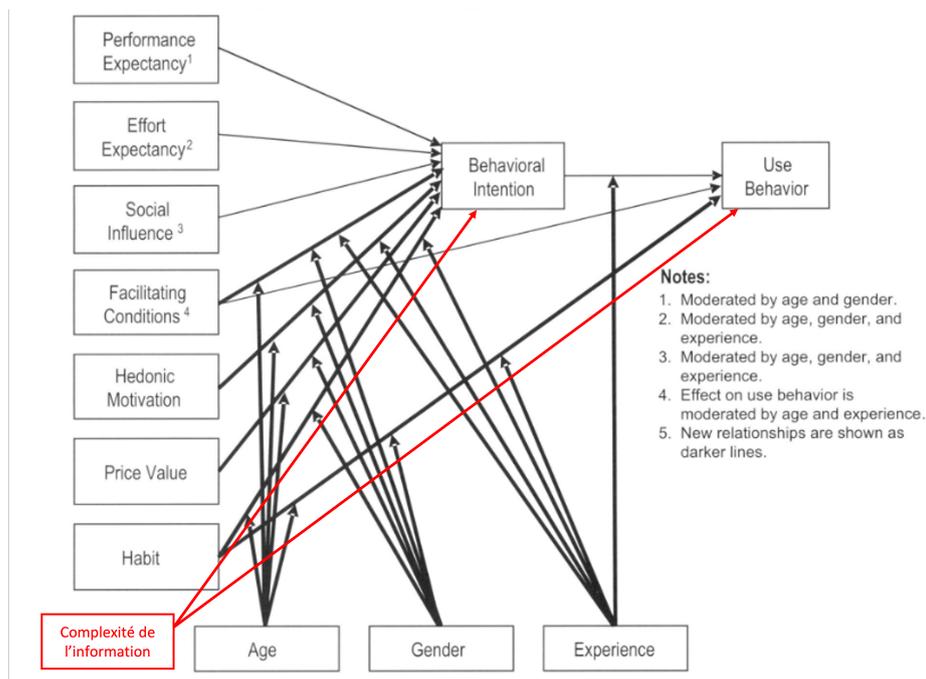


FIGURE 4.1 – Adaptation du modèle UTAUT2 [Venkatesh et al., 2012] à notre contexte de recherche. L’adaptation (en rouge) correspond à l’ajout de la complexité de l’information comme variable explicative de l’intention comportementale et du comportement d’utilisation.

#### 4.2.2 Les systèmes d’information durables

Pour encourager le changement de comportement, l’utilisation de SI verts (c’est-à-dire les systèmes d’information durables) a émergé. Les rôles joués par ces SI peuvent être classés en trois types : **automatiser, informer et transformer**. Même si dans la pratique, la majorité des SI ne peuvent remplir qu’un seul de ces trois rôles [Standing et al., 2008]. Au-delà des rôles spécifiques, ce sont les objectifs des SI verts centrés sur l’utilisateur qui restent à définir, ce domaine de recherche étant encore peu exploré [Brauer et al., 2016]. Les domaines d’application les plus courants de ces SI sont orientés vers la santé et l’éducation. Des études ont été menées avec des applications mobiles d’auto-suivi dans le domaine de la santé avec un feedback pour perdre du poids [Quero and Crié, 2020] ou pour réduire la consommation d’alcool [Lehto and Oinas-Kukkonen, 2015]. Dans le domaine de l’éducation, ces applications servent par exemple à promouvoir des méthodes d’apprentissage à l’aide de micro-contenus pour éviter la surcharge informationnelle [Bruck et al., 2012]. Les résultats de ces études montrent des interventions efficaces sur le comportement des utilisateurs.

Dans le contexte de notre recherche, à savoir les applications de mobilité durable, le rôle le plus employé est d’informer [Brauer et al., 2016] mais aussi de fournir des outils ou des systèmes de feedback [Flüchter et al., 2014]. Tulusan et al. proposent une

application mobile basée sur l'écofeedback pour réduire la consommation de carburant des employés pendant la conduite [Tulusan et al., 2012]. L'impact du feedback a été évalué et révèle que les conducteurs utilisant l'application économisent 3,23% de carburant. Il s'agit d'un impact qui peut sembler minime sur les 25 participants de l'étude mais qui peut avoir une signification sur une population plus large. Toujours dans l'objectif d'une mobilité plus responsable, la mesure de l'impact sur la mobilité individuelle d'une offre de transport plus respectueuse de l'environnement est intéressante [Froehlich et al., 2009]. Les résultats de cette étude préliminaire sur 13 volontaires montrent l'engagement dans l'application mais pas un changement de comportement. On peut noter que dans le contexte de la mobilité urbaine, les résultats de certaines études basées sur le retour d'expérience sont mitigés. On peut donc légitimement se poser la question de l'impact des informations transmises à l'utilisateur et de sa compréhension de celles-ci. Pour tenter de déterminer l'impact de l'information fournie sur les intentions et comportements de mobilité, nous pensons que la complexité de l'information peut avoir une influence. Nous avons donc créé notre propre application pour mesurer cette variable sur les intentions et les comportements (voir Section 4.3).



**À retenir :** Les systèmes d'information durables sont un domaine encore assez récent, en particulier ceux centrés autour de l'utilisateur. Et pourtant la majorité des SI proposés sont orientés autour de domaines avec une forte composante personnelle. Dans un contexte de mobilité urbaine, on note des résultats parfois décevants ou minimes. Nous pensons donc qu'un axe d'amélioration est possible concernant la compréhension de l'information et sa complexité.

### 4.2.3 Changement de comportement vers une consommation plus durable

Une fois le système d'information vert mis en place, il est important de s'interroger sur son utilisation à long terme. L'objectif de ces vecteurs d'information est d'engendrer un comportement (au sens d'utiliser) plus vertueux et dans la durée. Ce comportement permet de satisfaire les besoins actuels des personnes tout en limitant l'impact sur l'environnement [Trudel, 2019]. L'étude et la compréhension de ce changement de paradigme entre parfaitement en adéquation avec notre cadre de recherche sur une mobilité urbaine plus respectueuse de l'environnement. Dans sa méta-analyse, Trudel a synthétisé vingt années de recherches sur le comportement durable du consommateur a identifié quatre axes de recherche : les barrières cognitives, le soi, l'influence sociale, les caractéristiques des produits [Trudel, 2019]. On peut constater que les axes de recherche autour de cette question sont partagés de manière pluridisciplinaire.

Le premier frein à l'adoption d'un nouveau comportement repose sur la capacité des individus à prendre des décisions à court ou long terme. Trudel met en avant le fait que nous préférons dans la majorité des cas opter pour l'option la plus simple du point de vue cognitif. Pour cela, il se base sur le modèle psychologique de Daniel Kahneman [Morewedge and Kahneman, 2010] composé de deux systèmes de raisonnement appelés système

1 (S1) et système 2 (S2). Le premier (S1) est un système qui traite les informations de manière brute et rapide. Il est utilisé par défaut et est peu « coûteux » cognitivement. À l'inverse, le second système (S2) est associé à un traitement plus analytique, plus complexe, nécessitant plus de ressources cognitives. La perception psychologique des bénéfices d'une consommation plus durable sont distants, abstraits, incertains et plus difficiles à saisir pour les individus [Spence et al., 2012] et sont donc par conséquent plutôt traités par le système S2. Trudel préconise, pour augmenter l'adoption de comportements durables, de rendre les conséquences du comportement de consommation moins abstraites et plus tangibles. Cela démontre donc le besoin d'informer les gens sur leur comportement.

L'individualité de la décision, autrement dit « le soi » est à prendre en compte pour tendre vers une consommation plus durable. Cet aspect du comportement prend la forme d'une combinaison de plusieurs motivations comme :

- **l'auto-signalisation** qui est le sentiment positif qu'un individu se renvoie à lui-même quand il fait un choix de consommation durable ;
- **l'auto-identification** qui est un sentiment d'identification qui permet à un individu de se distinguer des autres. Contrairement à la motivation d'auto-signalisation, le jugement des autres individus est important ;
- **l'identification sociale** qui représente l'influence des attitudes, croyances ou comportements du groupe social auquel appartient l'individu sur ces propres choix de consommation.

Cela se traduit par le besoin de fournir aux autres et à soi-même une représentation d'un comportement en adéquation avec ces propres principes mais aussi avec des valeurs environnementales positives. Il est à noter que les dimensions sociale ou de groupe influencent ces comportements. En effet, un individu peut changer de comportement s'il y est contraint par les autres membres d'un même groupe social. Il est important de ne pas négliger ces normes, car elles existent dans presque tous les aspects du comportement humain [Cialdini, 1993]. L'utilisation de normes de type descriptives (par exemple, information sur la consommation moyenne) a été testée dans le domaine de la consommation énergétique et a prouvé son efficacité [Allcott, 2011, Schultz et al., 2007]. Par conséquent, il est utile d'intégrer ces normes pour informer les individus sur leur comportement de mobilité urbaine.

Le dernier axe de recherche, sur les caractéristiques du produit, montre que les individus font certaines associations entre les attributs qu'ils recherchent dans un produit et les caractéristiques de ce dernier (par exemple, les produits durables sont plus doux ou les produits traditionnels sont associés à la force) [Luchs et al., 2010].

Dans sa méta analyse, l'auteur précise également le manque de recherche sur les possibilités d'offrir aux individus la motivation de changer de comportement. De plus, il souligne que la majorité des recherches sont menées sur des comportements durables à faible impact écologique (par exemple, éteindre les lumières). Des recherches sur des comportements à fort impact écologique (par exemple la consommation d'énergie, la mobilité urbaine, etc.) permettrait une meilleure compréhension des mécanismes d'adoption de comportement durable. White et al. ont fixé un cadre global pour aider à influencer l'adoption de comportement durable, qui repose sur cinq axes psychologiques similaires

à ceux de l'analyse de Trudel : l'influence sociale, la formation d'habitudes, le moi individuel, les sentiments et la cognition, et la tangibilité [White et al., 2019].

Dans un contexte de mobilité, De Witte et al. ont contribué à la compréhension du choix modal en étudiant les déterminants de ce dernier. À partir des déterminants et de leur influence, ils ont pu modéliser les décisions prises lors du choix de moyen de transport. Dans l'étude, ces indicateurs du choix sont regroupés en trois catégories [Witte et al., 2013] :

- **Indicateurs socio-démographiques** comprenant l'âge, le genre, le niveau d'éducation, l'emploi, les revenus, la composition du foyer et la possession d'une voiture. Ce sont des variables propres à chaque individu.
- **Indicateurs spatiaux** comprenant la densité d'habitants, la diversité des moyens de transport, la proximité des services et des infrastructures de transport, la fréquence de passage des transports en commun et la disponibilité des parkings (lieux de stationnement). Ces variables représentent des caractéristiques de l'environnement de voyage, elles sont externes à l'individu.
- **Indicateurs des caractéristiques du trajet** comprenant la motivation du trajet, la distance, le temps de trajet, le coût du trajet, l'heure de départ, l'enchaînement des trajets, les conditions météorologiques, l'information sur les offres de transport et l'interconnexion entre différents moyens de transport telles que les correspondances. Ces variables environnementales ont un impact direct sur le choix des individus.
- **Indicateurs socio-psychologiques** comprenant les expériences, la familiarité (c'est-à-dire l'habitude de prendre différents moyens de transport), le mode de vie, les habitudes de transports et la perception du voyage. Ces variables individuelles sont dépendantes des options et des mécanismes de pensée créés par les indicateurs précédents.

De Witte et al. concluent leur analyse par le constat suivant : le choix du moyen de transport est un concept complexe et multifacteurs [Witte et al., 2013]. Les facteurs externes tels que la disponibilité des moyens de transport plus vertueux ne sont pas toujours contrôlables dans un contexte de recherche. Dans cet objectif d'atténuer les effets potentiellement négatifs de ces facteurs, il est important de maximiser l'efficacité de facteurs individuels comme l'information.



**À retenir :** L'examen complet de la littérature de ces vingt dernières années fait par Trudel permet de mieux comprendre et focaliser nos efforts sur des axes particuliers (individuels et collectifs) pour inciter à un comportement durable. Un cadre expérimental se basant sur des axes similaires a été proposé par White et al. Dans le cadre de la mobilité, différents facteurs internes et externes ont été définis pour mieux comprendre les choix des individus en matière de moyen de transport. Dans notre recherche, nous nous concentrons principalement sur des aspects cognitifs, donc individuels, du choix du moyen de transport. Nous essayons d'inciter les gens à changer de comportement au travers de l'information transmise. Il est tout de même pertinent de prendre en compte ces découvertes pour comprendre l'influence de l'information dans un contexte global.

## 4.3 Contribution : Application mobile (Eco-Mobilité)

### 4.3.1 Contexte

L'application de suivi individuel de la consommation de CO<sub>2</sub> dans les déplacements quotidiens est l'expérimentation centrale de ma thèse pour me permettre de confronter mes recherches théoriques aux conditions réelles. C'est dans ce cadre que nous avons, dès le début du projet Ecomob, mis en oeuvre des prototypes d'application de suivi. Au fur et à mesure des tests et des retours des bêta-testeurs, nous avons affiné notre protocole scientifique en évitant au maximum l'intégration de biais qui auraient affecté notre mesure de l'impact de la complexité sur le comportement. Certaines conditions de l'expérimentation ont été fixées avant même les premiers tests. Le public cible de notre application a été choisi en fonction des deux facteurs suivants : le premier correspond au niveau d'information de ce public sur les questions écologiques et la nécessité d'adopter un mode de transport plus vertueux (c'est-à-dire moins émetteur de CO<sub>2</sub> que la voiture). Le second, leur nombre potentiel, leur proximité géographique avec notre université et leur connaissance de leur environnement (par exemple, leur connaissance des moyens de transport en commun). En prenant en compte ces deux facteurs, notre choix s'est logiquement porté sur la population étudiante de La Rochelle Université. Pour avoir une diversité de comportements et un nombre statistiquement viable d'expérimentateurs, nous avons sélectionné les étudiants de toutes les disciplines à partir de la deuxième année de licence. Au total, nous avons contacté plus de 1000 étudiants pour leur présenter notre application, et leur avons proposé de prendre part à notre étude en conditions réelles de mobilité. Progressivement, grâce à l'avancée de nos recherches et des connaissances acquises, la nature et la forme des informations transmises et le design de l'application mobile ont évolué. Le développement et l'amélioration de l'application qui ont mené à la version présentée dans cette étude ont été guidés par les résultats des deux autres expérimentations majeures réalisées durant ma thèse (voir Section 2.3 et Section 3.3).

### 4.3.2 Développement et fonctionnement de l'application

Concernant le développement de notre application, nous avons choisi un framework hybride (Apache Cordova [Cordova Framework]) pour gagner du temps de développement afin de déployer Eco-Mobilité sur les plateformes Android et IOS. Après des premiers tests peu concluants avec une géolocalisation automatique restreinte qui s'active lorsque l'utilisateur se trouve dans des zones spécifiques pour préserver plus efficacement sa vie privée, nous avons choisi une base de développement déjà éprouvée, le projet E-mission [Shankari et al., 2018]. E-mission est une plateforme de mobilité open source développée à l'université de Berkeley qui a déjà fait ses preuves avec plusieurs déploiements dans le monde [Shankari et al.]. Notre application est donc basée sur cette architecture et le serveur est une application web python, les données sont stockées dans une base de données MongoDB, le client est une application Cordova pour Android et IOS. Nous avons complètement repensé l'interface utilisateur, avec l'intégration de notre protocole. Nous avons supprimé tous les services tiers (par exemple les services Google), développé une vue unique, appelé tableau de bord avec toutes les données de consommation de CO<sub>2</sub>.



FIGURE 4.2 – Écrans d'accueil avec les trois niveaux de complexité (simple, modéré et élevé, de gauche à droite).

#### 4.3.2.1 Suivi des utilisateurs

Dès la phase de conception de notre application nous avons voulu protéger au maximum les données personnelles de nos futurs utilisateurs, dans le but de maximiser le nombre de participants. Nous avons donc pris en compte le concept de *privacy by design*. Ce concept se traduit par différentes mesures dans notre application. La première question que nous avons traitée dans le développement est celle de la géolocalisation. Nous avons travaillé en collaboration avec la directrice de la protection des données de notre université, pour que le recueil des données de géolocalisation soit en conformité avec le Règlement Général sur la Protection des Données (RGPD) et respecte les règles suivantes :

- l'utilisateur doit activer la fonctionnalité de géolocalisation explicitement dans les paramètres de son téléphone ;
- l'utilisateur doit accorder l'autorisation de récolte des données de ce type ;
- enfin, et en accord avec l'article 21 du RGPD, l'utilisateur peut à tout moment faire valoir son droit d'opposition.

Pour mettre en oeuvre ces obligations dans notre application, il suffit à l'utilisateur de désactiver sa position et, dans ce cas, ses déplacements ne seront plus suivis. Pour respecter encore plus la vie privée de nos utilisateurs nous ne déclenchons la collecte des données de géolocalisation que deux minutes après la détection du mouvement par le téléphone. Ce délai d'activation permet de ne pas récolter l'adresse exacte du point de départ de l'utilisateur et ainsi de garantir la confidentialité de l'adresse de son domicile par exemple.

Le choix de la géolocalisation continue de nos utilisateurs, nous permet d'éviter de demander à l'utilisateur d'effectuer de nombreuses actions, comme remplir un questionnaire sur son voyage ou ses modes de transport, ce qui peut être fastidieux. Du point de vue du protocole scientifique, le fait de ne pas demander d'informations supplémentaires à l'utilisateur prévient l'introduction de potentiels biais dans notre expérience.

Afin de communiquer l'information la plus pertinente à nos utilisateurs, nous avons mis l'accent sur la détection de modes de transports multiples. Pour un calcul le plus précis possible de la consommation de CO<sub>2</sub>, l'application devait être capable de détecter un maximum de moyens de transport mais aussi une combinaison de plusieurs d'entre eux. La phase de détection d'un moyen de transport se fait en utilisant la précision des points GPS et l'horodatage associé à chaque point. Ces traitements sont également utilisés pour détecter les trajets de chaque utilisateur afin de calculer la consommation de CO<sub>2</sub> associée. L'application devait également pouvoir utiliser des valeurs de référence différentes dans le calcul des émissions de CO<sub>2</sub> pour chaque mode de transport, il est donc possible de changer ces valeurs dans notre application.

En complément de la détection des modes de transport, un soin particulier a été apporté au traitement de la géolocalisation. Dans un premier temps, l'algorithme élimine tous les points GPS de faible précision, puis élimine les points identiques. Ensuite, il considère tous les points des cinq dernières minutes et les 10 derniers points GPS associés à l'utilisateur. Il choisit l'ensemble des points contenant le point le plus ancien. Si tous les points de cet ensemble sont à moins de 100 mètres du dernier point trouvé, alors le

chemin est considéré comme complet. Ce traitement permet également de prendre en compte les faiblesses ou les éventuelles erreurs du signal GPS du téléphone.

#### 4.3.2.2 Interface utilisateur

L'interface utilisateur de l'application Eco-Mobilité (appelée tableau de bord), a été conçue à partir des résultats de notre étude préliminaire (voir Section 4.3.3.1).

Afin d'éviter certains biais et d'isoler notre variable correspondant à la complexité de l'information, nous avons décidé de supprimer certaines fonctionnalités habituellement présentes dans les applications d'autosuivi, en nous limitant au strict minimum. Ainsi notre application ne propose pas de personnaliser l'affichage des résultats ou d'avoir un profil détaillé de l'utilisateur avec un historique à long terme, une évolution de la consommation, etc. Il est également impossible de cartographier sa géolocalisation contrairement à la version développée par l'université de Berkeley. Enfin, les fonctions sociales permettant de partager sa consommation avec d'autres utilisateurs ou un public plus large ont été écartées afin d'améliorer la mesure de l'impact des informations sur le comportement. Toutes ces limitations ont pour but de ne pas introduire d'effets distractifs pour l'utilisateur qui pourraient faire varier l'impact de la complexité de l'information sur les intentions et les comportements. En revanche, ces limitations peuvent être perçues négativement par de potentiels utilisateurs comme étant un « manque de fonctionnalités ».

Le tableau de bord de l'application comprend des éléments communs aux trois niveaux de complexité (voir Figure 4.2). Dans la partie supérieure de l'écran, deux fonctionnalités sont présentées, dans le coin droit, le manuel de l'utilisateur qui explique la signification des données affichées dans les différents indicateurs (le symbole du point d'interrogation). Toujours dans le coin droit, une fonctionnalité de mise à jour des données (symbole de rafraîchissement) est proposée pour connaître sa consommation de CO<sub>2</sub> en quasi temps réel. Nous avons choisi dès le départ d'améliorer le retour d'information au maximum en informant le plus régulièrement possible l'utilisateur de sa consommation de CO<sub>2</sub>. Pour cela, nous avons également ajouté l'horodatage de l'exécution de l'analyse pour indiquer à l'utilisateur quand sa consommation de CO<sub>2</sub> est mise à jour.

Dans la partie inférieure de l'écran, les utilisateurs trouvent deux icônes pour naviguer dans notre application. La première (Dashboard) permet d'afficher les différentes données sur sa consommation de CO<sub>2</sub> en fonction du groupe de complexité attribué aléatoirement lors de l'inscription. La seconde (Profile) donne accès aux fonctionnalités de confidentialité telles que le bouton de déconnexion et la possibilité de télécharger ses propres données dans un format lisible (JSON). Cette page dispose également de la fonction « push up » pour envoyer les données du trajet au serveur.

Dans la partie centrale de l'écran tous les utilisateurs reçoivent deux informations sur leur consommation quotidienne de CO<sub>2</sub> : d'une part, une valeur numérique présentée en kilogrammes de CO<sub>2</sub> émis dans la journée et, d'autre part, un impact de CO<sub>2</sub> sur une échelle de 5 niveaux de A à E. Les différents niveaux de consommation représentent respectivement une fourchette de 0 kg émis jusqu'à 10 kg et plus de 10 kg pour le niveau E, avec un changement de lettre tous les 2,5 kg de CO<sub>2</sub> consommés. Cet indicateur est basé sur l'idée du Nutri-Score, également connu sous le nom de 5-Color Nutrition Label ou

5-CNL [Julia and Hercberg, 2017] et a déjà prouvé son efficacité sur les comportements de consommation [Dubois et al., 2021, Nabec et al., 2022]. En dessous de l'impact CO<sub>2</sub>, les utilisateurs des groupes modéré et complexe ont également leur consommation hebdomadaire moyenne. Enfin, dans le groupe complexe, les utilisateurs disposent de deux informations supplémentaires entre l'indicateur d'impact CO<sub>2</sub> et la consommation hebdomadaire moyenne : les valeurs de consommation minimale et maximale sur la semaine. Nous avons développé un affichage spécifique en fonction du groupe de complexité de l'utilisateur.

#### 4.3.2.3 Déploiement de l'application

À la suite de la phase de développement, des pré-tests techniques ont été conduits afin de pouvoir déployer dans les meilleures conditions notre application. Dans l'objectif de faciliter l'adoption de l'application par les utilisateurs nous avons déployé Eco-Mobilité sur AppStore et Play Store. Google nous ayant refusé l'accès à son store après de multiples tentatives, nous avons pris en main le déploiement complet pour la plateforme Android. Nous avons donc mis à disposition un fichier APK sur notre propre plateforme cloud sécurisée. Cette version est compatible avec Android 6 à Android 12. Pour les appareils IOS, Eco-Mobilité est compatible à partir d'IOS 8 jusqu'à IOS 13.

#### 4.3.3 Protocole expérimental

Notre protocole expérimental se compose de deux études complémentaires, mises en place à un an d'intervalle dû à la pandémie du COVID-19 et la limitations des déplacements de notre échantillon (voir Figure 4.3).

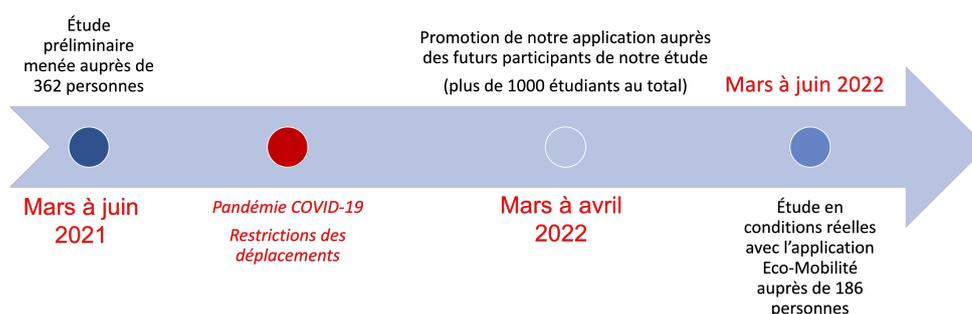


FIGURE 4.3 – Temporalité du déploiement de l'étude préliminaire et de celle en conditions réelles.

#### 4.3.3.1 Méthodologie de l'étude préliminaire

L'objectif principal de cette étude préliminaire est de tester nos premières hypothèses sur la perception de la complexité chez différents individus et ce, avant une intégration dans notre application mobile. Cette étude a été mise en place conjointement avec Ulysse Soulat, doctorant en sciences de gestion et membre du projet Ecomob. Cette étude illustre notre approche pluridisciplinaire. Les résultats obtenus serviront à enrichir à la fois la recherche en sciences de gestion et en informatique. Dans cette étude, nous avons sélectionné les différents visuels (interfaces avec différents niveaux de complexité) à présenter aux expérimentateurs. Ulysse Soulat a construit un questionnaire permettant de récolter les réponses des participants pour valider la perception de différents niveaux de complexité et leur influence sur l'adoption de la technologie. Le recrutement s'est fait par la méthode boule de neige sur le réseau social LinkedIn et par mail. Au final, 362 personnes (39% entre 15 et 20 ans, 46% entre 21 et 25 ans, 15% entre 26 et 35 ans, 49,5% de femmes) ont participé à cette étude. Les participants étaient tous volontaires et non-rémunérés. Pour assurer une validité externe de notre étude, le recrutement des participants s'est fait sur une trentaine de villes françaises de tailles différentes (par exemple Paris, Montpellier, Lille, Nantes, Bordeaux, etc.).

Après avoir visionné au hasard l'une des trois versions, les participants ont évalué leur acceptabilité de l'application en fonction de la complexité du stimulus à l'aide de 21 items (par exemple l'utilité perçue, la facilité d'utilisation perçue, l'influence sociale, les motivations hédoniques, l'habitude, ou l'intention d'utilisation) et de 7 dimensions basées sur le modèle UTAUT2 (plus fiable et plus récent) sur une échelle de 7 points (1 = « pas du tout d'accord », 7 = « tout à fait d'accord ») [Venkatesh et al., 2012], dont la moyenne a été calculée en un seul point de mesure.

Ils ont également évalué leur intention d'utiliser des moyens de transport émettant moins de CO<sub>2</sub> que la voiture sur une échelle de 4 items en 7 points (1 = « Pas du tout d'accord », 7 = « Tout à fait d'accord ») [Carrus et al., 2008], dont la moyenne a été calculée en une seule mesure. Les 4 items étaient les suivants :

- Après avoir utilisé l'application, j'aurai l'intention d'utiliser des moyens de transports plus écologiques que la voiture pour effectuer mes déplacements domicile-lieu d'étude/travail.
- Après avoir utilisé l'application, j'utiliserai des moyens de transports plus écologiques que la voiture pour effectuer mes déplacements domicile-lieu d'étude/travail.
- Après avoir utilisé l'application, j'aurai l'intention d'utiliser des moyens de transports qui émettent moins de CO<sub>2</sub> que la voiture afin d'avoir une faible empreinte carbone lors de mes déplacements domicile-lieu d'étude/travail.
- Après avoir utilisé l'application, j'utiliserai des moyens de transports qui émettent moins de CO<sub>2</sub> que la voiture afin d'avoir une faible empreinte carbone lors de mes déplacements domicile-lieu d'étude/travail.

### 4.3.3.2 Méthodologie de l'étude en conditions réelles

Notre expérience en conditions réelles suit un plan en deux étapes. Premièrement, les participants ont été invités à remplir un questionnaire avant d'installer l'application « Eco-Mobilité » permettant d'évaluer leur intention d'utiliser des modes de transport qui émettent moins de CO<sub>2</sub> que la voiture. Toutes les réponses du questionnaire ont été mesurées par des échelles de Likert à sept points. Un score a été attribué à chaque participant pour chaque période de trois jours et un score global pour l'ensemble des neuf jours. Une moyenne a été calculée en fonction du principal mode de transport enregistré par l'application (1 = marche ou vélo, 2 = bus ou train, 3 = voiture). Le score a ensuite été comparé aux données du questionnaire. Ensuite, les volontaires ont été répartis aléatoirement en trois groupes de complexité visuelle lors de leur inscription sur l'application. Après une courte présentation de l'expérience, les futurs utilisateurs ont été invités à utiliser régulièrement l'application (des rappels ont été envoyés toutes les 2 semaines) et se sont engagés à visualiser leur émission de CO<sub>2</sub> au moins une fois par jour.

Nous nous sommes concentrés sur les étudiants de La Rochelle Université. Ils ont été sollicités dans les salles de cours de l'université en mars et avril 2022 (voir Figure 4.3). Au total, 186 participants ont installé l'application sur 1000 demandes, soit un taux de pénétration de 18%. Sur ces 186 individus, 51 ont utilisé l'application pendant une période d'au moins 9 jours. Notre étude a porté sur ces 51 participants, en concentrant notre analyse sur les 9 premiers jours d'utilisation pour des raisons d'équité. Cela reflète également la difficulté de trouver des volontaires, même en effectuant des présentations de notre étude et de notre projet de recherche directement auprès des personnes. Les données recueillies par l'application Eco-Mobilité ont été stockées sur le serveur de l'université et analysées. L'échantillon réduit de 51 participants varie en termes de sexe (57% de femmes), d'âge (38 entre 15 et 20 ans, 12 entre 21 et 25 ans et 1 entre 26 et 30 ans) et d'année universitaire. Nous avons attribué de manière aléatoire les trois niveaux de complexité de l'application visuelle aux sujets mais la répartition parmi les personnes finalement retenues était la suivante : 22 sujets dans le groupe simple, 14 sujets dans le groupe modéré et 16 sujets dans le groupe complexe.

### 4.3.4 Hypothèses

Les hypothèses formulées lors de l'étude préliminaire se concentrent sur l'impact du niveau de complexité modéré par rapport à l'acceptation de la technologie (application mobile) et aux intentions comportementales, prémices à l'adoption d'un comportement durable. Les hypothèses concernant l'étude en conditions réelles avec notre application cherchent à savoir si une information de complexité modérée a un effet significatif sur l'adoption d'un comportement durable. Nous souhaitons ainsi tester si l'information transmise par notre application provoque une transformation de l'intention en un réel comportement sur le long terme (voir Figure 4.4).

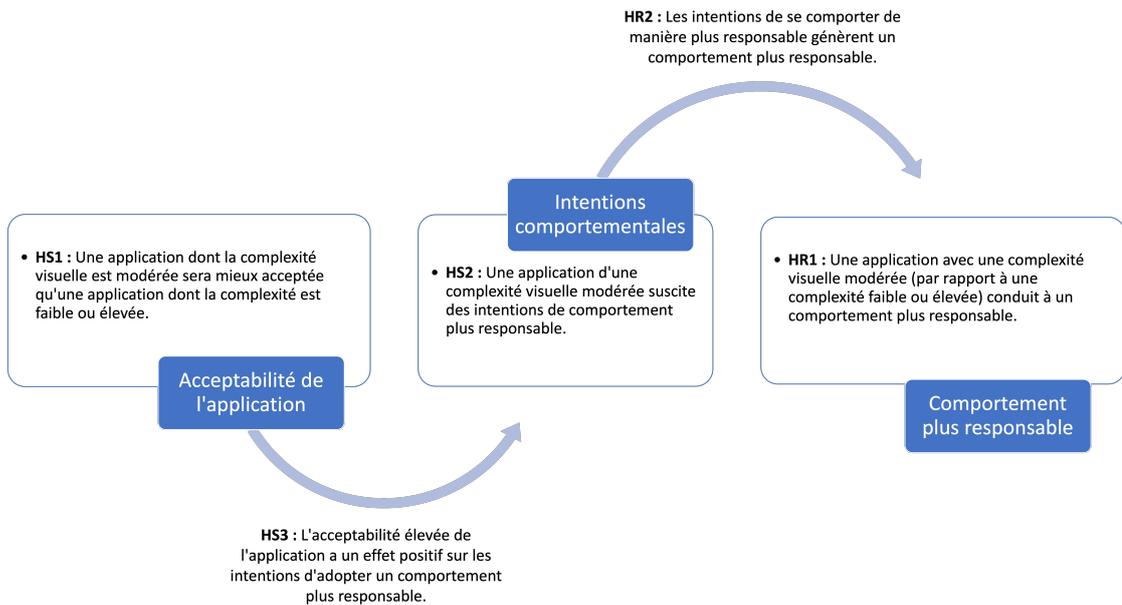


FIGURE 4.4 – Modèle d'hypothèses des différentes études.

#### 4.3.4.1 Étude préliminaire

Nous avons formulé les hypothèses suivantes sur l'étude préliminaire.

**HS1 : Une application dont la complexité visuelle est modérée sera mieux acceptée qu'une application dont la complexité est faible ou élevée.**

L'ensemble de la littérature mentionnée dans la Section 4.2 nous amène à considérer que le niveau de complexité pourrait déterminer l'impact de l'information sur l'acceptabilité d'une application de autosuivi puis sur le comportement.

**HS2 : Une application d'une complexité visuelle modérée suscite des intentions de comportement plus responsable.**

Outre l'efficacité d'une complexité modérée sur l'acceptabilité d'une technologie, nous supposons que ce niveau de complexité suscitera également des intentions comportementales plus favorables en fonction de l'objectif de l'application. Dans notre cas, il s'agit de l'intention d'utiliser un moyen de transport qui émet moins de CO<sub>2</sub> que la voiture.

**HS3 : L'acceptabilité élevée de l'application a un effet positif sur les intentions d'adopter un comportement plus responsable.**

Si la complexité de l'information produit cet effet, l'acceptabilité devrait à son tour influencer les intentions comportementales en faveur d'une plus grande responsabilité. Cette hypothèse est notamment étayée par les recherches (voir Section 4.2.1) examinant l'efficacité des outils d'auto-évaluation sur les attitudes et les intentions comportementales.

#### 4.3.4.2 Expérience en conditions réelles

Nous avons formulé les hypothèses suivantes sur l'expérience en conditions réelles.

**HR1 : Une application avec une complexité visuelle modérée (par rapport à une complexité faible ou élevée) conduit à un comportement plus responsable.**

Notre cadre théorique nous permet de supposer qu'un niveau de complexité sur un appareil numérique influence le comportement de mobilité urbaine. Nous supposons alors qu'une application à complexité modérée génère moins de comportements de mobilité urbaine émetteurs de CO<sub>2</sub>.

**HR2 : Les intentions de se comporter de manière plus responsable génèrent un comportement plus responsable.** Bien que la littérature soit au mieux exhaustive sur l'écart entre l'intention comportementale et le comportement réel, le lien persiste [Ajzen, 1991]. Notre hypothèse suppose que les intentions d'utiliser des transports moins émetteurs de CO<sub>2</sub>, grâce à l'application, se transforment en un comportement plus responsable.

### 4.3.5 Résultats

#### 4.3.5.1 Étude préliminaire

Nos résultats d'expériences ont été analysés afin de valider ou d'infirmer nos hypothèses par le biais d'une analyse statistique. Pour ce faire, nous avons utilisé le logiciel IBM SPSS. Nous avons analysé les effets d'une variable indépendante sur une variable dépendante par les analyses de variance (ANOVA). Les résultats nous indiquent si les moyennes des groupes proviennent de la même population et démontrent l'existence d'une influence entre les variables. Cet outil permet également d'effectuer des analyses statistiques de médiation. Les effets de médiation reflètent l'influence d'une variable sur une autre par l'intermédiaire d'une troisième variable. nous allons maintenant présenter les résultats de chaque hypothèse.

**HS1.** Un test ANOVA est effectué sur la variable d'acceptabilité de l'application (p-valeur=0,001). Les résultats montrent que la page contenant des stimuli modérément complexes génère une plus grande acceptabilité de l'application (moyenne=4,60, erreur-type=0,110) que les stimuli peu complexes (moyenne=3,99, erreur-type=0,125) ou très complexes (moyenne=4,18, erreur-type=0,100) (voir Table 4.1). Ceci est confirmé par un test de différence significative entre les moyennes de groupes ou test de Tukey : Complexe vs. Simple=0,184, p-valeur=0,468 ; Modéré vs. Simple=0,609, p-valeur=0,001 ; Modéré vs. Complexe=0,426, p-valeur=0,016 (voir Table 4.2, partie supérieure). Ainsi, l'hypothèse **HS1** est confirmée : les stimuli de complexité modérée génèrent une plus grande acceptabilité de l'application, par rapport aux stimuli de complexité élevée ou faible.

**HS2.** Un autre test ANOVA est réalisé sur l'intention de transport avec une p-valeur=0,080, Les résultats montrent que les stimuli modérément complexes ne génèrent pas significativement, de manière directe, une intention plus élevée de comportement

Complexité	Simple	Modérée	Complexe
Acceptabilité (Moyenne)	3,99	4,60	4,18
Erreur-type	0,125	0,110	0,100

TABLE 4.1 – Effet de la complexité de l'information sur l'acceptabilité de l'application. La première ligne correspond à la moyenne des valeurs issues de l'échelle de Likert associée. Il est à noter que les participants ont vu un seul et unique niveau de complexité. La deuxième ligne correspond à l'estimation de l'erreur type de chaque groupe.

de mobilité urbaine responsable avec l'application (moyenne=5,33, erreur-type=0,181) par rapport aux stimuli de faible complexité (moyenne=4,71, erreur-type=0,216) ou de forte complexité (moyenne=4,93, erreur-type=0,155). Ceci est confirmé par un test de différence significative de Tukey (Complexe vs. simple=0,221, p-valeur=0,659; Modéré vs. simple=0,615, p-valeur=0,071; Modéré vs. complexe=0,394, p-valeur=0,255). Néanmoins, la corrélation de Pearson a montré un effet significatif (p-valeur=0,027). Ainsi, l'hypothèse **HS2** est partiellement confirmée, il n'y a cependant pas de relation directe entre la complexité de l'information transmise et l'intention de transport.

**HS3.** Pour tester cette hypothèse, nous avons effectué une régression linéaire. Nous avons constaté que la régression de l'acceptabilité de l'application sur l'intention de prendre des transports plus responsables donnait un résultat positif (erreur-type=0,078, p-valeur=0,000) (voir Table 4.3, partie inférieure) validant ainsi notre troisième hypothèse. L'acceptabilité a un effet significatif sur l'intention comportementale de prendre des transports plus responsables. En d'autres termes, les participants perçoivent l'application d'une complexité modérée (par rapport à une complexité élevée ou faible) comme plus acceptable, ce qui influence l'intention d'utiliser un moyen de transport plus durable que la voiture.

Après avoir mené la première étude sur les intentions comportementales, il semblait approprié d'observer les effets sur le comportement dans des conditions réelles.



**À retenir :** L'objectif de cette première étude était de vérifier nos hypothèses sur la perception de la complexité et ses effets sur l'acceptabilité d'une application d'autosuiwi, ainsi que sur les intentions de transport induites par l'usage de cette dernière. Les résultats ont montré les stimuli de complexité modérée génèrent une plus grande acceptabilité de l'application. Les participants considèrent l'application d'une complexité modérée comme plus acceptable, ce qui influe sur l'intention d'utiliser un mode de transport plus durable que la voiture. Toutefois, il n'existe pas de relation directe entre la complexité de l'information au sein de l'application et l'intention de transport.

Complexité	Différence moyenne	Erreur-type	P-valeur
Modérée vs. complexe	0,426	0,153	0,016
Modérée vs. simple	0,609	0,171	0,001
Complexe vs. simple	0,184	0,156	0,468
Intergroupe	F	P-valeur	
	6,840	0,001	

TABLE 4.2 – Effet des différents niveaux de complexité de l’information les uns par rapports aux autres sur l’acceptabilité de l’application. La première colonne présente la différence entre la valeur moyenne d’un niveau de complexité en comparaison à un autre niveau. La deuxième colonne correspond à l’estimation de l’erreur-type de chaque groupe. La troisième colonne (P-valeur) indique le niveau de signification de nos données. La valeur  $F$  correspond à la valeur du test de Fisher qui correspond au ratio de la variance inter-classes sur la variance intra-classes.

Synthèse du modèle	R	R <sup>2</sup>	F	P-valeur
	0,358	0,148	6.689	0,000
Variable indépendante	B	Erreur-type	T	P-valeur
Acceptabilité	0,619	0,078	7,918	0,000

TABLE 4.3 – Effet de l’acceptabilité de l’application sur l’intention d’utiliser des moyens de transport qui émettent moins de CO<sub>2</sub> que la voiture. Les valeurs  $R$  et  $R^2$  représentent le coefficient de corrélation linéaire et le coefficient de détermination. La valeur  $F$  correspond à la valeur du test de Fisher qui représente au ratio de la variance inter-classes sur la variance intra-classes. La valeur  $B$  représente l’ordonnée à l’origine de l’équation du modèle. La valeur  $T$  ou test de Student représente la comparaison des moyennes des groupes entre complexité et acceptabilité.

#### 4.3.5.2 Étude en conditions réelles

Afin d’évaluer l’effet de la complexité visuelle sur le comportement de mobilité urbaine, le score global pour les 9 jours de l’étude est pris en compte. Une des explications possibles concernant cette période réduite d’analyse serait la lassitude des utilisateurs, ce qui entraîne une grande difficulté pour réaliser des recherches longitudinales. L’hypothèse **HR1** est testée à l’aide d’une analyse ANOVA à un facteur en considérant chaque niveau de complexité. Les résultats montrent un effet non significatif de la complexité sur le comportement de mobilité urbaine (comportement, p-valeur=0,252). Selon le score de mobilité, nous constatons que les participants ayant des visuels d’application de complexité modérée n’ont pas des comportements de mobilité urbaine plus responsables (moyenne=1,900, écart-type=0,5148) que les participants ayant des visuels d’application de complexité faible (moyenne=2,092, écart-type=0,6262) ou élevée (moyenne=2,286,

écart-type=0,6075) (voir Table 4.4). Un test de différence significative de Tukey est ensuite effectué pour confirmer ces résultats. Aucune différence n'est observée entre les groupes : Simple vs. modéré=0,1917, p-valeur=0,621 ; Simple vs. complexe=-0,194, p-valeur=0,599 ; Modéré vs. complexe=-0,3857, p-valeur=0,222 (voir Table 4.5). Notre hypothèse **HR1** est donc rejetée : le visuel de complexité modérée n'a pas conduit les participants qui l'ont utilisé à se comporter de manière plus responsable que les participants qui ont utilisé une application avec un visuel de complexité inférieure ou supérieure. Nous savons que la complexité a un impact sur les intentions comportementales [Wang et al., 2020] et qu'elle affecte le traitement de l'information de manière positive ou négative [Mosteller et al., 2014]. D'après la littérature, nous savons qu'une complexité visuelle modérée influence positivement les intentions, notre étude montre donc que dans notre contexte, le cadrage n'a pas d'effet sur le comportement responsable.

L'hypothèse **HR2** est testée par une analyse de régression. L'effet de l'intention comportementale sur un comportement de mobilité urbaine plus responsable est significatif et positif ( $B=-0,118$ , erreur-type=0,042, p-valeur=0,007), validant ainsi notre hypothèse (voir Table 4.6). L'étude montre que les intentions comportementales sont des antécédents des comportements de mobilité urbaine. Cependant, l'influence des comportements de mobilité urbaine doit être vérifiée avant l'expérience. Nous avons contrôlé l'influence des comportements de transport avant l'utilisation de l'application. L'analyse de la variable de contrôle montre un effet non significatif des habitudes de transport avant l'expérimentation à la fois sur l'effet de l'adoption de la technologie sur les intentions comportementales (erreur-type=0,4058, p-valeur=0,2147), l'effet de l'adoption de la technologie sur le comportement (erreur-type=0,1403, p-valeur=0,1484) et l'effet des intentions comportementales sur le comportement (erreur-type=0,0734, p-valeur=0,4528), confirmant ainsi notre hypothèse **HR2**.

La taille de notre échantillon de données (51 participants) et plus particulièrement la sélection de ces dernières sur une période de 9 jours doit être pris en compte dans les conclusions de ces hypothèses.

Variable dépendante	Simple	Modérée	Complexe
Comportement (Moyenne)	2,092	1,900	2,286
Écart-type ( $\sigma$ )	0,6262	0,5148	0,6075

TABLE 4.4 – Résultats de l'expérimentation en conditions réelles. Statistiques descriptives de l'effet de la complexité sur le comportement. La première ligne correspond à la moyenne des valeurs issues de l'échelle de Likert associée. Il est à noter que les participants ont vu un seul niveau de complexité. La deuxième ligne correspond à l'estimation de l'écart type de chaque groupe.

Complexité	Différence moyenne	Erreur-type	P-valeur
Simple vs. Modérée	0,1917	0,2049	0,621
Simple vs. Complexe	-0,1940	0,2001	0,599
Modérée vs. Complexe	-0,3857	0,2292	0,222
Intergroupe	F	P	
	1,417	0,252	

TABLE 4.5 – Résultats de l'expérimentation en conditions réelles. Effets de la complexité visuelle sur le comportement de mobilité urbaine. La première colonne présente la différence entre la valeur moyenne d'un niveau de complexité en comparaison à un autre niveau. La deuxième colonne correspond à l'estimation de l'erreur-type de chaque groupe. La troisième colonne P-valeur indique le niveau de signification de nos données.

Résumé du modèle	R	R <sup>2</sup>	F	P-valeur
	0,371	0,138	7,826	0,007
Variable indépendante	B	Erreur-type	T	P-valeur
Intention de comportement	-0,118	0,042	-2,798	0,007

TABLE 4.6 – Résultats de l'expérimentation en conditions réelles. Effet de l'intention comportementale sur le comportement de mobilité urbaine. Les valeurs  $R$  et  $R^2$  représentent le coefficient de corrélation linéaire et le coefficient de détermination. La valeur  $F$  correspond à la valeur du test de Fisher qui correspond au ratio de la variance inter-classes sur la variance intra-classes. La valeur  $B$  correspond à l'ordonnée à l'origine de l'équation du modèle. La valeur  $T$  ou test de Student correspond à une comparaison des moyennes des groupes entre complexité et l'intention comportementale.



**À retenir :** L'objectif de cette étude en conditions réelles était d'analyser le comportement des étudiants dans leur choix de moyen de transport dans un cadre de mobilité urbaine. Les résultats ont montré qu'une complexité visuelle modérée de l'information a un effet plus significatif sur l'intention d'adopter un comportement responsable en matière de mobilité. Cependant, la complexité de l'information n'a pas d'impact sur un éventuel comportement de mobilité plus douce. Il est nécessaire de relativiser ce résultat du fait de la réduction de notre échantillon de données, à 51 personnes sur une période de 9 jours consécutifs.

## 4.4 Conclusion

En examinant l'effet de la complexité de l'information sur l'intention et le comportement, nous contribuons à l'ensemble croissant de résultats de la littérature sur la complexité visuelle. Ce domaine de recherche tente de comprendre les effets de la complexité de l'information sur les individus [Althuizen, 2021, Wang et al., 2020] et nous apportons une meilleure compréhension des effets de la complexité de l'information sur l'intention et le changement de comportement. Nos résultats montrent qu'une complexité visuelle modérée a un effet plus significatif sur l'intention d'adopter un comportement responsable par rapport à une complexité plus faible ou plus élevée. Nous constatons cependant qu'une complexité visuelle modérée n'a pas d'effet plus significatif sur le comportement responsable qu'une complexité plus faible ou plus élevée. La complexité de l'information doit être prise en compte pour influencer sur la perception des individus. Mais en l'intégrant à un ensemble plus grand, composé de variables externes.

Le domaine de la consommation durable propose des axes de recherches cognitifs, sociaux et marketing à prendre en compte lors de la mise en place de tels dispositifs [Trudel, 2019]. L'intégration de variables environnementales (telles que la météo, les réseaux de transport public, la topographie du terrain) à une analyse permettrait en partie d'expliquer la différence entre l'intention et le comportement [Flüchter et al., 2014].

Notre expérimentation fournit des preuves supplémentaires de l'importance de l'acceptation de la technologie. Les résultats de notre étude montrent que l'acceptabilité de l'application influence positivement l'intention et le comportement en matière de transport. L'acceptation de l'application est également influencée par l'intention comportementale d'utiliser des modes de transport plus durable.

Au delà d'enrichir la littérature, notre recherche revêt un aspect applicatif fort. En se basant sur nos résultats, nous suggérons aux concepteurs de systèmes d'informations durables et persuasifs d'apporter une attention particulière aux aspects facilitant l'adoption de la technologie. Par conséquent, les concepteurs ne devraient pas seulement penser à l'information présentée, mais aussi considérer les facteurs qui peuvent encourager les utilisateurs à continuer à utiliser l'application. Cela peut passer par la ludification ou des fonctionnalités de partage et d'interactivité entre les utilisateurs basées sur les fonctionnalités des réseaux sociaux [Brauer et al., 2016].

**Contributions de ce chapitre :**

- Nous avons mené deux recherches complémentaires, une expérience avec mise en situation avec 302 participants et une expérience en situation réelle avec 189 participants sur une période de trois mois, afin de tester l'impact potentiel de la complexité de l'information sur le comportement et l'intention des personnes.
- Nous avons observé qu'une complexité visuelle modérée a un effet plus significatif sur l'intention d'adopter un comportement responsable en matière de transport que des pages d'accueil plus ou moins complexes.
- Nous constatons qu'une complexité visuelle modérée n'a cependant pas d'effet plus significatif sur le comportement de mobilité responsable que des pages d'accueil plus ou moins complexes.
- Nous avons montré l'influence positive de l'acceptation de la technologie sur les intentions et les comportements durables.



# Chapitre 5

## Conclusion

---

### Sommaire

---

<b>5.1</b>	<b>Problématiques et contributions . . . . .</b>	<b>104</b>
<b>5.2</b>	<b>Limites . . . . .</b>	<b>105</b>
<b>5.3</b>	<b>Perspectives . . . . .</b>	<b>106</b>
5.3.1	Amélioration de notre méthode de la mesure de la complexité .	106
5.3.2	Nouvelles études d'oculométrie . . . . .	107
5.3.3	Nouvelles versions de l'application Ecomobilité . . . . .	107
5.3.4	Vers une personnalisation de l'information . . . . .	109

---

Dans ce manuscrit de thèse, nous avons présenté un ensemble de travaux de recherche ayant pour objectif apporter de nouvelles solutions à trois questions centrées autour de l'information, de sa transmission et de son acceptabilité pour induire un changement de comportement en matière de mobilité durable. La spécificité de nos travaux repose sur deux piliers : notre approche pluridisciplinaire avec des chercheurs en sciences de gestion, en neurosciences et en informatique ; et un domaine d'application de nos travaux autour de la mobilité urbaine. En effet, la majorité des recherches menées sur des questions de changement de comportement plus vertueux et durable sont centrées sur une unique discipline, la synergie de plusieurs méthodologies et littératures y est alors absente. De plus, l'aide au changement de comportement est généralement centrée sur un intérêt individuel contrairement à celui, collectif, de l'écomobilité. Notre approche contribue à améliorer la communication auprès d'individus dans une volonté d'incitation au changement de comportement en travaillant spécifiquement sur les aspects suivants : la nature de l'information, la perception de l'information et la transmission de l'information.

## 5.1 Problématiques et contributions

Dans le chapitre 2, nous avons présenté notre approche de quantification de l'information suivant sa complexité algorithmique. Nous avons détaillé dans un état de l'art les principales approches de mesure de la complexité d'une information, à savoir l'entropie de Shannon, la complexité de Kolmogorov et la profondeur logique de Bennett. Cette étude des méthodes disponibles nous a permis de constater qu'elles ne répondaient pas totalement à notre problématique d'aide au changement. En effet, pour inciter à un changement de comportement, une seule information ne suffit pas. Une pluralité d'informations doit être transmise à travers plusieurs tests jusqu'à obtention d'une réponse satisfaisante de la part de l'individu (par exemple un comportement de mobilité vertueux). Pour pallier ces limitations, nous avons proposé une méthode générique, appelée Information Complexity Ranking ou ICR, permettant de mesurer et de classer n'importe quel type d'information, qu'il s'agisse d'images, de chaînes de caractères, ou autres. L'objectif de cette nouvelle méthode est de pouvoir contrôler la complexité de l'information transmise à un individu. Nous avons comparé nos résultats aux valeurs retournées par cinq algorithmes de référence dans l'évaluation de la complexité. Les résultats montrent que notre méthode atteint la meilleure performance pour certaines catégories et surpasse la majorité des algorithmes de pointe pour d'autres catégories. Pour les images comportant de nombreux éléments sémantiques, notre méthode n'est cependant pas aussi efficace que certains algorithmes de pointe. Cet outil donne une valeur précise de la complexité de chaque information transmise et, couplé à une analyse de la réponse cognitive d'une personne, il pourrait ainsi aider à individualiser automatiquement l'information et donc à améliorer l'aide au changement de comportement.

Dans le chapitre 3, nous avons présenté nos travaux en lien avec la perception et la sélection de l'information chez l'individu. L'objectif de ce chapitre était de comprendre comment l'information est susceptible d'être choisie et intégrée au processus de décision d'une personne. Cette étape est primordiale car si l'information n'est pas traitée par l'individu elle ne pourra en aucun cas être un déclencheur d'un comportement de mobilité plus vertueux. Pour débiter ce chapitre, nous avons présenté un travail bibliographique autour de la perception de l'information tels que l'attention visuelle et la perception de la complexité visuelle. Nous nous sommes également intéressés à la position de l'information pour savoir si cela influençait son intégration au sein du processus de décision. En outre, nous avons étudié le processus de décision pour comprendre comment le choix d'une information se construit. Cet état de l'art nous a permis de proposer une démarche expérimentale d'analyse du choix des individus parmi des informations de complexité visuelle différente. D'un point de vue théorique, cette étude à l'aide d'un outil d'oculométrie nous a permis de mieux comprendre la perception de la complexité visuelle d'une information. D'un point de vue applicatif, elle nous permet de savoir quel niveau de complexité visuelle de l'information est le plus plébiscité par les personnes. Les premiers enseignements de cette démarche montrent que la complexité modérée est préférée à des niveaux de complexité plus faibles ou plus élevées. De plus, la position centrale de l'information est une zone à privilégier pour maximiser le traitement de l'information et

son intégration au processus de décision des individus. Ces résultats nous ont permis de mieux intégrer l'information à notre application d'incitation au changement de mobilité et donc de maximiser l'efficacité de notre application mobile.

Dans le chapitre 4, nous nous sommes concentrés sur l'outil d'aide au changement et son acceptation chez le grand public, mais aussi sur les facteurs externes pouvant ou non conduire à des comportements plus durables. Pour cela, nous avons mobilisé la littérature sur les technologies d'autosuivi, sur les systèmes d'information durables (SI verts) et sur la consommation durable. Nous avons pu constater la polarisation de ces outils d'aide au changement autour de thématiques et d'enjeux individuels et non sur des problématiques collectives telles que la mobilité urbaine. Nous avons également intégré les résultats de notre recherche des chapitres précédents sur la complexité de l'information et son traitement pour tenter de limiter l'impact de facteurs externes (par exemple les conditions météorologiques, les infrastructures de transport, les habitudes...) sur le choix de mobilité. L'objectif de cette démarche est de faire de l'information un facteur central de ce dernier. Pour tester la validité de notre recherche en conditions réelles et pour tenter de combler le fossé entre mobilité durable et aide au changement, nous avons proposé une application mobile de suivi individuel d'émissions de CO<sub>2</sub>. Ce système d'information permet aux personnes volontaires de suivre en temps réel leur consommation de CO<sub>2</sub> en fonction de leurs déplacements. Cette application déployée auprès de 186 étudiants de notre université a permis de mettre en évidence l'influence de la complexité visuelle de l'information sur les intentions de changement de comportement mais pas sur le changement en lui-même. Ce premier résultat de recherche nous confirme l'importance de la complexité visuelle de l'information dans l'incitation au changement de comportement. À l'avenir, la perception de la complexité devra être prise en compte dans le développement de dispositifs similaires.

## 5.2 Limites

L'analyse du processus de décision pour présélectionner le niveau de complexité de l'information chez une majorité d'individus peut être améliorée. En effet, l'étude expérimentale d'oculométrie présentée dans le chapitre 3 ne nous permet pas de tirer des conclusions fortes concernant les différentes étapes du processus de décision. En analysant les résultats obtenus, nous avons remarqué certaines étapes du protocole à améliorer. Notamment, le fait que l'ordre de présentation des stimuli ne soit pas aléatoire peut entraîner des fixations ou des revisites sur un stimulus particulier du fait de sa position et non de sa complexité. L'écran de « réinitialisation » du regard entre chaque série, comporte une croix au centre du moniteur qui doit être fixée pendant une seconde, ce qui entraîne automatiquement des revisites et des fixations sur le stimulus positionné au centre sur la série suivante.

Notre application Ecomobilité présentée au chapitre 4 a été déployée exclusivement à La Rochelle Université auprès d'un public étudiant. Ces deux facteurs peuvent limiter l'impact de notre étude et renforcer l'influence de facteurs externes propres à la zone géographique ou au public ciblés. Même si la ville de La Rochelle fut pionnière dans

une offre diversifiée de mobilités douces, telles que les vélos et les voitures électriques en libre-service, tous les moyens de transport en commun ne sont pas disponibles. Un déploiement dans plus de villes en France permettrait de faire une comparaison et d'estimer plus précisément l'influence des infrastructures de transport dans les choix quotidiens de mobilité. De plus, la faible durée d'utilisation de l'application par la majorité des utilisateurs, seulement 51 d'entre eux l'ont utilisée plus de 9 jours consécutifs, montre le besoin d'améliorer l'attractivité globale de notre application. L'ouverture de notre application à un public plus diversifié avec des comportements différents permettrait d'enrichir notre analyse et de mieux comprendre l'influence de la complexité de l'information sur la mobilité urbaine. Actuellement, dans notre analyse, nous ne comparons pas les données obtenues avec d'autres sources comme les conditions météorologiques, ou encore les horaires des transports en commun. Ces données supplémentaires pourraient expliquer certains comportements de mobilité. Il convient de souligner que la pandémie de COVID-19 a retardé d'un an le déploiement initialement prévu, nous empêchant ainsi de mener une deuxième expérience.

## 5.3 Perspectives

### 5.3.1 Amélioration de notre méthode de la mesure de la complexité

Au vu des résultats présentés dans la Section 2.4, notre méthode ICR peut être améliorée dans sa détection d'informations avec un fort degré sémantique (par exemple des graphiques, des cartes, etc.). Si l'information présente un aspect sémantique, elle possède de fait une forme de structure qui devra être détectée. Un bon estimateur de la complexité organisée de l'information est la profondeur logique de Bennett. Nous suggérons donc d'intégrer ce nouvel aspect à notre méthode ICR. Cette nouvelle version reposera toujours sur la notion de similarité entre deux objets à l'aide de la distance de compression normalisée (NCD), mais la complexité de Kolmogorov sera remplacée par la profondeur logique. Une autre perspective serait de remplacer l'intégralité de la complexité de Kolmogorov dans NCD par la profondeur logique de Bennett pour avoir une homogénéité de la mesure de la complexité dans notre méthode.

Si l'on considère l'ensemble de l'information transmise dans notre application (par exemple du texte, des nombres, des icônes, etc.) comme une image, nous pouvons tester de nouvelles approches pour mesurer le contenu informationnel de cette dernière. Une piste de recherche intéressante serait d'utiliser la fréquence spatiale d'une image pour quantifier sa complexité. En effet, la fréquence spatiale permet de symboliser le niveau de détail d'une image. Si une image est composée de hautes fréquences, cela signifie qu'elle a un haut degré de détail et inversement si elle possède des basses fréquences. Cette notion de détail peut ensuite être corrélée avec la notion de complexité. Nous pourrions donc envisager d'utiliser des méthodes de détection de contours en nous basant sur les caractéristiques de couleur, de textures ou de luminance de l'image. Une évaluation de cette future métrique avec le dataset SAVOIAS ou une expérimentation avec l'oculomètre permettrait de savoir si cette approche orientée image, convient à notre contexte de

recherche.

### 5.3.2 Nouvelles études d'oculométrie

Pour pallier les limitations évoquées précédemment, nous pensons nécessaire de réaliser plusieurs nouvelles études d'oculométrie en faisant varier plusieurs paramètres tels que le nombre d'alternatives présentées, le temps alloué à chaque participant pour faire leur choix ou encore la forme de la présentation des stimuli.

Pour préciser l'influence de la position et de la complexité de chaque stimulus, la nouvelle étude devra les présenter dans un ordre aléatoire. Un nombre plus important d'alternatives devra également être proposé aux futurs participants pour analyser plus finement l'influence de la complexité. Selon la littérature présentée dans le chapitre 3, le nombre d'alternatives permettra également de confirmer ou d'infirmer la présence de certaines phases lors du processus de décision (par exemple la phase de vérification).

Un autre paramètre intéressant à faire évoluer serait le temps alloué aux participants. Actuellement, la limite de temps de trente secondes par série que nous avons imposée peut inciter les participants à précipiter leurs décisions et à éliminer différentes informations de leur processus de décision. Cette limite pourra être supprimée pour permettre une intégration de toutes les informations présentées et ainsi tester la construction du choix. Une autre expérimentation serait au contraire de proposer à chaque participant un nombre réduit d'alternatives avec un temps de présentation plus court, de dix secondes au maximum. L'objectif serait alors de travailler sur les propriétés de l'attention visuelle, en analysant les éléments qui attirent « mécaniquement » le regard.

La présentation des stimulus sous forme de ligne pourra être aussi modifiée pour mettre en évidence un travail oculomoteur plus important, ce qui traduirait plus précisément le choix de l'individu. Une solution possible est de présenter les stimulus sous forme de cercle, ce qui entraînera la suppression du biais induit par l'écran de « réinitialisation » actuel. Cette nouvelle disposition permettra également de supprimer l'influence de la position centrale.

En amont de nos études d'oculométrie, nous pourrions également implémenter le modèle d'Itti [Itti et al., 1998] pour tester les éléments de l'information les plus attractifs. Les cartes de saillance ainsi obtenues nous permettront de maximiser l'utilisation des éléments susceptibles d'attirer l'attention et d'être retenus dans le processus de décision des individus.

### 5.3.3 Nouvelles versions de l'application Ecomobilité

Au vu des résultats de notre expérimentation avec les étudiants de La Rochelle Université, nous pensons pertinent de créer une nouvelle version d'Ecomobilité possédant une fonctionnalité de personnalisation de l'affichage de l'information. Cette nouvelle fonctionnalité laisserait à l'utilisateur la liberté de choisir le niveau de complexité qu'il souhaite recevoir. L'ajustement du nombre et de la forme de l'information étant adopté par l'utilisateur, cela permettra d'analyser plus précisément l'impact de la complexité sur les

intentions comportementales. Cette fonctionnalité permettra également de tester si une complexité choisie engendre des comportements plus durables.

Compte tenu de nos recommandations dans la Section 4.4, une autre version d'Ecomobilité devra intégrer de nouvelles fonctionnalités pour maximiser l'acceptation de cette application. La première version d'Ecomobilité étant volontairement dépourvue de fonctionnalités supplémentaires, pour ne pas introduire d'éléments perturbateurs à la compréhension de l'impact de la complexité. Ces fonctionnalités comme le partage de sa consommation de CO<sub>2</sub> sur les réseaux sociaux ou encore un classement entre les utilisateurs permettront d'intégrer la notion de pression sociale, dont l'efficacité a été prouvée dans ce contexte (3ème caractéristique de la théorie unifiée de l'acceptation de la technologie [Venkatesh et al., 2012]). Une fonctionnalité de ludification ou des mécanismes de jeu pourront être intégrés pour éveiller une motivation à utiliser l'application régulièrement. Une fonctionnalité de cartographie de ses propres trajets pourrait également renforcer cette motivation. En se basant sur cette même fonctionnalité, l'application pourrait proposer à l'utilisateur des modes de transport pas ou peu carbonés en fonction de ses trajets quotidiens. Cependant, ce type de fonctionnalités devra être très encadré en particulier sur la complexité informationnelle qu'elle apportera, pour ne pas influencer sur les informations de mobilité délivrées par notre application. Ces deux nouvelles versions envisagées de notre application permettront de nous concentrer sur une première étape d'individualisation de l'information et de tester l'impact de certains facteurs sociaux sur l'adoption d'un comportement de mobilité douce.

D'un point de vue purement technique, le développement de ces deux nouvelles versions de l'application devra être à nouveau réalisé pour être déployé nativement sur les deux principales plateformes du marché du smartphone (Android et iOS). Cette nouvelle version native permettra d'améliorer les performances de géolocalisation d'Ecomobilité et de limiter les problèmes techniques dû au modèle de smartphone. Une dernière perspective pour notre application serait d'enrichir nos propres données avec des informations issues de sources de données ouvertes (opendata). L'objectif serait de réaliser une analyse plus fine des comportements des individus et d'acquérir une meilleure compréhension de leurs choix en fonction d'une situation donnée. L'utilisation et la promotion de données ouvertes permettrait de renforcer la motivation des utilisateurs en tant que contributeur d'un projet bénéficiant à tous.

#### 5.3.4 Vers une personnalisation de l'information

L'ensemble des perspectives précédemment présentées serviront à améliorer la personnalisation de l'information. À plus long terme, pour atteindre notre objectif d'incitation à un changement de comportement plus respectueux de l'environnement et pérenne, il sera indispensable d'individualiser l'information transmise. Au regard de la pluralité de comportements de mobilité, il est pertinent de communiquer des informations précises et conçues spécialement pour un individu. Ce mécanisme de personnalisation sera basé sur les données déjà récoltées, puis sera enrichi au fur à mesure avec de nouvelles données générées par de nouveaux utilisateurs. En pratique, cette tâche pourra être réalisée avec un scénario d'apprentissage par renforcement, afin d'optimiser la présentation d'informations susceptibles de conduire à un changement de moyen de transport.



## Publications issues de ces travaux de recherche

---

- Thomas Chambon, Jean-Loup Guillaume, Jeanne Lallement. Information Complexity Ranking : A New Method of Ranking Images by Algorithmic Complexity. *Entropy* 2023, 25, 439.
  
- Thomas Chambon, Ulysse Soulat, Jeanne Lallement, Jean-Loup Guillaume. The effect of visual information complexity on urban mobility intention and behavior. The 17th International Conference on Research Challenges in Information Science, RCIS 2023.
  
- Thomas Chambon, Jean-Loup Guillaume, Jeanne Lallement. Effects of information complexity on user decision making : an eye-tracking study to build a mobile application for carbon footprint estimation. **Article soumis.**



# Bibliographie

- Icek Ajzen. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2) :179–211, 1991. ISSN 0749-5978.  
Voir page 95.
- Hunt Allcott. Social norms and energy conservation. *Journal of Public Economics*, 95 (9) :1082–1095, 2011.  
Voir pages 80 et 85.
- Niek Althuisen. Revisiting berlyne’s inverted u-shape relationship between complexity and liking : The role of effort, arousal, and status in the appreciation of product design aesthetics. *Psychology & Marketing*, 38(3) :481–503, 2021.  
Voir page 100.
- C. H. Bennett. Logical depth and physical complexity. In *A Half-Century Survey on The Universal Turing Machine*, page 227–257, USA, 1988a. Oxford University Press, Inc. ISBN 0198537417.  
Voir page 45.
- C. H. Bennett. Logical depth and physical complexity. In *A Half-Century Survey on The Universal Turing Machine*, page 227–257, USA, 1988b. Oxford University Press, Inc. ISBN 0198537417.  
Voir page 30.
- C.H. Bennett, P. Gacs, Ming Li, P.M.B. Vitanyi, and W.H. Zurek. Information distance. *IEEE Transactions on Information Theory*, 44(4) :1407–1423, jul 1998.  
Voir pages 27 et 28.
- D. E. Berlyne. *Conflict, arousal, and curiosity*. McGraw-Hill Book Company, 1960. doi : 10.1037/11164-000. URL <http://content.apa.org/books/11164-000>.  
Voir pages 49 et 57.
- D. E. Berlyne. Novelty, complexity, and hedonic value. *Perception & Psychophysics*, 8 : 279–286, 9 1970. ISSN 0031-5117. doi : 10.3758/BF03212593. URL <http://link.springer.com/10.3758/BF03212593>.  
Voir page 57.

- Ali Borji and Laurent Itti. State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 :1395–1398, 12 2010. ISSN 0162-8828.  
Voir pages 5, 54 et 55.
- Benjamin Brauer, Carolin Ebermann, Björn Hildebrandt, Gerrit Remane, and Lutz Kolbe. Green by app : The contribution of mobile applications to environmental sustainability. In *Proceedings of 20th Pacific Asia Conference on Information Systems (PACIS 2016)*, 07 2016.  
Voir pages 79, 83 et 100.
- Peter A Bruck, Luvai Motiwalla, and Florian Foerster. Mobile learning with micro-content : A framework and evaluation. *BLED 2012 Proceedings*, 2012.  
Voir pages 11, 49 et 83.
- Gordon C. Bruner and Anand Kumar. Web commercials and advertising hierarchy-of-effects. *Journal of Advertising Research*, 40 :35–42, 1 2000. ISSN 0021-8499. doi : 10.2501/JAR-40-1-2-35-42. URL <http://www.journalofadvertisingresearch.com/lookup/doi/10.2501/JAR-40-1-2-35-42>.  
Voir pages 58 et 67.
- Pages Jeanne Brutel Chantal. Insee premiere. Technical Report 1835, Institut national de la statistique et des études économiques (Insee), 01 2021.  
Voir page 49.
- Giuseppe Carrus, Paola Passafaro, and Mirilia Bonnes. Emotions, habits and rational choices in ecological behaviours : The case of recycling and use of public transportation. *Journal of Environmental Psychology*, 28(1) :51–62, 2008. ISSN 0272-4944.  
Voir page 92.
- Pierre Chandon, J. Wesley Hutchinson, Eric T. Bradlow, and Scott H. Young. Does in-store marketing work ? effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing*, 73 :1–17, 11 2009. ISSN 0022-2429. doi : 10.1509/jmkg.73.6.1. URL <http://journals.sagepub.com/doi/10.1509/jmkg.73.6.1>.  
Voir pages 61 et 67.
- Min Chen. Improving website structure through reducing information overload. *Decision Support Systems*, 110 :84–94, 2018. ISSN 0167-9236.  
Voir page 58.
- Robert B. Cialdini. *Influence : the psychology of persuasion*. Quill/William Morrow, New York, 1993. ISBN 0688128165 ; 9780688128166.  
Voir page 85.
- Rudi Cilibrasi and Paul M.B. Vitányi. Clustering by compression. *IEEE Transactions on Information Theory*, 51(4) :1523–1545, 2005.

Voir pages 19, 27, 28 et 45.

Jesper Clement. Visual influence on in-store buying decisions : an eye-track experiment on the visual influence of packaging design. *Journal of Marketing Management*, 23 : 917–928, 11 2007. ISSN 0267-257X. doi : 10.1362/026725707X250395. URL <http://www.tandfonline.com/doi/abs/10.1362/026725707X250395>.

Voir pages 50 et 60.

D. Comaniciu and P. Meer. Mean shift : a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 :603–619, 5 2002. ISSN 01628828. doi : 10.1109/34.1000236. URL <http://ieeexplore.ieee.org/document/1000236/>.

Voir page 40.

Ioanna Constantiou, Alivelu Mukkamala, Mimmi Sjöklint, and Matthias Trier. Engaging with self-tracking applications : how do users respond to their performance data? *European Journal of Information Systems*, 0(0) :1–21, 2022.

Voir page 81.

Silvia Corchs, Francesca Gasparini, and Raimondo Schettini. No reference image quality classification for jpeg-distorted images. *Digital Signal Processing : A Review Journal*, 30 :86–100, 2014. ISSN 10512004. doi : 10.1016/j.dsp.2014.04.003.

Voir page 40.

Cordova Framework. Apache cordova framework. <https://cordova.apache.org/>.

Voir page 88.

Dena Cox and Anthony D. Cox. Beyond first impressions : The effects of repeated exposure on consumer liking of visually complex and simple product designs. *Journal of the Academy of Marketing Science*, 30(2) :119–130, 2002.

Voir page 59.

Muller Jorg Croutte Patricia. Baromètre du numérique - édition 2021. Technical report, Centre de Recherche pour l'Étude et l'Observation des Conditions de Vie (CRÉDOC), 07 2021.

Voir page 79.

Lindley Curtis, Carrie Edwards, Kristen L. Fraser, Sheryl Gudelsky, Jenny Holmquist, Kristin Thornton, and Kaye D. Sweetser. Adoption of social media for public relations by nonprofit organizations. *Public Relations Review*, 36 :90–92, 3 2010. ISSN 03638111. doi : 10.1016/j.pubrev.2009.10.003.

Voir page 82.

Hy Day. Evaluations of subjective complexity, pleasingness and interestingness for a series of random polygons varying in complexity. *Perception & Psychophysics*, 2 :281–286, 7 1967. ISSN 0031-5117. doi : 10.3758/BF03211042. URL <http://link.springer.com/10.3758/BF03211042>.

Voir page 57.

Agence de la transition écologique (ADEME). Impactco2. <https://impactco2.fr/>. En ligne ; dernier accès le 2 mai 2023.

Voir page 11.

Jean-Paul Delahaye. *Complexité aléatoire et complexité organisée*. Editions Quae, 2009. ISBN 9782759203208. doi : 10.3917/quae.delah.2009.01. URL <http://www.cairn.info/complexite-aleatoire-et-complexite-organisee--9782759203208.htm>.

Voir page 19.

Jean Paul Delahaye. MESURER LA COMPLEXITÉ DES OBJETS NUMÉRIQUES. *Bulletin de la société informatique de France*, pages 35–53, 2013.

Voir page 23.

Jean-Paul Delahaye and Clement Vidal. Organized Complexity : is Big History a Big Computation? *American Philosophical Association Newsletter on Philosophy and Computers*, 17(2) :49–54, sep 2016. URL <http://arxiv.org/abs/1609.07111>.

Voir pages 26 et 31.

Liqiong Deng and Marshall Scott Poole. Aesthetic design of e-commerce web pages – webpage complexity, order and preference. *Electronic Commerce Research and Applications*, 11(4) :420–440, 2012. ISSN 1567-4223.

Voir pages 58 et 59.

W. Deng, Q. Liu, H. Cheng, and Z. Qin. A malware detection framework based on kolmogorov complexity. *Journal of Computational Information Systems*, 7 :2687–2694, 08 2011.

Voir page 45.

Pierre Dubois, Paulo Albuquerque, Olivier Allais, Céline Bonnet, Patrice Bertail, Pierre Combris, Saadi Lahlou, Natalie Rigal, Bernard Ruffieux, and Pierre Chandon. Effects of front-of-pack labels on the nutritional quality of supermarket food purchases : evidence from a large-scale randomized controlled trial. *Journal of the Academy of Marketing Science*, 49 :119–138, 1 2021. ISSN 0092-0703.

Voir pages 49 et 91.

Andreas Eckhardt, Sven Laumer, and Tim Weitzel. Who influences whom? analyzing workplace referents' social influence on it adoption and non-adoption. *Journal of Information Technology*, 24(1) :11–24, 2009.

Voir page 82.

R.K. Pachauri et L.A. Meyer. Rapport de synthèse. contribution des groupes de travail i, ii et iii au cinquième rapport d'évaluation du groupe d'experts intergouvernemental sur l'évolution du climat. Technical report, GIEC, Genève, Suisse, 2014.

Voir page 9.

- Robert L. Fantz. Pattern vision in young infants. *The Psychological Record*, 8 :43–47, 4 1958. ISSN 0033-2933. doi : 10.1007/BF03393306. URL <http://link.springer.com/10.1007/BF03393306>. Utilisation de la longueur relative des fixations visuelles. Voir page 65.
- Marie Ferbus-Zanda. Kolmogorov Complexity in perspective. Part II : Classification, Information Processing and Duality. *Synthese*, page 00, 2010. URL <https://hal.archives-ouvertes.fr/hal-00525505>. 43 pages. Voir pages 27 et 28.
- Marie Ferbus-Zanda and Serge Grigorieff. Kolmogorov Complexity in perspective. Part I : Information Theory and Randomnes. *Synthese*, page 00, 2010. URL <https://hal.archives-ouvertes.fr/hal-00525504>. 40 pages. Voir pages 25 et 27.
- Kristina Flüchter, Felix Wortmann, and Elgar Fleisch. Digital commuting : The effect of social normative feedback on e-bike commuting – evidence from a field study. In *ECIS 2014 Proceedings - 22nd European Conference on Information Systems*, 04 2014. Voir pages 83 et 100.
- The Python Software Foundation. gzip module. <https://docs.python.org/3/library/gzip.html>, 2023a. [En ligne ; dernier accès le 10 Avril 2023]. Voir page 43.
- The Python Software Foundation. pstats module. <https://docs.python.org/3/library/profile.html#module-pstats>, 2023b. [En ligne ; dernier accès le 10 Avril 2023]. Voir page 43.
- The Python Software Foundation. zlib module. <https://docs.python.org/3/library/zlib.html#module-zlib>, 2023c. [En ligne ; dernier accès le 10 Avril 2023]. Voir page 43.
- Carl M. Francolini and Howard E. Egeth. On the nonautomaticity of “automatic” activation : Evidence of selective seeing. *Perception & Psychophysics*, 27 :331–342, 7 1980. ISSN 0031-5117. doi : 10.3758/BF03206122. URL <http://link.springer.com/10.3758/BF03206122>. Voir page 51.
- Jon Froehlich, Tawanna Dillahunt, Predrag Klasnja, Jennifer Mankoff, Sunny Consolvo, Beverly Harrison, and James A. Landay. Ubigreen. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1043–1052. ACM, 4 2009. ISBN 9781605582467. Voir pages 49 et 84.
- Silvia Gabrielli, Rosa Maimone, Paula Forbes, Judith Masthoff, Simon Wells, Laura Primerano, Laura Haverinen, Giancarlo Bo, and Marco Pompa. Designing motivational

- features for sustainable urban mobility. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 1461–1466. Association for Computing Machinery, 2013.  
Voir page 81.
- Gary L. Geissler, George M. Zinkhan, and Richard T. Watson. The influence of home page complexity on consumer attention, attitudes, and purchase intent. *Journal of Advertising*, 35 :69–80, 5 2006. ISSN 0091-3367.  
Voir pages 58 et 67.
- Mackenzie G. Glaholt and Eyal M. Reingold. Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, and Economics*, 4 :125–146, 5 2011. ISSN 1937321X. doi : 10.1037/a0020692.  
Voir pages 61 et 68.
- Klaus G. Grunert, Lisa E. Bolton, and Monique M. Raats. Processing and acting on nutrition labeling on food. *Transformative consumer research for personal and collective well-being*. New York : Routledge, pages 333–51, 2012.  
Voir page 34.
- Peter Grunwald and Paul Vitanyi. Shannon Information and Kolmogorov Complexity, 2004.  
Voir page 25.
- Steven A. Hillyard and Thomas F. Münte. Selective attention to color and location : An analysis with event-related brain potentials. *Perception & Psychophysics*, 36 :185–198, 3 1984. ISSN 0031-5117. doi : 10.3758/BF03202679. URL <http://link.springer.com/10.3758/BF03202679>.  
Voir page 51.
- James E. Hoffman. Spatial attention in vision. *Psychological Research*, 48 :221–229, 12 1986. ISSN 0340-0727. doi : 10.1007/BF00309086. URL <http://link.springer.com/10.1007/BF00309086>.  
Voir page 51.
- James E. Hoffman, Billie Nelson, and Michael R. Houck. The role of attentional resources in automatic detection. *Cognitive Psychology*, 15 :379–410, 7 1983. ISSN 00100285. doi : 10.1016/0010-0285(83)90013-0. URL <https://linkinghub.elsevier.com/retrieve/pii/0010028583900130>.  
Voir page 51.
- David Huffman. A Method for the Construction of Minimum-Redundancy Codes. *Proceedings of the IRE*, 40(9) :1098–1101, sep 1952. ISSN 0096-8390.  
Voir page 21.

- Glyn W. Humphreys. Feature confirmation in object perception : Feature integration theory 26 years on from the treisman bartlett lecture. *Quarterly Journal of Experimental Psychology*, 69 :1910–1940, 10 2016. ISSN 17470226. doi : 10.1080/17470218.2014.988736.  
Voir page 53.
- Ira E. Hyman, S. Matthew Boss, Breanne M. Wise, Kira E. McKenzie, and Jenna M. Caggiano. Did you see the unicycling clown? inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology*, 24 :597–607, 10 2009. ISSN 08884080. doi : 10.1002/acp.1638.  
Voir page 50.
- L. Itti. Automatic foveation for video compression using a neurobiological model of visual attention. *IEEE Transactions on Image Processing*, 13 :1304–1318, 10 2004. ISSN 1057-7149. doi : 10.1109/TIP.2004.834657. URL <http://ieeexplore.ieee.org/document/1331443/>.  
Voir page 55.
- Laurent Itti and Pierre Baldi. Bayesian surprise attracts human attention. *Vision Research*, 49 :1295–1306, 2009. ISSN 00426989.  
Voir pages 49 et 55.
- Laurent Itti and Christof Koch. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40 :1489–1506, 6 2000. ISSN 00426989.  
Voir pages 49 et 55.
- Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 :1254–1259, 1998. ISSN 01628828. doi : 10.1109/34.730558. URL <http://ieeexplore.ieee.org/document/730558/>.  
Voir pages 5, 55, 56 et 107.
- Robert J. K. Jacob. The use of eye movements in human-computer interaction techniques. *ACM Transactions on Information Systems*, 9 :152–169, 4 1991. ISSN 1046-8188. doi : 10.1145/123078.128728. URL <https://dl.acm.org/doi/10.1145/123078.128728>.  
Voir page 65.
- Daoyan Jin, Hallgeir Halvari, Natalia Maehle, and Anja H. Olafsen. Self-tracking behaviour in physical activity : a systematic review of drivers and outcomes of fitness tracking. *Behaviour & Information Technology*, 41(2) :242–261, 2022.  
Voir pages 79 et 81.
- Chantal Julia and Serge Hercberg. Nutri-score : evidence of the effective-ness of the french front-of-pack nutrition label. *Ernahrungs Umschau*, 64 :181–187, 2017.  
Voir pages 11 et 91.

- D. Kahneman and A. Treisman. *Changing views of Attention and Automaticity*. Academic Press, Inc., San Diego, CA, 1984.  
Voir page 51.
- Christof Koch and S. Ullman. Shifts in selective visual attention : towards the underlying neural circuitry. *Human neurobiology*, 4 :219–227, 1985.  
Voir pages 49, 54 et 55.
- Timo Koivumäki, Annu Ristola, and Manne Kesti. The perceptions towards mobile services : an empirical analysis of the role of use facilitators. *Personal and Ubiquitous Computing*, 12(1) :67–75, 2008. doi : 10.1007/s00779-006-0128-x. URL <https://doi.org/10.1007/s00779-006-0128-x>.  
Voir page 82.
- A. N. Kolmogorov. Three Approaches to Information. *Problemy Peredachi Informatsii*, 1(1) :3–11, 1965.  
Voir pages 19, 24 et 25.
- Ian Krajbich and Antonio Rangel. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108 :13852–13857, 8 2011. ISSN 0027-8424. doi : 10.1073/pnas.1101328108. URL <https://pnas.org/doi/full/10.1073/pnas.1101328108>.  
Voir pages 61 et 67.
- Arni Kristjansson. Simultaneous priming along multiple feature dimensions in a visual search task. *Vision Research*, 46 :2554–2570, 8 2006. ISSN 00426989. doi : 10.1016/j.visres.2006.01.015. URL <https://linkinghub.elsevier.com/retrieve/pii/S0042698906000514>.  
Voir page 49.
- Arni Kristjansson and Howard Egeth. How feature integration theory integrated cognitive psychology, neurophysiology, and psychophysics. *Attention, Perception, and Psychophysics*, 82 :7–23, 1 2020. ISSN 1943393X. doi : 10.3758/s13414-019-01803-7.  
Voir pages 5 et 52.
- Jeanne Lallement, Sylvain Dejean, Florence Euzéby, and Carole Martinez. The interaction between reputation and information search : Evidence of information avoidance and confirmation bias. *Journal of Retailing and Consumer Services*, 53 :101787, 3 2020. ISSN 09696989. doi : 10.1016/j.jretconser.2019.03.014.  
Voir page 68.
- Mostafa Langarizadeh, Malihe Sadeghi, Atefeh As’habi, Pakzad Rahmati, and Abbas Sheikhtaheri. Mobile apps for weight management in children and adolescents; an updated systematic review. *Patient Education and Counseling*, 104(9) :2181–2188, 2021.  
Voir page 11.

- Tuomas Lehto and Harri Oinas-Kukkonen. Examining the persuasive potential of web-based health behavior change support systems. *AIS Transactions on Human-Computer Interaction*, 7 :126–140, 9 2015. ISSN 19443900.  
Voir pages 11, 49 et 83.
- Ming Li and Paul Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Texts in Computer Science. Springer International Publishing, 2019. ISBN 978-3-030-11297-4.  
Voir pages 23, 25 et 27.
- Ming Li, Xin Chen, Xin Li, Bin Ma, and Paul M.B. Vitányi. The similarity metric. *IEEE Transactions on Information Theory*, 50(12) :3250–3264, 2004.  
Voir page 28.
- Huiying Liu, Shuqiang Jiang, Qingming Huang, and Changsheng Xu. A generic virtual content insertion system based on visual attention analysis. In *Proceedings of the 16th ACM international conference on Multimedia*, pages 379–388. ACM, 10 2008. ISBN 9781605583037. doi : 10.1145/1459359.1459410. URL <https://dl.acm.org/doi/10.1145/1459359.1459410>.  
Voir page 55.
- Gerald L. Lohse and Eric J. Johnson. A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes*, 68 :28–43, 10 1996. ISSN 07495978. doi : 10.1006/obhd.1996.0087. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749597896900874>.  
Voir page 66.
- Michael G. Luchs, Rebecca Walker Naylor, Julie R. Irwin, and Rajagopal Raghunathan. The sustainability liability : Potential negative effects of ethicality on product preference. *Journal of Marketing*, 74(5) :18–31, 2010.  
Voir page 85.
- Arien Mack and Irvin Rock. *Inattentional Blindness*. MIT Press, 1998.  
Voir page 50.
- Per Martin-Löf. The definition of random sequences. *Information and Control*, 9(6) : 602–619, 1966.  
Voir page 26.
- Daniel Memmert. The effects of eye movements, age, and expertise on inattentional blindness. *Consciousness and Cognition*, 15 :620–627, 9 2006. ISSN 10538100. doi : 10.1016/j.concog.2006.01.001.  
Voir page 50.
- Hiroshi Miura. py7zr module. <https://pypi.org/project/py7zr/>, 2023. [En ligne ; dernier accès le 10 Avril 2023].  
Voir page 43.

- Carey K. Morewedge and Daniel Kahneman. Associative processes in intuitive judgment. *Trends in Cognitive Sciences*, 14(10) :435–440, 2010. ISSN 1364-6613. doi : <https://doi.org/10.1016/j.tics.2010.07.004>. URL <https://www.sciencedirect.com/science/article/pii/S1364661310001713>.  
Voir page 84.
- Bruce John Morrison and Marvin J. Dainoff. Advertisement complexity and looking time. *Journal of Marketing Research*, 9 :396–400, 11 1972. ISSN 0022-2437.  
Voir page 57.
- Steven B. Most, Daniel J. Simons, Brian J. Scholl, Rachel Jimenez, Erin Clifford, and Christopher F. Chabris. How not to be seen : The contribution of similarity and selective ignoring to sustained inattentive blindness. *Psychological Science*, 12 :9–17, 1 2001. ISSN 0956-7976. doi : 10.1111/1467-9280.00303.  
Voir page 50.
- Jill Mosteller, Naveen Donthu, and Sevgin Eroglu. The fluent online shopping experience. *Journal of Business Research*, 67(11) :2486–2493, 2014. ISSN 0148-2963.  
Voir pages 58 et 98.
- Lydiane Nabec, Nathalie Guichard, Valérie Hémar-Nicolas, and Florence Durieux. The role of nutri-score front-of-pack labels on children’s food products in informing parents : An analysis of the branding effect. *Décisions Marketing*, 106(2) :11–30, 2022.  
Voir pages 49 et 91.
- Mary Jo Nissen. Accessing features and objects : Is location special. *Attention and performance XI*, pages 205–219, 1985.  
Voir page 53.
- Behrad Noudoost, Mindy H Chang, Nicholas A Steinmetz, and Tirin Moore. Top-down control of visual attention. *Current Opinion in Neurobiology*, 20 :183–190, 4 2010. ISSN 09594388. doi : 10.1016/j.conb.2010.02.003. URL <https://linkinghub.elsevier.com/retrieve/pii/S0959438810000255>.  
Voir page 49.
- World Health Organization. Who global air quality guidelines. particulate matter (pm2.5 and pm10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. executive summary. Technical report, World Health Organization, Genève, Suisse, 2021.  
Voir pages 9 et 10.
- World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health>, 2023. [En ligne ; dernier accès le 08 mai 2023].  
Voir page 9.
- Jacob L. Orquin and Simone Mueller Loose. Attention and choice : A review on eye movements in decision making. *Acta Psychologica*, 144 :190–206, 9 2013. ISSN 00016918. doi : 10.1016/j.actpsy.2013.06.003.

Voir page 59.

Jonathan W. Palmer. Web site usability, design, and performance metrics. *Information Systems Research*, 13(2) :151–167, 2002.

Voir page 58.

Rik Pieters, Michel Wedel, and Rajeev Batra. The stopping power of advertising : Measures and effects of visual complexity. *Journal of Marketing*, 74(5) :48–60, 2010.

Voir page 58.

Projet Ecomob. Projet ecomob. <https://projet-ecomob.fr/>, 2023. [En ligne ; dernier accès le 08 mai 2023].

Voir page 11.

Christelle Quero and Dominique Crié. Un modèle d'expérience consommateur et de création de valeur dans un contexte de bien-être : le cas du quantified-self dans un but de régulation pondérale. *Management & Avenir*, N° 119 :15–38, 10 2020. ISSN 1768-5958. doi : 10.3917/mav.119.0015. URL <https://www.cairn.info/revue-management-et-avenir-2020-5-page-15.htm?ref=doi>.

Voir page 83.

Philip T. Quinlan. Visual feature integration theory : Past, present, and future. *Psychological Bulletin*, 129 :643–673, 9 2003. ISSN 1939-1455. doi : 10.1037/0033-2909.129.5.643. URL <http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-2909.129.5.643>.

Voir page 53.

Keith Rayner. The 35th sir frederick bartlett lecture : Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62 :1457–1506, 8 2009. ISSN 1747-0218. doi : 10.1080/17470210902816461. URL <http://journals.sagepub.com/doi/10.1080/17470210902816461>.

Voir page 61.

D. Regan. Visual information channeling in normal and disordered vision. *Psychological Review*, 89 :407–444, 1982. ISSN 1939-1471. doi : 10.1037/0033-295X.89.4.407. URL <http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-295X.89.4.407>.

Voir page 51.

Elena Reutskaja, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel. Search dynamics in consumer choice under time pressure : An eye-tracking study. *American Economic Review*, 101 :900–926, 4 2011. ISSN 0002-8282. doi : 10.1257/aer.101.2.900. URL <https://pubs.aeaweb.org/doi/10.1257/aer.101.2.900>.

Voir pages 61, 66 et 67.

Hannah Ritchie, Max Roser, and Pablo Rosado. Energy. *Our World in Data*, 2022. <https://ourworldindata.org/energy>.

Voir pages 5 et 10.

- Brett Rosen and Ted Czotter. cProfile module. <https://docs.python.org/3/library/profile.html>, 2023. [En ligne ; dernier accès le 10 Avril 2023].  
Voir page 43.
- Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. Measuring visual clutter. *Journal of Vision*, 7 :17, 8 2007. ISSN 1534-7362. doi : 10.1167/7.2.17. URL <http://jov.arvojournals.org/article.aspx?doi=10.1167/7.2.17>.  
Voir page 40.
- Ruth Rosenholtz, Amal Dorai, and Rosalind Freeman. Do predictions of visual perception aid design ? *ACM Transactions on Applied Perception*, 8 :1–20, 1 2011. ISSN 1544-3558. doi : 10.1145/1870076.1870080. URL <https://dl.acm.org/doi/10.1145/1870076.1870080>.  
Voir page 55.
- E. J. Russo. Eye fixations can save the world : A critical evaluation and a comparison between eye fixations and other information processing methodologies. *Advances in Consumer Research*, 5 :561–570, 1978. URL <https://www.acrwebsite.org/volumes/9480/volumes/v05/NA-05>.  
Voir pages 61 et 67.
- J. Edward Russo and Barbara A. Doshier. Strategies for multiattribute binary choice. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 9 :676–696, 1983. ISSN 1939-1285. doi : 10.1037/0278-7393.9.4.676. URL <http://doi.apa.org/getdoi.cfm?doi=10.1037/0278-7393.9.4.676>.  
Voir page 66.
- J. Edward Russo and France Leclerc. An eye-fixation analysis of choice processes for consumer nondurables. *Journal of Consumer Research*, 21 :274, 9 1994. ISSN 0093-5301. doi : 10.1086/209397. URL <https://academic.oup.com/jcr/article-lookup/doi/10.1086/209397>.  
Voir pages 50, 59, 60, 61 et 67.
- J. Edward Russo and Larry D. Rosen. An eye fixation analysis of multialternative choice. *Memory and Cognition*, 3 :267–276, 5 1975. ISSN 0090-502X. doi : 10.3758/BF03212910. URL <http://link.springer.com/10.3758/BF03212910>.  
Voir pages 61 et 67.
- Elham Saraee, Mona Jalal, and Margrit Betke. Savoias : A diverse, multi-category visual complexity dataset, 2018.  
Voir pages 7, 37, 41, 43 et 45.
- P. Wesley Schultz, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vidas Griskevicius. The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5) :429–434, 2007.  
Voir pages 80 et 85.

- K. Shankari, Mohamed Amine Bouzaghrane, Samuel M. Maurer, Paul Waddell, David E. Culler, and Randy H. Katz. E-mission deployment. <https://emission.readthedocs.io/en/latest/use/deployments/>.  
Voir page 88.
- K. Shankari, Mohamed Amine Bouzaghrane, Samuel M. Maurer, Paul Waddell, David E. Culler, and Randy H. Katz. e-mission : An open-source, smartphone platform for collecting human travel data. *Transportation Research Record*, 2672(42) :1–12, 2018.  
Voir page 88.
- C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3) :379–423, jul 1948.  
Voir page 19.
- Shinsuke Shimojo, Claudiu Simion, Eiko Shimojo, and Christian Scheier. Gaze bias both reflects and influences preference. *Nature Neuroscience*, 6 :1317–1322, 12 2003. ISSN 10976256.  
Voir pages 49 et 65.
- Daniel J Simons and Christopher F Chabris. Gorillas in our midst : Sustained inattentive blindness for dynamic events. *Perception*, 28 :1059–1074, 9 1999. ISSN 0301-0066. doi : 10.1068/p281059. URL <http://journals.sagepub.com/doi/10.1068/p281059>.  
Voir page 50.
- R. J. Solomonoff. A formal theory of inductive inference. Part I and Part II. *Information and Control*, 7(1) :1–22, 1964.  
Voir page 25.
- Alexa Spence, Wouter Poortinga, and Nick Pidgeon. The psychological distance of climate change. *Risk Analysis*, 32 :957–972, 6 2012. ISSN 02724332. doi : 10.1111/j.1539-6924.2011.01695.x.  
Voir page 85.
- Craig Standing, Paul Jackson, Adela J.w. Chen, Marie Claude Boudreau, and Richard T. Watson. Information systems and ecological sustainability. *Journal of Systems and Information Technology*, 10 :186–201, 11 2008. ISSN 17588847.  
Voir page 83.
- Julie S. Stevenson, Gordon C. Bruner, and Anand Kumar. Webpage background and viewer attitudes. *Journal of Advertising Research*, 40 :29–34, 1 2000. ISSN 0021-8499.  
Voir pages 58 et 67.
- Louise Viard-Guillot Stéphane Legleye, Amandine Nougaret. Insee focus. Technical Report 259, Institut national de la statistique et des études économiques (Insee), 01 2022.  
Voir pages 14 et 79.

- Richard H Thaler and Cass R Sunstein. *Nudge : Improving decisions about health, wealth, and happiness*. Yale University Press, 2008. ISBN 978-0-300-12223-7 (Hardcover).  
Voir page 10.
- Anne Treisman. Features and objects : The fourteenth bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology Section A*, 40(2) :201–237, 1988. doi : 10.1080/02724988843000104.  
Voir page 53.
- Anne Treisman. Search, similarity, and integration of features between and within dimensions. *Journal of Experimental Psychology : Human Perception and Performance*, 17 :652–676, 1991. ISSN 1939-1277. doi : 10.1037/0096-1523.17.3.652. URL <http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-1523.17.3.652>.  
Voir page 53.
- Anne Treisman. Spreading suppression or feature integration? a reply to duncan and humphreys (1992). *Journal of Experimental Psychology : Human Perception and Performance*, 18 :589–593, 5 1992. ISSN 1939-1277. doi : 10.1037/0096-1523.18.2.589. URL <http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-1523.18.2.589>.  
Voir page 53.
- Anne Treisman and Garry Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12 :97–136, 1980. ISSN 13036521.  
Voir pages 49 et 50.
- Anne Treisman and Janet Souther. Search asymmetry : A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology : General*, 114 : 285–310, 1985. ISSN 1939-2222. doi : 10.1037/0096-3445.114.3.285. URL <http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-3445.114.3.285>.  
Voir page 53.
- Remi Trudel. Sustainable consumer behavior. *Consumer Psychology Review*, 2(1) :85–96, 2019.  
Voir pages 80, 84 et 100.
- Tokuhiro Tsukiyama, Yoshie Kondo, Katsuharu Kakuse, Shinpei Saba, Syoji Ozaki, and Kunihiro Itoh. Method and system for data compression and restoration, Apr 1986.  
Voir page 26.
- Johannes Tulusan, Thorsten Staake, and Elgar Fleisch. Providing eco-driving feedback to corporate car drivers : What impact does a smartphone application have on their fuel efficiency ? In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp '12*, page 212–215, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450312240.  
Voir pages 49 et 84.

- A. M. Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1) :230–265, 1937.  
Voir page 25.
- Venkatesh, James Y. L. Thong, and Xin Xu. Consumer acceptance and use of information technology : Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1) :157–178, 2012. ISSN 02767783.  
Voir pages 6, 81, 82, 83, 92 et 108.
- Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. User acceptance of information technology : Toward a unified view. *MIS Quarterly*, 27(3) : 425–478, 2003. ISSN 02767783.  
Voir page 81.
- Jef C. Verhoeven, Dirk Heerwegh, and Kurt De Wit. Information and communication technologies in the life of university freshmen : An analysis of change. *Computers and Education*, 55 :53–66, 8 2010. ISSN 03601315. doi : 10.1016/j.compedu.2009.12.002.  
Voir page 82.
- Hsiu-Feng Wang and Chia-Hsiung Lin. An investigation into visual complexity and aesthetic preference to facilitate the creation of more appropriate learning analytics systems for children. *Computers in Human Behavior*, 92 :706–715, 2019. ISSN 0747-5632.  
Voir pages 58 et 59.
- Qiuzhen Wang, Da Ma, Hanyue Chen, Xuhong Ye, and Qing Xu. Effects of background complexity on consumer visual processing : An eye-tracking study. *Journal of Business Research*, 111 :270–280, 2020. ISSN 0148-2963.  
Voir pages 98 et 100.
- Douglas H. Wedell and Stuart M. Senter. Looking and weighting in judgment and choice. *Organizational Behavior and Human Decision Processes*, 70 :41–64, 4 1997. ISSN 07495978. doi : 10.1006/obhd.1997.2692. URL <https://linkinghub.elsevier.com/retrieve/pii/S0749597897926923>. PAGE 45.  
Voir pages 50 et 60.
- Katherine White, Rishad Habib, and David J. Hardisty. How to shift consumer behaviors to be more sustainable : A literature review and guiding framework. *Journal of Marketing*, 83 :22–49, 5 2019. ISSN 0022-2429. doi : 10.1177/0022242919825649. URL <http://journals.sagepub.com/doi/10.1177/0022242919825649>.  
Voir page 86.
- Astrid De Witte, Joachim Hollevoet, Frédéric Dobruszkes, Michel Hubert, and Cathy Macharis. Linking modal choice to motility : A comprehensive review. *Transportation Research Part A : Policy and Practice*, 49 :329–341, 3 2013. ISSN 09658564. doi : 10.1016/j.tra.2013.01.009. URL <https://linkinghub.elsevier.com/retrieve/pii/S0965856413000165>.

Voir page 86.

Jeremy M Wolfe. Forty years after feature integration theory : An introduction to the special issue in honor of the contributions of anne treisman. *Attention, Perception, & Psychophysics*, 82 :1–6, 1 2020. ISSN 1943-3921. doi : 10.3758/s13414-019-01966-3. URL <http://link.springer.com/10.3758/s13414-019-01966-3>.

Voir page 53.

Jeremy M. Wolfe and Todd S. Horowitz. What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5 :495–501, 2004. ISSN 1471003X. doi : 10.1038/nrn1411.

Voir page 51.

Kewen Wu, Julita Vassileva, Yuxiang Zhao, Zeinab Noorian, Wesley Waldner, and Ifeoma Adaji. Complexity or simplicity? designing product pictures for advertising in online marketplaces. *Journal of Retailing and Consumer Services*, 28 :17–27, 2016. ISSN 0969-6989.

Voir page 58.

H P Young. Condorcet’s Theory of Voting. *American Political Science Review*, 82(4) : 1231–1244, 1988.

Voir page 40.

Hector Zenil. *UNE APPROCHE EXPERIMENTALE A LA THEORIE ALGORITHMIQUE DE LA COMPLEXITE*. PhD thesis, Université de Lille 1, Université de Lille 1, 6 2011.

Voir page 42.

Hector Zenil, Jean Paul Delahaye, and Cédric Gaucherel. Image characterization and classification by physical complexity. *Complexity*, 17 :26–42, 2012. ISSN 10990526. doi : 10.1002/cplx.20388.

Voir pages 26, 30, 31 et 42.



## **Effets de la complexité de l'information sur les intentions et les comportements de mobilité urbaine : Construction d'un outil d'aide au changement de comportement.**

**Résumé :** Les travaux de recherche menés dans le cadre de cette thèse ont pour objectif de contribuer à trois problématiques liées à la nature de l'information, à sa perception et à sa transmission à l'aide d'une application mobile comme outil d'aide au changement. Nos contributions sur la nature de l'information et plus particulièrement sa complexité sont destinées à faciliter la compréhension de l'information. Nous avons proposé une nouvelle méthode de classement basée sur la théorie algorithmique de l'information (complexité de Kolmogorov). Ces travaux nous ont permis de mieux comprendre l'impact de la complexité de l'information chez les individus. Nous avons également porté une attention particulière à la perception de l'information en analysant les caractéristiques exerçant une influence sur le processus de décision chez l'humain. À partir de cette analyse, nous avons proposé une étude exploratoire d'oculométrie permettant de mettre en lumière les effets de la complexité et de la position de l'information sur le choix d'un individu dans une tâche de mobilité urbaine. Notre recherche sur la création d'un outil d'aide au changement de comportement nous a mené à la conception et au développement d'une application mobile de suivi de sa propre consommation de CO<sub>2</sub>. L'objectif de cette application est de comprendre les facteurs d'acceptation de ces outils et l'influence de la complexité de l'information sur ces facteurs. Cette contribution constitue également une expérimentation en conditions réelles de l'ensemble de nos travaux dans un contexte de recherche encore peu exploité, la mobilité urbaine.

**Mots clés :** mobilité urbaine, comportement éco-responsable, théorie algorithmique de l'information, complexité de Kolmogorov, oculométrie, application mobile d'autosuiti, systèmes d'information durables.

### **Effects of information complexity on urban mobility intentions and behaviors: Construction of a tool to help change behavior.**

**Abstract:** The research work conducted in the framework of this thesis aims to contribute to three issues related to the nature of information, its perception and its transmission using a mobile application (tool for change). Our contributions on the nature of information and more particularly its complexity are intended to facilitate the understanding of information. Based on the algorithmic theory of information and more particularly the complexity of information (Kolmogorov complexity), a new classification method has been proposed. This work has allowed us to better understand the impact of information complexity on individuals. We also paid particular attention to the perception of information by analyzing the characteristics that influence the decision process in humans. From this analysis, an exploratory eye-tracking study was carried out to highlight the effects of information complexity and position on an individual's choice in an urban mobility task. Our research on the creation of a behavior change tool led us to the design and development of a mobile application to monitor one's own CO<sub>2</sub> consumption. The objective of this application is to understand the factors of acceptance of these tools and the influence of the complexity of the information on these factors. This contribution is also an experimentation in real conditions of all our work, in a research context still little exploited, the urban mobility.

**Keywords:** urban mobility, eco-responsible behavior, algorithmic information theory, Kolmogorov complexity, eye tracking, mobile self-tracking application, green information systems.

**Laboratoire Informatique, Image, Interaction  
Institut LUDI - La Rochelle Université  
Avenue Michel Crépeau  
17042 LA ROCHELLE CEDEX 1**