



HAL
open science

L'attention comme vecteur d'ajustement de la difficulté dans les jeux vidéo

Zahen Malla Osman

► **To cite this version:**

Zahen Malla Osman. L'attention comme vecteur d'ajustement de la difficulté dans les jeux vidéo. Interface homme-machine [cs.HC]. Conservatoire national des arts et métiers - CNAM, 2015. Français. NNT : 2015CNAM1031 . tel-01359625

HAL Id: tel-01359625

<https://theses.hal.science/tel-01359625>

Submitted on 2 Sep 2016

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.

École Doctorale Informatique, Télécommunications et Électronique
(EDITE - Paris)

Equipe ILJ - Groupe MIM - Laboratoire CEDRIC

THÈSE DE DOCTORAT

présentée par

Zahen Malla Osman

soutenue le : **09 décembre 2015**

pour obtenir le grade de : **Docteur du Conservatoire National des Arts et Métiers**

Discipline / Spécialité : **Informatique**

L'attention comme vecteur d'ajustement de la difficulté dans les jeux vidéo

THÈSE DIRIGÉE PAR

M. CUBAUD Pierre

Professeur - CNAM

M. DUPIRE Jérôme

Maitre de conférences - CNAM

RAPPORTEURS

M. RAMALHO Geber

Professeur - Université fédérale de Pernambuco - Brésil

M. ARCHAMBAULT Dominique

Professeur - Université Paris 8

EXAMINATEURS

M. NATKIN Stéphane

Professeur - CNAM

M. BRAITMAN Alexis

Psychiatre - Clinique gérontopsychiatrique de Rochebrune

à ma charmante femme

Pour ton amour merveilleux, ton cœur gracieux
Pour ton aide immense et ton soutien formidable
Pour ces onze ans du bonheur

à mon père, ma mère et mes frères

Pour votre soutien sans faille
Pour votre confiance et votre encouragement
Sans vous rien n'aurait été possible
Sans vous, je ne serais pas ce que je suis aujourd'hui

Remerciements

Tout d'abord, je tiens à remercier très fortement mon directeur de thèse, Pierre Cubaud, pour son soutien, ses grandes qualités scientifiques et humaines, sa patience et son encouragement à terminer ce travail de thèse.

J'exprime mes profonds remerciements à mon encadrant, Jérôme Dupire. Cette personne est la première qui a cru en moi et en mon aptitude d'intégrer le milieu de la recherche scientifique. Pour son aide immense, ses précieux conseils et sa confiance accordée depuis plusieurs années, je tiens à lui remercier du fond de cœur.

Je tiens à remercier tous ceux qui, de près ou de loin, ont contribué à la réalisation de ce travail.

Delphine Soriano, Clément Pillias, Katharine Neil, Lubna Odeh, Stéphane Natkin, Viviane Gal, Guillaume Levieux, Stéphanie Mader, Simon Chauvin, Thomas Constant, et toute l'équipe ILJ, sans oublier Alexandre Topol

Dominique Archambault et Geber Ramalho

Godefroy Clair

Laurence Battais et toute l'équipe de Carrefour Numérique² (de la cité des sciences et de l'industrie)

Résumé

Dans ce travail de thèse, nous proposons d'étudier les méthodes d'évaluation des indicateurs attentionnels ainsi que leur implication dans la gestion de la difficulté dans le contexte du jeu vidéo. La prise en compte des processus attentionnels joue un rôle important dans la construction des interfaces homme-machine. L'évaluation de l'attention pendant la phase de conception d'un jeu vidéo peut être extrêmement précieuse pour analyser les effets de choix de design. La perturbation de ces processus peut, symétriquement, servir de variable d'ajustement des difficultés rencontrées dans les challenges proposés par les game designers.

Les méthodes de mesure courantes utilisent des matériels dédiés, allant de l'oculométrie au dispositifs d'électroencéphalographie. Bien que ces types de capteur fournissent des résultats tout à fait exploitables, leur déploiement est complexe et coûteux. Il nécessite par ailleurs un environnement dédié et une expertise significative dans leur maniement. Dans un contexte de jeu sérieux et afin de répondre à certaines contraintes d'accessibilité et de facilité d'acquisition de ce type de mesure, nous avons proposé une alternative articulée autour de la surveillance des temps de réaction de l'utilisateur. L'algorithme développé permet, par une analyse des trajectoires du pointeur contrôlé par le joueur, de détecter des valeurs anormales de temps de réaction. Celles-ci peuvent alors être considérées comme des indicateurs d'une possible perte ou diminution de l'attention, temporaire ou durable.

Cette première approche des phénomènes attentionnels dans un contexte de jeu vidéo nous a amené à étudier dans quelle mesure la quantité d'information présentée à l'écran pouvait altérer les performances d'un joueur. Nous nous sommes en particulier intéressés aux effets produits par une variation du champ de vision (field of view - FOV) choisi sur la caméra virtuelle dans les jeux à la première personne. Nous avons pu déterminer expérimentalement des valeurs limites du FOV pour lesquelles les performances de joueurs,

même très expérimentés, diminuaient.

Dans la suite de ce travail, nous nous sommes intéressés aux nouveaux usages liés à la prolifération des dispositifs mobiles (téléphones, tablettes). Nous avons donc étudié dans quelle mesure la distribution de l'attention d'un utilisateur sur plusieurs dispositifs pour réaliser une simple tâche pouvait modifier les performances comparativement à un dispositif unique. Cette étude montre que l'utilisation de deux dispositifs en parallèle (vs. un seul) provoque une augmentation significative de la difficulté. Ces données originales pourront servir les réflexions autour de la gestion de la difficulté par les game designers dans les jeux destinés à ce type de dispositif multiple (type Wii-U).

Mots clés : Attention, difficulté, performance, jeux vidéo, pointage, caméra virtuelle, écrans multiples.

Abstract

In this thesis, we study attention evaluation methods and their application to the management of difficulty within the context of video games. The consideration of attentional processes plays an important role in the development of human-computer interfaces. The disruption of these attentional processes can be used symmetrically as a variable for adjusting game difficulty.

Our first approach was to measure user reaction time, which we did in the context of a serious game. By analyzing the trajectories of a mouse cursor controlled by the player, our approach allows us to detect abnormal reaction time values.

This led us to consider how the amount of information displayed on screen could alter player performance. Our second approach, therefore, was to measure the effects produced by variations in the field of view of the virtual camera in a 3D first-person game.

Finally, we studied how the distribution of user attention across multiple devices for the execution of a simple task could affect performance relative to the use of a single device. This study showed that the use of two devices in parallel causes a significant increase in difficulty.

Keywords : Attention, difficulty, performance, video games, pointing, virtual camera, multiple screens.

Table des matières

1	Introduction	21
1.1	Contexte et motivation	21
1.2	Problématique et objectif	22
1.3	Organisation du manuscrit	23
2	État de l’art	25
2.1	Qu’est ce que l’attention ?	25
2.2	Testes d’évaluation de l’attention	30
2.3	L’attention dans les jeux vidéo	34
2.4	L’évaluation de l’attention en utilisant des capteurs spécifiques	38
2.5	Modéliser la performance de l’utilisateur à partir de son temps de réaction	40
2.5.1	Le temps de réaction	40
2.5.2	Les caractéristiques des mouvements dans les tâches de pointage	44
2.5.3	Synthèse	47
2.6	De l’oculométrie à la camera virtuelle	48
2.6.1	Les comportements oculaires	49
2.6.2	L’application de l’oculométrie en interaction homme machine et en jeux vidéo	52
2.6.3	Des mouvement de l’œil aux mouvements de la caméra	53
2.6.4	La tâche de recherche comme une activité de performance	54

TABLE DES MATIÈRES

2.6.5	Synthèse	56
2.7	L'interaction en présence d'écrans multiples	56
2.7.1	L'impact d'utiliser plusieurs écrans sur l'attention	57
2.7.2	Classification des types d'interaction avec les jeux vidéo	63
2.7.3	Synthèse	68
3	Relation attention-difficulté dans un serious game	71
3.1	Contexte de l'étude	71
3.2	Approche géométrique des phénomènes attentionnels	74
3.2.1	Construction des profils utilisateurs	74
3.2.2	Estimation du temps de réaction par analyse géométrique	76
3.2.2.1	Mesure d'une référence contextualisée	80
3.2.2.2	Valeurs limites de la tolérance	82
3.2.3	Discussion	82
3.3	Conclusion	83
4	Gestion de la difficulté par les outils d'interface logicielle	87
4.1	Contexte de l'étude	87
4.2	Le modèle comportemental de la caméra virtuelle	90
4.2.1	Variables descriptives	90
4.2.2	Le champ de vision de la caméra virtuelle	91
4.2.3	L'environnement virtuel 3D	92
4.2.4	Présentation et analyse des résultats	94
4.2.5	Discussion	99
4.3	Conclusion	100
5	Gestion de la difficulté par les outils d'interface matérielle	103

TABLE DES MATIÈRES

5.1	Contexte de l'étude	103
5.2	Hypothèse de l'étude	105
5.3	Expériences	106
5.3.1	Test préliminaire	106
5.3.2	Procédure	108
5.3.3	Présentation et analyse des résultats	112
5.3.3.1	Vocabulaires descriptives	112
5.3.3.2	Visualisation des données	113
5.3.3.3	Histogrammes des temps de réaction	114
5.3.3.4	Comparaison selon la difficulté de la tâche	114
5.3.3.5	Comparaison selon la configuration des écrans	116
5.3.3.6	Comparaison selon l'ordonnancement des stimuli	118
5.3.3.7	Synthèse générale	119
5.3.4	Discussion	123
5.4	Conclusion	126
6	Perspectives et conclusion	127
6.1	Relation attention-difficulté dans un serious game	127
6.2	Gestion de la difficulté par les outils d'interface logicielle	128
6.3	Gestion de la difficulté par les outils d'interface matérielle	129
6.4	Conclusion	130
	Bibliographie	131

TABLE DES MATIÈRES

Liste des tableaux

2.1	Classification des types d'interaction avec les jeux vidéo	68
3.1	Expérience de <i>pointage complexe</i> - La moyenne des moyennes de temps réaction (en ms) pour tous les sujets	79
4.1	Les mesures extraites à partir de la camera virtuelle	91
4.2	Contrôler la camera virtuelle en utilisant le clavier et la souris	94
4.3	Résultats de l'expérimentation : les données numériques présentées dans les colonnes correspondants aux différents angles de champ de vision représentent les logarithmes des moyennes des valeurs mesurés pour chaque sujet, le dernier colonne (p) représente les résultats de l'analyse de variance (ANOVA) .	95
5.1	Statistiques descriptives pour l'ensemble des expériences en fonction de la difficulté	115
5.2	Statistiques descriptives pour l'ensemble des expériences en fonction de la configuration	117
5.3	Statistiques descriptives pour l'ensemble des expériences en fonction de l'ordonnement des stimuli	118

LISTE DES TABLEAUX

Table des figures

2.1	Le corridor de Sundstedt et al., 2005.	28
2.2	Les déplacement oculaires sur une image, (Yarbus, 1967).	28
2.3	L'effet de Stroop, 1935.	30
2.4	Le test D2 de Brickenkamp, 1966.	30
2.5	Le test de Treisman, 1982.	31
2.6	L'expérimentation de Wolfe, 1998.	32
2.7	Un jeu vidéo adaptatif avec un système de décision (Andrade et al., 2006). . .	36
2.8	La procédure de triage dans le jeu Code Red Triage (Spek et al., 2010). . . .	37
2.9	Capteur électroencéphalographie	38
2.10	L'environnement de Mikropoulos, 2001.	39
2.11	Schéma du dispositif expérimental (Gueye et al., 2002) : TR est le temps de réaction et TM est le temps de mouvement	41
2.12	L'expérimentation de Cinaz et al., 2011.	42
2.13	Trajectoire tracée lors d'une tâche de pointage (Hourcade et al., 2008). . . .	45
2.14	La décomposition d'un profil de vitesse (Wonner et al., 2011).	46
2.15	Illustration de relations structurelles entre chaque segments sur une trajec- toire réalisée lors d'une activité de dessin (Renau et al., 2011).	46
2.16	Analyse de regard lors d'une tâche de lecture en utilisant une oculométrie (Kunze, 2013).	51

TABLE DES FIGURES

2.17	Un environnement virtuel 3D avec une tâche de recherche, Hillaire et al., 2010	55
2.18	Le sous-marin virtuel (environnement dynamique, en haut) et une galerie virtuelle (environnement statique, en bas) (Lee et al., 2007).	55
2.19	<i>Gazemarks</i> (Kern et al., 2010)	58
2.20	La représentation de trois techniques d'interaction de Khan et al., 2005	59
2.21	L'installation de Brown et al., 2014.	60
2.22	Les trois types d'interaction de Bobeth et al., 2014.	61
2.23	Les heat maps de l'attention visuelle pour les neuf configurations d'écrans, Vatavu et al., 2014	61
2.24	Les prototypes de jeux - Emmirch et al., 2014 : <i>Catch a Thief</i> (a), <i>Data Theft Algorithm</i> (b), <i>The Mole Rush</i> (c).	62
2.25	Différents types de contrôleurs des jeux vidéo	65
3.1	Captures d'écran du <i>Village aux Oiseaux</i> : Le château (a), la carte générale (b), photos en jeu (c, d).	72
3.2	Le mécanisme de protocole de pointage dynamique	77
3.3	Algorithme d'analyse géométrique	78
3.4	Expérience de <i>pointage complexe</i> : Les distributions des moyennes de temps de réaction pour tous les sujets	79
3.5	Les TR(s) obtenus dans l'expérimentation de <i>pointage complexe</i> et les TR(s) obtenus dans l'expérimentation de <i>référence-bis</i> , respectivement pour chaque utilisateur.	81
4.1	Exemples de jeux vidéo : (a) <i>Half Life</i> , (b) <i>Assassins Creed</i> .	88
4.2	Différence entre le comportement oculaire et celui de la caméra virtuelle : (a) capture d'écran de <i>Halo-2</i> , (b) trajectoires du regard via un système d'oculométrie, (c) caméra virtuelle son cône de vision.	89

TABLE DES FIGURES

4.3	Visualisation d'une scène en utilisant deux différentes tailles de FOV : (a) secteur large, (b) secteur réduit	91
4.4	Notre environnement virtuel 3D	92
4.5	Quatre différents points de vue de l'environnement pour une même position de la caméra virtuelle : (a) FOV de 10°, (b) FOV de 30°, (c) FOV de 80°, (d) FOV de 110°.	94
4.6	La distribution des mesures en fonction du FOV.	96
4.7	Comportements des joueurs	97
4.8	Comportements des non joueurs	98
4.9	La durée totale TD de chaque tâche, pour les joueurs (VGP) et les non-joueurs (NVGP)	100
5.1	La <i>Wii U</i> de Nintendo (a), le <i>Xbox SmartGlass</i> de Microsoft (b).	104
5.2	Configurations Mono-écran fixe - Test préliminaire	107
5.3	Configurations Mono-écran mobile - Test préliminaire	107
5.4	Configurations Multi-écran fixe - Test préliminaire	107
5.5	Configurations Multi-écran mobile - Test préliminaire	108
5.6	Configurations Multi-écran fixe et mobile - Test préliminaire	108
5.7	L'épreuve Go/No-Go : 2 stimuli dont 1 cible (a), 5 stimuli dont 2 cibles (b).	109
5.8	Les différentes configurations d'écrans durant la phase d'expérimentation	110
5.9	Design expérimental	110
5.10	L'installation de l'expérimentation à la cité des sciences et de l'industrie - Carrefour Numérique ²	111
5.11	Spécification des comportements de l'utilisateur face aux stimuli présentés	113
5.12	Histogrammes des temps de réaction pour l'ensemble des expériences	114
5.13	Relation entre la difficulté et le temps de réaction	115
5.14	Relation entre la difficulté et les omissions	116

TABLE DES FIGURES

5.15	Relation entre la difficulté et les fausses	116
5.16	Distribution des TR en fonction de la configuration	117
5.17	Distribution des fausses en fonction de la configuration	118
5.18	Distribution des TR en fonction de l'ordonnancement des stimuli	119
5.19	Distribution des TR moyens et régressions linéaires pour chaque configura- tion d'écran dans le cas facile (gauche) et difficile (droite).	120
5.20	Distribution du nombre d'erreur pour chaque configuration d'écran dans le cas facile (gauche) et difficile (droite).	121
5.21	Distribution des TR moyens par type d'ordonnancement des stimuli pour les conditions facile (gauche) et difficile (droite)	122
5.22	Distribution du nombre d'erreur par type d'ordonnancement des stimuli pour les conditions facile (gauche) et difficile (droite)	122
5.23	Distribution des TR moyens en fonction du nombre d'écrans	123
5.24	Distribution des TR moyens en fonction de la difficulté	124
5.25	Distribution des TR moyens en fonction de l'ordonnancement de l'information	125

Chapitre 1

Introduction

1.1 Contexte et motivation

Aujourd'hui, les jeux vidéo dominant notre société et occupent une part importante de nos activités de divertissement. Ces formes de loisir n'ont pas seulement dévoré notre temps libre, mais elles ont également impacté nos relations socioculturelles et influencé nos interactions quotidiennes. Ces outils de plaisir ont ouvert de nouveaux horizons à la créativité, au delà de leur fonction ludique, pouvant aujourd'hui être utilisés pour intégrer d'autres services, comme la réhabilitation cognitive ou l'apprentissage.

Nous observons actuellement que les jeux vidéo ont déjà ouvert la voie vers une émergence croissante de nouveaux travaux scientifiques, afin d'étudier les comportements de joueurs et leur implications dans la conception de ces applications fortement interactives. La nécessité d'avoir des outils d'évaluation dynamique de l'attention et de la performance de joueurs, offrent la possibilité de fournir des bases de validation concrètes, qui peuvent s'intégrer dans les phases de la conception de jeux vidéo.

Pour cela, nous avons besoin de comprendre les stratégies employées par les joueurs dans la réalisation de leurs tâches. Analyser ces stratégies, peut fournir des indicateurs pour alimenter des éventuelles conceptions des jeux vidéo, plus innovants et plus amusants. La difficulté d'un jeu vidéo est une variable majeure pour la conception d'expériences vidéo ludiques. Il semble nécessaire de fournir des outils d'évaluation dynamique de l'attention, pour déterminer la performance de joueur. Cela peut être impliqué afin de gérer plus précisément la relation de la difficulté avec l'expérience de joueur.

Le contexte de cette thèse, se décline suivant trois axes, les jeux vidéo à visée pédagogique ou thérapeutique (serious game), les jeux vidéo 3D en vue subjective et les jeux vidéo utilisant une interface répartie sur plusieurs écrans.

1.2 Problématique et objectif

L'attention est une fonction cognitive complexe qui a été largement étudié dans plusieurs domaines comme la psychologie et la neuropsychologie. Sa complexité est due au fait qu'elle couvre un ensemble de comportements et d'activités mentales. L'analyse des processus attentionnels est complexe et une simple tâche du pointage, au sein de laquelle nous demandons au sujet de répondre le plus vite possible à un stimulus projeté sur un écran, implique au moins cinq actions où le sujet doit être capable (1) d'orienter son attention vers la source de stimulation, (2) d'identifier le stimulus et de le reconnaître comme pertinent, (3) de décider d'y réagir, (4) de sélectionner la réponse appropriée et (5) d'activer le programme moteur correspondant [MS04].

Dans un contexte de jeux vidéo, les paramètres de l'attention peuvent donner des indicateurs sur le comportement de l'utilisateur, son état de fatigue, sa motivation, sa rapidité et sa performance. Ils peuvent être mesurés de plusieurs façons : d'un point de vue pratique, nous pouvons fournir des outils non-intrusifs, facilement conçus afin d'être intégrés dans des expériences de jeux vidéo pour évaluer dynamiquement la performance des joueurs ; D'un point de vue théorique, ces outils d'évaluation peuvent répondre aux questions liées à la conception d'un jeu, en affinant la relation entre la difficulté du jeu d'une part, et la performance d'un joueur d'autre part.

En outre, les systèmes à base de capteurs (comme l'oculométrie ou l'électroencéphalographie) ont été largement utilisés pour mesurer l'attention depuis de nombreuses années. Ces systèmes ont pu analyser les comportement des utilisateurs, afin d'évaluer leur attention dans plusieurs domaines. Ils ont également fourni un très bon moyen d'évaluation de la conception d'applications interactives. Par contre, le déploiement de ce type de capteur, peut être complexe et coûteux. Ce qui rend leur utilisation difficile de notre point de vue.

Au cœur de ce travail de thèse se trouvent trois expérimentations articulées autour

de la relation entre l'attention et la difficulté, de la gestion de la difficulté par les outils d'interface logicielle et de la gestion de la difficulté par les outils d'interface matérielle. Ces expérimentations servent à produire de nouvelles connaissances sur les rapports entre l'attention et la performance dans le contexte des jeux vidéo.

1.3 Organisation du manuscrit

Nous débutons cette thèse par un état de l'art, afin de définir l'attention et de décrire les méthodes d'évaluation couramment utilisées, suivant les trois axes principales de thèse. Cette première étape permettra d'offrir une synthèse de différents travaux réalisés dans ce contexte afin de construire ensuite nos propres expériences.

Au chapitre 3, nous présentons une première démarche qui a pour objectif d'intégrer des indicateurs attentionnels dans la gestion d'un serious game appelé *Le Village aux Oiseaux*. La motivation principale de cette étude provient de l'importance d'adapter dynamiquement des scénarios pour les compétences de chaque joueur. La solution que nous avons développée est une méthode non-invasive qui évalue l'attention dans une tâche de pointage. Pour cela, nous proposons un algorithme qui permet d'analyser les trajectoires du pointeur contrôlé par le joueur.

Nous abordons au chapitre 4, la question de l'impact de certains choix de design d'un environnement virtuel sur la performance de l'utilisateur qui y évolue. Pour évaluer cet impact, notre démarche consiste à transposer les mesures habituellement appliquées à l'œil en oculométrie vers la caméra virtuelle de cet environnement.

Au chapitre 5, nous présentons une approche qui a pour objectif d'étudier les effets de la répartition de l'attention entre différents dispositifs matériels. La principale motivation de cette étude, est de comprendre comment la présence de dispositifs à affichages multiples peut perturber l'utilisateur. Pour cela, nous proposons d'évaluer les performances d'un sujet qui manipule une interface répartie sur plusieurs écrans.

Nous présentons enfin les perspectives et les travaux futurs envisagés suite à cette thèse.

1.3. ORGANISATION DU MANUSCRIT

Chapitre 2

État de l'art

2.1 Qu'est ce que l'attention ?

Dans notre vie quotidienne, nous faisons référence à la fonction d'attention pour exprimer plusieurs actions. Par exemple, le mot *Attention!* est utilisé pour avertir une personne d'un danger imminent. L'expression *faire attention* est utilisée pour révéler plusieurs actions comprenant la fixation de l'esprit sur quelque chose comme regarder quelque chose (fixer, surveiller) ou tenir compte de (prêter attention à quelque chose).

Le concept d'attention est beaucoup plus ancien que l'étude académique de l'informatique, de la psychologie ou de la neuropsychologie. Dès 400 Av. J-C, les orateurs Grecs (les conférenciers Grecs) ont cherché à améliorer leurs compétences en considérant que l'attention est une condition préalable de la mémoire, par exemple :

« Si vous payez attention, le jugement va mieux percevoir les choses qui se passent à travers lui. »

Frances A. Yates [Yat66] (1966)

Cette citation suggère que l'attention est un processus contrôlé consciemment, qui améliore la perception.

L'étude de l'attention dans le contexte de l'interaction homme-machine n'est pas nouvelle. En septembre 2015, la bibliothèque d'ACM¹, a montré plus de 228 000 résultats qui intègrent le terme de l'attention. Cela relève l'importance de la fonction d'attention et sa

1. Association for Computing Machinery - <http://dl.acm.org>

2.1. QU'EST CE QUE L'ATTENTION ?

variabilité d'être appliquée dans ce contexte, afin de trouver des solutions pour fournir une meilleure conception et une excellent ergonomie de l'interface homme-machine.

Définir l'attention, ses composantes, son rôle et ses interactions est loin d'être simple. Nous ne pouvons pas proposer une seule définition de l'attention. Les premières définitions de l'attention sont apparues à la fin de XIX siècle. Selon *William James* :

« L'attention est la prise de possession par l'esprit, sous une forme claire et vive, d'un objet ou d'une suite e pensées parmi plusieurs qui semblent possibles. La focalisation, la concentration, de la conscience sont de son essence. Elle implique le retrait de certains objets afin de traiter plus efficacement les autres. »

William James [[Jam81](#)] (1881)

Malgré que la définition proposée par *William James* a insisté sur le fait que l'attention est une fonction qui repose sur les aspects de sélectivité, elle implique aussi d'autres notions de l'attention, telles que l'intensité et la perception.

En effet, *William James* a divisé l'attention en deux types : l'attention sélective et l'attention soutenue. Selon lui, l'attention sélective exprime la focalisation de l'esprit sur un objet, un événement ou une idée avec l'exclusion d'autres. Tandis que l'attention soutenue sera appliquée sur un objet, un événement ou une idée lorsque nous maintenons celle-ci pour une durée de temps suffisante, comme dans la lecture ou l'apprentissage.

Van Zommeren et Brouwer [[vZB94](#)] ont classifié l'attention en deux dimensions : l'intensité, qui comprend les aspects d'alerte et l'attention soutenue ; la sélectivité, qui comprend l'attention focalisée et l'attention divisée.

Selon cette classification, la sélectivité intervient donc dans deux situations : (1) situation d'attention focalisée, qui correspond à la capacité à concentrer son attention sur un stimulus prédéterminé (la cible) et à ignorer les autres (rejetés et inhibés), afin d'éviter toute interférence ; (2) situation d'attention divisée, lorsqu'il y a plusieurs cibles ou tâches à traiter. Cela correspond à la capacité à partager ses ressources attentionnelles entre plusieurs stimuli pertinents présentés simultanément. C'est donc ici qu'intervient la capacité à effectuer deux choses en même temps.

En effet, ces définitions de l'attention ont ouvert des nouvelles horizons pour une explo-

2.1. QU'EST CE QUE L'ATTENTION ?

ration immense aux travaux ultérieurs, concernant les aspects de sélectivité de l'attention :

« *L'attention est une référence très signifiante pour signaler la présence d'une information importante et la nécessité de se focaliser sur aspect particulier de l'objectif qu'il s'agit d'atteindre.* »

Bernard Cadet [Cad98] (1998)

« *L'attention est un processus cognitif de se concentrer sur un aspect de l'environnement tout en ignorant d'autres choses.* »

John R. Anderson [And05] (2005)

« *L'attention sélective décrit le mécanisme qui consiste à donner la priorité au traitement d'une catégorie de stimuli par rapport à une autre.* »

Thierry Meulemans et Xavier Seron [MS04] (2014)

En conséquence, dans l'état actuel de nos connaissances, la sélectivité prend en compte le nombre d'informations qu'un sujet peut considérer, par exemple, il/elle peut sélectionner un stimulus pertinent, parmi les stimuli non pertinents.

Dans cette thèse, nous nous intéressons à la modalité visuelle de l'attention. Celle-ci décrit en effet la capacité d'un système de vision, qu'il soit humain ou artificiel, à sélectionner rapidement les informations les plus pertinentes de l'environnement dans lequel il opère [Oue03].

En effet, la perception d'un stimulus peut être soit active, soit passive. Dans le premier cas, l'attention est portée volontairement sur le stimuli, et dans le second cas, c'est le stimuli qui attire l'attention automatiquement (involontairement).

Il existe donc deux mécanismes généraux de l'attention visuelle sélective : le volontaire *top-down*, qui est contrôlée par nos attentes et l'involontaire *bottom-up* qui dépend du milieu extérieur [CEY04].

Dans le mécanisme *bottom-up*, le traitement de l'attention visuelle se fait automatiquement sans aucun contrôle volontaire [IKN98] [CEY04]. Les caractéristiques *bottom-up* influant sur l'attention visuelle sont : le contraste, la taille, la forme, la couleur, la luminosité, l'orientation, les brodeurs, et les mouvements [SDL⁺05] (Fig. 2.1).

2.1. QU'EST CE QUE L'ATTENTION ?



FIGURE 2.1 – Le corridor de Sundstedt et al., 2005.

En revanche, le mécanisme *top-down* se concentre sur l'objectif de l'observateur et il dépend de la tâche. C'est lui qui intervient principalement lorsque nous cherchons un objet dans un environnement [CEY04]. Par exemple, lorsque nous demandons à un observateur de regarder une image pendant que ses mouvements oculaires seront enregistrés, nous trouvons que l'œil de l'observateur se concentre beaucoup plus sur certains éléments de l'image, tandis que les autres éléments recevaient peu d'attention. Le déplacement de son regard ne se fait pas donc de manière aléatoire sur une image, mais il suit les éléments les plus pertinents [Yar67] (Fig. 2.2).



FIGURE 2.2 – Les déplacements oculaires sur une image, (Yarbus, 1967).

2.1. QU'EST CE QUE L'ATTENTION ?

Ces deux mécanismes de l'attention visuelle sont donc étudiés dans ce travail de thèse.

En outre, il existe plusieurs tests d'évaluation de l'attention, offrant la possibilité de souligner non seulement les aspects d'attention sélective, mais aussi l'attention divisée, l'alerte, ou l'attention soutenue. Ces tests sont caractérisés selon le type d'attention à évaluer :

- Tests d'évaluation de l'attention sélective :
 - *Test des cloches* [GDJ89]
 - *Digit Cancellation Test* [SBPS95]
 - *Go-No/Go* [ZF94] [LP07]
 - *Stroop Effect* [Str35].
- Tests d'évaluation de l'attention divisé :
 - *Double tâche Baddeley* [Bad97]
 - *Épreuve en double tâche* [LP07]
 - *Épreuve d'attention divisée* [ZF94].
- Tests d'évaluation de l'attention soutenue² :
 - *Brief test of attention* [SBB96]
 - *Batterie d'Attention William Lennox - BAWL* [LP07]
 - *Test d'Evaluation de l'Attention - TEA* [ZF94].

Nous avons cherché à étudier les différents types de tests qui peuvent s'occuper d'évaluer l'attention sélective. Nous avons observé que les tests (les plus communs) reposants sur la modalité visuelle, consistent généralement à effectuer des tâches de recherche. Le sujet est donc invité à repérer des stimuli-cibles parmi un ensemble de stimuli-non-cibles. Les mesures utilisées correspondent le plus souvent au temps de réaction et au taux d'erreurs. Nous présentons donc par la suite quelques exemples de ces tests.

2. "L'attention soutenue réclame un effort volontaire de concentration pour maintenir d'une part l'alerte sur une durée importante, afin de préserver la stabilité des performances malgré la routine des stimuli, et d'autre part la focalisation en inhibant les stimuli non-cibles" [vZB94].

2.2 Testes d'évaluation de l'attention

John Ridley Stroop [Str35], a évalué l'attention sélective ou les capacités d'inhibition. Dans son test, les sujets devaient lire des mots colorés, alors que leur tâche consistait à dénommer la couleur des mot (Fig. 2.3). Selon *Stroop*, le temps nécessaire pour identifier la couleur d'un mot est plus long lorsque la couleur d'un mot est différent (le mot "bleu" écrit en rouge) que lorsque la couleur est identique (le mot "rouge" écrit en rouge).

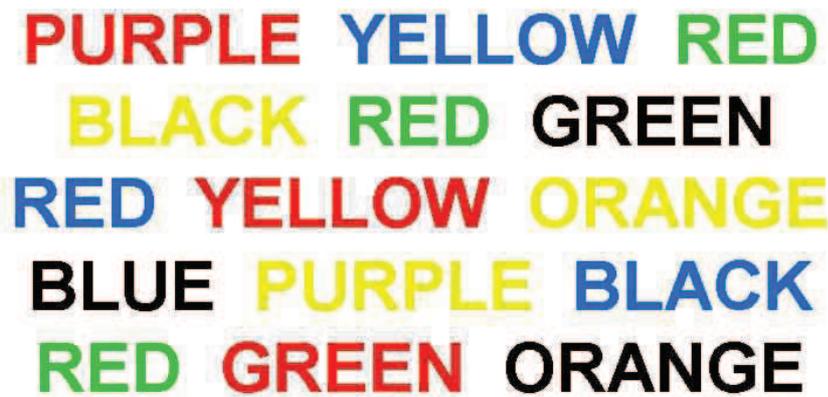


FIGURE 2.3 – L'effet de Stroop, 1935.

Rolf Brickenkamp [Bri66] a proposé aussi un test d'évaluation de l'attention sélective. Dans son test, le sujet doit rayer, le plus rapidement possible, des lettres "d", lorsqu'elles sont associées à deux apostrophes placées soit l'une au-dessus et l'autre en dessous, soit les deux au-dessus ou les deux en dessous de la lettre (Fig. 2.4). Ce test permet non seulement d'évaluer l'attention sélective, mais aussi d'analyser la performance des sujets en terme de rapidité.

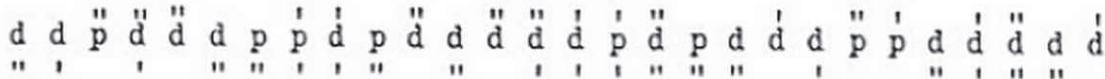


FIGURE 2.4 – Le test D2 de Brickenkamp, 1966.

Ann Treisman [TG80] a développé une tâche de recherche visuelle, qui consistait à repérer un stimulus-cible parmi un ensemble des stimuli-non-cibles. La cible pouvait se distinguer par une ou plusieurs caractéristiques comme la couleur, la forme ou l'orientation

2.2. TESTES D'ÉVALUATION DE L'ATTENTION

(Fig. 2.5). Les résultats de ce test ont montré qu'en cas de recherche disjonctive (la cible diffère d'une seule caractéristique), les sujets ont pu trouver la cible quel que soit le nombre des stimuli. Cependant, lorsque la recherche a été conjonctive (la cible diffère par plusieurs caractéristiques), le temps de réaction s'est prolongé (à mesure que le nombre des stimuli s'augmente).

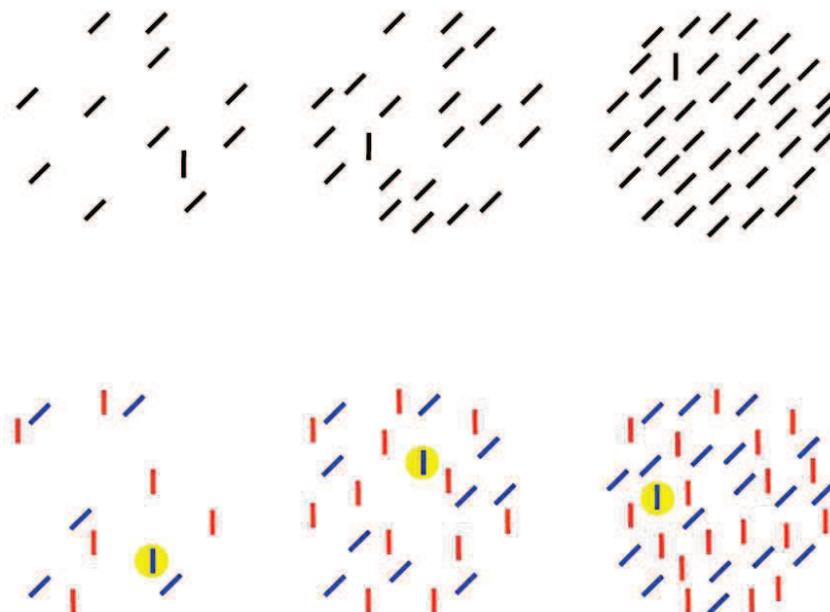


FIGURE 2.5 – Le test de Treisman, 1982.

Jeremy M Wolfe [Wol98] a effectué une expérimentation pour étudier l'attention sélective dans une tâche de recherche visuelle pourrait (Fig. 2.6). Si nous focalisons l'astérisque, nous trouverons probablement le «X» immédiatement parmi les douze éléments. Toutefois, si nous sommes invités à trouver le «T», nous ne pouvons pas le repérer jusqu'à une sorte de traitement supplémentaire sera effectuée (un laps de temps supplémentaire est nécessaire). Le traitement visuel de tous les éléments à la fois, fournit par conséquent suffisamment d'informations pour nous permettre de différencier un « X » parmi des « L », mais non le « T » parmi des « L ». Le temps de réaction moyen nécessaire à la détection du « X » est donc plus court que celui du « T ».

Nous constatons donc que la modalité de la recherche visuelle est donc indispensable dans l'étude de l'attention.

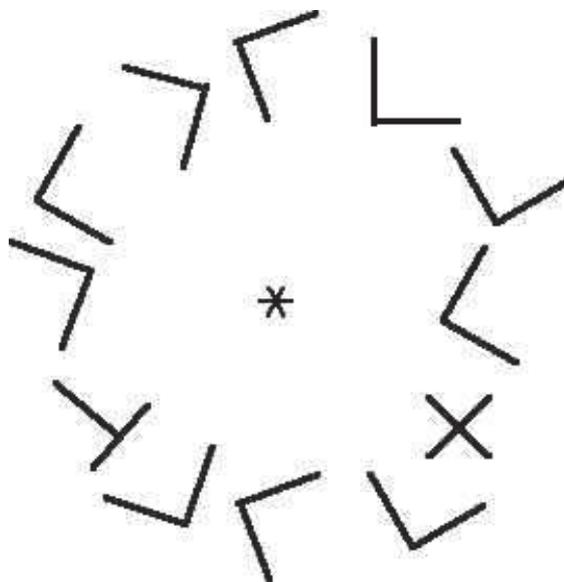


FIGURE 2.6 – L'expérimentation de Wolfe, 1998.

En effet, nous nous intéressons dans notre présente travail de deux ensembles de tests, qui sont le BAWL et le TEA :

La batterie BAWL [LP07] est un outil informatisé d'évaluation de différentes composantes attentionnelles. Elle a été élaborée dans une visée neuropsychologique. Elle peut toutefois être utilisée à d'autres fins : en psychiatrie, en psychologie cognitive, en gériatologie ou en interaction homme-machine. Cette batterie comprend un total de 10 épreuves, pour évaluer le temps de réaction.

Les épreuves de cette batterie comprennent différentes tâches, ne requérant de la part du sujet que la détection de signaux élémentaires sur le plan perceptif et auxquels il/elle doit réagir, le plus rapidement possible, en actionnant un bouton de réponse. Par exemple, le sujet doit réagir à l'apparition d'un carré noir présenté systématiquement au centre de l'écran. Une autre épreuve dont des stimuli cibles (carré noir) apparaissent à des endroits différents de l'écran, ce qui permet de calculer des temps de réponses en fonction de leur topographie de présentation.

Concernant les épreuves TEA [ZF94], qui ont été développées en tenant compte des besoins spécifiques du diagnostic en neuropsychologie, le développement de ces épreuves, favorise des tâches ne sollicitant que des réactions facilement identifiables et ne requérant

qu'une réponse sélective, consistant à appuyer sur une touche réponse. L'ensemble des épreuves de TEA se compose de 13 épreuves.

Par exemple, l'épreuve Go/No-Go de TEA, consiste à inhiber une réaction suscitée par un stimulus externe, au profit d'une réponse comportementale contrôlée. Cette épreuve comprend deux formes : (1) 2 stimuli dont 1 cible, (2) 5 stimuli dont 2 cibles. L'utilisateur doit appuyer, le plus rapidement possible, dès qu'une cible apparaît.

Morooka et al. [MOT⁺12] ont évalué l'attention sélective en utilisant une tâche Go/No-Go dérivée de [BBM⁺03] avec quelques modifications. Ils ont utilisé 3 stimuli dans leur tâche (2 cibles, 1 non-cible). Chaque stimuli a été présenté pour une durée de 500 ms. L'intervalle inter-stimuli varie aléatoirement entre 1000, 1500 et 2000 ms (moyenne 1500 ms). Trois blocs dont chacun entre eux est composé d'une série de 100 stimuli (85 cible, 15 non-cibles). Une pause de 10 secondes est prise en compte entre chaque exécution de bloc. *Barbarotto et al.* [BLF⁺98] ont aussi utilisé une tâche Go/No-Go (2 stimuli dont 1 cible). Ils ont utilisé 70 stimuli dont 35 sont des cibles et 35 sont des non-cibles. *Schulz et al.* [SFM⁺07] ont utilisé une tâche Go/No-Go de 96 stimuli dont 72 sont des cibles et 24 sont des non-cibles. Chaque stimuli a été présenté pour une durée de 500 ms. L'intervalle inter-stimuli varie aléatoirement entre 1250, 1500 et 1750 ms. D'autres recherches ont aussi évalué l'attention sélective en utilisant une tâche Go/No-Go (2 stimuli dont 1 cible) de TEA comme [LMG⁺14] [GWR⁺11]. Ils ont utilisé 40 stimuli dont 20 sont des cibles et 20 sont des non-cibles. Ils ont calculé la moyenne du temps de réaction et le nombre total d'erreurs. *Buechner et al.* [BMKZ05] ont évalué l'attention sélective en utilisant le test Go/No-Go (5 stimuli dont 2 cibles) de TEA, ils ont utilisé 60 stimuli dont 24 sont des cibles et 36 sont des non-cibles.

Dans un contexte interactif comme les jeux vidéo, les joueurs effectuent des tâches de recherche afin de sélectionner les stimuli cibles (des personnages ennemis, des objets à ramasser) parmi l'ensemble des stimuli proposés et ce, à l'intérieur d'une fenêtre de temps réduite. Ces caractéristiques correspondantes à notre contexte, peuvent être étudiées par le biais des tests mentionnés ci-dessous. Nous nous sommes donc inspirés de ces tests afin de construire nos outils de mesure dans cette thèse, puisque ces tests sont pratiques, adaptables et répondent pleinement à notre contexte.

2.3 L'attention dans les jeux vidéo

L'industrie du jeu vidéo est l'une des industries la plus rentable dans le monde. Un jeu vidéo est l'un des formes de divertissement le plus populaire depuis le début des années 1970. Le marché mondial du jeu vidéo, qui comprend les jeux vidéo sur console, les jeux vidéo en ligne, les jeux mobiles et PC, a atteint 93 milliards de dollars en 2013, contre 79 milliards de dollars en 2012. Ce marché, soutenu par la vente très forte des jeux mobiles et des jeux vidéo sur console, devrait atteindre 111 milliards de dollars en 2015, selon [vdMR13].

Aujourd'hui, il y a 34 millions de joueurs aux États-Unis qui dépensent 22 heures par semaine (en moyenne) pour jouer aux jeux vidéo [Gro14]. Les utilisateurs de jeux vidéo ne sont pas uniquement les enfants mais aussi des adultes. Par exemple, plus de la moitié des adultes américains jouent aux jeux vidéo, et quatre sur cinq jeunes jouent aussi. Chez les adultes, l'ordinateur est le dispositif le plus populaire pour jouer, par contre, chez les jeunes c'est les consoles de jeu qui leur attirent plus [LJM08], [Ass13]. Les jeux vidéo sont l'un des outils ludiques qui attire un grand nombre des utilisateurs dans le monde entier. Jouer aux jeux vidéo est une forme de passe-temps (pastime), qui est de-plus-en-plus répandu [GB03].

Il est essentiel de préciser ce que nous entendons par le terme *jeux vidéo* et en quoi ils sont différents par rapport à d'autres médias (livres numériques, télévisions ou films par exemple). Les jeux vidéo sont un sort de multimédia interactif, qui relie l'utilisateur en temps réel avec un monde virtuel derrière un écran à l'aide des images, des animations et des sons. C'est là où l'interaction de l'utilisateur prend place dans l'histoire du jeu vidéo formant l'expérience du joueur [Wol01]. La caractéristique la plus essentielle de jeux vidéo, ce est qu'ils sont interactifs ; les joueurs ne peuvent pas se rendre passivement à l'histoire d'un jeu. Au lieu de cela, les jeux vidéo sont conçus pour les joueurs qui doivent s'engager activement avec ces systèmes. Il y a des millions de jeux vidéo, avec des très différents thèmes et objectifs. Ces jeux peuvent être joués en coopération ou en compétition, seul, avec d'autres joueurs ; et ils sont joués sur différents appareils comme des consoles (Wii, *Playstation* ou *Xbox*) ou des ordinateurs, ou bien en utilisant des téléphones mobiles [GLE14].

Selon *Bavaliar et al.* [BGH⁺11], les changements dans le fonctionnement neural peuvent être un moyen, par lequel les compétences cognitives acquises à travers des jeux vidéo peuvent se généraliser à des contextes à l'extérieur des jeux. Les jeux vidéo peuvent fournir un cadre pour tester de nombreux types de mesures de l'attention [BKS⁺08], par exemple, jouer à des jeux vidéo, tels que Pac Man, peut améliorer le temps de réaction des personnes âgées [CLR87]. Un jeu d'action peut améliorer les compétences visuelles / attentionnelles [GB03], un jeu de stratégie peut améliorer les compétences de contrôle de l'exécutif et jeu de puzzle peut améliorer certaines compétences "spatiales" [BKS⁺08]. Nous pouvons donc aujourd'hui, surveiller et analyser l'état cognitif d'un utilisateur de jeux vidéo, afin de savoir s'il/elle est motivé(e), ennuyé(e), sommé(e) ou bien déprimé(e). Cela nous peut être représenter des indicateurs pour améliorer la conception de jeux vidéo afin d'offrir des versions de jeux plus efficaces avec une bonne qualité aux utilisateurs.

Andrade et al. [ACL⁺06] ont mis en place un jeu vidéo adaptatif avec un système de décision, afin de diagnostiquer les processus du déficit d'attention / hyperactivité³. Le jeu (Fig. 2.7) est en effet un labyrinthe qui doit être traversé par le joueur qui doit ramasser des éléments présentés dans une liste de courses. Les règles qui doivent être suivies par le joueur sont : de commencer par l'entrée du client et se terminer à la caisse, tous les éléments énumérés doivent être achetés, un point est gagné pour l'achat de chaque élément énuméré et le chemin doit être traversé en une seule fois (sinon le joueur perd un point). Des données initiales, comme la capacité de planification du joueur et la capacité d'exécution, ont été utilisées comme des entrées pour le système de décision, afin d'adapter la difficulté de jeu. Le système évalue les points (scores) du joueur et le temps écoulé. Cela permet de vérifier si les tâches exécutées dans les phases initiales ont été acceptables ou il est nécessaire de construire des nouvelles phases adaptées aux résultats précédents. Le processus d'adaptation se termine lorsque l'utilisateur remplit les objectifs souhaités dans un délai raisonnable à une quantité minimum d'erreurs. Nous pouvons observé que l'utilisation du système de score et la rapidité des actions pendant l'exécution de tâche, peut être utile afin d'évaluer l'attention et d'identifier la performance du joueur pendant une session de jeu.

3. Le déficit d'attention / hyperactivité est l'un des troubles de la diversité des dysfonctionnements.



FIGURE 2.7 – Un jeu vidéo adaptatif avec un système de décision (Andrade et al., 2006).

Spek et al. [vdSvOWA10] ont développé un jeu dit (*Code Red Triage*). Dans ce jeu en 3 dimensions, les joueurs doivent identifier des victimes d'un attentat, afin de les classer chacune selon son état en attendant l'équipe d'urgence. Le jeu commence dans une station de métro, où il y a eu une attaque terroriste dans une station adjacente. Les joueurs doivent donc trouver le quai du métro où il a eu l'attaque, afin d'examiner et de catégoriser les victimes, chacune selon son état : non urgent, urgent modérée, urgent vitale ou mort (Fig. 2.8). Les joueurs reçoivent des indices sonores pendant l'exécution du jeu afin de leur indiquer l'état de la victime (comme son état de mobilité ou ses vois respiratoires). Chaque victime bien classée, donne au joueur un score de 100 points. Le score total offre une représentation exacte sur le comportement du joueur pendant le jeu. *Spek et al.* ont trouvé que les indices sonores seuls, ne suffisent pas pour aider les joueurs à bien effectuer une tâche dans un jeu. Cette recherche nous montre l'importance d'étudier l'attention visuelle dans les jeux vidéo. Si le joueur doit apprendre une certaine technique dans un jeu vidéo, les indices visuels seront plus explicites que les indices sonores.

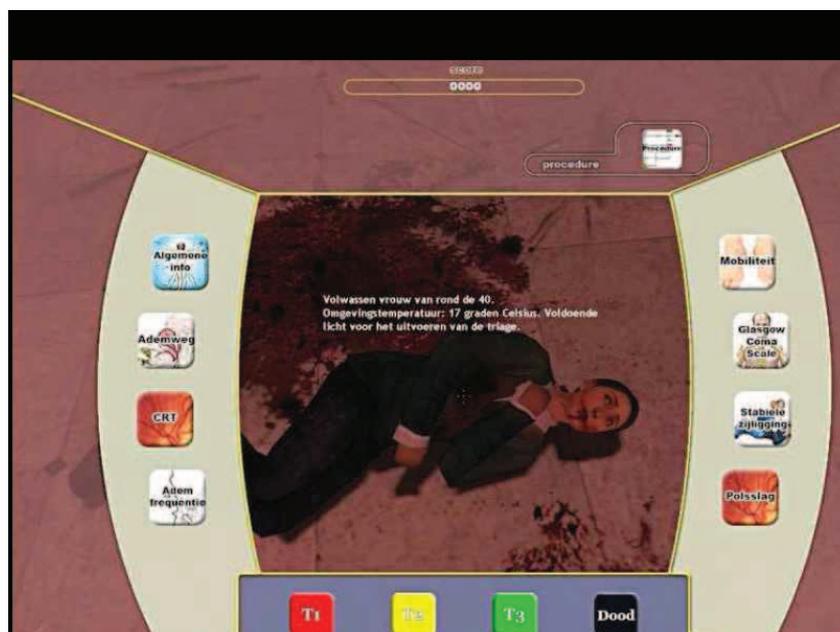


FIGURE 2.8 – La procédure de triage dans le jeu Code Red Triage (Spek et al., 2010).

Jouer aux jeux vidéo peut améliorer la performance des personnes âgées [CLR87], [DES+92]. L'identification précoce du déclin cognitif est très utile (chez les personnes âgées) pour se référer à la déficience fonctionnelle imminent [SZN+04], [Bra11]. *Jimison et al.* [JPMP04] ont montré l'importance de la détection précoce de déclin cognitif. Ils ont adapté le jeu dit *FreeCell* (un jeu de solitaire, gratuit sur Internet), pour mesurer la performance de l'utilisateur, qui a été représentée par la moyenne de l'efficacité par unité de temps. Selon *Jimison et al.*, les utilisateurs âgés de plus de 75 ans sont risqués de tomber dans des problèmes cognitifs, donc la détection précoce de ces problèmes peut fournir des indicateurs pour une intervention clinique. Nous pouvons observer que cette généralité peut être intéressant, bien qu'il y a des nombreuses maladies dont il n'existe pas, à nos jours, un traitement efficace et même l'intervention clinique ne retarde pas la progression de maladie.

En synthèse, l'analyse du comportement des joueurs peut permettre de donner des indicateurs sur leur performance. Ce comportement peut fournir des informations qui peuvent être utiles pour les concepteurs de jeux, leur permettant d'améliorer le gameplay, de gérer la difficulté de l'environnement de jeux et d'optimiser la répartition des ressources visuelles.

2.4 L'évaluation de l'attention en utilisant des capteurs spécifiques

Par définition, l'interaction homme-machine se réfère à l'ensemble des phénomènes cognitifs, matériels, logiciels et sociaux mis en jeu dans l'accomplissement de tâches sur support matériel (machine, système). Aujourd'hui, les chercheurs en interaction homme-machine ont peu d'intérêt pour les sciences cognitives, c'est pour cela, nous ne trouvons pas beaucoup de modèles psychologiques adaptés à l'interaction homme-machine.

Par conséquent, le traitement et l'analyse de l'attention, dans le domaine d'interaction homme machine, présente une challenge très intéressante pour les chercheurs dans ce domaine, en vue d'étudier le comportement de l'utilisateur d'une part et d'améliorer la conception des systèmes interactifs d'autre part.

En effet, les manifestations physiologiques et comportementales sont mesurables à travers d'un certain nombre de capteurs spécifiques permettant d'analyser : l'activité électrodermale, la fréquence cardiaque, la pression artérielle, la fréquence respiratoire, l'activité électromyographique ou l'activité encéphalographique. La prise en compte des ces signaux physiologiques et comportementales, permettent d'offrir une simulation exacte de l'état d'un(e) individu(e). L'utilisation de ces capteurs spécifiques comme l'électroencéphalographie (Fig. 2.9), est très courante, afin d'évaluer l'attention dans le domaine d'interaction homme-machine. Nous pouvons aujourd'hui utiliser ces capteurs pour acquérir plusieurs sortes de signaux (mesures, indicateurs) offrant une vue très proche sur le statut attentionnel d'un utilisateur d'une application interactive comme les jeux vidéo.



FIGURE 2.9 – Capteur électroencéphalographie

2.4. L'ÉVALUATION DE L'ATTENTION EN UTILISANT DES CAPTEURS SPÉCIFIQUES

Chen et al. [CWC⁺10] ont réalisé une classification de l'attention soutenue et non-soutenue, en se basant sur la variabilité du rythme cardiaque. Leurs sujets portent un capteur cardiaque appelé *My-ECG*, avec une fréquence de 500 Hz. Ce capteur a pour objectif d'enregistrer les signaux électrocardiographe (ECG) en temps réel. Pour induire l'attention, chaque utilisateur a effectué une tâche de performance. Dans cette tâche, des alphabets ont été affichés sur l'écran pour une durée de 200 millisecondes (ms) de manière aléatoire, l'utilisateur a été invité de bouger la souris si un non-X alphabet a été apparaît sur l'écran. La méthode utilisée dans cette recherche, a montré l'importance de détecter les états de motivation/fatigue chez les sujets en utilisant un capteur cardiaque.

Mikropoulos [Mik01] a mesuré l'activité cérébrale d'un utilisateur. Les utilisateurs ont été invités à réaliser une tâche en deux fois. La première fois dans un environnement réel et la deuxième fois dans un environnement virtuel. La méthode se base sur l'utilisation d'une analyse des signaux EEG. L'utilisateur a dû naviguer dans l'environnement virtuel (Fig. 2.10), en traversant le chemin entre la chaise et le pot de fleurs, puis faire un cercle autour du pot et revenir au point de lancement de navigation. Les résultats ont montré que les utilisateurs ont été plus attentifs lorsqu'ils ont réalisé la tâche demandée en utilisant l'environnement virtuel.



FIGURE 2.10 – L'environnement de Mikropoulos, 2001.

En synthèse, les capteurs spécifiques sont largement utilisés pour mesurer l'attention et la performance. Ce type de systèmes peut coûter cher et son déploiement peut être com-

plexe. Dans cette thèse, nous cherchons donc à mettre en place des méthodes non-invasives pour évaluer l'attention, sans l'utilisation de dispositif invasif, ni capteur particulier.

Suivant le contexte de cette thèse, nous allons détailler par la suite les travaux qui ont été élaboré dans trois différentes catégories : (1) la modélisation de comportement d'utilisateur à partir de son temps de réaction, (2) la transposition de l'oculométrie vers la camera virtuelle dans les environnements virtuels et (3) l'interaction en présence d'écrans multiples et son impact sur notre attention.

2.5 Modéliser la performance de l'utilisateur à partir de son temps de réaction

Dans cette thèse, nous utilisons le temps de réaction comme un indicateur sur l'état attentionnel d'un sujet. L'utilisation du temps de réaction est une approche commune (sollicitée) pour évaluer l'attention et la performance dans le contexte interaction homme-machine.

2.5.1 Le temps de réaction

En effet, l'association des mesures fournis par la chronométrie mentale⁴ avec ceux fournis en psychophysiologie, a bien offert des avantages pour les sciences cognitives [MOIY88]. Comme le traitement de l'information humain a progressé depuis le milieu du dix-neuvième siècle, les psychologues expérimentaux ont porté une attention considérable sur la dynamique de la cognition et de l'action. Il a souvent été supposé que les processus mentaux se manifestent par certaines mesures comportementales, y compris le temps de réaction, la précision de la réponse, et les compromis vitesse-précision [LLB79], [Pac73], [PM82], [Smi68], [Tay76], [Woo38]. Ces citations montrent l'importance d'utiliser le temps de réaction comme un indicateur sur les comportements d'utilisateurs.

Paul Fraisse [Fra68] a réalisé une expérimentation pour calculer le temps de réaction. Dans son expérimentation, les sujets ont été invités à appuyer sur une touche de réponse pour répondre à un stimulus sonore intercalé entre deux sons (dans une suite rythmique).

4. "La chronométrie mentale cherche à mesurer l'évolution dans le temps pour des opérations mentales dans notre système nerveux" [Pos78].

2.5. MODÉLISER LA PERFORMANCE DE L'UTILISATEUR À PARTIR DE SON TEMPS DE RÉACTION

Fraisse a montré que plus le moment où l'utilisateur doit répondre à une cible est proche de l'apparition de la cible, plus le temps de réaction est plus court. Selon lui, les études de la vigilance se multiplient dans les applications qui impliquent des tâches de surveillance, où il faut réagir à des séquences de stimuli. Cela peut être effectué en utilisant deux critères : le pourcentage de signaux détectés et la latence de la réaction à l'apparition d'un signal (le temps de réaction). Nous observons que cette recherche permet donc d'estimer l'évolution du niveau de la vigilance, en fonction de temps de réaction.

Gueye et al. [GLV⁺02] ont étudié le temps de réaction lors d'un mouvement de pointage. (Fig.2.11). Trois sources sonores ont été placées à gauche, à droite et au dessus d'une plateforme conçue pour leur expérimentation. Elles délivrent un stimulus sonore (signal préparatoire), destiné à orienter l'attention vers la zone de l'apparition de la cible visuelle. Cette cible peut apparaître en deux emplacements situés sur la plate-forme. Le temps correspond à l'intervalle situé entre l'apparition d'un stimulus et le début du mouvement identifié par le relâchement de cette plateforme, caractérise le temps de réaction.

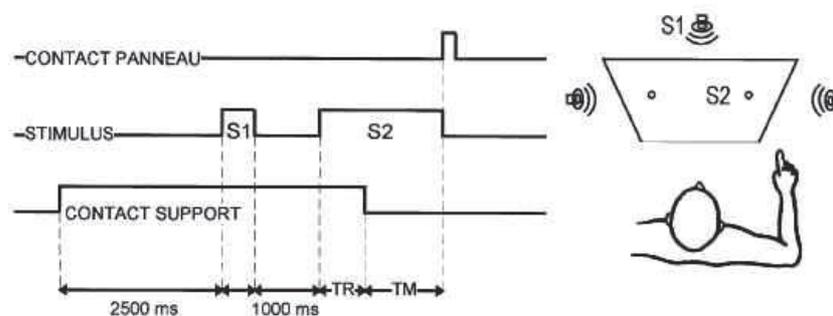


FIGURE 2.11 – Schéma du dispositif expérimental (Gueye et al., 2002) : TR est le temps de réaction et TM est le temps de mouvement

Cinaz et al. [CAT11] ont développé des tests de temps de réaction, qui peuvent être exploités à travers la vie quotidienne, au moyen de dispositifs portables. L'objectif principal était de comparer les temps de réaction obtenus par un capteur portable avec ceux qui sont obtenus par une approche bureautique (desktop-based). *Cinaz et al.* ont utilisé l'épreuve TOVA⁵. Il s'agit d'un carré noir qui apparaît sur un écran blanc. Si le carré apparaît dans la partie supérieure de l'écran, il est considéré comme une cible, dans le cas

5. TOVA s'agit d'un simple jeu vidéo, qui permet de mesurer les réponses à des stimuli visuels ou auditifs. <http://www.tovatest.com>, accès en septembre, 2015

2.5. MODÉLISER LA PERFORMANCE DE L'UTILISATEUR À PARTIR DE SON TEMPS DE RÉACTION

inverse, il est considéré comme une non-cible (Fig. 2.12). Dans un premier temps, l'utilisateur doit répondre à l'apparition de la cible en appuyant sur la touche espace du clavier (approche bureautique). Dans une seconde partie, il/elle doit répondre à la présentation de la cible en effectuant un mouvement de poignet (approche portable). L'utilisateur porte un bracelet, avec un capteur à trois axes d'accélération, placé au niveau du poignet. Lorsque l'accélération du capteur dépasse un seuil prédéterminé, il s'agit d'une réponse de la part d'utilisateur. Les résultats montrent que le temps de réaction est bien plus long dans l'approche portable par rapport à l'approche bureautique. Ce qui nous permet de dire que les caractéristiques d'un mouvement parcouru pour effectuer une réponse, doivent être pris en considération lorsque nous voulons mesurer le temps de réaction.

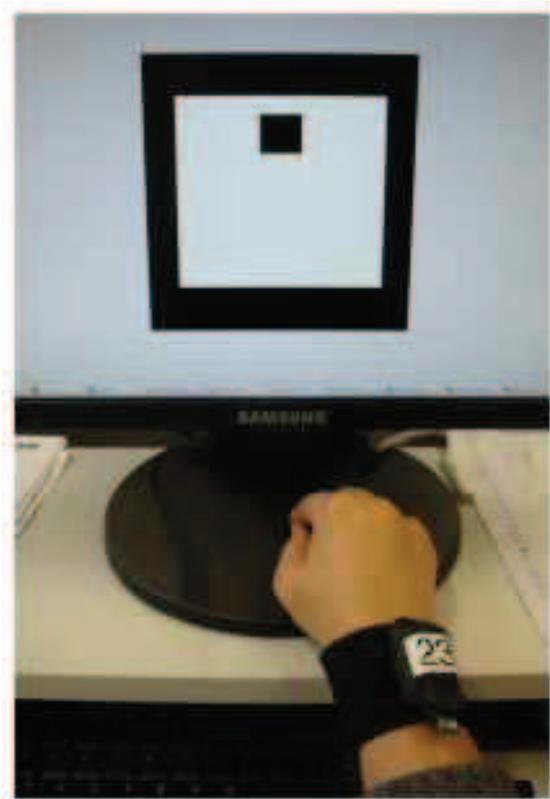


FIGURE 2.12 – L'expérimentation de Cinaz et al., 2011.

En synthèse, le temps de réaction reflète la durée du traitement de l'information nécessaire à l'élaboration d'une action. Cette variable dépend à l'attention, à la motivation ou à la complexité de la tâche demandée. Il est donc un élément très important dans la déter-

2.5. MODÉLISER LA PERFORMANCE DE L'UTILISATEUR À PARTIR DE SON TEMPS DE RÉACTION

mination du statut attentionnel d'un utilisateur. Par exemple, nous pouvons mesurer son temps de réaction d'un(e) individu(e) afin de savoir si celui/celle-ci est motivé ou fatigué, lors d'une exécution d'une tâche donnée.

En outre, l'outil que nous allons détailler son mise en place au chapitre 3, fait l'objet de l'étude des capacités attentionnelles dans le vieillissement ou dans la maladie d'Alzheimer. Dans la littérature, nous avons trouvé que les patients atteints de la maladie d'Alzheimer obtiennent des résultats hétérogènes aux épreuves attentionnelles [CVdL02].

Au niveau de la sélectivité, *Nebes et Brady* [NB89] ont montré une étude où des ensembles de lettres étaient présentés aux sujets. La moitié des essais se compose de six lettres noires (condition non indiquée) et l'autre moitié, de deux lettres noires et de quatre rouges (condition indiquée). La tâche des sujets consiste à détecter la présence d'une lettre-cible parmi l'ensemble, sachant que cette lettre-cible est identifiée par une couleur noire. Les patients d'Alzheimer montrent une réduction des temps de réaction semblable à ceux des sujets normaux, lorsque l'indice de couleur est disponible. Cela confirme que les patients d'Alzheimer sont capables de focaliser leur attention sur l'information pertinente de façon efficace.

Foster et al. [FBS99] ont également exploré les capacités d'attention sélective chez les patients d'Alzheimer, en utilisant une épreuve du temps de réaction. Les sujets doivent détecter le plus rapidement possible, en appuyant sur une touche de réponse, l'apparition d'un stimulus-cible (un cercle plein parmi des cercles vides). Cette épreuve se compose de deux conditions ; dans la première, le stimuli-cible diffère par une seule caractéristique (la cible est un cercle plein). Dans la deuxième condition, le stimuli-cible diffère par la conjonction de deux caractéristiques (la cible est un cercle plein, les autres stimuli sont des cercles vides et des rectangles pleins). Les résultats indiquent que le temps de réaction est beaucoup plus long dans la deuxième condition (condition de conjonction) chez les patients d'Alzheimer, alors que leur performance est relativement normale dans la première condition (condition simple). De plus, le nombre d'erreur chez les patients d'Alzheimer est faible (similaire à celui des sujets normaux) dans la condition simple, mais il augmente de façon plus importante lors de la recherche des caractéristiques conjointes. Les patients d'Alzheimer présentent donc des troubles spécifiques d'attention sélective lorsque la tâche

nécessite de déplacer l'attention entre différents attributs d'un stimulus.

En conséquence, nous constatons qu'il faut proposer des tâches assez simples afin d'évaluer avec précision le temps de réaction des patients d'Alzheimer.

L'outil que nous allons proposer dans le chapitre 3 vise à évaluer l'attention d'un sujet dans une tâche de pointage. Cette tâche de pointage devra être effectuée sur le plan d'écran en utilisant une souris. Pour cela, nous avons étudié les différents types d'évaluations qui peuvent donner des indicateurs sur les processus que nous devons appliquer, afin de calculer le temps de réaction dans ce type de tâche. Nous présentons donc par la suite les travaux menés afin de caractériser les mouvements d'utilisateur et leur relations avec le temps de réaction dans les tâches de pointage.

2.5.2 Les caractéristiques des mouvements dans les tâches de pointage

Les tâches de pointage sont parmi les éléments qui présentent un intérêt dans la conception d'une interface interactive. Dans ce type de tâche, l'utilisateur doit identifier et sélectionner un élément parmi d'autres sur la surface de l'interface. La conception de celle-ci, doit prendre en compte plusieurs facteurs afin de préserver son utilisation, comme par exemple le niveau de conscience d'utilisateur, la simplicité des actions et la présentation fluide des données.

Analyser les trajectoires sur un mouvement réalisé par l'utilisateur en se basant sur le profil de vitesse de ce mouvement, est une approche commune [VT82], afin de comprendre les étapes réalisées par l'utilisateur pour produire ce mouvement, en vue d'analyser sa performance.

Hourcade et al. [HPS08] ont proposé un outil dit *PointAssist*, proposant d'effectuer une tâche de pointage en utilisant une souris. Cet outil (Fig. 2.13) a fourni une solution pour faciliter l'exécution d'une tâche de pointage, puisque qu'il utilise un mécanisme qui s'occupe de ralentir la vitesse du curseur de souris, lorsque l'utilisateur est en train de pointer une cible donnée. Afin de déclencher ce ralentissement, l'outil a analysé les trajectoires réalisées par l'utilisateur afin de déterminer les sous-mouvements qui s'occupent de ralentir la vitesse. Ces sous-mouvements ont été détectés en fonction de changement de direction et

2.5. MODÉLISER LA PERFORMANCE DE L'UTILISATEUR À PARTIR DE SON TEMPS DE RÉACTION

d'accélération. Un sous-mouvement a été caractérisé par sa longueur, sa durée et sa vitesse.

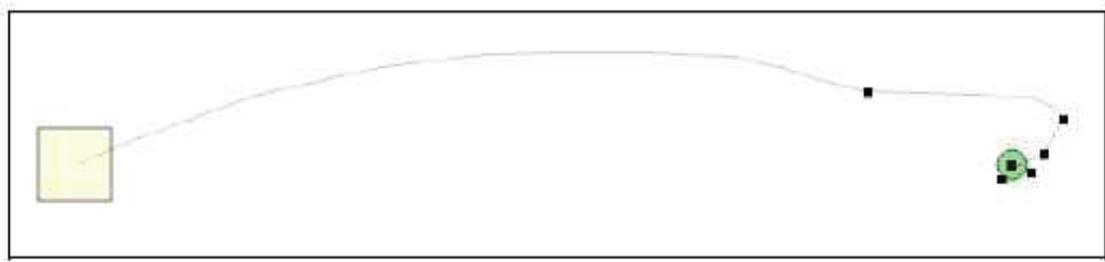


FIGURE 2.13 – Trajectoire tracée lors d'une tâche de pointage (Hourcade et al., 2008).

Khalid et al. [KYA10] ont réalisé une expérimentation dont le but est d'extraire les caractéristiques dynamiques d'un mouvement, lors d'une activité de dessin. Leur objectif était d'identifier les utilisateurs souffrant des difficultés dans l'écriture. La tâche proposée était de dessiner des carrés à l'aide d'une tablette avec un stylet. La méthode s'est basée sur une analyse de la vitesse des mouvements effectués par les utilisateurs (si le nombre de pics de vitesse était supérieur à un, cela traduisait une perturbation lors de l'exécution de mouvements). Les utilisateurs qui ont des problèmes de mains, présentent plus qu'un seul pic dans leur profil de vitesse [Pla91].

Wonner et al. [WGCB11] ont proposé une méthode dite *SPEED* dont le but était d'analyser les caractéristiques cinématiques des mouvements, dans une tâche de sélection de cibles. Leur modèle s'est reposé sur la séparation du mouvement de sélection en deux phases : une phase d'accélération et une phase de décélération (Fig. 2.14). Cette séparation avait pour but d'estimer la cible visée par l'utilisateur. L'analyse de la première phase a permis d'identifier un pic de vitesse qui marque le début de la seconde phase. L'environnement a été présenté à l'utilisateur par une grille de cases. Une des cases était en rouge, représentant la case de départ et l'autre était en vert, représentant la cible à sélectionner. Les utilisateurs ont utilisé une souris pour effectuer la tâche de pointage. Une sélection a consisté à un "click" sur la case de départ, puis un "click" sur la cible.

Nous constatons, à partir des travaux cités ci-dessus, que l'utilisation du profil de vitesse, offre une technique très intéressante dans le cadre de l'analyse des caractéristiques des mouvements. Nous pouvons ainsi constater l'importance d'analyser les trajectoires dans un outil de pointage, afin de traduire des intentions, pour déterminer le début d'un

2.5. MODÉLISER LA PERFORMANCE DE L'UTILISATEUR À PARTIR DE SON TEMPS DE RÉACTION

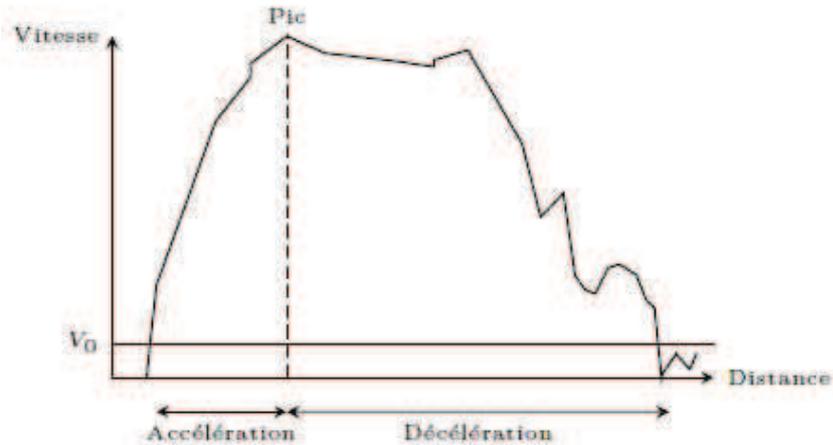


FIGURE 2.14 – La décomposition d'un profil de vitesse (Wonner et al., 2011).

mouvement vers une cible donnée dans une tâche de pointage.

Renau-Ferrer et al. [RFR11b], [RFR11a], [RFR12] ont étudié la façon de réaliser un dessin (étapes produites par l'utilisateur afin de dessiner une forme donnée), en utilisant une tablette avec un stylet. Leur méthode dite *OPSRL* dérivée par [YS05], se base sur l'extraction de caractéristiques procédurales des mouvements, pour déterminer les phases d'exécution d'un dessin. Parmi ces caractéristiques, l'orientation de chaque segment sur l'ensemble d'une trajectoire et leur relations structurelles (Fig. 2.15).

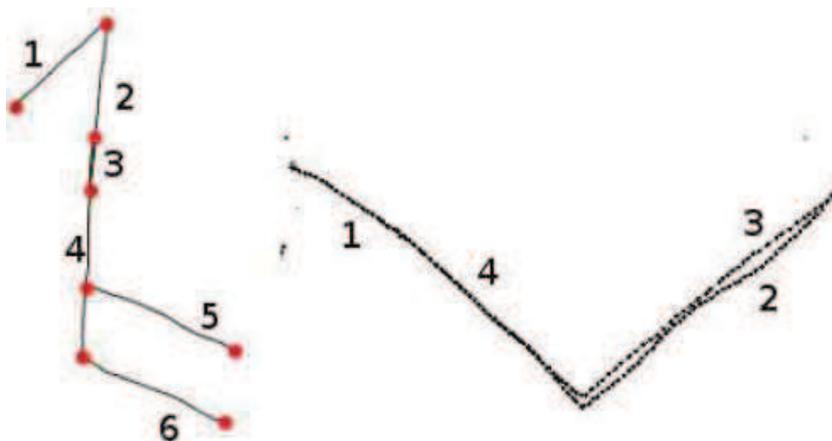


FIGURE 2.15 – Illustration de relations structurelles entre chaque segments sur une trajectoire réalisée lors d'une activité de dessin (Renau et al., 2011).

L'utilisation d'une algorithme similaire à celle de *Renau-Ferrer et al.*, permet d'ajou-

ter d'autres éléments aux travaux ultérieurs, visant à étudier les mouvements procédés lors d'une tâche de pointage. Nous pouvons donc constater que l'utilisation de cette méthode peut permettre de déterminer les relations structurelles entre chaque segments sur l'ensemble d'une trajectoire réalisée par un individu, lors d'une tâche de pointage. Nous pouvons donc comprendre la nature de mouvements d'utilisateur.

2.5.3 Synthèse

Le temps de réaction est une variable sollicitée dans les expériences interactives, puisqu'il peut être utilisé comme un indicateur sur la performance d'utilisateur. Les méthodes du temps de réaction, sont souvent utilisées dans les procédures de recherche rapide. Ces procédures portent sur les tâches habituellement exécutées dans l'ordre des millisecondes ou, au maximum, de quelques secondes [Cha78] cité par [MGB+84].

Dans une tâche de pointage, la procédure d'une réponse à l'apparition d'un stimulus, implique au moins quatre étapes de traitement successives :

- la réception du signal sensoriel et la transmission des informations qu'il contient vers des structures nerveuses supérieures ;
- l'analyse et l'identification du signal ;
- le choix d'une réponse ;
- l'initiation de l'action constituant la réponse représentant la partie motrice de la tâche.

Nous notons donc que le temps de réaction est une mesure très pertinente, afin d'évaluer la performance d'un(e) individu(e). Au chapitre 3, nous détaillerons donc la mise en place d'un outil d'évaluation de l'attention au sein d'une tâche de pointage. Le développement de notre outil a commencé lors la phase de développement du serious game appelé *Le Village aux Oiseaux*.

Selon *Guardiola et al.* [GNS+12], "Un serious game est un système formel fondé sur des règles avec une variable et les résultats quantifiables, où des résultats différents sont affectés des valeurs différentes ; le joueur exerce effort en vue d'influencer le résultat, le joueur se sent attaché au résultat. Mais un jeu sérieux est combiné avec un objectif de vie

réel défini". Cette définition est basée sur la définition du jeu classique de *Juul* [Juu03].

Les serious games sont avant tout des logiciels pour apprendre, qui mettent en place des outils permettant de contrôler l'apprentissage pour donner en suite un jeu qui s'adapte en fonction du parcours du joueur.

Dans *le village aux oiseaux*, nous devons calculer le temps de réaction des utilisateurs afin d'évaluer leur attention. La tâche proposée au sein de ce jeu, consiste à photographier (cliquer sur) des cibles affichées sur l'écran, en utilisant une souris.

L'analyse des trajectoires du curseur de souris d'une manière géométrique, peut offrir des indicateurs pouvant être considérés comme des intentions dans le but de calculer les temps de réaction.

Nous pouvons définir le temps de réaction par l'intervalle qui sépare la date de l'apparition d'une cible et la date de début du mouvement volontaire.

2.6 De l'oculométrie à la camera virtuelle

Pendant les interactions que nous faisons chaque jours dans notre vie quotidienne, nos yeux traduisent beaucoup d'informations qui reflètent nos états émotionnels et mentaux. En principe, les yeux nous transmettent une vision active de la réalité autour de nous, en faisant des déplacements réguliers dans notre environnement. Les commandes envoyées du cerveau aux muscles de l'œil, obligent les yeux à réaliser des mouvements de manière rapide.

Ces mouvements dits *saccades* permettent de diriger la fovéa sur un objet ou une région d'intérêt. La trajectoire réalisée par nos yeux sur une scène donnée, est appelée le chemin de balayage (*scanpath*). Celui-ci correspond à une série de mouvements successifs des yeux, séparés par des durées de temps, où nos yeux prennent des petites pauses pour encoder les informations perçues. Ces pauses sont appelées des *fixations*.

Les saccades oculaires sont nécessaires dans un grand nombre de comportements. Elles peuvent être volontaires, par exemple lorsque nous demandons à quelqu'un de regarder à droite ; elles peuvent être de type réflexe lorsque le regard est attiré par un événement inattendu, par exemple lorsque quelqu'un nous appelle. Ces caractéristiques sont largement

utilisées pour concevoir et décrire des systèmes qui se basent sur les processus attentionnels.

Il est courant que notre attention peut être déplacée de manière indépendante de notre regard [Rem80], [KF89], mais dans certaines situations, il semble que nos comportements oculaires ne puissent pas se déplacer sans être précédés d'un déplacement de l'attention [GG89]. Analyser donc les comportements oculaires (fixations, saccades) permet d'évaluer l'attention et la performance d'un utilisateur dans une application interactive comme un jeu vidéo.

Dans le chapitre 4, nous nous intéressons à évaluer la performance d'utilisateur dans une interface interactive. Nous supposons que la caractérisation des paramètres intrinsèques de cette interface, peut permettre de gérer sa difficulté en fonction de la performance de l'utilisateur. Pour cela, nous allons détailler la mise en place d'un modèle d'évaluation de la performance d'utilisateur, en transposant les comportements de la vision humaine (fixations, saccades) vers une vision synthétique. Ces comportements de la vision humaine sont habituellement mesurés par une oculométrie. Pour cela, nous allons présenter par la suite les travaux menés dans ce contexte afin d'étudier et caractériser ces comportements.

2.6.1 Les comportements oculaires

L'oculométrie (*Eye Tracker*) est devenue l'une des techniques la plus pertinente dans le but d'analyser les processus visuels et cognitifs au cours des dernières années. Elle est le processus qui enregistre les mouvements des yeux, ce qui nous permet de déterminer la zone à laquelle nos yeux sont fixés. L'oculométrie permet d'identifier la direction du regard d'un observateur, ce qui donne une vue précise sur son attention. Autrement dit, elle permet de connaître ce qu'une personne regarde, ce qu'elle ne regarde pas, mais aussi, ce à quoi elle prête attention et ce à quoi elle ne prête pas attention.

La plupart de systèmes d'oculométrie fonctionne avec des micro-caméras qui concentrent leurs objectifs sur les deux yeux et enregistrent leurs mouvements lorsque la personne fixe un stimulus. Pour le faire, elles captent le contraste entre la rétine et la pupille à l'aide d'une projection de lumière infra-rouge, en analysant la lumière réfléchiée par la cornée de l'œil humain. Par alternance des réflexions contrastées de la pupille (clair / foncé), il devient donc possible d'accrocher le mouvement de l'œil. Ainsi, dès qu'une première ca-

libration est effectuée, l'oculométrie fixe des repères sur trois axes dimensionnels et peut analyser les mouvements de la pupille de haut en bas, comme de gauche à droite (voir <http://trinity-advise.com/> pour plus d'informations).

L'oculométrie a été déjà largement utilisée en psychologie, en neurosciences et en interaction homme-machine. De nombreux recherches visent à étudier l'attention visuelle en utilisant des méthodes qui se basent sur l'utilisation de l'oculométrie. Par exemple, l'oculométrie a été employé pour étudier le comportement dans des domaines tels que la conduite routière [LL94], la lecture [Ray95] et l'interaction homme-machine [MD02], [KKS07].

Les informations dans l'environnement qui attirent nos yeux, sont beaucoup plus grandes que la capacité de notre cerveau à les traiter. Pour cela, nos yeux utilisent de comportements spécifiques (les fixations et les saccades) pour percevoir ces informations. La saccade est donc un outil essentiel pour l'attention visuelle sélective. Il est souvent considéré qu'une saccade oculaire se produit toujours en ligne droite et que nous ne "voyons" pas pendant ces mouvements. Nous pouvons donc considérer simplement que nous voyons souvent le monde au moyen de fixations.

La fixation et la saccade, dépendent sur la tâche dont l'utilisateur doit effectuée. La saccade survient entre les fixations avec des durées d'environ 150-200 ms [MR05]. La fixation est une courte durée de temps qui se traduit par une pause dans le mouvement des yeux (la saccade) d'une durée minimale de 100 ms [FRBK89] ou 200 ms [SSWR08]. Généralement la saccade doit durer moins de 100 ms et la durée de fixation commence à partir de 100 ms et peut être prolongée jusqu'à 500-600 ms (moyenne 200-500 ms [HE98]). Par exemple, dans une tâche de lecture, la durée moyenne de fixation est 250 ms [Ray95]. Par conséquent, nous pouvons définir la saccade et la fixation en se basant sur [JK03], [HE98], [Viv91], [FRBK89] par la suite :

- La *saccade* est un mouvement rapide des yeux, sous un contrôle volontaire. Le but de ces mouvements est d'apporter des images à partir des zones particulières de l'univers visuel vers la région fovéa.
- La *fixation* est le moment où les yeux sont relativement en pause court entre les saccades. Le but de ces pauses est d'encoder les informations perçues. Lors de la fixation, une cible est analysée, le champ périphérique est échantillonné et la saccade

prochaine est préparée.

En conséquence, la perception des objets visuels se produit lors des fixations, par contre, les signales oculaires sont inhibés lors des saccades [Wan95], [Gre97], [War00].

La durée de fixations est un sujet de recherche important en ce qui concerne l'information visuelle de l'observateur qui participe ainsi à son état cognitive. Une longue durée de fixation peut être dû à la difficulté d'extraire de l'information [JC76]. La durée du regard (gaze duration) est le temps passé à regarder un élément visuel [JK03]. Quand nous regardons un objet visuel, nous pouvons faire beaucoup de fixations sur cet objet et le nombre de ces fixations représente parfois l'importance de la zone d'affichage, mais un grand nombre de fixations traduit une interface mal conçue.

La figure 2.16 montre une analyse de regard lors d'une tâche de lecture, à partir d'une oculométrie. L'utilisateur qui lit un texte (à gauche) ; les yeux de l'utilisateur éclairés par la lumière infrarouge de l'oculométrie (au centre) ; le regard de l'utilisateur projeté sur le document (à droite). Les lignes représentent les saccades et les cercles sont les fixations.

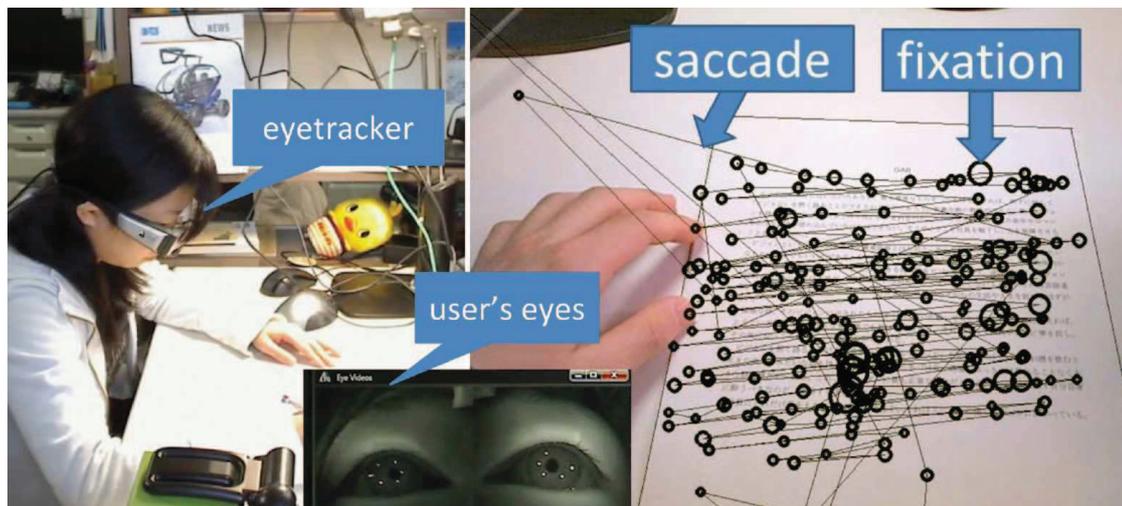


FIGURE 2.16 – Analyse de regard lors d'une tâche de lecture en utilisant une oculométrie (Kunze, 2013).

2.6.2 L'application de l'oculométrie en interaction homme machine et en jeux vidéo

Analyser le comportement oculaire d'un(e) individu(e) dans le contexte de l'interaction homme-machine [GB09], [GSL⁺02], [LKC07], [YWK06], [LM04], [BWA⁺10], a permis de fournir de nombreux indicateurs qui peuvent être utilisés pour évaluer l'attention de l'utilisateur d'une part, et améliorer la conception d'une interface interactive d'autre part.

Smith et al. [SG06] ont aussi comparé l'utilisation d'une oculométrie avec celle d'une souris dans le contexte de jeux vidéo. Leur sujets ont joué aux plusieurs types de jeux vidéo, tels que : *Quake* (FPS), *Neverwinter Nights* (roleplay) et *Lunar Command* (action). Ils/elles ont montré des meilleures performances dans *Lunar Command* en utilisant la souris comme dispositif d'interaction, mais aucune différence significative a été détectée pour *Quake* ou *Neverwinter Nights*. Nous pouvons dire que l'utilisation d'une oculométrie pour interagir avec une interface interactive peut diminuer nos performances et complexifier nos tâches.

Plusieurs travaux ont été conduits pour analyser les comportements oculaires dans les jeux vidéo. Les joueurs de jeux vidéo en vue subjective (First Person Shooter, FPS), focalisent leur attention en une seule région de l'environnement du jeu. Par exemple, le joueur paie plus d'attention au centre de l'écran autour du curseur de la souris, parce qu'il doit tirer sur les ennemis ; En revanche, la zone d'attention est plus grande dans un jeu d'aventure, car l'attention du joueur n'est pas contrainte par une zone spécifique de l'écran [ENY06]. Kenny et al. [KKD⁺03] ont montré que le joueur d'un jeu FPS, focalise son attention au centre de l'écran environ 88% du temps total de la session.

Sundstedt et al. [SSWR08] ont mené une étude expérimentale pour analyser le comportement oculaire des joueurs dans un jeu de labyrinthe (puzzle), en utilisant une oculométrie. Le modèle repose sur la détection des fixations (plus de 200 ms), puis le calcul de la distance entre ces fixations et les pixels situés dans un cercle (radius 43 pixel), afin de déterminer les objets où l'utilisateur regarde. Les résultats montrent que les mouvements et les fixations, sont principalement influencés par la tâche utilisée dans le jeu. Nous pouvons donc constater que l'analyse des comportements oculaires des utilisateurs d'une application interactive, peut être utile dans les phases des conceptions. Cela permet de valider

l'utilisation de l'application avant de fournir le produit final.

Gibbes et al. [GB09] ont réalisé une expérimentation, dont le but de déterminer si le comportement oculaire d'un utilisateur diffère par le type de sites web. Ils ont étudié les mouvements oculaires des utilisateurs, en utilisant une oculométrie. *Gibbes et al.* ont utilisé plusieurs variables donnant une mesure intéressante de l'attention, comme le nombre de fixations, la durée de fixation, la durée de regard et le taux de saccade. Nous pouvons observer que les mesures de *Gibbs et al.* peuvent être appliquées dans d'autres contextes, comme les jeux vidéo, afin de caractériser l'attention dans une utilisation interactive.

En synthèse, l'utilisation de l'oculométrie (pour analyser les comportements oculaires) pendant la phase de conception d'un jeu vidéo, peut être extrêmement précieux pour comprendre où les joueurs se concentrent leur attention, par rapport à l'objectif du jeu. Sur cette base, nous pouvons proposer que ces comportements oculaires peuvent être transposer vers des comportements synthétisés, afin de valider la conception d'une interface interactive.

2.6.3 Des mouvement de l'œil aux mouvements de la caméra

Les utilisateurs interagissent avec les environnements virtuels 3D en visualisant la scène à travers les "yeux" d'une caméra virtuelle. Une vue 3D virtuelle de la caméra, est définie par des plusieurs paramètres comme par exemple, la position de la caméra, la vecteur de direction, et le champ de vision. La caméra virtuelle joue un rôle essentiel dans les jeux vidéo 3D car elle affecte directement le comportement de joueur et la satisfaction par rapport au jeu [PWS08].

L'utilisation de cette caméra virtuelle peut donner des résultats intéressants pour l'étude non-intrusive et la caractérisation de la performance des utilisateurs. Le comportement d'un joueur décrit la façon dont celui/celle-ci réalise les tâches dans le jeu, par exemple, combien de déplacements il/elle effectue pour atteindre son objectif, alors que le comportement de la caméra virtuelle décrit comment celui/celle-ci la déplace pour atteindre son but [Bur13].

La modélisation et l'analyse du comportement de la camera virtuelle, en utilisant une combinaison de fixation et de saccade (habituellement utilisé en oculométrie), peut per-

mettre de comprendre comment un utilisateur manipule une application pour accomplir une tâche donnée.

Il est important de dire que les chercheurs en interaction homme-machine ont commencé à employer d'autres moyens que l'oculométrie, en vue d'évaluer la performance dans les environnements virtuels.

Stellmach et al. [SND10] ont mis en place un modèle comparant le chemin de balayage d'une oculométrie et le chemin de balayage d'une camera virtuelle, pour analyser le comportement d'utilisateur dans un environnement virtuel 3D. La comparaison est réalisée en dessinant les trajectoires entre les mouvements oculaires extraits par une oculométrie et les objets où l'utilisateur regarde en utilisant la camera virtuelle. Les résultats montrent que les trajectoires de caméra fournissent des informations précieuses sur la façon dont un utilisateur navigue dans un environnement virtuel 3D.

Yannakakis al. [YMJ10] ont étudié l'impact de la caméra sur le comportement de joueur et ont construit un modèle pour prédire cet impact. Cette étude a montré l'existence d'une relation entre les émotions des joueurs, les signaux physiologiques et la camera virtuelle. Le comportement oculaire peut être similaire à celui d'une caméra virtuelle [HLBC09].

En conséquence, comprendre le comportement d'un utilisateur dans un jeu vidéo, peut servir à comprendre sa performance, en analysant les données fournies par la camera virtuelle.

2.6.4 La tâche de recherche comme une activité de performance

Pour étudier le comportement d'utilisateur dans un environnement virtuel, une approche commune est de demander aux utilisateurs de réaliser une tâche de recherche au sein de cet environnement.

Hughes et al. [Hug06] ont proposé un modèle qui se base sur l'utilisation d'un "joystick" pour naviguer dans un environnement 3D. Le modèle demande aux utilisateurs de prendre en photo tous les objets observés dans la scène et il calcule ensuite la durée de fixation sur chaque objet. Une fois que l'utilisateur a pris l'objet en photo, il peut appuyer sur le "joystick" pour faire un zoom sur l'objet observé, afin de trouver un numéro dessiné sur

cet objet et le prendre aussi en photo.

Hillaire et al. [HLRC⁺10] ont montré que l'utilisation d'une tâche de recherche est très importante pour engager les utilisateurs à interagir avec un environnement virtuel 3D. Leur sujets doivent trouver des clés cachées dans cet environnement (Fig. 2.17) et collecter le maximum de ces clés, pour réussir la tâche.

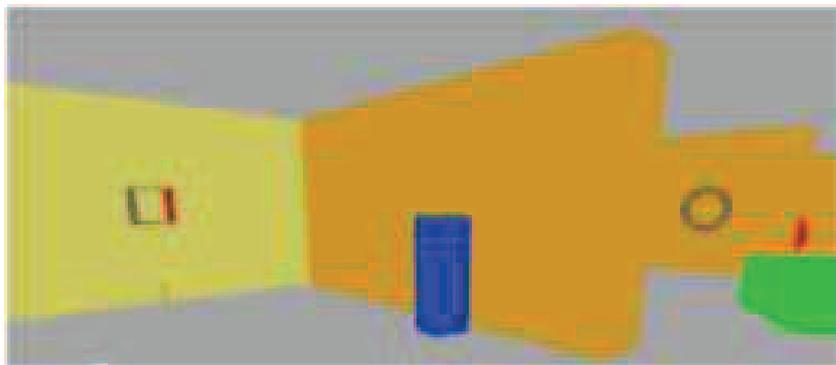


FIGURE 2.17 – Un environnement virtuel 3D avec une tâche de recherche, Hillaire et al., 2010

Lee et al. [LKC07] ont proposé un modèle pour évaluer l'attention dans un environnement virtuel, en utilisant une oculométrie. Les utilisateurs naviguent dans deux environnements (Fig. 2.18), afin de trouver des objets énumérés. Le modèle consiste à identifier les régions d'intérêt et calculer la durée du regard.

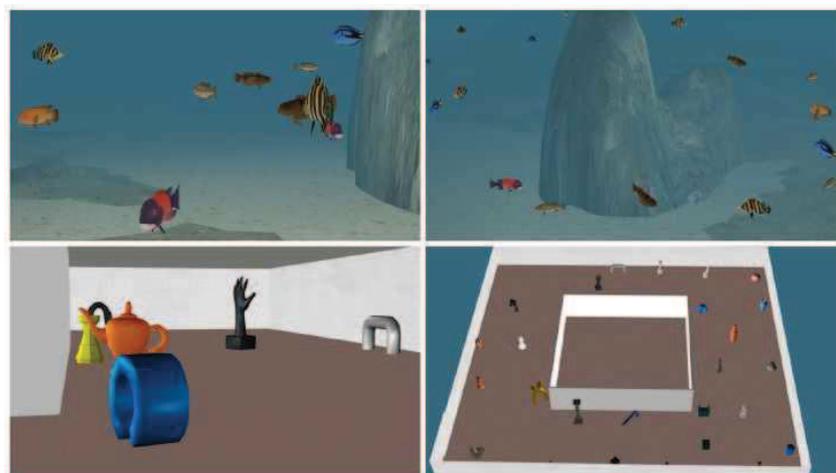


FIGURE 2.18 – Le sous-marin virtuel (environnement dynamique, en haut) et une galerie virtuelle (environnement statique, en bas) (Lee et al., 2007).

En conséquence, nous nous basons dans notre étude sur les caractéristiques (3D, prise d'information, tâche simple) de l'expérience de Lee et al. qui correspondent pleinement à notre contexte.

2.6.5 Synthèse

Aujourd'hui, nous pouvons utiliser d'autres moyens que l'oculométrie, pour évaluer la performance d'utilisateur. Quelques chercheurs ont déjà commencé à progresser dans ce contexte [HLBC09], [SND10].

Les paramètres d'une caméra virtuelle dans un jeu vidéo, sont d'habitude contrôlées par le système de jeu. L'étude et la caractérisation de comportement des joueurs, en analysant la façon dont celui-ci contrôle la caméra virtuelle, peut fournir des informations pour les concepteurs des jeux vidéo, leur permettent de comprendre la relation entre les mouvements de la caméra virtuelle et le comportement de joueur.

Nous supposons que les paramètres de la caméra virtuelle (comme le champ de vision) d'un jeu vidéo peut être utilisées pour ajuster la difficulté du jeu, selon les performances de joueurs.

Au chapitre 4, nous étudions la question de l'impact de certains choix de design d'un environnement virtuel 3D sur la performance de l'utilisateur. Pour cela, des indicateurs (fixation, saccade) traditionnellement utilisés en oculométrie vont être appliqués pour caractériser les comportements de la caméra virtuelle contrôlée par les joueur.

2.7 L'interaction en présence d'écrans multiples

Au cours des dernières années, l'utilisation d'écrans multiples a bien été augmentée. *Robertson et al.* [RCB⁺05] ont rapporté en 2001 qu'au moins de 20% des utilisateurs de système *Windows*, utilisent plusieurs écrans pour effectuer leur tâches. Il est probable que ce pourcentage est beaucoup plus élevé aujourd'hui. Une étude réalisé par Yahoo! et Nielsen [AWM⁺10] a indiqué que, en 2010, 79% des personnes utilisent Internet à partir d'un smartphone tout en regardant la télévision.

La multiplications des écrans a créé des opportunités pour faire différentes applications

facilement accessibles. La répartition de l'attention sur plusieurs écrans, est devenu une partie de notre travail quotidien. Par exemple, il est devenu courant de regarder des vidéos, consulter des mails et lire des documents sur les téléphones mobiles et les tablettes, en parallèle avec l'utilisation des ordinateurs de bureau.

Jonathan Grudin [Gru01] a observé que l'utilisation d'un écran "secondaire" (un écran utilisé en parallèle avec un écran principal "téléviseur") est souvent mise en jeu afin d'effectuer des activités secondaires, en particulier, celles de la communication (le courrier électronique ou la messagerie instantanée). *Gonzalez et Mark* [GM04] ont découvert que les utilisateurs passent en moyenne trois minutes de travail sur un événement unique, avant de passer à un autre.

Nous nous intéressons de ce travail de thèse, d'évaluer les effets de la distribution de l'attention sur plusieurs écrans, dans un contexte très interactive comme les jeux vidéo. Pour cela, nous avons étudié tout d'abord les travaux proposant une synthèse générale sur les méthodes d'évaluation de l'attention en présence d'écrans multiples.

2.7.1 L'impact d'utiliser plusieurs écrans sur l'attention

L'attention est un processus qui demande l'accès et l'exploitation des informations d'une manière rapide et efficace. Elle est une concentration sélective sur un seul objet parmi d'autres [Hyr06], [Ray98]. L'attention est un processus cognitif qui permet d'interpréter de manière sélective des sous-ensembles d'information, tout en ignorant les autres, c'est à dire, de se concentrer de manière sélective sur un seul aspect de l'environnement [And05]. Par définition, l'attention permet aux gens de focaliser leur intérêt sur une seule tâche à un moment donné. Cette notion d'attention peut être impliqué dans plusieurs contextes, notamment dans le contexte d'écrans multiples.

Kern et al. [KMS10] ont étudié les moyens de réduire les coûts de la répartition de l'attention sur plusieurs écrans. Ils ont développé un outil dit *Gazemarks*, afin de créer automatiquement des "espaces visuels réservés" en se basant sur le regard de l'utilisateur. En effet, dans les tâches où l'utilisateur est obligé de distribuer son attention d'un écran principal vers un écran secondaire, la récupération de l'information présentée sur l'écran principal peut être complexe et prend quelques durées de temps. Pour cela, textitGazemarks

2.7. L'INTERACTION EN PRÉSENCE D'ÉCRANS MULTIPLES

met en évidence la dernière zone où l'utilisateur a fixé son regard sur l'un des écrans. Les utilisateurs ont été placés devant deux écrans (8 pouces, 42 pouces). Ils/elles ont été invités à effectuer une tâche de recherche sur l'écran de 8 pouces, où il a fallu trouver des lettres placés sur une carte géographique. L'autre écran a été utilisé pour fournir des questions, dont le but est de distraire l'attention des participants. Une oculométrie a été utilisée pour enregistrer les fixations et les saccades des participants sur le petit écran (voir Fig. 2.19). Les participants qui ont trouvé une lettre, devaient détourner leur regard vers le grand écran où deux questions ont été présentées. Après avoir répondu à ces deux questions, les participants devaient revenir sur le petit écran, pour retrouver la même lettre à nouveau et alerter l'expérimentateur. Les résultats indiquaient que le temps pour compléter ce type de tâche est devenu plus rapide en utilisant une indice rappelant les utilisateurs de leur dernière fixation "intérêt" sur les écrans.



FIGURE 2.19 – *Gazemarks* (Kern et al., 2010)

Khan et al. [KMF05] ont présenté une nouvelle technique d'interaction dite *spotlight*, dont le but est d'orienter l'attention lors d'une visualisation des données sur des grands écrans. Leur expérimentation a consisté de présenter des nombres aléatoires sur plusieurs

2.7. L'INTERACTION EN PRÉSENCE D'ÉCRANS MULTIPLES

écrans, où les participants devaient trouver un nombre spécifique avec un des trois techniques d'interaction : un simple curseur, un *spotlight* et *searchlight* (voir Fig. 2.20). Une fois que le participant a trouvé le nombre précis, il devait l'entrer en utilisant un clavier. Le temps entre l'affichage de curseur et le moment d'appuyer sur la bon touche de clavier a été calculé afin d'analyser la performance d'utilisateur. Le résultat indiquait que l'utilisateur a eu une meilleur performance, en terme du temps, lorsqu'il a utilisé le technique *spotlight*. La fonction de temps est bien utilisée ici comme indicateur sur la performance d'utilisateur.



FIGURE 2.20 – La représentation de trois techniques d'interaction de Khan et al., 2005

Forlins et al. [FSWB06] ont observé que le temps de réaction est plus long lorsque nous utilisons plusieurs écrans afin d'effectuer une tâche de recherche sur une image statique.

Rashid et al. [RNQ12] ont exploré le coût d'orienter l'attention entre un petit écran et un grand écran. Ils ont comparé plusieurs configuration d'affichage : téléphones mobiles, téléviseurs et configurations mixtes. Ils ont indiqué que l'utilisation d'écrans multiples (la configuration mixte) peut relever des temps de réactions plus longs afin de réaliser une tâche de recherche visuelle. Ce type de configuration peut affecter et diminuer donc la performance d'utilisateur, tandis que les configuration qui utilisent un seul écran, offrent une utilisation plus confortable aux utilisateurs.

Holmes et al. [HJC12] ont utilisé une oculomètre pour évaluer l'attention dans le contexte d'écrans multiples. Leur sujets ont été invités de regarder une émission sur un téléviseur et de manipuler en parallèle une application qui tournait sur une tablette. Leur résultats ont signalé que les utilisateurs passaient plus de temps pour regarder la télévision que la tablette (63% à la télévision, 30% sur la tablette). Ils ont également noté que l'attention se déplace vers la tablette lorsque le téléviseur montre une publicité.

Brown et al. [BEJ⁺14] ont utilisé deux oculométries pour surveiller les regards des utilisateurs qui distribuait leur attention entre plusieurs écrans. Leur sujets ont été invité

2.7. L'INTERACTION EN PRÉSENCE D'ÉCRANS MULTIPLES

aussi de regarder une émission sur un téléviseur et de manipuler en parallèle une application qui tournait sur une tablette. (voir Fig. 2.21). Les oculométries ont été installés sur les deux appareils. Les résultats indiquaient aussi que les utilisateurs ont mis plus de temps en regardant le téléviseur que la tablette.



FIGURE 2.21 – L'installation de Brown et al., 2014.

Bobeth et al. [BSD⁺14] ont présenté une étude comparant trois types d'interaction avec un téléviseur : tablette, geste à distance et télécommande (voir Fig. 2.22). Leur sujets ont été des personnes âgées et des jeunes. *Bobeth et al.* ont analysé la performance d'utilisateur en utilisant le temps requis pour compléter les tâches demandées. Les résultats montraient que les trois types d'interaction ont eu une influence significative sur la performance d'utilisateur. La tablette a montré une meilleure performance et a été préférée par les deux groupes des utilisateurs. Les personnes âgées ont été globalement plus lents par rapport aux jeunes. Nous constatons donc que les manipulations directes pour effectuer des tâches de sélection semble être avantageuses dans ce contexte, en particulier pour les personnes âgées. L'utilisation de la télécommande comme moyen d'interaction est donc difficile et n'est pas optimal pour les personnes âgées. Cela peut être dû au grand nombre de boutons intégrés dans la télécommande, le *Wimote* serait un moyen d'interaction plus facile parce que le nombre de boutons intégrés est visiblement beaucoup plus réduit.

2.7. L'INTERACTION EN PRÉSENCE D'ÉCRANS MULTIPLES



FIGURE 2.22 – Les trois types d'interaction de Bobeth et al., 2014.

Vatavu *et al.* [VM14] ont proposé plusieurs mesures pour caractériser l'attention d'utilisateur lors d'une utilisation d'écrans multiples. Ils ont appliqué ces mesures en utilisant neuf configurations d'affichage (voir Fig. 2.23). Leur expérimentation consistait à tester les mesures d'attention, en fonction de deux facteurs principales : le nombre d'écrans (de 2 à 4) et leur mise en place (tous les écrans ont la même taille ou un des écrans est plus grand que les autres). Les participants ont été invités à regarder des films pour chaque combinaison de facteurs. Les résultats indiquaient que le nombre d'écrans et leur mise en place peuvent affecter l'attention et la charge cognitive d'un utilisateur.

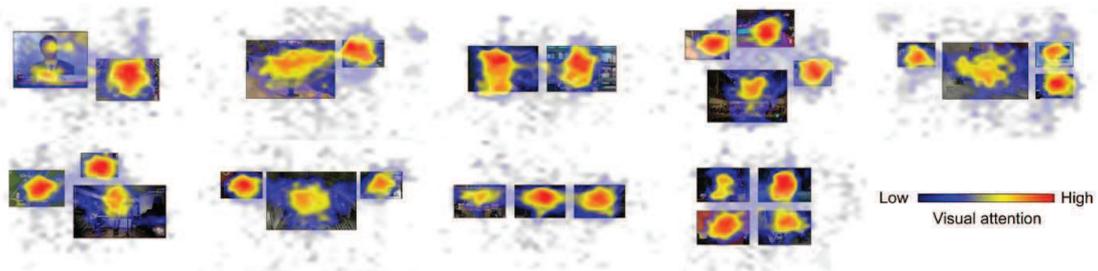


FIGURE 2.23 – Les heat maps de l'attention visuelle pour les neuf configurations d'écrans, Vatavu et al., 2014

En synthèse, nous constatons à partir des travaux mentionnés précédemment, que la fonction de temps (réaction, requis) est l'une des fonctions la plus utilisée dans les travaux de l'évaluation de la distribution de l'attention sur plusieurs écrans. Cette évaluation peut donner donc des indicateurs sur la façon dont un utilisateur manipule l'interface répartie sur ces écrans.

Emmerich *et al.* [ELM14] ont étudié la tendance utilisant les écrans secondaires dans le contexte des jeux vidéo. Ils ont défini l'écran secondaire comme un champ de solutions pour la conception de jeu, où plus d'un écran est utilisé pour afficher le contenu du jeu. Ils ont

montré trois jeux. Dans le premier jeu dit *Catch a Thief* (Fig. 2.24a), un joueur prend le rôle d'un garde de sécurité dont sa tâche, est de protéger le trésor d'art d'un musée, tandis que les autres joueurs, sont des voleurs et tentent de ramasser autant des produits, cachés dans les salles d'exposition. Sur leurs écrans privés (tablettes), les joueurs peuvent voir tout ce qui appartient à la salle d'exposition, tandis que la carte complète de la musée est affichée sur l'écran principal (téléviseur). Dans le deuxième jeu dit *Data Theft Algorithm*, les joueurs doivent recueillir des objets qui apparaissent sur l'écran principal. L'écran secondaire est utilisé ici pour afficher des informations privées comme les classements (voir Fig. 2.24b). Les joueurs doivent se concentrer dans la plupart de temps sur les événements affichés sur l'écran principal pour accomplir leurs tâches. Pour cela, les informations affichées sur l'écran secondaire doivent être très simples et visualisées en utilisant des grands symboles, afin d'être facilement reconnaissable. Dans le troisième jeu dit *The Mole Rush* (voir Fig. 2.24c), les joueurs doivent recueillir des objets et trouver des entrées spécifiques cachés dans le sol, qui est visible sur l'écran secondaire du joueur. L'écran principal affiche l'ensemble de l'environnement du jeu y compris les objets à ramasser.

La recherche de *Emmerich et al.* montre des scénarios importants en ce qui concerne l'utilisation d'un écran supplémentaire avec un écran principal afin d'offrir aux joueurs différents informations. Cependant, il ne montre pas des mesures sur les comportements des utilisateurs. Cela nous ouvre la porte d'effectuer nos travaux dans ce champs d'applications, afin d'évaluer les effets de la distribution de l'attention sur une interface répartie sur plusieurs écrans.

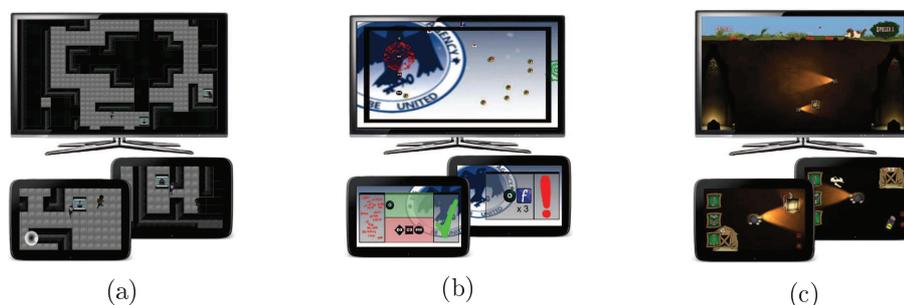


FIGURE 2.24 – Les prototypes de jeux - Emmirch et al., 2014 : *Catch a Thief* (a), *Data Theft Algorithm* (b), *The Mole Rush* (c).

Nous constatons donc que l'utilisation des écrans multiples a vraiment commencé à

attirer l'attention des chercheurs en interaction homme-machine d'une part, et les games designers d'autre part. Tout cela, pour offrir des nouvelles expériences de jeu, plus innovants et ludiques.

Par conséquent, l'évaluation de la répartition de l'attention sur plusieurs écrans dans un contexte de jeu vidéo, peut être un moyen qui offre des informations utiles sur le comportement d'utilisateurs, ce qui peut permettre aux concepteurs de jeux d'améliorer la conception de ce type des jeux.

Nous constatons aussi que ce type d'évaluation peut être en considération dans un contexte de jeux vidéo, où les joueurs peut s'en servir d'un écran complémentaire afin d'enrichir l'expérience de jeu.

Pour cela, nous avons classifié les types d'interaction avec les jeux vidéo, dans le but de déterminer les fonctionnalités de l'écran secondaire comme étant un dispositif d'interaction.

2.7.2 Classification des types d'interaction avec les jeux vidéo

Les jeux vidéo sont des logiciels qui fournissent une interaction à l'aide de certaines interfaces d'utilisateur de style d'entrée et/ou de sortie. Certains jeux vidéo sont joués en utilisant différents dispositifs, qui sont généralement : un contrôleur de jeu, une console principale et une unité d'affichage visuel. Récemment, les contrôleurs de jeux vidéo ont commencé à évoluer assez rapidement. De 1972 jusqu'au milieu des années 2000, tous les contrôleurs étaient très semblables. Ils sont tous présentés avec un certain nombre de boutons. Depuis 2006, lorsque la *Wii* de Nintendo a été présenté, nous remarquons aussi un grand changement dans le style des contrôleurs de jeux vidéo, comme l'utilisation des écrans tactiles ou des mouvements corporels/gestuels pour jouer.

Un contrôleur de jeu est un dispositif utilisé avec des jeux ou des systèmes de divertissement, généralement pour manipuler un objet ou une personnage dans le jeu. Il est d'habitude relié à une console de jeu ou un ordinateur au moyen d'un fil, bien que, depuis le milieu des années 2000, les contrôleurs sans fil sont devenus très répandus et très utilisables. Généralement, les contrôleurs originaux de jeux vidéo étaient des pièces de matériel qui ont été conçus pour un jeu de manière individuelle. Aujourd'hui, la plupart des contrô-

leurs modernes suivent une approche plus générale. Au lieu d'être spécialisé pour un jeu particulier, ils offrent suffisamment d'options afin de couvrir une variété de jeux divers.

Un des premiers contrôleurs de jeu vidéo était la *Magnavox Odyssey 100*. Le Magnavox a été introduit en 1972, il avait la capacité de contrôler les curseurs pour un jeu de ping pong. Le jeu ne nécessite qu'un mouvement vertical et horizontal pour manipuler les palettes pong, donc le contrôleur a été très limitée dans ses capacités et sa conception. En 1977, Atari a publié l'*Atari 2600*, et introduit un nouveau type de contrôleur de jeu, le joystick. Alors que le joystick a été d'abord utilisé en 1969 dans le jeu d'arcade *Sega's Missile*, Atari était la société qui l'a rendu plus populaire et plus accessible. Le joystick publié par Atari était assez simple. Il a présenté un manche numérique 4-directionnelle et 1 bouton. Contrairement au *Magnavox Odyssey 100*, l'*Atari 2600* était capable de manipuler plus qu'un match, le joystick a offert des mouvements plus précis, avec un simple design et une facilité à apprendre. Le contrôleur *Nintendo 64* est sorti en 1996, 24 ans après le *Magnavox Odyssey 100*. Il avait 14 boutons et un joystick, menant à de nombreux types d'actions allant de contrôler *Mario* dans un monde 3D au manipuler un vaisseau de *Star Fox* à travers un champ d'astéroïdes. Cette manette de jeu a donné à l'utilisateur la capacité de se déplacer d'une façon précise. Nous pouvons aussi distinguer les contrôleurs qui intègrent des fonctionnalités à détecter les mouvements. Ce type de contrôleur est capable à lire les mouvements physiques du joueur et lui permettre en suite d'interagir avec les jeux plus facilement. En 2006, la *Wii* a été publiée qui utilise cette technique de contrôle et qui a été le premier contrôleur qui utilise des mouvements gestuels pour contrôler un jeu vidéo au moyen d'une télécommande appelée la Wii remote la *wiimote* (Fig. 2.25a et 2.25d). Le *PlayStation Move* de Sony (Fig. 2.25b et 2.25e) a aussi des capacités à détecter les mouvements (similaire à la *Wiimote*). La télécommande Wii *Wiimote* dispose d'une caméra infrarouge qui est utilisé pour déterminer où la télécommande Wii pointe, par rapport à une source de lumière infrarouge. Les contrôleurs modernes de jeux vidéo ne sont pas toujours des périphériques physiques. En 2010, Microsoft a publié la *Kinect*, un dispositif qui a transformé le corps humain pour être un contrôleur de jeu. Ce dispositif de détection de mouvement utilise des caméras stéréoscopiques pour détecter des gestes, ainsi que d'un microphone pour l'entrée vocale. Désormais, un joueur peut jouer aux jeux vidéo

2.7. L'INTERACTION EN PRÉSENCE D'ÉCRANS MULTIPLES

simplement en bougeant son corps grâce à la *Kinect* (Fig. 2.25c et 2.25f).



FIGURE 2.25 – Différents types de contrôleurs des jeux vidéo

Les contrôleurs de jeu ont été beaucoup évolués au fil de temps, aujourd'hui, nous pouvons distinguer des contrôleurs plus avancés comme par exemple : les manettes qui intègrent plusieurs boutons, les contrôleurs qui utilisent la détection de mouvement et les écrans tactiles. Nintendo est un des leaders qui fabriquent des consoles de jeu avec l'utilisation d'un écran tactile. Le premier système de jeu vidéo qui utilise des écrans tactiles était la *Nintendo DS*, sorti en 2004. Depuis cette date, les appareils à écran tactile (les smartphones, les tablettes, la *Wii U* de Nintendo) ont dominé les marchés de jeux vidéo. En conséquence, l'introduction des smartphones et des tablettes au marché, a offert aux utilisateurs la possibilité de changer leur style afin d'effectuer leur tâches, grâce à la grande capacité de ces dispositifs mobiles, qui sont conçus pour faire des fonctions multiples avec une très bonne puissance de traitement, tout en nécessitant moins d'énergie et en offrant une large choix d'applications. La domination des smart phones et des tablettes a provoqué aussi une augmentation dans le nombre de jeux profitant de ce type des dispositifs, le succès de jeux mobiles comme *Angry Birds* (Rovio Entertainment Ltd, 2009), *Fruit Ninja*

(Halfbrick Studios, 2010), *Subway Surfers* (Kiloo, 2012) ou *Candy Crush Saga* (King, 2012) ouvre le chemin des éventuels conceptions pour les chercheurs et les game designers.

La plupart des contrôleurs de jeu (comme les manettes "standard", la *Wii mote*, le *Kinect* et les claviers/souris) permettent uniquement de commander le jeu, sans la possibilité d'avoir un rendu visuel sur les actions réalisées (affichage des contenus sur leur surfaces). La différence entre ces types de contrôleurs apparaît plus profondément dans la manière dont l'utilisateur doit commander le jeu. Autrement dit, le type d'interaction dont l'utilisateur possède pour manipuler le jeu, par exemple en utilisant : des gestes comme le *Kinect* et la *Wii mote* ou des configurations traditionnels (qui se basent principalement sur l'utilisation des boutons) comme les manettes de jeu. Tandis que les smartphones, les tablettes et la *Wii U* nous permet de non seulement commander un jeu, mais aussi d'avoir un rendu visuel de nos actions sur leur surfaces d'affichage. Cela nous ramène à l'importance de faire une classification de type d'interaction avec les jeux vidéo.

Cox et al. [CWJB12] ont mesuré l'efficacité de certains contrôleurs de jeux dans le contexte de télévision interactive. Ils ont utilisé quatre méthodes d'interaction : le **Kinect** de Microsoft, le *Wii mote* de Nintendo et 2 méthodes utilisant une tablette Android. Les résultats montrent que le *Wii mote* est le meilleur contrôleur de jeu en termes de performance. Ce qui nous intéresse dans cette étude est la classification rapport au type d'interaction avec la télévision. Selon les auteurs, le *Wii mote* et le *kinect* constituent une interface d'interaction naturelle. D'autre part, une tablette constitue une interface d'interaction tactile. Les interfaces d'interaction naturelles étendent le paradigme de la manipulation directe, en permettant aux utilisateurs d'interagir avec le téléviseur avec des mouvements ressemblant davantage à celles que nous utilisons dans la vie réelle.

Brik et al. [BM13] ont étudié l'effet de trois contrôleurs : le *Kinect*, le *PlayStation Move* et le *Xbox GamePad*. Selon les auteurs, nous pouvons distinguer trois types d'interaction avec les jeux : (1) le traditionnel comme les manettes de jeux, la souris et le clavier ; (2) le positionnel comme le *Wii mote* et le *PlayStation Move* ; (3) le gestuel comme le *kinect*.

Oshita et al. [OI12] ont comparé deux types d'interaction avec les jeux vidéo : les manettes de jeux et l'écran tactile. Les résultats montrent qu'un écran tactile a obtenu des meilleurs performances par rapport à la manette. Selon les auteurs, nous pouvons aussi

distinguer deux types d'interaction avec les jeux vidéo : (1) les interfaces tactiles comme les tablettes, les smartphones, le *PlayStation Vita*, le Nintendo *2DS* et *3DS* ; (2) les interfaces manettes comme la manette *Xbox* et la manette de *PlayStation*.

Selon *Pham et al.* [PT12], il existe trois principaux types d'interaction avec les jeux vidéo : (1) les contrôleurs à base de boutons comme la manette de *Xbox* et de *PlayStation* ; (2) les contrôleurs gestuels comme le *kinect* ; (3) les contrôleurs mixés (gestuel et boutons) comme le *Wimote*. Selon *Pillias et al.* [PRBL14], nous pouvons aussi distinguer deux types d'interaction avec les jeux vidéo : (1) l'interaction continue en suivant le corps du joueur comme le *kinect* ; (2) l'interaction en utilisant la main du joueur comme la souris et le *Wimote*. Selon *Sajjadi et al.* [SCGTD14] et *Klochek et al.* [KM06], un contrôleur comme la souris et le clavier constitue une interface d'interaction standard. Selon *Natapov et al.* [NCM09], les manettes de jeu constituent une interface d'interaction traditionnelle.

Certains contrôleurs de jeu ont exploré l'utilisation des interfaces tactiles (comme la *Wii U*, les smartphones et les tablettes). Dans le cas de la *Wii U* qui peut être considérée comme un écran secondaire couplé avec un téléviseur, l'utilisateur peut non seulement manipuler le jeu à partir de l'écran secondaire, mais aussi avoir un rendu visuel sur celui-ci. Dans ce cas d'interaction particulière, nous entendons parler des expériences dites *multi-écrans*.

En se basant sur les recherches mentionnées précédemment dans cette partie, nous avons mis en place une classification des types d'interaction avec les jeux vidéo. Cette classification nous offre la possibilité d'introduire plusieurs fonctionnalités de l'écran secondaire, ce qui nous permet ensuite d'évaluer les effets de la répartition de l'attention en utilisant plusieurs écrans.

En effet, nous pouvons distinguer deux dimensions caractérisant le type d'interaction avec les jeux vidéo. La première dimension est le style de commande et la deuxième est la surface d'affichage. L'utilisateur peut manipuler le jeu, soit en utilisant des gestes, des commandes tactiles ou des actions physiques traditionnels (boutons) en le style de commande. La surface d'affichage permet d'envoyer un rendu visuel sur nos actions. Nous pouvons aussi distinguer deux modes d'affichage : le mono-écran et le multi-écrans. Le premier se compose d'une seule surface d'affichage, ce qui permet à un utilisateur d'avoir un rendu visuel sur un seul écran, tandis que le multi-écrans se compose de, au moins, deux ou plusieurs

2.7. L'INTERACTION EN PRÉSENCE D'ÉCRANS MULTIPLES

surfaces d'affichage, ce qui oblige l'utilisateur de distribuer son attention sur ces surfaces d'affichage simultanément (voir le tableau 2.1 pour plus d'informations).

Style de commande	Surface d'affichage	
	Mode Mono-écran	Mode Multi-écrans
Geste : Kinect de Microsoft, Wii mote de Nintendo, Playstation Move de Sony	Un seul surface d'affichage	Deux ou plusieurs surfaces d'affichage
Tactile : Tablette, Smart phone, Nintendo DS		
Traditionnel : Manette de jeu, clavier, souris		

TABLE 2.1 – Classification des types d'interaction avec les jeux vidéo

2.7.3 Synthèse

L'utilisation d'un écran secondaire couplé avec un écran principal, peut se faire donc en utilisant plusieurs et différentes manières, tout en dépendant de la fonctionnalité du second écran et sa relation avec l'écran principal. Par exemple, pour effectuer une tâche en utilisant cette approche, comme la vidéo surveillance ou le montage vidéo, l'utilisateur doit regarder plusieurs écrans simultanément. Il/elle peut non-seulement manipuler son application affichée sur un écran principal, mais aussi avoir des rendus visuels sur l'écran secondaire.

Pour simplifier cela, nous introduisons trois fonctions principales de l'écran secondaire (en se basant sur notre classification mentionnée précédemment) :

- *Surface de commande* : dans cette fonction, l'écran secondaire prend le rôle d'un dispositif de commande (comme une manette, un télécommande ou une gyroscope). Il sert à manipuler une interface sans avoir un rendu visuel sur sa surface (il n'affiche aucun contenu). En plus, l'utilisateur n'est pas souvent obligé de regarder l'écran secondaire lorsqu'il effectue son travail. Une fois qu'il/elle a bien repéré les positions de commande (des boutons par exemple), il/elle peut se concentrer sur l'écran principal.
- *Surface d'affichage* : cette fonction permet l'écran secondaire d'avoir un contenu visuel, sans avoir par contre la possibilité de manipuler l'interface (le commande se

- fait à partir d'utiliser un autre moyen, comme des boutons intégrés physiquement avec l'écran). Dans cette fonction, un écran secondaire peut prendre le rôle d'un miroir qui affiche le même contenu de l'interface (déjà affiché sur l'écran principal) ou par contre, il peut afficher un contenu différent/supplémentaire à celui affiché sur l'écran principal. En plus, l'utilisateur doit impérativement regarder l'écran secondaire, à condition que celui-ci affiche un contenu différent à celui affiché sur l'écran principal.
- *Surface de commande/d'affichage* : dérivée de deux fonctions ci-dessus. L'écran secondaire agit comme un contrôleur de l'interface et joue aussi le rôle d'une surface visuelle (affiche des contenus différents/supplémentaires) simultanément. Cette situation oblige l'utilisateur à regarder simultanément les deux écrans, afin d'accomplir ses tâches. Cela divise l'attention de l'utilisateur pour le simple fait que celui/celle-ci doit observer deux différents contenus repartis sur deux surfaces d'affiche distincts.

Ce qui nous intéresse dans notre étude est la fonction de surface de commande/d'affichage (comme la Wii U ou le Xbox smartglass), puisque cette fonction/configuration permet aux utilisateurs de jouer activement aux jeux vidéo, en utilisant plusieurs surfaces d'interaction en parallèle.

En outre, l'utilisateur peut rencontrer des difficultés lorsqu'il/elle réalise une tâche répartie sur plus d'un écran. Cette difficulté est due à la présence de plusieurs écrans qui s'occupent d'afficher le contenu de cette tâche simultanément. Pour cela, l'utilisateur doit observer les écrans de façon continue, pour ne pas avoir une perte d'information, et donc son attention doit être dirigée vers plusieurs écrans afin d'accomplir la tâche demandée. Cette situation peut être fatigante d'une part et elle demande beaucoup d'efficacité d'autre part.

L'évaluation de l'attention dans le contexte de jeux vidéo en utilisant des écrans multiples, est un nouvel sujet de recherche, qui n'est pas encore très exploité par les chercheurs en informatique. Cette évaluation peut rendre des informations très importantes aux concepteurs des jeux vidéo, afin qu'ils puissent les prendre en compte lors du développement des jeux, pour adapter la difficulté des jeux en fonction de la performance d'utilisateur. Pour cela, nous visons dans le chapitre 5 de cette thèse, à étudier et évaluer les effets de répartition de l'attention sur des écrans multiples, dans le contexte des jeux

2.7. L'INTERACTION EN PRÉSENCE D'ÉCRANS MULTIPLES

vidéo.

Chapitre 3

Relation attention-difficulté dans un serious game

3.1 Contexte de l'étude

Dans ce chapitre, nous présentons nos travaux qui ont pour objectif d'évaluer le niveau attentionnel d'un utilisateur dans un serious game. Le jeu qui sert de support à notre travail se nomme *Le Village aux Oiseaux* (Fig. 3.1). Il a été développé dans le contexte du projet SGCogR¹, un projet financé par la DGCIS pendant 2 ans (2009-2011) à la suite de l'appel à projet serious game, dit "NKM". Le consortium de ce projet est un partenariat entre quatre sociétés (TEKNEO, Seaside Agency, SpirOps, Neofactory) et deux laboratoires de recherche (CEDRIC, INSERM). *Le village aux oiseaux* est un serious game destiné aux personnes présentant les premiers symptômes de la maladie d'Alzheimer.

L'hypothèse principale de ce projet, proposée par les neuropsychologues de l'INSERM, est que l'entraînement de l'attention chez ces patients peut réduire les symptômes de la maladie d'Alzheimer (voir [GB03] pour plus de détails). L'objectif est donc de leur permettre d'entraîner leurs ressources attentionnelles dans le but d'améliorer leur quotidien et de diminuer la perte cognitive due à cette maladie [MNL12]. Dans cette perspective, le *Village aux Oiseaux* a été développé pour fournir : (1) un environnement d'entraînement de l'attention, basé sur une conception du jeu pertinente, (2) un outil qui soit capable de s'adapter aux réactions de l'utilisateur, notamment dans le cas de fluctuations de son

1. SGCogR : Serious Game for Cognitive Rehabilitation

3.1. CONTEXTE DE L'ÉTUDE



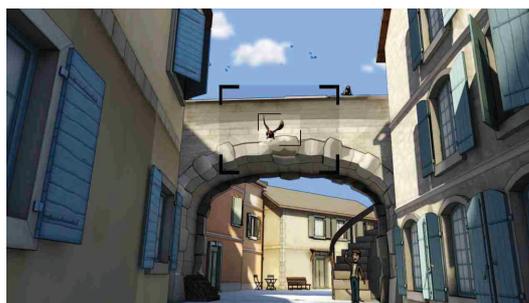
(a) Le château



(b) La carte générale



(c) photo en jeu



(d) photo en jeu

FIGURE 3.1 – Captures d'écran du *Village aux Oiseaux* : Le château (a), la carte générale (b), photos en jeu (c, d).

attention. Notre travail a consisté au développement de ces outils de mesure de l'attention, qui seront utilisés par le système pour adapter la difficulté du jeu ou remotiver l'utilisateur par le déclenchement d'événements ad-hoc par le moteur du jeu. Par ailleurs, ce système de mesure permet d'enregistrer les traces de l'activité des utilisateurs, les rendant ainsi disponibles en cas de besoin dans un contexte médical (évaluation de l'évolution de la symptomatologie, par exemple).

Le Village aux Oiseaux a pour vocation à être joué à domicile, sans assistance. Pour cela, cet outil devait répondre à des contraintes supplémentaires : (1) ne pas être invasif, (2) ne pas nécessiter de compétences informatiques, (3) coûter aussi peu que possible.

Dans *Le Village aux Oiseaux*, le joueur incarne un photographe qui doit prendre en

3.1. CONTEXTE DE L'ÉTUDE

photos des oiseaux qui apparaissent aléatoirement dans l'environnement du jeu. *Le Village aux Oiseaux* utilise un mécanisme similaire à celui utilisé dans un jeu de tir à la première personne². Le déplacement de la caméra virtuelle³ du jeu est géré automatiquement en fonction de trajectoires prédéfinies par les concepteurs.

Le gameplay utilisé dans *Le Village aux Oiseaux* repose sur des tâches de pointage, qui sont effectuées dans une interface en deux dimensions (dans le plan de l'écran), en utilisant une souris. Cette situation correspond à une activité classique dans l'utilisation de l'outil informatique et permet de ne pas demander un apprentissage trop important et/ou des compétences particulières.

Notre tâche au sein de ce projet était de concevoir et d'implémenter un dispositif permettant d'adapter automatiquement les challenges rencontrés par l'utilisateur dans le jeu. Ceci répond au besoin de maintenir le joueur à un niveau de motivation optimal [Che07] afin de maximiser le temps d'utilisation du jeu (et, par conséquent, le temps de l'action thérapeutique). Afin de respecter les contraintes citées plus haut, nous avons proposé d'étudier une solution basée sur l'analyse des données fournies par le joueur, en l'occurrence, celles de son activité de pointage. En effet, nous pensons qu'il est possible d'extraire des caractéristiques des trajectoires du pointeur les intentions qui sous-tendent à ces mouvements.

Afin de caractériser les performances réalisées dans les tâches de pointage, il est courant, dans le contexte de l'interaction homme-machine, de mesurer le temps de réaction (TR). Celui-ci représente le temps écoulé entre l'apparition de la cible et le début du mouvement volontaire de l'utilisateur vers cette cible. Cette variable de TR est intéressante pour notre étude, puisqu'elle va nous permettre d'établir un profil d'utilisateur, qui peut s'avérer utile dans le cas d'une étude longitudinale. Dès lors, nous sommes en mesure, d'une part, de détecter les éventuels décrochages attentionnels, par l'interprétation de l'allongement du TR (en temps réel, pendant une session) et, d'autre part, de caractériser l'évolution de la symptomatologie par une augmentation significative dans les TR enregistrés (entre différentes sessions).

2. Jeu de tir en vue subjective, c'est-à-dire que le joueur voit l'action à travers les yeux du protagoniste.

3. La caméra virtuelle permet de situer un point de vue, au travers duquel l'utilisateur perçoit l'environnement du jeu.

Dans une telle tâche de pointage, il est nécessaire d'identifier le début du mouvement volontaire, réalisé en réaction à la présentation d'une cible. La fin de ce mouvement, qui permet le calcul du temps total pris par l'utilisateur pour effectuer cette tâche, est relativement facile à détecter. Il est validé par la sélection de l'utilisateur (en prenant une photo) via le bouton de la souris. Le début du mouvement volontaire est plus difficile à détecter. Dans le contexte du jeu vidéo, l'utilisateur a la liberté de déplacer son curseur de souris partout sur l'écran, indépendamment de la présence de cible, ce qui rend difficile l'identification des mouvements qui ont été motivés par les cibles et ceux qui ne l'ont pas été.

Dans la suite de ce chapitre, nous décrivons les différentes étapes qui ont été menées afin de répondre à cet objectif. Nous avons tout d'abord développé un logiciel permettant de calculer le TR, pour un utilisateur donné, dans une tâche de pointage contrôlée. Les données fournies par cet outil nous permettront de produire des valeurs de référence des TR des utilisateurs et d'en dresser des profils. Nous avons développé ensuite un second outil, s'appuyant sur un algorithme original, qui permet de caractériser les mouvements du pointeur d'un utilisateur afin d'en extraire ceux en réaction à une présentation de cible. Cet outil nous permettra de calculer des TR dans des contextes où les conditions initiales du mouvement ne sont pas contraintes. Enfin, nous avons validé cet outil par son intégration dans *le Village aux Oiseaux*.

3.2 Approche géométrique des phénomènes attentionnels

3.2.1 Construction des profils utilisateurs

Pour concevoir un test d'évaluation de l'attention, nous nous sommes inspiré dans nos expérimentations de la batterie de tests BAWL (Batterie Attention William Lennox) [LP07]. Cette batterie est un outil informatisé d'évaluation de différentes composantes attentionnelles, qui comprend un total de 10 situations permettant de mesurer des TR. L'ensemble de ces test ne requiert cependant que l'utilisation d'un simple bouton. Afin de nous conformer au contexte du *Village aux Oiseaux*, nous avons transposé le formalisme de présentation des stimuli de BAWL à notre situation de pointage.

L'application que nous avons développée intègre une tâche de pointage présentée sur un écran PC. Les utilisateurs manipulent la souris pour sélectionner des cibles affichées les unes après les autres. Cette première expérimentation (que nous appellerons *pointage simple*) est conçue afin de calculer des valeurs moyennes du TR pour chaque utilisateur et les utiliser ensuite comme référence. La construction de cette référence a pour but de valider les valeurs du TR produites dans une tâche de pointage qui ne dispose pas de contraintes géométriques (l'expérimentation suivante).

Nous avons mis en place une condition initiale forte qui oblige l'utilisateur à maintenir sa souris immobile. Cette condition est un pré-requis à l'apparition de la cible, qui intervient alors dans un intervalle de temps situé entre 1 et 3 secondes. Le TR correspond à l'intervalle de temps entre le moment d'apparition de la cible et celui du début de mouvement volontaire de l'utilisateur. L'équation suivante décrit le TR :

$$TR = T_0 - T_{target} \quad (3.1)$$

où :

TR : le temps de réaction,

T_{target} : le moment d'apparition de cible,

T_0 : le moment du début de mouvement.

Nous avons utilisé C++ avec la librairie OpenGL pour créer notre outil. Deux fichiers de code ont été produits (représentant au total 700 lignes de codes), afin de gérer l'interaction, le recueil et l'analyse des données.

Procédure

Dans cette expérimentation, nous donnons les consignes suivantes aux sujets :

1. Garder leur curseur de souris sans aucun mouvement, dans une zone matérialisée situé au centre de l'écran.
2. Dès que la cible (un cercle blanc avec un rayon de 100 pixels) apparaît, positionner le curseur de souris dessus et cliquer aussi vite que possible.

La cible est affichée aléatoirement sur l'écran. La position de la cible est calculée de manière à ce qu'elle ne soit pas trop près de position initiale du curseur de souris (elle apparaît sur l'écran à une distance minimale de 400 pixels du curseur).

L'utilisateur doit réaliser cette procédure quinze fois, ce qui permet de calculer 15 TR par utilisateur. Le TR a été calculé en temps réel pour chaque essai. Toutes les trajectoires du curseur sont enregistrées (positions et temps) dans un fichier. L'application ne dessine aucune trajectoires, qui restent donc invisibles pour l'utilisateur pendant l'exécution de la tâche.

Quinze volontaires (9 hommes et 6 femmes) ont participé à cette expérimentation. Leurs âges varient entre 25 et 40 ans, avec un moyenne de 30,2 ans. Tous les participants sont droitiers et en bonne santé. Les expérimentations ont été réalisées sur une station de travail classique (Intel 3 GHz Core 2 Duo, 4 GB RAM) sous *Windows 7* professionnel avec un écran LCD (1920 × 1080 px). Les trajectoires sont échantillonnées à une fréquence de 100 Hz.

Résultats

La moyenne générale de TR que nous avons obtenu pour tous les utilisateurs est égale à 342 ms. L'utilisation du test Shapiro-Wilk [SW65] nous a permis de vérifier que ces résultats suivent une loi normale. Cela nous permet d'identifier un intervalle de confiance qui se situe entre 293 et 391 ms (2σ).

3.2.2 Estimation du temps de réaction par analyse géométrique

Afin de nous rapprocher des conditions finales du *Village aux Oiseaux*, nous proposons ci-après une technique d'estimation du TR dans un contexte où la valeur du T_0 de l'équation 3.1 ne peut être identifiée ni immédiatement, ni facilement. Nous envisageons donc de considérer une approximation de cet instant. Celle-ci serait déterminée par la proximité des directions de deux vecteurs : le vecteur de déplacement instantané (DVW) et le vecteur entre la position courante et la cible (DVMT).

Grâce à la prise en compte d'un système de tolérance angulaire entre ces vecteurs, il nous a été possible d'approcher plus finement l'instant à partir duquel l'utilisateur a décidé

3.2. APPROCHE GÉOMÉTRIQUE DES PHÉNOMÈNES ATTENTIONNELS

de déplacer son curseur vers la cible. Ce secteur de tolérance est construit autour du DVMT auquel on ajoute un angle de $+\alpha/2$ et $-\alpha/2$ (alpha étant fixé par ailleurs).

Le DVW est construit en utilisant une fenêtre glissante sur l'ensemble de la trajectoire effectuée par l'utilisateur. Si l'angle (β) entre les deux vecteurs DVMT et DVW est supérieur à la valeur $\alpha/2$ (état A sur la figure 3.2) nous considérons que le mouvement d'utilisateur n'est pas en direction vers la cible. Si l'angle (β) entre les deux vecteurs DVMT et DVW est inférieur à la valeur $\alpha/2$, nous considérons que le mouvement du curseur est en direction de la cible (état B sur la figure 3.2). La première occurrence de ce dernier cas, suite à la présentation d'une cible, nous permet alors de déterminer (une approximation de) la date du début du mouvement volontaire.

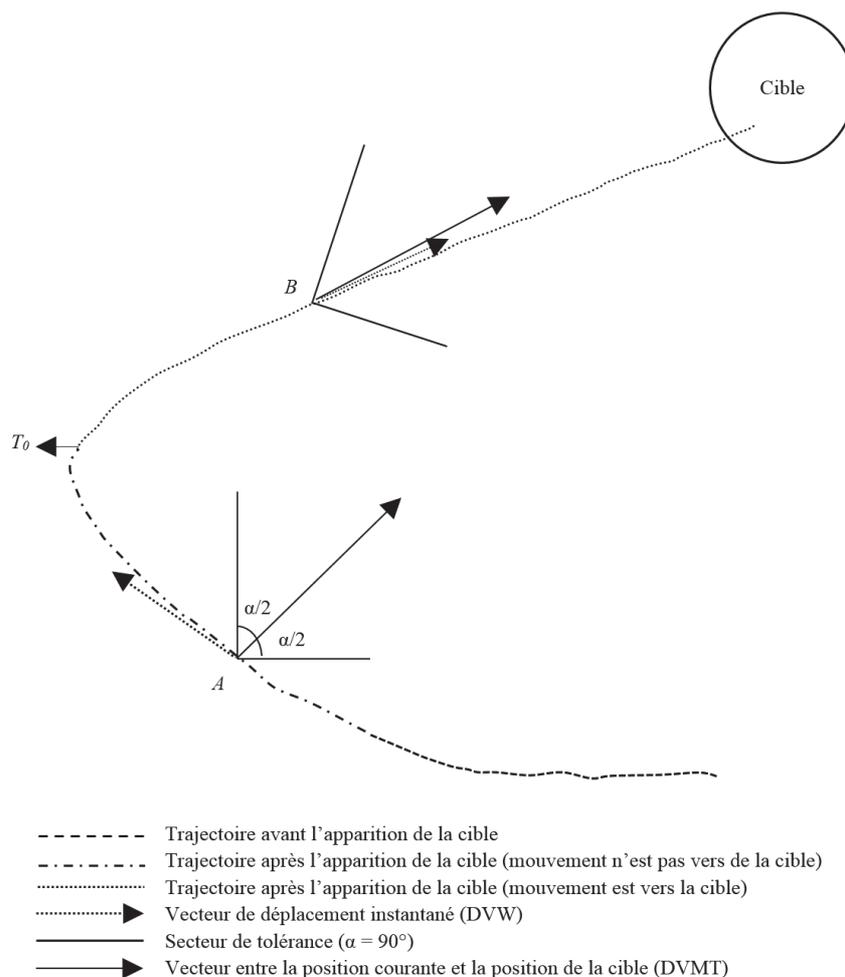


FIGURE 3.2 – Le mécanisme de protocole de pointage dynamique

Nous avons développé une application qui intègre cet algorithme (voir 3.3), dans le cadre d'une tâche de pointage. Les utilisateurs manipulent toujours une souris afin de sélectionner les cibles affichées les unes après les autres. Cette seconde situation expérimentale (que nous appellerons *pointage complexe*) doit nous permettre de tester différentes valeurs d' α afin de se rapprocher des valeurs de référence mesurées dans la condition *pointage simple*.

Algorithm 1 Calculer TR - Pseudocode

Require: target's position, trajectory's positions, α , T_{target} , $time$
for Trajectory's positions : Last position to Start position **do**
 Calculate \overleftarrow{DVW} and \overleftarrow{DVM} , and the angle (β) between them
 if $\beta < \alpha/2$ **then**
 $T_0 \leftarrow time$
 else
 Break the loop
 end if
end for
return $TR = T_0 - T_{target}$

FIGURE 3.3 – Algorithme d'analyse géométrique

L'application a été développée en C++ avec les fonctions 2D de la librairie OpenGL. Deux fichiers ont été produits (représentant au total 1000 lignes de codes), afin de gérer les interactions, le recueil et l'analyse des données.

Procédure

Contrairement à la situation expérimentale précédente, nous forçons cette fois-ci le sujet à maintenir sa souris en mouvement. Cette condition initiale est un pré-requis à l'apparition de la cible, qui intervient alors dans un intervalle de temps situé entre 1 et 3 secondes. Le T_0 est alors calculé comme indiqué précédemment. Dans cette expérimentation, nous donnons les consignes suivantes aux sujets :

1. Bouger le curseur de souris, afin de déclencher l'apparition de la cible sur l'écran.
2. Dès que la cible est visible, positionner le curseur de souris dessus et cliquer aussi vite que possible.

Nous avons gardé les mêmes réglages que ceux utilisés dans l'étape précédente, en ce

qui concerne l'interface, l'intervalle du temps, le positionnement des cibles, le non affichage des trajectoires et le nombre d'essais. Ce sont les mêmes sujets qui ont effectué le *pointage simple* qui ont participé à cette expérimentation.

Résultats

Les enregistrements nous ont permis de calculer 4 TR par utilisateur pour chaque essai, correspondant à quatre valeurs expérimentales de l'angle du secteur de tolérance : 45°, 60°, 90°, 120°(voir la figure 3.4 et le tableau 3.1).

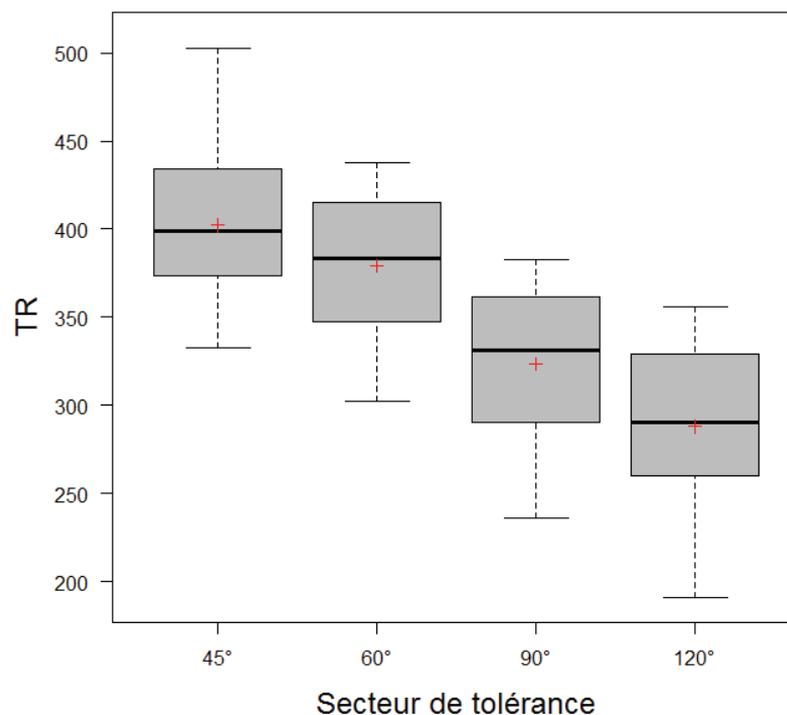


FIGURE 3.4 – Expérience de *pointage complexe* : Les distributions des moyennes de temps de réaction pour tous les sujets

Secteur de tolérance	45°	60°	90°	120°
Moyenne	402	379	324	288
Ecart-type	44	42	47	49

TABLE 3.1 – Expérience de *pointage complexe* - La moyenne des moyennes de temps réaction (en ms) pour tous les sujets

L'angle du secteur de tolérance permet de minimiser l'écart entre le TR de référence (l'expérimentation précédente) et le TR obtenu par le calcul géométrique (l'expérimentation courante). Nous avons effectué une analyse de variance (ANOVA) entre les valeurs de TR de référence (*pointage simple*) et celles de l'expérimentation courante (*pointage complexe*). Nous avons constaté que nous pouvons obtenir des valeurs acceptables du TR à partir d'un secteur de tolérance de 90° ($p = 0.186$), les valeurs de TR obtenues en utilisant un secteur de tolérance de 45° ($p < 0.0001$ ***), 60° ($p = 0.00647$ **) et 120° ($p = 0.000668$ ***) sont significativement différentes par rapports à celles de *pointage simple*.

En raison de la différence entre les situations de *pointage simple* et de *pointage complexe*, il s'est avéré nécessaire de renouveler la mesure d'une référence temporelle dans des conditions proches de cette dernière configuration. Nous avons donc conduit une troisième expérimentation (voir 3.2.2.1), ayant le même but que celle de *pointage simple*. Elle est cependant réalisée en utilisant des conditions initiales similaires à celles qui sont en vigueur dans la situation de *pointage complexe*, i.e. la souris doit être en mouvement pour que la cible apparaisse.

3.2.2.1 Mesure d'une référence contextualisée

Pour valider les valeurs de TR obtenues précédemment, nous avons mis en place un nouveau protocole (que nous appellerons *référence-bis*).

Dans cette situation, nous donnons les consignes suivantes aux sujets :

1. Bouger le curseur de souris, afin de provoquer l'apparition de la cible, aléatoirement sur l'écran
2. Dès l'apparition de cible, lâcher la souris aussi vite que possible

Le TR est calculé ici entre la date d'apparition de la cible et la date de l'arrêt des mouvements du curseur de souris.

Nous avons développé cette troisième application en C++ et OpenGL (représentant au total 700 lignes de codes), afin de gérer l'interaction, le recueil et l'analyse des données.

Résultats

La moyenne générale de TR que nous avons obtenu pour tous les utilisateurs est égale à 355 ms, avec un intervalle de confiance qui se situe entre 298 et 413 ms ($2*\sigma$).

Nous avons effectué une analyse de variance (ANOVA) entre les valeurs de TR de l'expérimentation courante (*référence-bis*) et celles de l'expérimentation précédente (*pointage complexe*). Nous avons constaté que nous pouvons obtenir des valeurs acceptables du TR à partir d'un secteur de tolérance de 60° ($p = 0.0852$), les valeurs de TR obtenues en utilisant un secteur de tolérance de 45° ($p = 0.00211$ **), 90° ($p = 0.0322$ *) et 120° ($p < 0.0001$ ***) sont significativement différentes par rapports à celles de *référence-bis*.

La figure (3.5) montre la distribution du TR. *d1* représente la condition *pointage complexe* avec un secteur de tolérance de 60° , *d2* représente la condition *référence-bis*, l'acronyme *u* représente l'utilisateur. Nous pouvons donc observer que les utilisateurs ont eu des valeurs proches du TR dans les deux expériences.

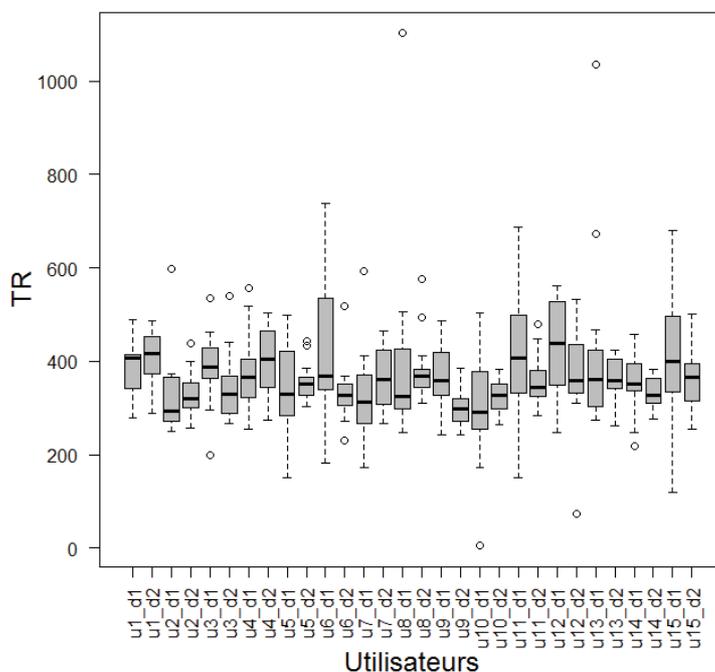


FIGURE 3.5 – Les TR(s) obtenus dans l'expérimentation de *pointage complexe* et les TR(s) obtenus dans l'expérimentation de *référence-bis*, respectivement pour chaque utilisateur.

3.2.2.2 Valeurs limites de la tolérance

Nous avons cherché à savoir jusqu'à quelle valeur d'angle de tolérance la mesure des TR demeurerait viable, ce afin d'en maximiser la précision. Nous avons donc à nouveau analysé les données obtenues lors de l'expérimentation de *pointage complexe*, en appliquant cette fois-ci des secteurs de tolérance qui vont de 60° jusqu'à 90° d'une part, et d'autre part, de 60° jusqu'à 45°.

En analysant la variance entre les valeurs de TR obtenues à partir de chaque angle parmi [45°-90°] et celles de TR de référence-bis, nous avons pu construire un intervalle d'angle situant entre 58° ($p = 0.0562$) et 88° ($p = 0.101$). Cet intervalle nous permet donc d'obtenir des valeurs correctes du TR à partir de notre algorithme d'analyse géométrique.

3.2.3 Discussion

Les mesures effectuées dans notre première expérimentation (condition *pointage simple*) nous ont permis de définir des temps de réaction moyens des sujets. L'outil développé a pu calculer que la valeur moyenne du TR était égale à 342 ms. Cette première référence a validé l'utilisation de notre outil d'analyse géométrique. En effet, nous avons pu constater que les valeurs de TR obtenus grâce à cet outil (condition *pointage complexe*), ne sont pas significativement différentes par rapport à celles calculées dans la condition *pointage simple*, pour une valeur d'angle de tolérance de 90° (moyenne à 324 ms).

Cette référence à la condition *pointage simple* a semblé néanmoins perfectible et a motivé la mise en place d'une nouvelle condition de référence (condition *référence bis*), qui intégrait les conditions initiales particulières (souris en mouvement) de la condition *pointage complexe*. Les nouvelles valeurs de référence calculées s'articulaient autour d'une moyenne de 355 ms.

Ces données nous permettent donc de considérer sous un regard nouveau les secteurs de tolérance qui peuvent être considérés comme viables. En effet, nous avons constaté que les valeurs de TR obtenus grâce à la condition *pointage complexe*, ne sont pas significativement différentes par rapport à celles calculées dans la condition *référence-bis*, pour une valeur d'angle de tolérance de 60° (moyenne à 379 ms).

3.3. CONCLUSION

Dès lors, notre limite basse n'étant plus définie, nous avons poussé nos calculs vers les valeurs de secteur de tolérance qui vont de 60° jusqu'à 90° d'un côté et de 60° jusqu'à 45° d'autre côté (i.e. $[45^\circ-90^\circ]$). Nos résultats montrent un intervalle d'angle ($[58^\circ-88^\circ]$) permettant d'obtenir des valeurs correctes du TR à partir de notre algorithme d'analyse géométrique.

3.3 Conclusion

Nous avons décrit dans ce chapitre notre démarche afin de surveiller les variations de l'attention d'un joueur de jeu vidéo. Notre démarche s'inscrivait plus précisément dans le contexte des serious games destinés à des personnes présentant des troubles attentionnels légers.

Afin de répondre à cet objectif, nous avons développé un algorithme d'analyse géométrique des trajectoires du curseur déplacé par le joueur. L'outil développé sur la base de cet algorithme nous a permis de calculer des approximations satisfaisantes du temps de réaction des joueurs, sans dispositif invasif, ni capteur particulier.

L'intégration dans une version simplifiée du *Village aux Oiseaux* a doré et déjà été effectuée et testée sur des joueurs sains. Les utilisateurs continuaient de devoir photographier tous les oiseaux présentés dans cet environnement de jeu mais, dans cette version particulière, un seul oiseau apparaissait sur l'écran à la fois. Chaque cible (un oiseau avec un rayon de 100 pixels) apparaissait de façon aléatoire sur l'écran de jeu après une période de temps qui variait entre 1 et 3 secondes. Les utilisateurs devaient cliquer sur la cible, aussi vite que possible, dès son apparition. Chaque sujet réalisait alors une session de jeu impliquant la présentation de 20 cibles.

Très encourageante, l'expérimentation réalisée avec cette version du *Village aux Oiseaux* se doit maintenant d'être étendue à une version plus complète du jeu, dans lequel de nombreux oiseaux peuvent apparaître en même temps, voler ou marcher et même disparaître avant que le joueur n'ait eu le temps de les prendre une photo. Par ailleurs, nous envisageons à court terme d'intégrer des sujets/testeurs dont les profils cognitifs correspondent à ceux de la cible initialement visée (déficit attentionnel). Il sera en effet important d'éprou-

3.3. CONCLUSION

ver la viabilité d'une telle démarche thérapeutique, originale et innovante, d'un point de vue strictement médical.

Sans se limiter à ce champs d'application thérapeutique, les outils développés dans ce contexte de serious game pourront être utilisés dans des contextes plus larges du jeu vidéo. Ils pourraient permettre notamment aux game designers de récupérer des mesures originales, utiles à des fins de réglage de la difficulté des challenges qu'ils conçoivent.

Articles publiés

Zahen Malla Osman, Jérôme Dupire, Stéphanie Mader, Pierre Cubaud and Stéphane Natkin. *Monitoring Player Attention :A Non-Invasive Measurement Method Applied to Serious Games*. International Journal of Entertainment Computing (2015).

3.3. CONCLUSION

Chapitre 4

Gestion de la difficulté par les outils d'interface logicielle

4.1 Contexte de l'étude

Dans un environnement virtuel en 3 dimensions (3D), la caméra virtuelle donne un point de vue du joueur au travers duquel il/elle perçoit cet environnement. La caméra virtuelle est utilisée dans les jeux vidéo, quel que soit leur type, pour déterminer les informations à présenter visuellement au joueur. La caméra virtuelle est notamment employée dans les mondes virtuels en 3D suivant une vue à la première personne (Fig. 4.1a) ou à la troisième personne (Fig. 4.1b).

Par exemple, les jeux de tir à la première personne (First Person Shooter, FPS), sont des jeux en vue subjective, où le joueur voit l'action à travers les yeux du personnage-joueur, tandis que les jeux à la troisième personne (Third Person Shooter, TPS), sont des jeux en vue objective, où le joueur voit le personnage-joueur de manière externe. Un jeu à la première personne permet une vue de l'environnement du jeu, sans aperçu du personnage-joueur au premier plan, tandis qu'un jeu à la troisième personne, montre l'action du personnage prise dans le dos.

Si la caméra virtuelle peut être gérée automatiquement par le système (voir chapitre précédent), elle est cependant souvent contrôlée directement par le joueur. Dans un FPS, les utilisateurs explorent l'environnement du jeu en contrôlant l'orientation et la position de la caméra virtuelle. En déplaçant/orientant cette caméra, l'utilisateur peut accéder à

4.1. CONTEXTE DE L'ÉTUDE



FIGURE 4.1 – Exemples de jeux vidéo : (a) Half Life, (b) Assassins Creed.

toutes les informations disponibles dans l'environnement du jeu pour réaliser une tâche requise. Sa capacité à prendre ces informations va donc dépendre directement de deux facteurs : la qualité/quantité des mouvements de caméra qu'il effectue et les caractéristiques intrinsèques de la caméra virtuelle (notamment, son champ de vision - FOV).

Dans le cas où l'utilisateur contrôle la caméra virtuelle, les mouvements de celle-ci sont les conséquences directes de stratégies cognitives adoptées par le joueur. Leur analyse peut donc nous permettre a posteriori de décrypter ces stratégies ou nous donner des indicateurs chiffrés sur les performances ou les contraintes subies par le joueur, qui sont autant de représentations de la difficulté vécue par lui [Bur13].

Il serait possible de savoir, par exemple, combien de déplacements il/elle effectue pour réaliser une tâche donnée, d'en calculer un niveau de performance/difficulté et d'ajuster le contexte de cette tâche pour minimiser l'écart entre cette difficulté mesurée et la difficulté visée.

Le contexte de ce travail n'est cependant pas aussi simple que décrit précédemment. En effet, un utilisateur peut percevoir la scène 3D d'un environnement virtuel via deux modalités : (1) en déplaçant/orientant la caméra virtuelle pour générer des points de vue différents ou (2) en déplaçant son regard sur l'écran afin d'explorer la scène virtuelle courante (sans modification de celle-ci).

Prenons par exemple la figure 4.2, nous pouvons voir une scène en 3D du jeu "Halo 2"

4.1. CONTEXTE DE L'ÉTUDE

4.2a, nos comportements oculaires permettent de visualiser cette scène sur le plan d'écran (voir Fig. 4.2b), tandis que les comportements de la caméra virtuelle permettent de changer le point du vue dans cette scène (4.2c).

Dans notre travail, nous nous limiterons à l'étude de la première modalité, qui doit nous permettre, à terme, d'extraire des indicateurs et/ou des leviers pour agir sur la difficulté d'un jeu/niveau de jeu.

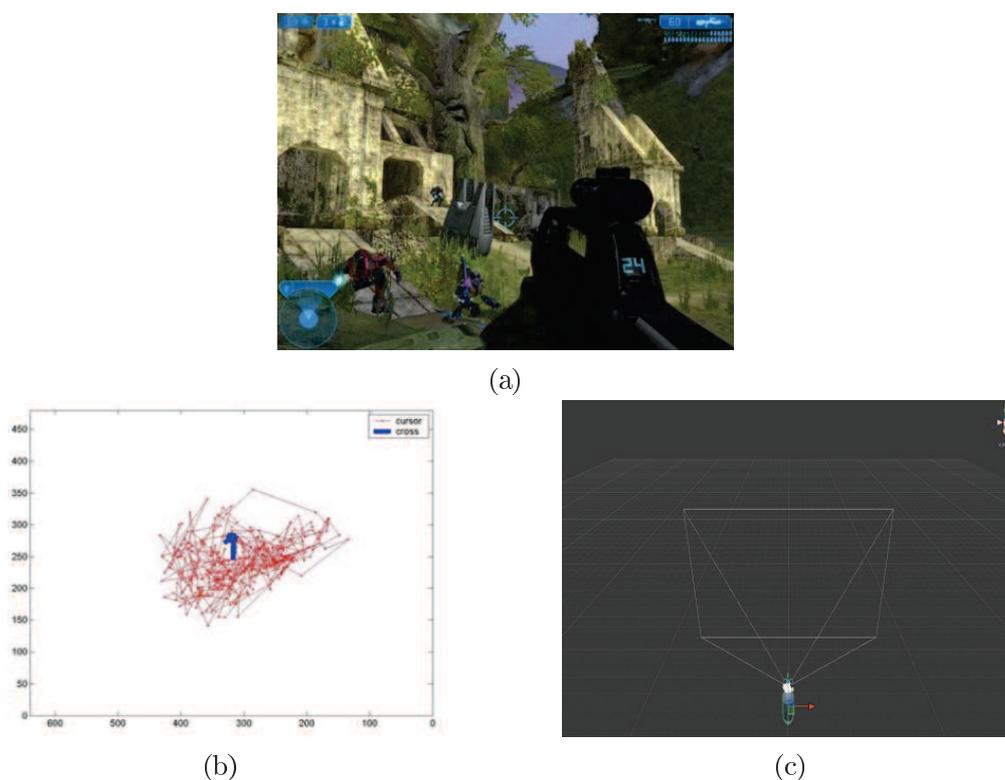


FIGURE 4.2 – Différence entre le comportement oculaire et celui de la caméra virtuelle : (a) capture d'écran de Halo-2, (b) trajectoires du regard via un système d'oculométrie, (c) caméra virtuelle son cône de vision.

Le cœur de notre démarche repose sur le rapprochement comportemental entre la vision humaine (médiatisée par l'œil) et la vision synthétique (médiatisée par la caméra virtuelle). Nous proposons donc de transposer les méthodes d'analyse de la vision humaine (oculométrie) vers la vision synthétique. Ainsi, des indicateurs traditionnellement utilisés en oculométrie vont être appliqués pour caractériser les comportements d'une caméra virtuelle contrôlée par les joueur.

Dans la suite de ce chapitre, nous présentons en détail notre démarche d'évaluation de la difficulté par la caractérisation des comportements de la caméra virtuelle à partir de techniques utilisées en oculométrie. Nous décrivons ensuite l'expérience que nous avons menée afin de moduler la difficulté d'une tâche en agissant uniquement sur les paramètres intrinsèques de la caméra. Nous discuterons enfin les résultats et de leur utilisation dans un contexte de conception de jeu vidéo.

4.2 Le modèle comportemental de la caméra virtuelle

4.2.1 Variables descriptives

Nous avons sélectionné plusieurs mesures de l'attention visuelle, employées par Gibbs et al. [GB09], afin d'étudier le comportement d'utilisateur. Gibbs et al. ont utilisé un système d'oculométrie afin de déterminer si le comportement oculaire d'un utilisateur diffère entre l'usage des sites web de journaux et des sites web de télévision. Parmi les mesures sélectionnées par l'étude de Gibbs et al. est le nombre de fixations (NF) et la durée moyenne des fixations (FD). Nous reprendrons ces critères en modifiant la terminologie afin de bien nous détacher des aspects de vision humaine (voir le tableau 4.1).

Nous avons également introduit des nouvelles mesures visant à prendre en compte les spécificités de la caméra virtuelle (voir le tableau 4.1). En effet, la façon dont la caméra virtuelle est manipulée par l'utilisateur diffère de la manière dont celui-ci contrôle ses mouvements oculaires. Par exemple, la durée d'une saccade oculaire est relativement courte et stable entre occurrences alors que les mouvements de la caméra virtuelle peuvent être beaucoup plus longs et variables.

Les mesures que nous avons ajoutées sont donc : la durée moyenne des regards (GD)¹, le nombre de pauses (NG), le nombre de mouvements (NM), la durée moyenne des mouvements (MD), la somme des durées des fixations effectuées par tâche (STDF), la somme des durées des pauses effectuées par tâche (STDG), la somme des durées des mouvements effectués par tâche (STDM) et la durée totale de chaque tâche (TD).

1. Les fixations sont des pauses dans les mouvements oculaires se déroulant sur une durée minimale. Les regards sont des pauses prolongées, utilisées pour regarder un élément [GB09] [JK03].

Mesure	Comportement de la caméra
Fixation	Immobile entre 100 et 300 ms.
Pause	Immobile pendant plus de 300 ms.
Mouvement	Déplacement supérieur à 100 ms ² .

TABLE 4.1 – Les mesures extraites à partir de la camera virtuelle

4.2.2 Le champ de vision de la caméra virtuelle

Le champ de vision (Field of View, FOV) de la caméra virtuelle joue un rôle essentiel dans la conception d'un jeu 3D. La taille de FOV est l'une des paramètres qui semble pouvoir influencer de manière significative l'extraction et l'identification des informations [BBR⁺88],[MS09], parce qu'elle impacte directement la densité visuelle de la surface sur laquelle l'utilisateur porte son attention.

Les choix de réglage de cet objet posent un certain nombre de questions fondamentales : Doit-on copier le champ de vision de l'œil humain et l'appliquer à la caméra virtuelle ? Si non, un secteur de FOV large (voir Fig. 4.3a) va-t-il aider les utilisateurs à récupérer plus d'informations et effectuer leurs tâches plus rapidement ? Au contraire, cette prolifération d'information peut-elle noyer l'information importante et dégrader la performance ? Inversement, comment un réglage de secteur réduit de FOV (voir 4.3b) va-t-il modifier la perception du joueur et potentiellement affecter ses performances ?

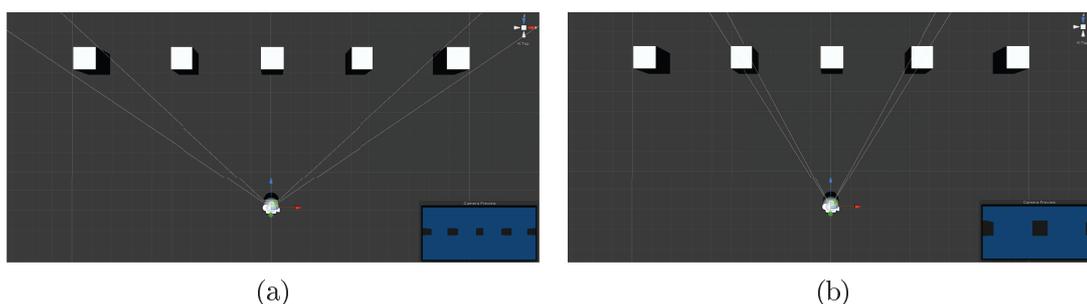


FIGURE 4.3 – Visualisation d'une scène en utilisant deux différentes tailles de FOV : (a) secteur large, (b) secteur réduit

Pour cela, nous avons conçu et mené une expérience dans laquelle nous étudions les effets de la variation de FOV sur les performances des utilisateurs, en utilisant le modèle précédemment décrit pour évaluer et comparer les différentes conditions.

4.2.3 L'environnement virtuel 3D

Nous avons construit un environnement virtuel en 3 dimensions (3D) (Fig. 4.4), afin de mener notre expérimentation. Cet environnement a été inspiré par les travaux de Lee et al. [LKC07]. Celui-ci proposait un environnement virtuel dans lequel les utilisateurs avaient pour tâche de retrouver des nombres imprimés sur les objets de la scène (dans une métaphore de galerie d'art, voir 2.6.4 pour plus d'informations). Nous avons repris les caractéristiques de cette expérience qui correspondait pleinement à notre contexte (3D, prise d'information, tâche simple). Nous avons cependant remplacé la recherche de nombre par une recherche de bouton que les sujets devaient trouver et actionner. Chaque objet 3D de la scène contenait un bouton caché. Un compteur affiché en haut et à gauche de l'écran indiquait le nombre de boutons restant à valider.



FIGURE 4.4 – Notre environnement virtuel 3D

Nous avons utilisé le moteur de jeu *Unity3D*³ pour construire notre environnement, y

3. <https://unity3d.com/>

compris les interfaces graphiques et les objets 3D. Ce moteur de développement puissant et flexible, permet en effet de créer des jeux vidéo et des expériences interactives de manière rapide et assez simplifiée. Nous avons donc utilisé C# (intégrée en Unity3D) pour gérer la partie de programmation. Trois scripts ont été élaboré (800 lignes de codes), afin de calculer les mesures considérées, gérer l'environnement et le contrôle de la caméra virtuelle.

Les participants ont réalisé notre expérimentation en passant par trois étapes :

1. Tout d'abord, ils/elles ont été invité(e)s à remplir un court questionnaire pour fournir des informations comprenant le nom, l'âge, le sexe et leur expertise en jeu vidéo de type FPS (expérimenté ou débutant).
2. Ensuite, ils/elles bénéficiaient de 2mn pour naviguer librement dans l'environnement. Cette étape a été créée pour apprendre à manipuler la caméra virtuelle, qui étaient réglée sur un FOV de 80°.
3. Enfin, ils/elles effectuaient une tâche de recherche pour trouver et valider des objets repartis de façon aléatoire dans l'environnement. Les mesures précédemment décrites, temporelles et cinématiques, ont alors été enregistrées.

Il a été demandé aux participants de répéter cette tâche six fois. Nous avons changé les positions de tous les objets, ainsi que la taille de FOV de la caméra virtuelle à chaque essai (10° pour le premier essai, 20° pour le deuxième, 30° pour le troisième, 50° pour le quatrième, 80° pour le cinquième, et 110° pour le sixième essai, voir la figure 4.5).

Le but de modifier la taille de FOV est pour découvrir comment le FOV influe le comportement d'utilisateur. Le FOV classiquement utilisé dans un FPS varie de 75° à 110°.

Les participants utilisaient une souris et un clavier (*AZERTY*) pour évoluer dans la scène 3D comme ils le feraient dans un FPS (voir le tableau 4.2). Dans cette configuration adoptée comme un standard dans les FPS, l'utilisateur contrôle directement la position de la caméra virtuelle ainsi que son orientation : les déplacements sont réalisés grâce aux touches du clavier (à gauche, à droite, en avant et en arrière) et les rotations sur les plans horizontal et sagittal (*yaw* et *pitch*) en utilisant la souris.

Un total de 14 volontaires (10 hommes et 4 femmes) ont participé à notre expérimentation. Leurs âges varient entre 25 et 42 ans, avec une moyenne de 30. Tous les participants

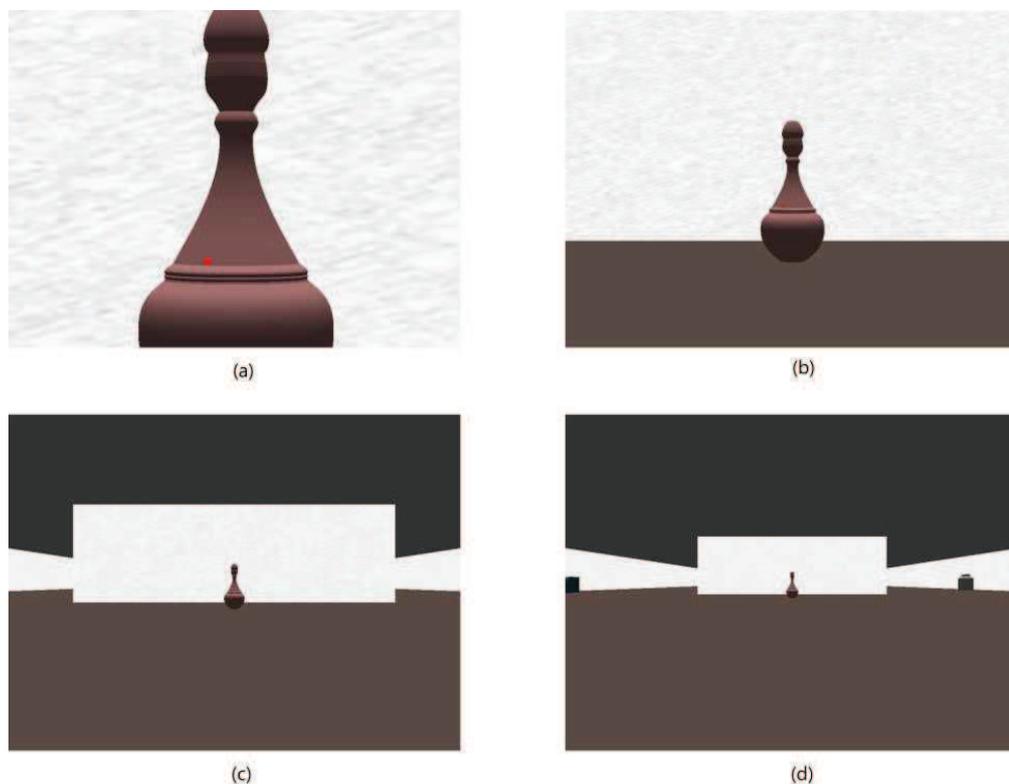


FIGURE 4.5 – Quatre différents points de vue de l’environnement pour une même position de la caméra virtuelle : (a) FOV de 10° , (b) FOV de 30° , (c) FOV de 80° , (d) FOV de 110° .

étaient droitiers et en bonne santé. Les expérimentations ont été réalisées sur une station de travail classique (Intel 3 GHz Core 2 Duo, 4 GB RAM) sous *Windows 7* professionnel avec un écran LCD d’une résolution de 1920×1040 pixels.

Contrôle	Actions
Touches : Q, D, Z, S	Déplacer l’avatar (sur les axes X,Y)
Touche : Espace	Cliquer sur un objet (valider)
Mouvement de la souris	Orienter la camera virtuelle (sur l’axe Z)

TABLE 4.2 – Contrôler la camera virtuelle en utilisant le clavier et la souris

4.2.4 Présentation et analyse des résultats

Nous avons tout d’abord testé la normalité de nos mesures en utilisant le test de Shapiro-Wilk [SW65]. Les résultats ont montré que les mesures ne suivaient pas une loi normale. Nous avons donc transformé nos mesures par le biais d’un logarithme de base 10. Ce passage

4.2. LE MODÈLE COMPORTEMENTAL DE LA CAMÉRA VIRTUELLE

ne modifie pas la nature de nos mesures et permet d'éliminer les bruits produits dans les mesures et d'obtenir des résultats qui peuvent suivre une loi normale. Nous avons donc ré-testé la normalité de nos mesures (après cette transformation) en utilisant le test de Shapiro-Wilk et confirmé que nos résultats suivaient une loi normale, ce qui nous a ensuite permis d'effectuer une analyse de variance (ANOVA).

Le tableau 4.3 montre les résultats de l'analyse de variance sur l'ensemble de nos mesures⁴ en fonction du FOV.

	10°	20°	30°	50°	80°	110°	<i>p</i>
NF	2.15	1.84	1.78	1.66	1.60	1.57	<i><0.0001</i> ***
NG	1.93	1.66	1.50	1.45	1.43	1.40	<i><0.0001</i> ***
NM	2.35	2.06	1.96	1.87	1.82	1.79	<i><0.0001</i> ***
FD	2.23	2.23	2.21	2.22	2.22	2.21	<i>0.358</i>
GD	2.90	2.95	2.93	2.89	2.88	2.90	<i>0.489</i>
MD	2.44	2.53	2.49	2.51	2.57	2.47	<i>0.838</i>
STDF	4.37	4.07	3.99	3.89	3.82	3.78	<i><0.0001</i> ***
STDG	4.83	4.61	4.43	4.33	4.31	4.30	<i><0.0001</i> ***
STDM	4.79	4.59	4.45	4.37	4.39	4.26	<i><0.0001</i> ***
TD	5.20	4.99	4.85	4.76	4.75	4.69	<i><0.0001</i> ***

TABLE 4.3 – Résultats de l'expérimentation : les données numériques présentées dans les colonnes correspondants aux différents angles de champ de vision représentent les logarithmes des moyennes des valeurs mesurés pour chaque sujet, le dernier colonne (*p*) représente les résultats de l'analyse de variance (ANOVA)

L'analyse de variance présentée dans le tableau précédent, nous a permis de noter qu'il existe une différence significative en ce qui concerne les mesures : NF, NG, NM, STDF, STDG, STDM et TD, en fonction du FOV.

A partir de cela, nous avons observé que ces mesures diminuent lorsque le FOV augmente (voir Fig. 4.6). Ce résultat nous permet de constater que les utilisateurs ont déplacé la caméra virtuelle moins fréquemment pendant l'exécution de la tâche requise et qu'ils/elles

4. NF : le nombre de fixations, NG : le nombre des pauses, NM : le nombre de mouvements, FD : la durée moyenne des fixations, GD : la durée moyenne des pauses, MD : la durée moyenne des mouvements, STDF : la somme des durées des fixations effectuées par tâche, STDG : la somme des durées des pauses effectuées par tâche, STDM : la somme des durées des mouvements effectués par tâche, TD : la durée totale pour compléter une tâche requise(TD).

avaient besoin de moins de temps pour accomplir cette tâche quand ils/elles utilisent une grande taille de FOV.

Une explication possible réside dans un équilibre dans les stratégies des utilisateurs, qui ont à leur disposition deux modalités complémentaires pour prendre de l'information dans l'environnement : d'une part, explorer la vue courante (mouvements oculaires sur l'écran) et, d'autre part, générer une autre vue de la scène 3D (mouvement manuels/digitaux). Un FOV plus important permet d'avoir visuellement accès à des zones plus étendues, qui devraient, dans le cas d'un FOV plus réduite, être explorées par des déplacements successifs de la caméra virtuelle.

Par ailleurs, nous avons constaté visuellement que la distribution de nos mesures semblait converger à partir d'un FOV de 30° dans le sens des FOV plus larges (Fig. 4.6).

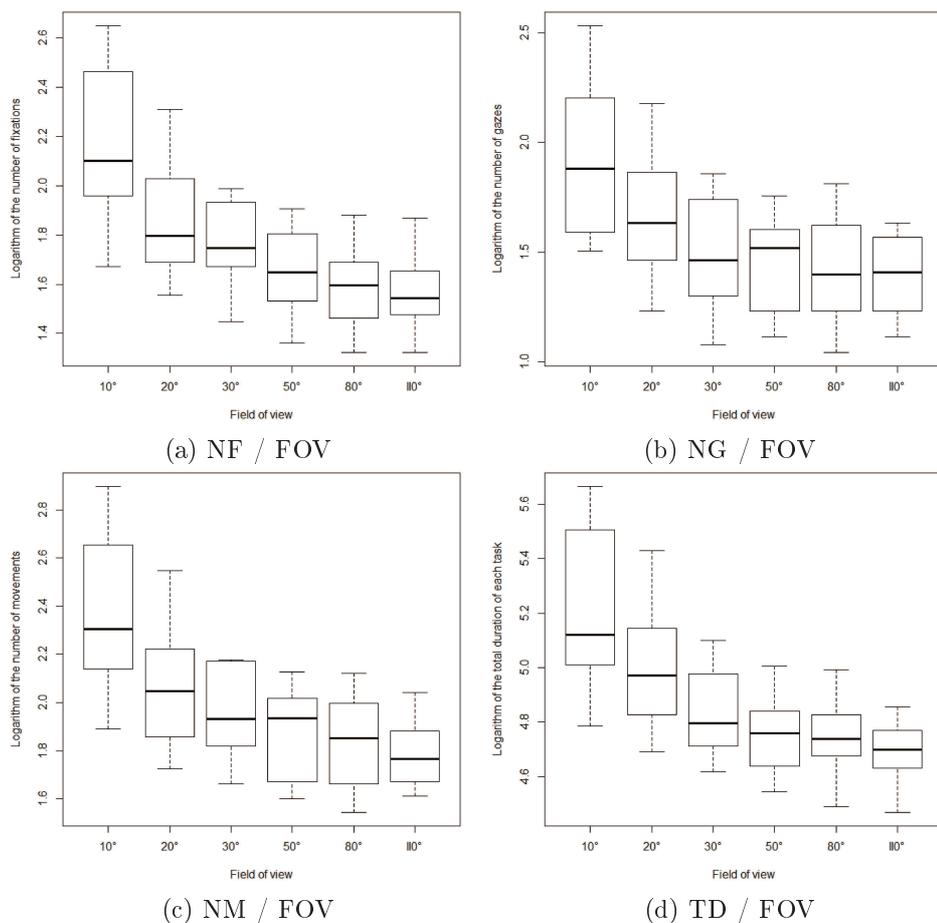


FIGURE 4.6 – La distribution des mesures en fonction du FOV.

4.2. LE MODÈLE COMPORTEMENTAL DE LA CAMÉRA VIRTUELLE

Cependant, il n'existe pas de différence significative entre les mesures de FD, GD et MD. Cela se traduit par des durées stables de fixations, pauses et mouvements lors de l'exécution de la tâche de recherche visuelle et indépendantes du FOV. Cette indépendance de ces durées vis à vis du FOV pourrait s'expliquer par le fait que l'exploration d'une scène sur une surface physique (écran) occupe un temps globalement identique. Cette exploration semble donc ne pas dépendre de ce qui est présenté (densité, type) mais plutôt de la taille de la surface à parcourir. La surface correspondante, dans notre configuration, est constituée par l'écran de l'ordinateur.

Nous avons ensuite exploré la répartition des différents comportements (Fixation, Pause et Mouvement) utilisés dans la réalisation d'une tâche. Pour cela, nous avons considéré deux groupes de sujets en tenant compte de leur expertise en jeu vidéo de type FPS : d'une côté, les *joueurs* et de l'autre, les *non joueurs*.

La figure 4.7 montre les pourcentages de temps passé dans chaque comportement par les joueurs pour effectuer la tâche requise, en fonction du FOV. Ces pourcentages sont exprimés par les durées des fixations, des pauses et des mouvements (STDF, STDG et STDM).

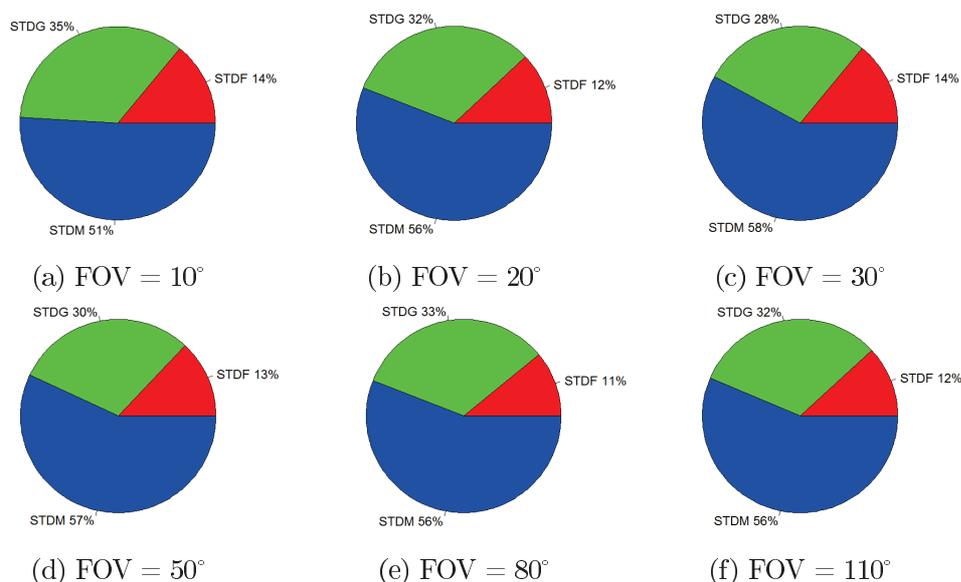


FIGURE 4.7 – Comportements des joueurs

Nous avons observé que le comportement "mouvement" domine (temporellement) au

4.2. LE MODÈLE COMPORTEMENTAL DE LA CAMÉRA VIRTUELLE

sein de chaque tâche. Cette stratégie qui favorise les mouvements au détriment de fixation et de pauses peuvent se comprendre par l'habitude et la connaissance suffisante pour contrôler la caméra virtuelle (en utilisant la souris et le clavier) et leur capacité à déplacer/orienter la vue tout en réalisant en parallèle la tâche de recherche sur les images affichées sur l'écran (parallélisme des tâches). Cette attitude de "recherche active" se retrouve dans les comportements usuels des joueurs de FPS, pour lesquels une immobilité est souvent synonyme de vulnérabilité, d'une part et de perte de temps, d'autre part.

L'analyse de ces comportements dans la population de non joueurs a donné des profils très différents. Les stratégies utilisées, contrairement aux joueurs, ne sont pas dominées par les mouvements (Fig. 4.8).

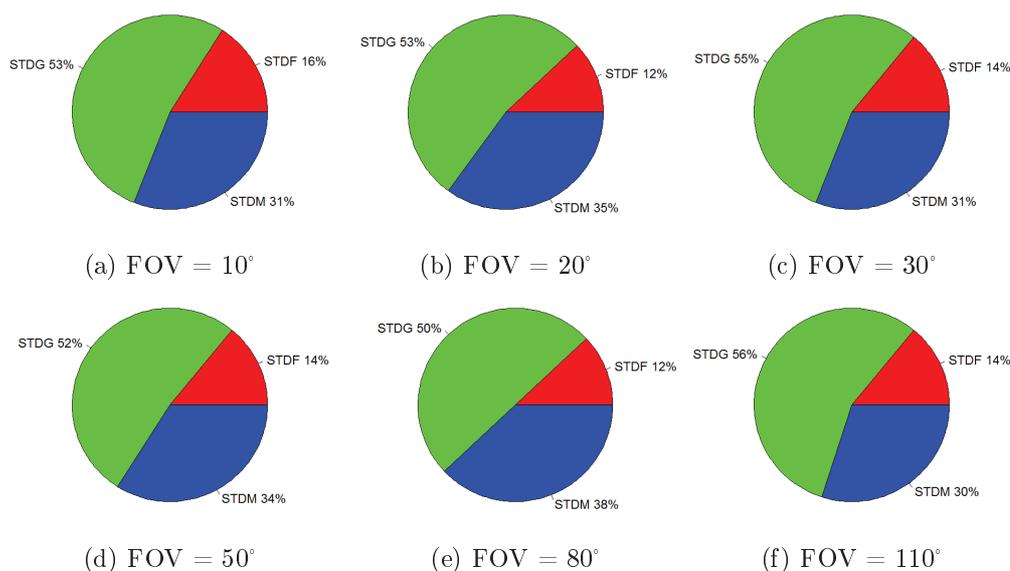


FIGURE 4.8 – Comportements des non joueurs

Nous observons qu'ils/elles ont effectué plus de pauses au sein de chaque tâche. Si les temps de fixation (STDF) ne sont pas significativement différents entre les deux population ($p = 0.26$ pour le FOV de 10° , $p = 0.74$ pour le FOV de 20° , $p = 0.64$ pour le FOV de 30° , $p = 0.13$ pour le FOV de 50° , $p = 0.58$ pour le FOV de 80° et $p = 0.36$ pour le FOV de 110°), nous notons une inversion flagrante des ratios des pauses (STDG) et des mouvements (STDM). Ce phénomène semble pouvoir être expliqué par une moins bonne dextérité quant à la manipulation de la souris couplée à la caméra et la conséquence directe

qui est la réalisation séquentielle des tâches de déplacement/orientation et de recherche. La stabilité de l'image semble donc recherchée par les non joueurs pour pouvoir analyser l'image et identifier les informations, contrairement aux joueurs.

4.2.5 Discussion

Les mesures effectuées dans notre expérimentation nous ont permis de savoir que la quantité d'informations présentée à l'écran peut perturber l'utilisateur pendant l'exécution d'une tâche de recherche, dans un environnement virtuel type FPS. Nous avons constaté qu'il existe une différence significative dans la performance d'utilisateur en fonction de la variation du FOV de la caméra virtuelle.

Nous avons observé que les utilisateurs qui utilisent un FOV réduit, rencontrent des difficultés lorsqu'ils/elles effectuent les tâches demandées. Par contre, en utilisant des valeurs plus large de FOV, les utilisateurs ont pu réaliser les tâches demandées de manière plus rapide.

Les mesures de FD, GD et MD sont indépendantes du FOV et semblent plutôt dépendre de la taille de la surface (écran) à parcourir. Il serait utile d'aller plus loin et de tester cette hypothèse, en menant une expérimentation dans laquelle nous proposons plusieurs tailles d'écran, afin de vérifier si celles-ci modifient de manière significative ces mesures. Ces résultats peuvent être aussi utiles dans la compréhension d'un contexte où l'utilisateur doit chercher et trouver des informations complexes (comme dans la lecture numérique).

Enfin, nous avons aussi réalisé une comparaison entre les joueurs et les non joueurs au regard de la durée de la tâche (TD, temps total de chaque session). Nous constatons sur la figure 4.9 que les joueurs sont plus rapides à réaliser les tâches requises que les non joueurs, peu importe le FOV considéré. Cette différence va de 70% plus rapide que les non joueurs (avec un FOV de 10°) à 19% (FOV de 110°). Cette différence a donc tendance à se réduire pour des valeurs de FOV plus grande et semblent converger vers des valeurs, sinon identiques, proches.

Ces différentes données apportent un éclairage sur la manière donc les caractéristiques de la caméra virtuelle influencent la difficulté d'une tâche donnée. Du point de vue des

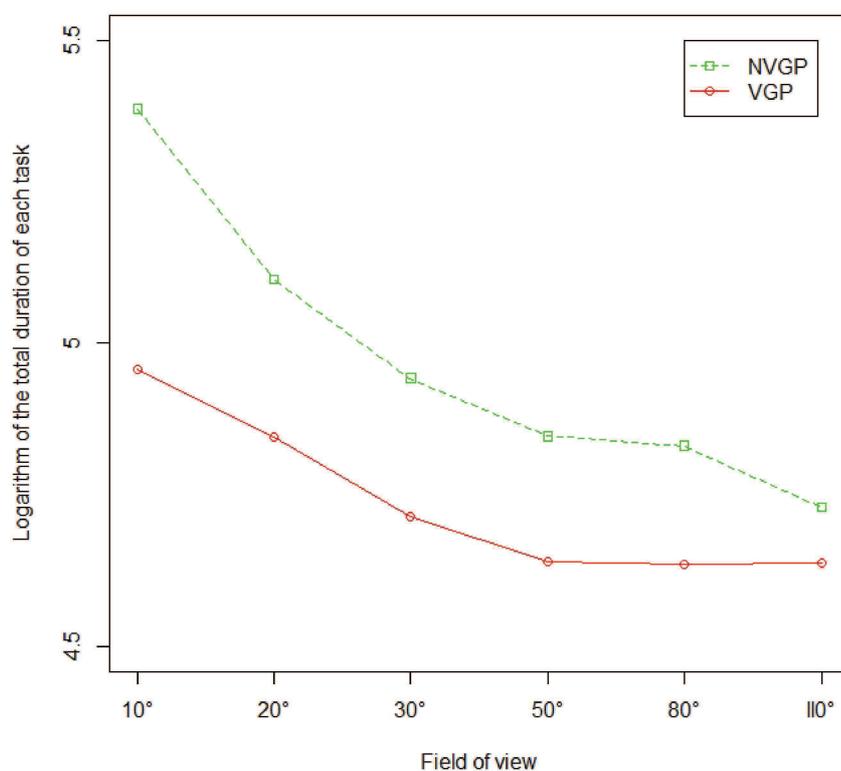


FIGURE 4.9 – La durée totale TD de chaque tâche, pour les joueurs (VGP) et les non-joueurs (NVGP)

concepteurs de jeu vidéo, ces informations peuvent avantageusement être utilisées pour, d'une part, ajuster un niveau de difficulté sans modifier une tâche ou un challenge particulier du jeu et, d'autre part, prendre en compte les effets de la caméra virtuelle sur des populations aux compétences différentes.

4.3 Conclusion

Nous avons décrit dans ce chapitre notre démarche d'évaluation de la performance d'un sujet dans les jeux vidéo. Notre démarche s'inscrivait plus particulièrement dans le contexte des expériences interactives utilisant des environnements virtuels avec une vue à la première personne.

L'idée forte à la base de notre démarche était d'effectuer un rapprochement entre les comportements rencontrés en vision naturelle (au travers de l'œil, notamment) et la vision synthétique (au travers d'une caméra virtuelle). Ce rapprochement nous a permis de

4.3. CONCLUSION

construire des modèles comportementaux de joueur, basés sur leurs manipulation de la caméra virtuelle.

Nous avons pour cela transposé les mesures traditionnellement utilisés en oculométrie, afin de caractériser le comportement de la caméra virtuelle. Nous avons pu donc analyser la façon dont l'utilisateur contrôle cette caméra, afin d'étudier comment la quantité d'information présentée à l'écran pouvait modifier les performances d'un joueur.

Bien que perfectible, notre modèle constitue d'ores et déjà une base exploitable par les *game designers* pour comprendre les comportements et/ou ajuster des paramètres de jeu vidéo en 3D.

Articles publiés

- Zahen Malla Osman, Jérôme Dupire, Alexandre Topol and Pierre Cubaud. *Non Intrusive Measures for Determining the Minimum Field of View for User Search Task in 3D Virtual Environments*. Proc. of the Seventh International Conference on Advances in Computer-Human Interactions, pp. 208-213, 2014.
- Zahen Malla Osman, Jérôme Dupire, Alexandre Topol and Pierre Cubaud. *A Non Intrusive Method for Measuring Visual Attention Designed for the Study and Characterization of Users' Behavior in Serious Games*. International Journal on Advances in Internet Technology, vol(7), 3&4, pp. 262-271, 2014.

Chapitre 5

Gestion de la difficulté par les outils d'interface matérielle

5.1 Contexte de l'étude

La *Wii U* de Nintendo (Fig. 5.1a) est l'un de des systèmes de jeux vidéo qui propose une périphérique doté d'un écran tactile intégré, connecté au système classique (console, gamepad). Cet écran interactif "secondaire" peut être synchronisé avec d'autres périphériques (comme un téléviseur). Cette configuration innovante offre une autre dimension des expériences interactives, dans laquelle est exploitée une relation asymétrique : les joueurs ne disposent pas nécessairement des mêmes informations sur l'écran périphérique et sur l'écran principal.

Dans la même perspective que la *Wii U*, le *Xbox SmartGlass* de Microsoft nous permet d'utiliser des dispositifs mobiles (tablettes, smartphones) en les connectant à la console *Xbox* pour les inclure dans l'expérience interactive (Fig. 5.1b). Ces systèmes déplacent les frontières traditionnellement établies entre les joueurs, les jeux et les plateformes de jeu.

La plupart des jeux conçus pour utiliser ces systèmes considèrent l'écran secondaire comme étant "complémentaire de l'interaction" à un écran principal. Dans ce type des jeux, comme *ZombiU* (Ubisoft, 2012), nous avons la possibilité d'effectuer certaines tâches en utilisant l'écran secondaire, tout en maintenant l'interaction principale au travers de l'écran classique. Ces tâches peuvent consister en une visualisation des résultats statistiques, une consultation de cartes ou une ouverture des options. Elle peuvent aussi aller plus loin

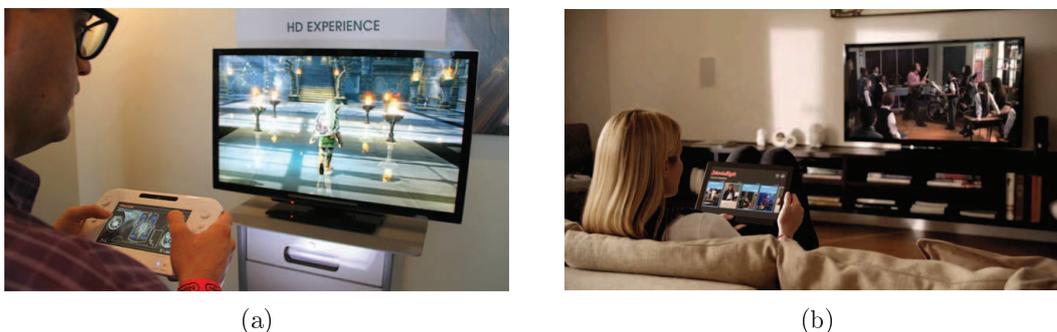


FIGURE 5.1 – La *Wii U* de Nintendo (a), le *Xbox SmartGlass* de Microsoft (b).

dans le sens d'une interaction plus riche comme la manipulation de la caméra virtuelle (déplacement, orientation), le lancement de projectile ou le pointage d'objet.

De plus, les dispositifs mobiles (smartphones, tablettes) dotés des dimensions relativement réduites, sont conçus pour être généralistes disposant, aujourd'hui, d'une très bonne puissance de traitement, d'espace de stockage confortables, tout en étant moins consommateur d'énergie. Il est devenu courant de regarder des vidéos, consulter des emails, surfer sur internet et même travailler sur des documents grâce à ces plateformes récentes.

Cette apparition récente dans notre quotidien a provoqué de nouveaux comportements et usages, liés à ces nouvelles opportunités. Il n'est en effet plus rare de manipuler son téléphone portable autour d'une (voire plusieurs applications) tout en regardant une émission télévisée (voire plusieurs) ou utiliser son PC. En effet, regarder l'écran d'un téléviseur est une action relativement passive, qui laisse nos mains libres pour tenir et manipuler un autre appareil (une tablette par exemple).

La question de l'attention et de sa distribution est, en revanche, beaucoup plus complexe à appréhender. En effet, cette configuration nous oblige à considérer deux écrans de manière simultanée, tout au moins de devoir mettre en place un arbitrage pour déterminer, à chaque instant, lequel sera l'objet de l'attention. Cet arbitrage repose sur des probabilités, des pondérations établies a priori et/ou des événements perçus en temps réel.

Cette action apparaît comme étant plus sollicitante de nos processus cognitifs, et donc, plus difficile et plus fatigante. La question de l'altération possible des performances de l'utilisateur se pose donc légitimement face à ces nouveaux usages. Qu'est-ce qui se passe

sur les plans de l'attention et de la performance en utilisant deux surfaces d'affichage ?
Peuvent-elles perturber l'exécution de nos tâches ? Y a-t-il des effets de la répartition de l'attention sur la performance ?

5.2 Hypothèse de l'étude

Plusieurs facteurs peuvent être considérés afin de caractériser une expérience interactive engageant plusieurs dispositifs.

Notre hypothèse est que la multiplication des dispositifs (à la fois en termes d'affiche et de commande) altère nos performances et amène à des temps de réaction plus longs, voire à des échecs.

Dans ce contexte, nous considérons trois facteurs : la configuration des écrans (nombre et mise en place), la difficulté de la tâche demandée et l'ordonnancement des informations sur les écrans.

L'accessibilité des écrans (prise en main ou posé sur la table) peut être aussi considérée pour évaluer la performance de l'utilisateur. Ce facteur se repose en effet sur de critères de confort, propres aux utilisateurs, comme l'emplacement des deux écrans l'un par rapport à l'autre.

La difficulté de la tâche demandée présente, elle aussi, une autre dimension dans notre hypothèse. Plus la tâche est difficile, plus son traitement risque d'être prolongé dans le cadre d'une configuration complexe avec plusieurs dispositifs (vs. une situation classique impliquant un dispositif unique).

La probabilité d'apparition des informations (ordonnancement) est aussi à prendre en compte lorsque nous développons une interface répartie sur plusieurs écrans (comme la distribution des éléments *gameplay* dans le contexte des jeux vidéo). Cet ordonnancement peut être soit anticipable, soit aléatoire. Dans ce dernier cas, l'utilisateur devra suivre les écrans en parallèle, ce qui peut rendre l'identification d'information plus complexe.

5.3 Expériences

5.3.1 Test préliminaire

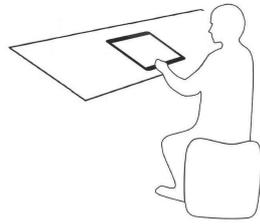
Nous avons tout d'abord réalisé un test préliminaire dont le but de déterminer quelles configurations d'écran doivent être utilisées dans notre expérimentation. Nous avons choisi de tester seize configurations en fonction de nombre d'écrans et de leur disposition. L'idée sous jacente était de simuler le nombre maximum de configurations à deux écrans, utilisés dans les systèmes des jeux vidéo. Nous avons groupé les configurations en cinq groupes :

- Configurations Mono-écran fixe (Fig. 5.2)
- Configurations Mono-écran mobile (Fig. 5.3)
- Configurations Multi-écrans fixe (Fig. 5.4)
- Configurations Multi-écrans mobile (Fig. 5.5)
- Configurations Multi-écrans fixe et mobile (Fig. 5.6)

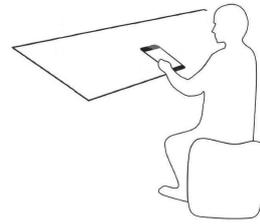
Pour cela, nous avons utilisé un écran PC (LCD, 27 puces), deux tablettes tactiles (Samsung, 8 puces) et un smartphone (iPhone 4, 3.5 puces). Cinq participants ont participé à ce test préliminaire, ils/elles ont été installés sur un poste de travail constitué d'une chaise et d'une table. Leur tâche a consisté de toucher l'écran (ou l'un des écrans, dans le cadre d'une configuration multi-écran) lorsqu'un stimuli (carré blanc) apparaît sur l'écran. Nous leur avons demandé de noter, sur un feuille, leur degré de satisfaction par rapport à chacune des configurations proposées. Notre référentiel se basait sur une échelle [Lik31] permettant d'évaluer de un à dix chacune des configurations proposées¹. Les résultats obtenus nous ont permis de ne retenir que les quatre configurations ayant obtenu les scores maximum :

- Une configuration *mono-écran* : un smartphone pris en main
- Trois configurations *multi-écrans* :
 - deux tablettes posées l'une à côté de l'autre sur la table (en orientation portrait).
 - une tablette posée (en orientation landscape) sur la table et un smartphone pris en main.
 - une tablette prise en main + un écran pc (16/9).

1. 1-2 : Pas du tout satisfait, 3-4 : Pas satisfait, 5-6 : Ni satisfait ni pas satisfait, 7-8 : Satisfait, 9-10 : Tout à fait satisfait.

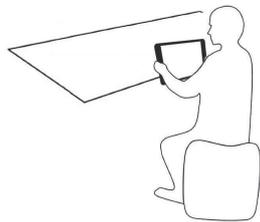


(a) Tablette posée sur table

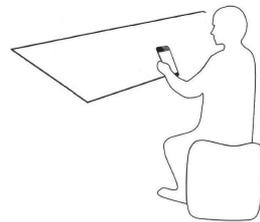


(b) Smartphone posé sur table

FIGURE 5.2 – Configurations Mono-écran fixe - Test préliminaire

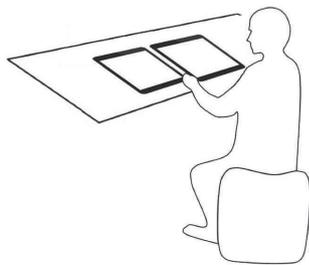


(a) Tablette prise en main

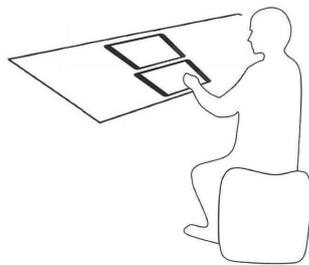


(b) Smartphone pris en main

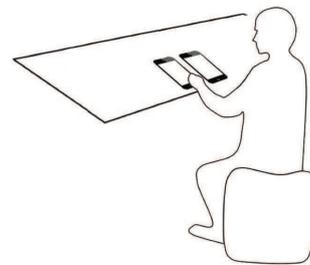
FIGURE 5.3 – Configurations Mono-écran mobile - Test préliminaire



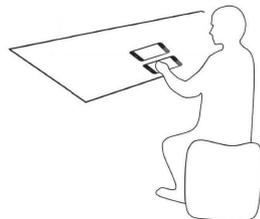
(a) Deux tablettes (portrait) posées sur table



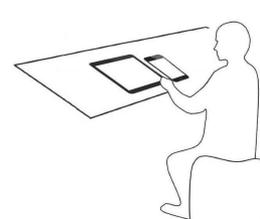
(b) Deux tablettes (landscape) posées sur table



(c) Deux smartphones (portrait) posés sur table

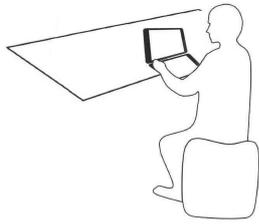


(d) Deux smartphones (landscape) posés sur table

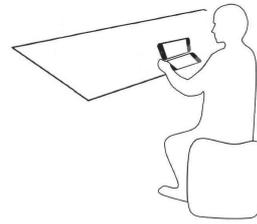


(e) Tablette et Smartphone posés sur table

FIGURE 5.4 – Configurations Multi-écran fixe - Test préliminaire

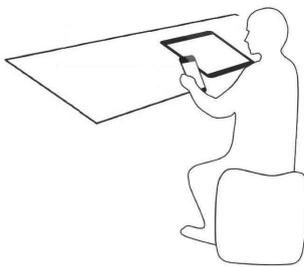


(a) Deux tablettes prises en main

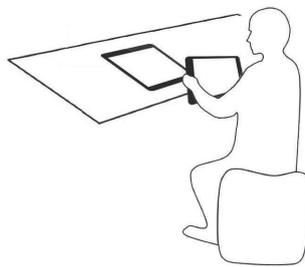


(b) Deux smartphones pris en main

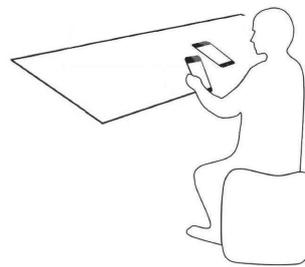
FIGURE 5.5 – Configurations Multi-écran mobile - Test préliminaire



(a) Tablette posée sur table + Smartphone pris en main



(b) Deux tablettes (posée sur table, prise en main)



(c) Deux smartphones (posé sur table, pris en main)



(d) Tablette prise en main + écran pc



(e) Smartphone pris en main + écran pc

FIGURE 5.6 – Configurations Multi-écran fixe et mobile - Test préliminaire

5.3.2 Procédure

Nous avons choisi d'adapter l'épreuve dite *Go/No-Go* de [ZF94] pour notre expérimentation. Cette épreuve invite l'utilisateur à identifier les stimuli cibles parmi l'ensemble des stimuli proposés et ce, à l'intérieur d'une fenêtre de temps réduite.

Dans notre version, nous avons mis en place deux niveaux de difficulté :

- *Niveau facile* (2 stimuli dont 1 cible) (Fig. 5.7a) : une croix correspondant soit au signe de l'addition (+), soit à celui de la multiplication (x) apparaissent au centre

de l'écran dans un ordre aléatoire. L'un des stimuli (x) est une cible, et l'autre (+) est une non-cible. Les participants doivent toucher le plus vite possible l'écran (ou l'un des écrans, dans le cadre d'une configuration multi-écran) lorsqu'une cible (x) apparaît sur l'écran.

- *Niveau difficile* (5 stimuli dont 2 cibles) (Fig. 5.7b) : cinq stimuli apparaissent au centre de l'écran dans un ordre aléatoire. Chaque stimuli possède son propre motif. Les participants doivent mémoriser les motifs de deux cibles et doivent toucher l'écran (resp. un des écrans) le plus vite possible lorsqu'une des deux cibles apparaît.

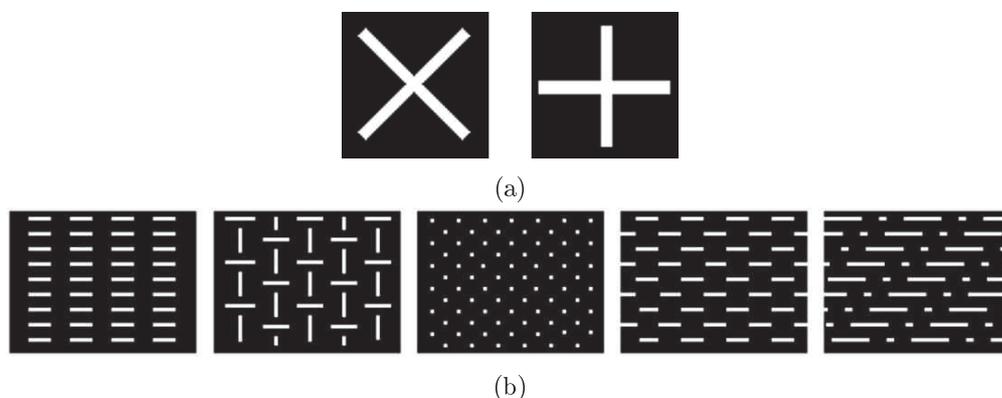


FIGURE 5.7 – L'épreuve Go/No-Go : 2 stimuli dont 1 cible (a), 5 stimuli dont 2 cibles (b).

Aucune action n'est requise pour les non-cibles. Chaque stimuli est présenté pendant une durée de 500 ms. L'intervalle inter-stimuli prend aléatoirement la valeur de 1000, 1500 ou 2000 ms. Les quatre configurations multipliées par les deux niveaux de difficulté nous amènent à une session composée de 8 passages (Fig. 5.8). Chaque passage correspond à un couple {configuration, difficulté}.

La figure (5.9), montre l'arbre complet des configurations de notre expérimentation. Nous remarquons que pour un ordonnancement des stimuli donné, soit aléatoire ou alternée, l'utilisateur doit effectuer le test Go/No-Go huit fois : deux fois pour la configuration mono-écran (l'ordonnancement n'est évidemment pas pris en considération dans ce cas), six fois pour les configurations multi-écrans (2 fois en utilisant deux tablettes, 2 fois en utilisant une tablette couplée avec un smartphone et 2 fois en utilisant une tablette couplée avec un écran PC). La répétition de chaque configuration provient du fait que l'utilisateur doit tester chaque configuration suivant deux niveaux de difficulté (facile et difficile).

5.3. EXPÉRIENCES

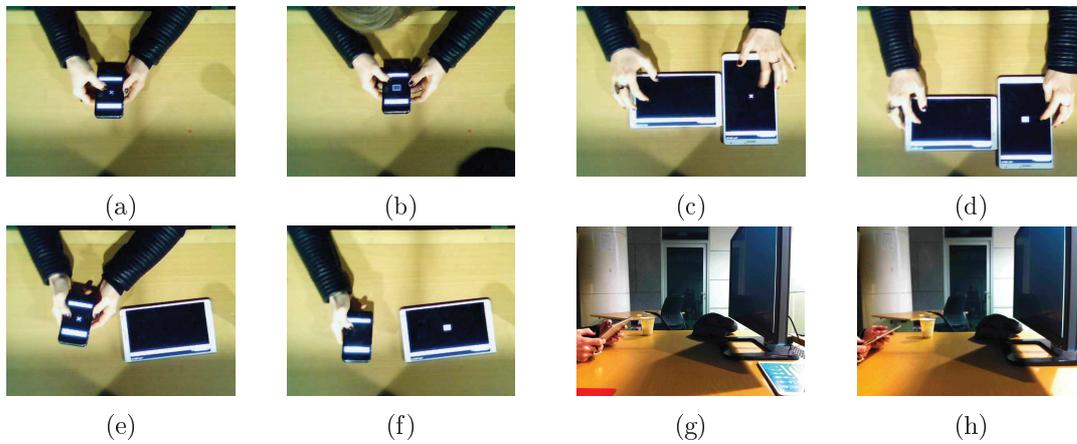


FIGURE 5.8 – Les différentes configurations d’écrans durant la phase d’expérimentation

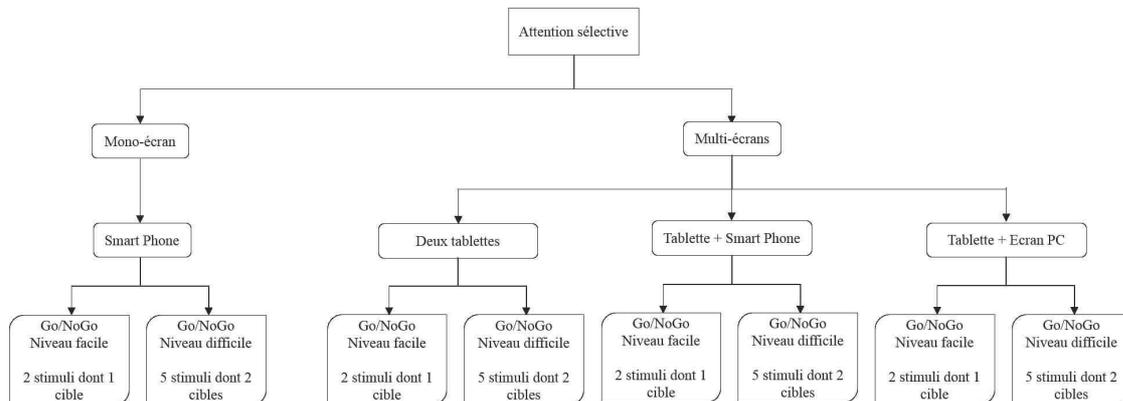


FIGURE 5.9 – Design expérimental

Chaque passage présente aléatoirement 40 stimuli (20 cibles, 20 non-cibles) au sujet. Les stimuli présentés sont différents à chaque passage. Un délai de 10 secondes est imposé entre chaque passage. L’ordre des passages est aussi aléatoire pour chaque sujet. La durée totale d’une session complète, sans les consignes, ni les essais, est d’environ 8 min.

Le temps de réaction, le nombre de réponses correctes, omises, fausses, ainsi que le taux d’erreur sont calculés pour chaque sujet.

Nous avons utilisé HTML/CSS/JavaScript pour créer et adapter la tâche Go/No-Go, ainsi que NodeJS² pour gérer la communication et la synchronisation des données entre les clients (les écrans).

2. NodeJS est une plateforme logicielle libre en JavaScript, comprenant une bibliothèque de serveur HTTP intégrée, ce qui rend possible de faire tourner un serveur web sans avoir besoin d’un logiciel externe comme Apache. <https://nodejs.org/>

5.3. EXPÉRIENCES

Pour chaque sujet, la procédure a été la suivante :

- Tout d’abord, il/elle a été invité(e) à remplir un court questionnaire afin de fournir des informations sur son âge, sexe, niveau scolaire, milieu professionnel, et son expertise en jeu vidéo.
- Ensuite, il/elle a regardé une vidéo (1 min), qui présentait les différentes configurations/tâches à effectuer pendant le test.
- Enfin, il/elle passait le test à proprement parler (les 8 passages de la session).

Chaque passage débute par l’affichage de son niveau de difficulté (facile ou difficile).

La moitié de sujets a effectué l’expérimentation en utilisant un ordonnancement aléatoire des stimuli, tandis que l’autre moitié l’a effectuée en utilisant un ordonnancement alterné. L’ordonnancement n’est évidemment pas pris en considération dans le cas de la configuration mono-écran.

Nous avons réalisé notre expérimentation dans les locaux de la cité des sciences et de l’industrie - Carrefour Numérique² (Fig. 5.10). Nous avons accueilli 100 participants (44 femmes, 56 hommes) avec un âge moyen de 27.77 ans.



(a)



(b)

FIGURE 5.10 – L’installation de l’expérimentation à la cité des sciences et de l’industrie - Carrefour Numérique²

Les participants ont été installés, assis face à un bureau, dans une salle calme et isolée.

Les dispositifs matériels utilisés dans notre expérimentation, comprennent un PC (Dell XPS, Intel core i7-4712 HQ, 2.30 GHz, 16 Go RAM), un écran PC (LG, LCD, 27 puces), deux tablettes tactiles (Samsung, 8 puces) et un smartphone (iPhone 4, 3.5 puces). Une webcam a été utilisée pour photographier les mains de nos participants (avec leur accord), dans le but d'observer leurs comportements face aux configurations proposées.

5.3.3 Présentation et analyse des résultats

5.3.3.1 Vocabulaires descriptives

Nous présentons tout d'abord quelques termes pour éclairer la lecture de cette partie :

- Une *tâche* est la requête de base demandée à un utilisateur, soit un certain comportement suite à l'apparition d'un stimulus sur l'écran. Ce comportement peut être une action (toucher l'écran) ou au contraire une inhibition (ne pas toucher l'écran).
- Une *expérience* est un ensemble de 40 tâches demandées à un utilisateur dans :
 - une configuration parmi :
 - ★ *phone* pour un seul smartphone (mono-écran)
 - ★ *tab* pour deux tablettes (multi-écran)
 - ★ *tabphone* pour une tablette + un smartphone (multi-écran)
 - ★ *pctab* pour une tablette + un écran PC (multi-écran)
 - un ordonnancement parmi :
 - ★ *non-pertinent* pour la configuration mono-écran
 - ★ *alternée*
 - ★ *aléatoire*
 - un niveau de difficulté parmi :
 - ★ *facile*
 - ★ *difficile*

Les actions de l'utilisateur sont décrites par les termes suivants (Fig. 5.11) :

- Une *omission* est une absence d'action de la part de l'utilisateur lorsque celui/celle-ci devait pourtant réagir (par exemple, lorsqu'il ne touche pas l'écran lorsqu'une

- cible(Go) est présentée).
- Un *correct.go* est une action correcte de la part de l'utilisateur (toucher l'écran lorsqu'une cible(Go) est présentée).
 - Une *fausse* est une action de la part d'utilisateur lorsque celui-ci ne devait pas agir (toucher l'écran lorsqu'une non-cible(No-Go) est présentée).
 - Un *correct.nogo* est une inhibition correcte (ne pas toucher l'écran lorsqu'une non-cible(No-Go) est présentée).

	Toucher	Ne pas toucher
Go	<i>Correct.go</i>	<i>Omise</i>
No-Go	<i>Fausse</i>	<i>Correct.nogo</i>

FIGURE 5.11 – Spécification des comportements de l'utilisateur face aux stimuli présentés

5.3.3.2 Visualisation des données

Nous avons utilisé le logiciel d'analyse R^3 pour analyser nos résultats. Nous avons aussi employé la librairie *ggplot2*⁴ pour présenter nos données. Cette librairie R de visualisation de données développée par *Hadley Wickham*, se base sur les principes de *The Grammar of Graphics* de *Leland Wilkinson*.

Chaque graphique présenté dans cette partie en dehors de l'histogramme repose sur une représentation des données par des nuages de points. Chaque point correspond à une expérience. Son abscisse permet de représenter la valeur d'une caractéristique de l'expérience parmi la configuration, l'ordonnancement et la difficulté (par exemple pour un graphique portant sur la difficulté, les tâches "faciles" ont une abscisse différente des tâches "difficiles"). L'ordonnée varie elle selon le graphique, cela peut être le temps de réaction (TR), le nombre de fausses, le nombre d'omissions ou le nombre d'erreurs.

Les segments bleus sont tracés sur la base d'une régression linéaire calculée avec les données des expériences des 2 nuages de points que cette droite joint (par la méthode des moindres carrés). Nous pouvons donc interpréter la pente de cette droite comme l'effet de la configuration sur la mesure réalisée. La bande en gris foncé autour des segments représente

3. <https://www.r-project.org/>

4. <http://ggplot2.org/>

5.3. EXPÉRIENCES

l'intervalle de confiance à 95% des coefficients pour une distribution normale.

Afin d'éviter la superposition dans les nuages de points, nous avons ajouté un effet "Jitter" à l'abscisse, c'est-à-dire, une valeur comprise entre -1 et 1 tirée au hasard selon une loi uniforme. A cela, nous ajoutons un coefficient de transparence qui permet aussi de mieux prendre en compte la superposition des données (une grande densité de points dans une zone renforce son opacité).

5.3.3.3 Histogrammes des temps de réaction

La figure 5.12 montre les histogrammes des temps de réaction (TR) pour l'ensemble des expériences (médian = 597.4 ms, moyenne = 602.3, min = 356.1 et max = 1050.3). Nous avons vérifié que nos résultats suivaient une loi normale, ce qui nous a permis de calculer différentes régressions.

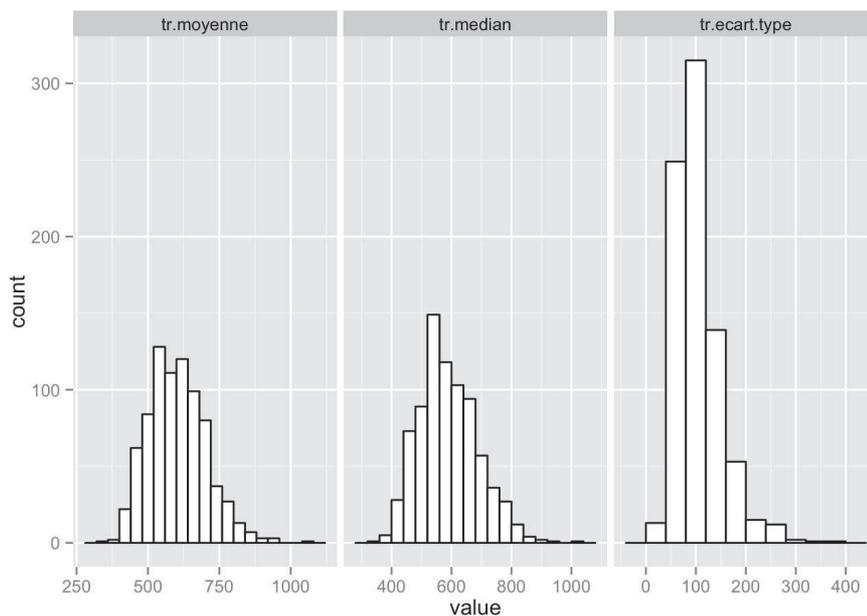


FIGURE 5.12 – Histogrammes des temps de réaction pour l'ensemble des expériences

5.3.3.4 Comparaison selon la difficulté de la tâche

Le tableau 5.1 montre quelques statistiques descriptives des expériences pour chaque niveau de difficulté de tâche. Nous observons une augmentation du TR d'environ 100 ms,

5.3. EXPÉRIENCES

soit environ +17% pour les tâches ayant un niveau de difficulté élevé. Au niveau des erreurs commises, nous observons une légère diminution de (-4.1%). Cette diminution est due aux fausses actions (-25%), alors que les omises augmentent avec la difficulté (+200%).

	tr.moyenne	tr.ecart.type	tr.median	corrects.totals	corrects.go	corrects.nogo	omises	fausses
facile	554.3432	100.6886	540.8288	37.9225	19.795	18.1275	0.205	1.8725
difficile	650.3120	108.3505	637.3575	38.0075	19.420	18.5875	0.580	1.4125

TABLE 5.1 – Statistiques descriptives pour l’ensemble des expériences en fonction de la difficulté

L’augmentation des omises et la baisse des fausses dans le cas des tâches difficiles nous permet d’avancer quelques interprétations : il semblerait que nous arrivions mieux à inhiber mais moins bien à agir lorsque le niveau de difficulté d’une tâche augmente. Nous faisons moins des fausses, parce que nous savons déjà que la difficulté augmente (les consignes ont été proposé avant le commencement de chaque tâche). Nous avons donc la tendance à ne pas réagir (toucher l’écran) lorsque la difficulté est plus complexe (nous prenons des mesures de vigilance en quelques sorte). Nous constatons donc que l’augmentation du niveau de la difficulté favorise l’inhibition de la tendance à réagir. Par conséquent, plus la difficulté augmente, plus les actions omises et le temps de réaction augmentent (voir les figures 5.13 et 5.14). Par contre, plus la difficulté augmente, plus les fausses actions diminuent (voir la figure 5.15).

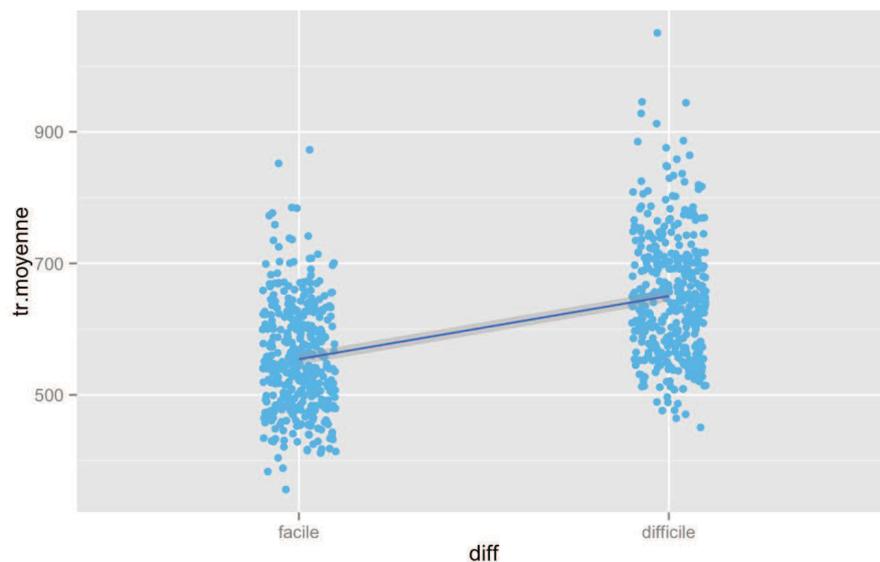


FIGURE 5.13 – Relation entre la difficulté et le temps de réaction

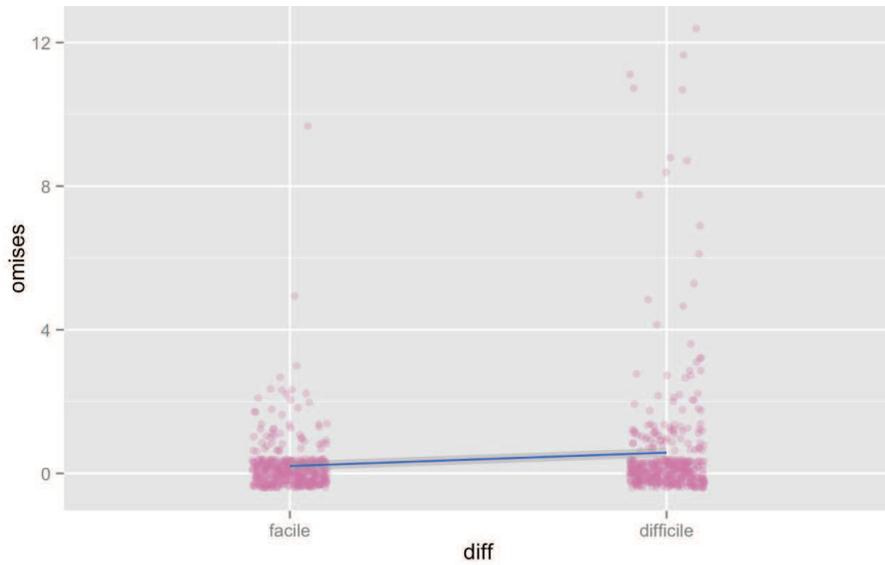


FIGURE 5.14 – Relation entre la difficulté et les omissions

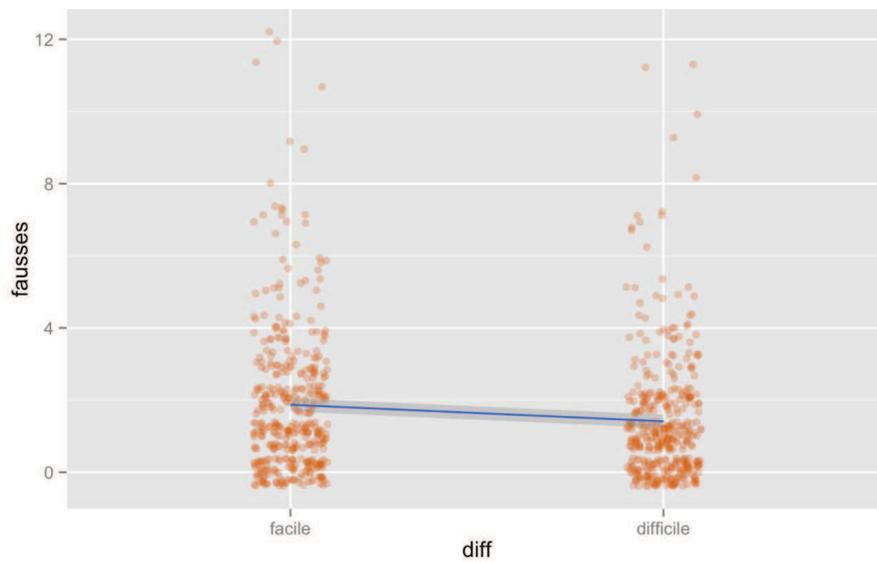


FIGURE 5.15 – Relation entre la difficulté et les fausses

5.3.3.5 Comparaison selon la configuration des écrans

Le tableau 5.2 présente les statistiques descriptives de l'ensemble des expériences en fonction de la configuration des écrans. A la vue des TR moyens et médians, nous pouvons identifier deux groupes. La configuration "phone" d'un côté (TR médian égal à 516 ms) et les autres configurations (TR médian compris entre 603 et 627 ms). Concernant

5.3. EXPÉRIENCES

le nombre d'erreurs, nous remarquons que la configuration "pctab" est la plus marquée (3.14 erreurs en moyenne contre 1.5 à 1.9 pour les autres configurations). La configuration d'écran ne semble donc pas influencer l'écart entre fausses et omises (sauf dans la configuration "pctab"), mais elle affecte visiblement le TR (un écran vs plusieurs), le TR est systématiquement supérieur pour les tâches réalisées sur deux écrans.

Configuration	tr.moyenne	tr.ecart.type	tr.median	corrects.totals	corrects.go	corrects.nogo	omises	fausses
phone	522.9045	75.34885	515.8625	38.510	19.725	18.785	0.275	1.215
tab	636.2230	95.48842	627.7500	38.415	19.815	18.600	0.185	1.400
tabphone	624.0931	116.47336	609.1425	38.075	19.775	18.300	0.225	1.700
pctab	626.0900	130.76772	603.6175	36.860	19.115	17.745	0.885	2.255

TABLE 5.2 – Statistiques descriptives pour l'ensemble des expériences en fonction de la configuration

La figure 5.16 montre le TR en fonction de la configuration. Nous pouvons voir que le TR augmente lorsque nous utilisons plus d'un écran. La figure 5.17 montre les fausses actions en fonction des configurations d'écran. Nous pouvons observer que nous faisons plus de fausses lorsque nous utilisons la configuration "pctab". Cela peut être expliqué par la distance entre les écrans (plus la distance est importante, plus nous faisons des fausses).

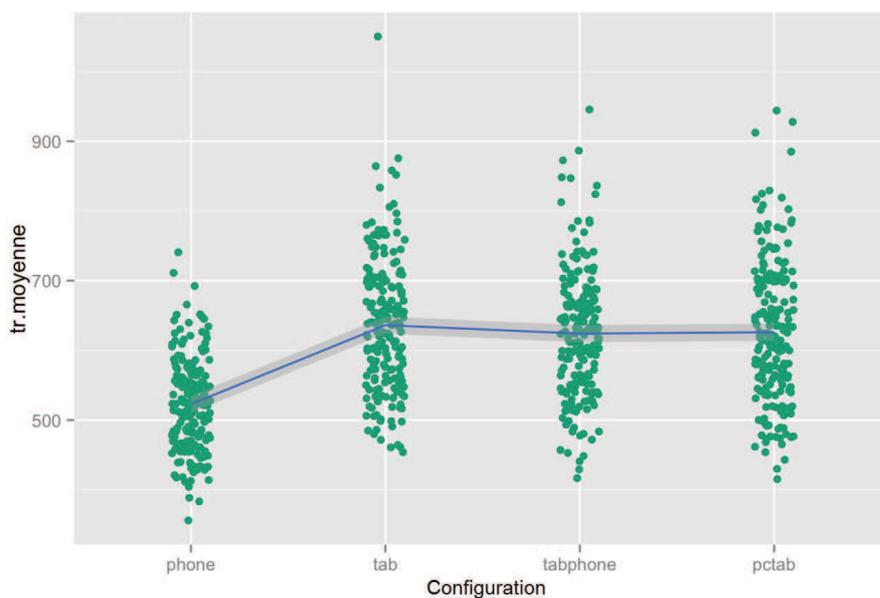


FIGURE 5.16 – Distribution des TR en fonction de la configuration

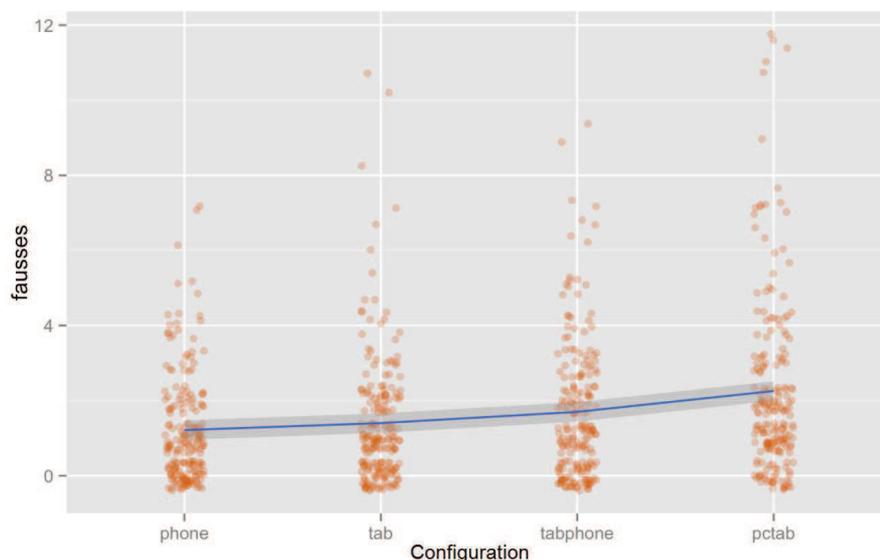


FIGURE 5.17 – Distribution des fausses en fonction de la configuration

5.3.3.6 Comparaison selon l'ordonnement des stimuli

Le tableau 5.3 présente les statistiques descriptives de l'ensemble des expériences en fonction de l'ordonnement des stimuli. Nous remarquons que l'ordonnement a un effet fort sur le TR et le nombre d'erreur.

Nous constatons que la médiane du TR est égale à 516 ms pour l'ordonnement non-pertinent⁵, 581 ms pour l'ordonnement alterné et 645 ms pour l'ordonnement aléatoire.

	tr.moyenne	tr.ecart.type	tr.median	corrects.totals	corrects.go	corrects.nogo	omises	fausses
non.pertinent	522.9045	75.34885	515.8625	38.51000	19.72500	18.78500	0.2750000	1.215000
alternees	597.4574	107.92622	581.3450	37.84667	19.57000	18.27667	0.4300000	1.723333
aleatoires	660.1466	120.56012	645.6617	37.72000	19.56667	18.15333	0.4333333	1.846667

TABLE 5.3 – Statistiques descriptives pour l'ensemble des expériences en fonction de l'ordonnement des stimuli

Nous remarquons que le TR moyen augmente d'environ 65 ms pour les tâches ayant un ordonnancement alterné par rapport à celles dont l'ordonnancement est non-pertinent. Il augmente aussi d'environ 129 ms pour les tâches aléatoires. Le TR moyen est aussi bien supérieur pour les tâches alternées par rapport à celles aléatoires, pour environ 64 ms. Nous observons également qu'il y a plus de fausses que d'omises en fonction de l'ordonnement

⁵. Cet ordonnancement utilisé juste dans le cas mono-écran

(ordonnancement non-pertinent : moyenne des omises = 0.27, moyennes des fausses = 1.21 - pour ordonnancement alterné : moyenne des omises = 0.43, moyenne des fausses = 1.72 - pour ordonnancement aléatoire : moyenne des omises = 0.43, moyenne des fausses = 1.84).

Nous pouvons donc observer un effet significatif sur le TR moyen d'utilisateur en fonction de l'ordonnancement des stimuli. Nous avons tendance à effectuer la tâche plus rapidement lorsque l'ordonnancement d'affichage est connu (cas de l'alternance) avant l'exécution de la tâche (Fig. 5.18).

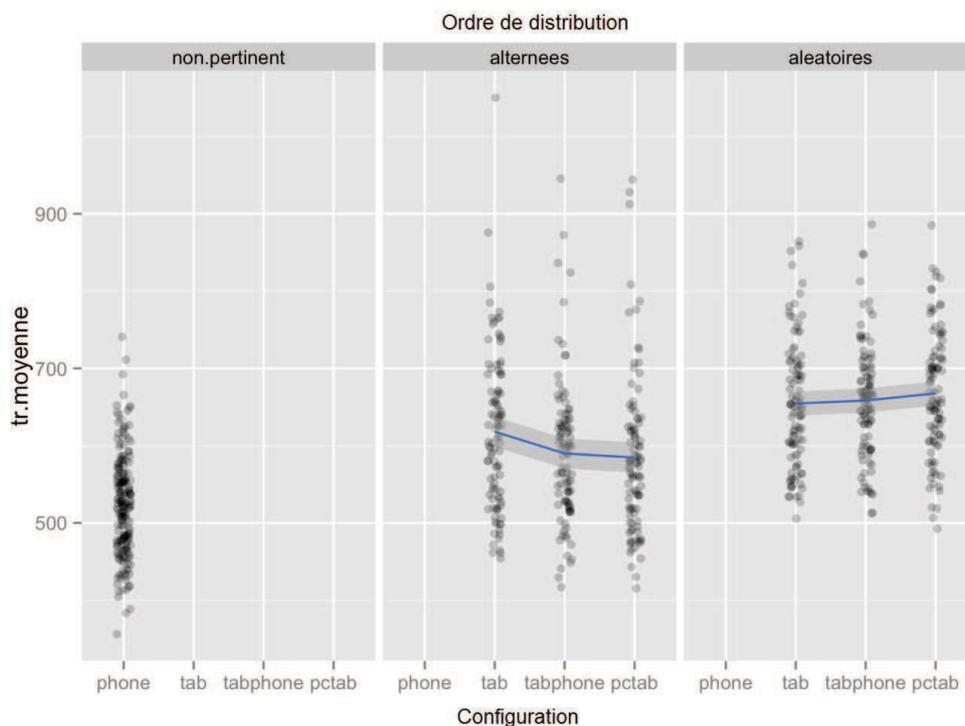


FIGURE 5.18 – Distribution des TR en fonction de l'ordonnancement des stimuli

5.3.3.7 Synthèse générale

En différenciant l'ensemble des expériences en fonction du niveau de difficulté de la tâche et de la configuration des écrans :

- Nous avons observé une augmentation du TR moyen, expliquée par le niveau de difficulté d'une part et par le nombre d'écrans d'autre part. Nous notons que l'augmentation du nombre des écrans (configuration mono-écran vs multi-écrans) réduit

la performance (en terme de TR moyen). Le niveau de la difficulté de tâche impacte également en ce sens la performance.

Nous constatons donc que nous pouvons réaliser une tâche plus rapidement lorsque nous utilisons un seul écran. L'ordre de grandeur du TR moyen dans une tâche difficile réalisée en utilisant un seul écran correspond à celui du TR dans une tâche facile effectué en utilisant plusieurs écrans (Fig. 5.19).

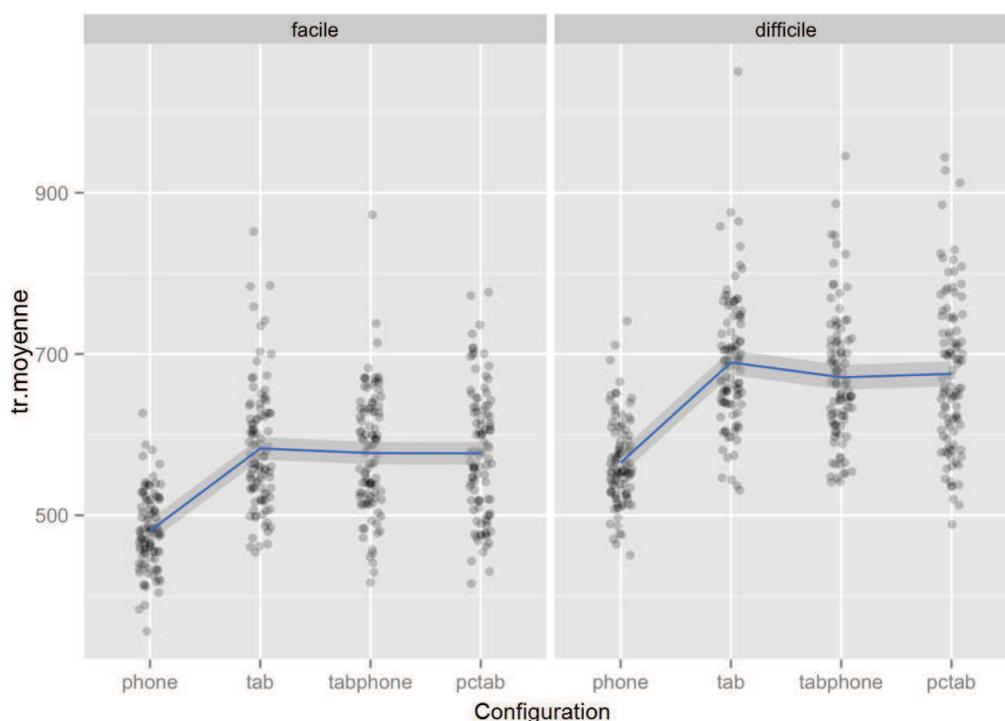


FIGURE 5.19 – Distribution des TR moyens et régressions linéaires pour chaque configuration d'écran dans le cas facile (gauche) et difficile (droite).

- Nous avons remarqué que les configurations semblent ne pas avoir d'effet sur le nombre d'erreurs, quelque soit le niveau de difficulté (Fig. 5.20).

Cette observation est cependant moins vraie pour la configuration "pctab". Cette originalité nous amène à penser que la gestion de l'attention sur des objets positionnés dans des plans différents (un écran pc + une tablette) peut être plus complexe que dans le cas où les objets sont placés sur un même plan (deux tablettes, un smartphone + une tablette).

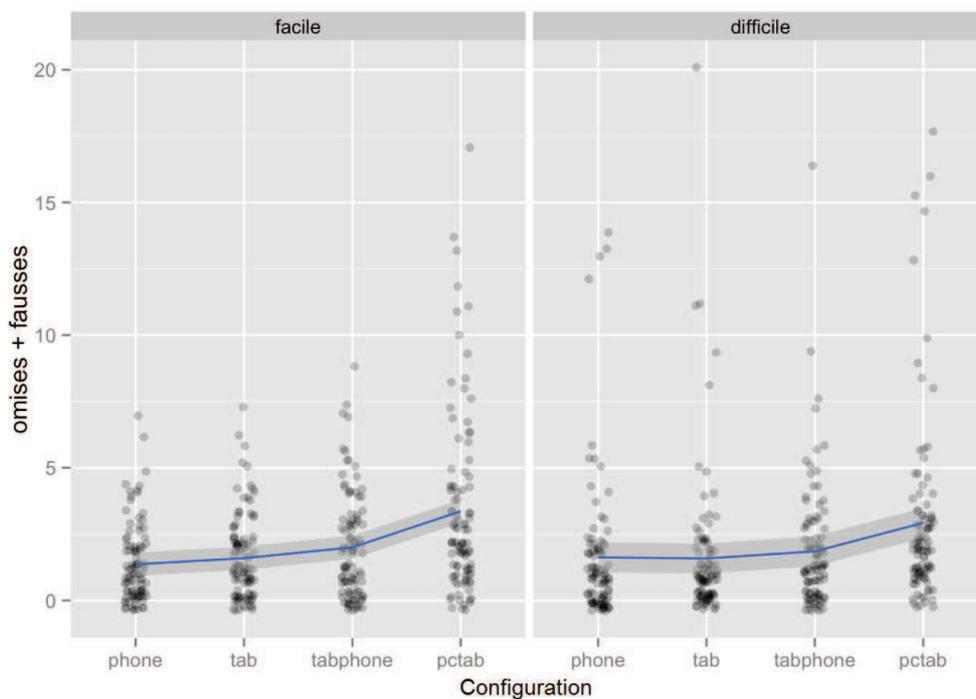


FIGURE 5.20 – Distribution du nombre d’erreur pour chaque configuration d’écran dans le cas facile (gauche) et difficile (droite).

Nous avons maintenant considéré l’ensemble des expériences en fonction du niveau de difficulté et de l’ordonnancement des stimuli :

- Nous avons constaté une augmentation du TR moyen, expliqué par le niveau de la difficulté d’une part et par l’ordonnancement des stimuli d’un autre part (Fig. 5.21). Nous notons que l’utilisation d’un ordonnancement aléatoire des stimuli réduit la performance (en terme de TR moyen) afin d’effectuer une tâche donnée. Il en est de même lorsque la tâche demandée est plus difficile.

Par ailleurs, nous pouvons constater que l’ordre de grandeur du TR moyen dans la condition difficile / ordonnancement alterné, correspond à celui du TR moyen dans la condition facile / ordonnancement aléatoire.

- Nous avons remarqué que l’ordonnancement semble ne pas avoir d’effet sur le nombre d’erreurs, quelque soit le niveau de difficulté (5.22), même si la condition aléatoire / facile semble déroger à cette observation.

5.3. EXPÉRIENCES

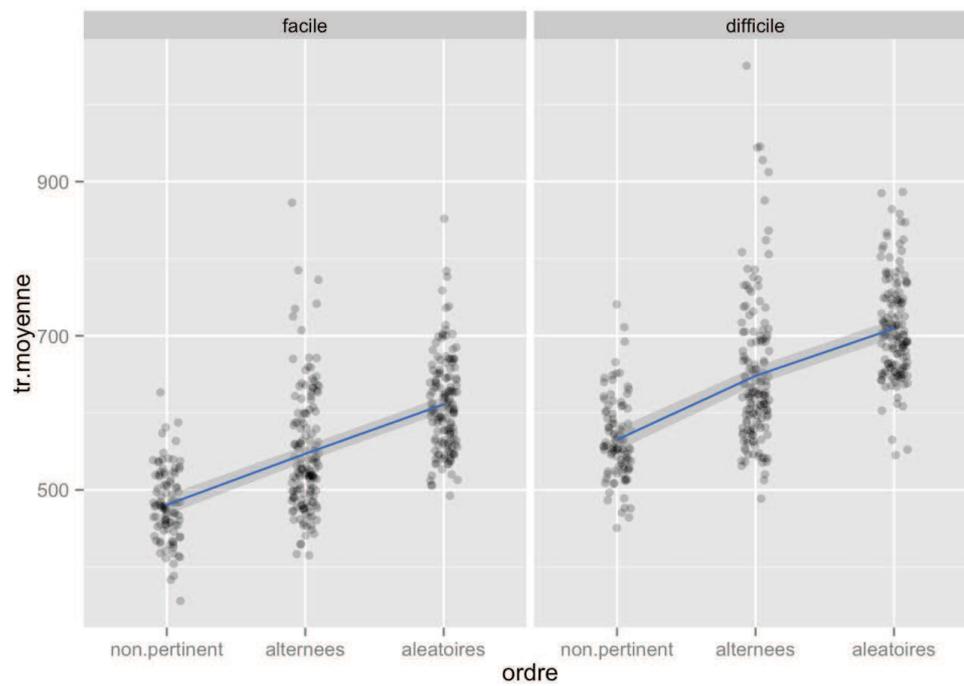


FIGURE 5.21 – Distribution des TR moyennes par type d'ordonnement des stimuli pour les conditions facile (gauche) et difficile (droite)

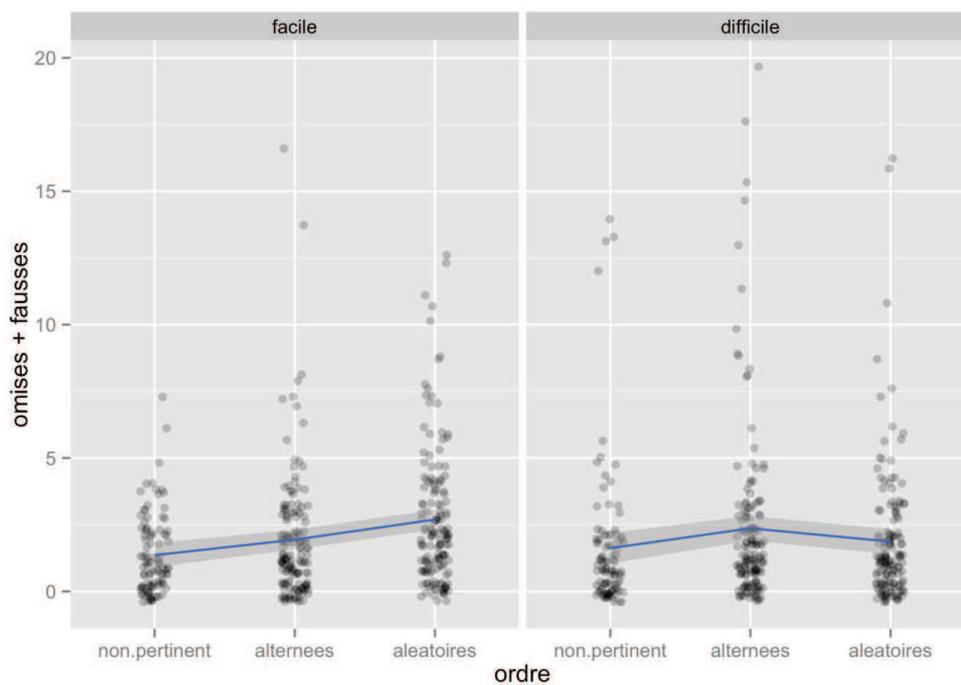


FIGURE 5.22 – Distribution du nombre d'erreur par type d'ordonnement des stimuli pour les conditions facile (gauche) et difficile (droite)

5.3.4 Discussion

A partir de nos paramètres, il semble que le temps de réaction soit plus important lorsque nous devons appréhender une interface répartie sur plusieurs écrans, tandis que nos réactions sont plus rapides face à une interface centrée sur un seul écran (voir Fig. 5.23). Nous pouvons dire que la surveillance de plusieurs écrans de manière concurrente, peut perturber notre système attentionnel, entraînant ainsi de latences majorées quant à nos réactions.

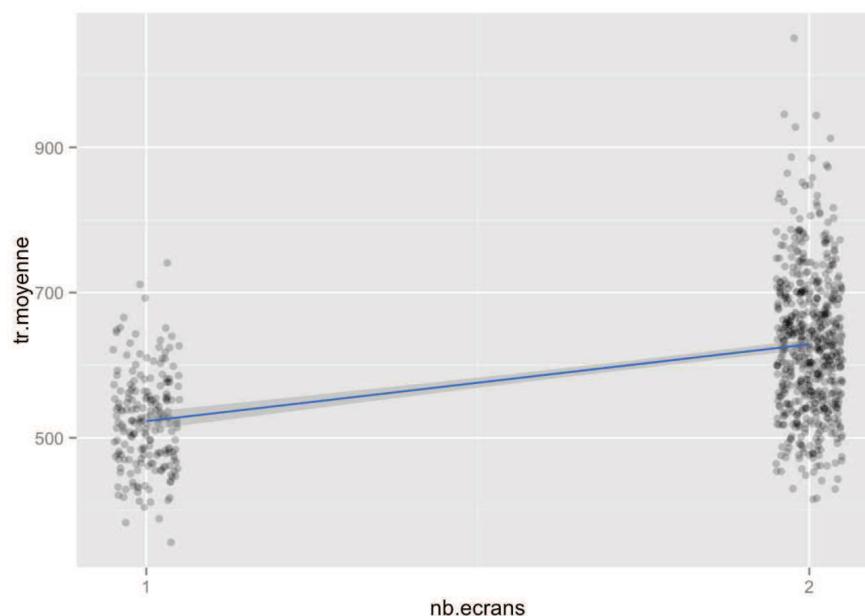


FIGURE 5.23 – Distribution des TR moyens en fonction du nombre d'écrans

Nous avons remarqué qu'en augmentant la difficulté de la tâche, le temps de réaction augmente (voir Fig. 5.24). L'augmentation concomitante du nombre d'écrans accentue cette difficulté et, par conséquent, les latences dans le TR. Cette analyse peut avantageusement être utilisée par les concepteurs d'application (notamment de jeu vidéo), pour intégrer les effets propres au matériel dans la gestion de la difficulté au sein de leurs produits.

Par ailleurs, nous avons pu mettre en évidence une correspondance quantitative entre les difficultés d'une même tâche, réalisée sur un dispositif mono écran d'une part et sur un dispositif multi écran, d'autre part. Ainsi, le fait d'utiliser plusieurs écrans (vs. un seul) augmente la difficulté de la tâche demandée. Ce résultat est directement exploitable dans

5.3. EXPÉRIENCES

un contexte de réglage de la difficulté d'un même jeu vidéo, en permettant d'envisager le design des challenges différemment, selon que ce soit la version PS Vita⁶ (mono écran) ou 3DS (multi écran)⁷.

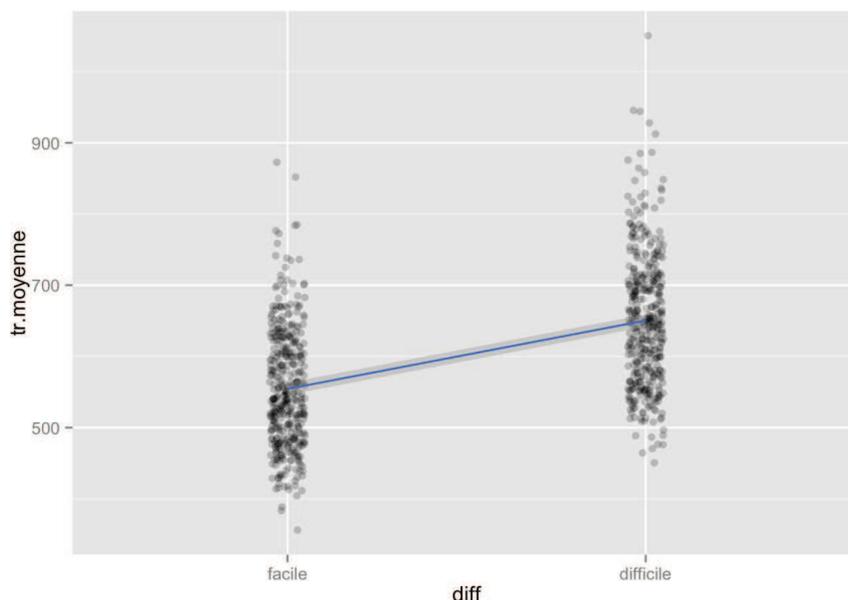


FIGURE 5.24 – Distribution des TR moyens en fonction de la difficulté

Nous avons remarqué que le temps de réaction augmente lorsque l'ordonnement des stimuli est aléatoire (voir Fig. 5.25). Nous avons constaté que le temps de réaction dans une tâche d'un niveau élevé de difficulté, utilisant un ordonnancement alterné des stimuli, correspond à celui dans une tâche d'un niveau bas de difficulté, utilisant un ordonnancement aléatoire.

Nous avons remarqué aussi que les sujets font plus de fausses que d'omises. Il semble donc y avoir une tendance plus forte à se tromper lorsqu'il ne faut pas agir que lorsqu'il faut agir. Par contre, nous avons observé que les fausses diminuent et les omissions augmentent lorsque la difficulté est plus complexe. Nous constatons donc que l'augmentation du niveau de la difficulté favorise l'inhibition de la tendance à réagir. Ceci peut également se traduire en terme de stratégie cognitive, qui aurait tendance à favoriser l'inhibition (face à l'action) dans des situations de difficulté plus importante.

6. une console de jeux vidéo portable comprend un écran

7. une console de jeux vidéo portable comprend deux écrans

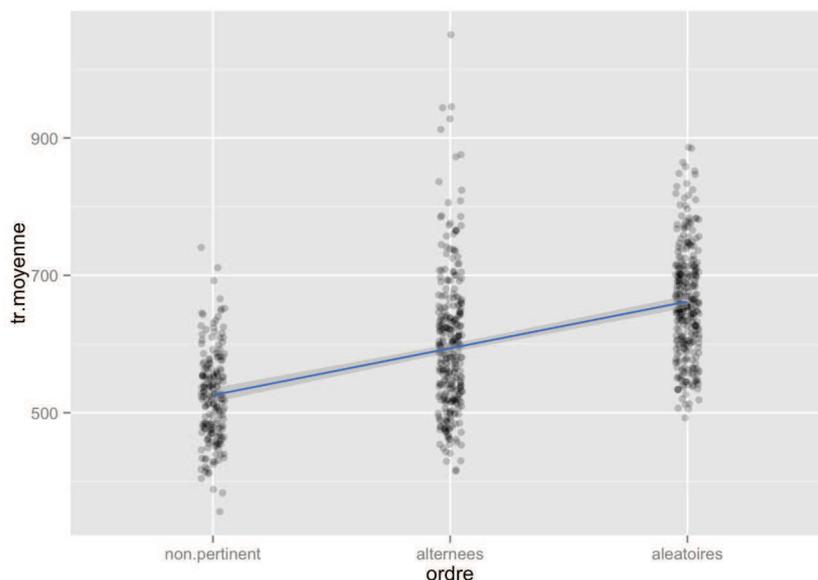


FIGURE 5.25 – Distribution des TR moyens en fonction de l’ordonnancement de l’information

Nous avons montré qu’il y a plusieurs effets de la répartition de l’attention sur plusieurs écrans. Cela peut être expliqué suivant la définition de focus d’attention, qui désigne le centre de l’attention. Dans le cas où nous devons gérer un seul écran, le focus d’attention se porte sur un seul surface d’affichage. Par contre, ce focus doit être déplacé dans le cas où nous devons surveiller plusieurs surfaces en parallèle. Le passage d’un surface d’affichage à un autre peut donc entraîner des augmentations de nos temps de réaction, causées par ces mouvements complémentaires et l’instabilité qu’ils entraînent.

Nous notons que la configuration {phone} (mono-écran) a permis aux utilisateurs d’avoir des temps de réaction plus courts (en comparant avec ceux obtenus en utilisant plusieurs écrans), parce que les utilisateurs ont manipulé le smartphone en utilisant leur pouce. La distance donc entre l’écran et le doigt était très petite et permettait donc de réagir rapidement. Dans le cas où les utilisateurs ont utilisé {deux tablettes} ou {tablette/phone} (positionnés horizontalement), ils/elles ont effectué leurs tâches en utilisant leur index. Dans le cas {tablette/pc}, les utilisateurs ont également utilisé leur pouce. Par contre, la surveillance des stimuli dans le cas {tablette/pc} n’a pas permis d’avoir des réactions aussi rapides que dans le cas du téléphone seul.

Nous avons observé que les utilisateurs ont effectué plus d'erreurs, lorsqu'ils/elle ont utilisé la configuration {tablette/pc}. Cela peut être dû à la distance entre les deux surfaces {tablette/pc} et aux types de mouvements engagés pour adapter le focus d'attention (verticaux versus horizontaux dans les autres cas). Il sera sans doute important d'expérimenter l'effet de la distance pour voir l'évolution de la performance de l'utilisateur. Ce taux d'erreur pourrait augmenter en éloignant, de plus en plus, la tablette par rapport à l'écran pc.

En conclusion, la difficulté d'une tâche est maximale (ainsi que les temps de réaction moyens) dans le cas d'une distribution aléatoire des stimuli et d'une architecture matérielle possédant de multiples surfaces.

5.4 Conclusion

Nous avons décrit dans ce chapitre notre démarche pour évaluer les effets de la répartition de l'attention sur les performances des utilisateurs de dispositifs à affichages multiples.

Notre hypothèse reposait sur le fait que la multiplication des écrans (à la fois en termes d'affiche et de commande) peut entraîner des variations de nos performances et amène à des temps de réaction plus longs.

Pour cela, nous avons mis en place une expérimentation qui avait pour objectif d'étudier les effets de la distribution de l'attention, dans différentes configuration matérielles.

Nous avons montré que l'utilisation de deux écrans provoque une augmentation de la difficulté (i.e. une diminution des performances). Cette donnée originale peut être donc prise en compte dans la conception d'expériences interactives (jeux vidéo ou autres), soit pour adapter la difficulté d'une tâche à des contextes matériels différents, soit afin de moduler la difficulté d'une application en utilisant uniquement les outils d'interface matérielle.

Chapitre 6

Perspectives et conclusion

6.1 Relation attention-difficulté dans un serious game

Dans le chapitre 3, nous avons présenté nos démarches, dans le cadre du projet *SGCogR*, afin d'évaluer le niveau attentionnel d'un sujet dans un serious game dit *Le Village aux Oiseaux*. Ce jeu, destiné à des personnes présentant des troubles attentionnels légers, a été développé pour fournir un environnement d'entraînement de l'attention. Dans ce jeu le joueur prend le rôle d'un photographe, qui doit photographier (cliquer sur) des oiseaux (cibles) distribués dans l'environnement du jeu. Le gameplay de ce jeu consiste donc à effectuer des tâches de pointage dans sur interface en deux dimensions, en utilisant la souris.

Notre travail a consisté à développer un outil de mesure de l'attention, qui sera utilisé par le système pour adapter la difficulté du jeu.

Sachant que ce jeu doit être joué à domicile et sans assistance, nous avons dû développer notre outil de manière à ce qu'il ne soit pas invasif, qu'il ne nécessite pas de compétences informatiques et coûte aussi peu que possible. Pour cela, nous avons proposé d'étudier une solution basée sur l'analyse des données du joueur fournies par le système.

Nous avons mis en place un algorithme qui nous a permis d'analyser les trajectoires du pointeur du joueur. Cela nous a permis de calculer des approximations satisfaisantes du temps de réaction des joueurs.

L'outil d'évaluation que nous avons proposé, est suffisant pour fournir des données,

pouvant être appliquées pour tester des versions futures de jeux vidéo. L'analyse de temps de réaction peut être avantage pour les concepteurs des jeux, leur permettant d'améliorer la gestion de la difficulté des jeux et de la motivation des joueurs.

Une limitation de notre démarche est que, faute de temps, le financement du projet s'est achevé, nous n'avons expérimenté notre outil qu'avec des cibles non-mobiles. Comme le résultat est prometteur, nous devons maintenant, évaluer cet outil dans une version complète du *Village aux oiseaux* (i.e. incluant également des cibles mobiles). Dans ce jeu, de nombreux oiseaux peuvent apparaître en même temps. Ils peuvent voler ou marcher et même disparaître avant que le joueur ait eu la possibilité de les prendre en photo. Par conséquent, nous devons expérimenter notre outil dans ces conditions d'évaluation pour éprouver notre outil dans un tel environnement moins contrôlé.

Nous devons aussi tester notre outil avec des patients initialement visés dans le projet, c'est à dire présentant les symptômes de la maladie d'Alzheimer, afin de montrer comment notre modèle peut être utilisé dans le service de rééducation cognitive. Ceci pourrait ouvrir la voie pour l'inscription de tels outils dans les processus utilisés par les professionnels de santé, pour évaluer le temps de réaction des patients à la place ou en complétant des tests d'évaluation d'attention actuels, comme l'Attention Network Test (ANT) par exemple.

6.2 Gestion de la difficulté par les outils d'interface logicielle

Dans le chapitre 4, nous avons présenté nos travaux dont le but était de réaliser un rapprochement entre le comportement de la vision humaine et le comportement de la vision synthétique. Nous avons donc transposé les méthodes d'analyse de la vision humaine vers la vision synthétique, en appliquant les mesures utilisés en oculométrie aux comportements de la caméra virtuelle. Ces mesures ont été employées pour caractériser les comportements de la caméra virtuelle contrôlée par les utilisateurs.

Notre étude a donné des indicateurs sur les stratégies employées par les utilisateurs pour effectuer une navigation dans les environnements virtuel 3D de type FPS. En analysant la façon dont l'utilisateur contrôle la caméra virtuelle de l'environnement, nous avons pu fournir des informations utiles pour le concepteurs de jeux vidéo. De nouvelles options ont

6.3. GESTION DE LA DIFFICULTÉ PAR LES OUTILS D'INTERFACE MATÉRIELLE

vu le jour, notamment celle d'utiliser uniquement les paramètres intrinsèques de la caméra, afin de gérer la difficulté dans les jeux vidéo.

Nous avons choisi de reproduire le code couleur de notre environnement virtuel 3D, inspiré des travaux de Lee et al. [LKC07], qui ont utilisé une couleur sombre avec des contrastes importants. Une piste future serait d'étudier comment la couleur pourrait affecter à son tour les performances dans ce type de contexte.

Nous souhaitons également tester notre modèle dans un environnement en 2 dimensions, comme dans le cas de la lecture numérique. Cela peut apporter des informations utiles sur le comportement d'utilisateur, en diminuant le champ de vision sur les régions de la lecture (comprendre si cela peut affecter son comportement, l'aider à s'organiser).

Nous prévoyons également de tester notre modèle dans le cas du *Village aux Oiseaux*, expliqué précédemment. Nous pouvons intégrer notre modèle au sein de ce jeu pour voir les effets qui peuvent être apportés sur le niveau attentionnel d'un utilisateur, en agissant sur les paramètres propres de la caméra virtuelle.

Dans notre expérience, nous avons utilisé la caméra virtuelle pour produire le comportement de la vision synthétique. Nous pouvons en conséquence intégrer d'autres moyens pour caractériser ce comportement, en utilisant par exemple casque de réalité virtuelle (RV). Ce dispositif dispose d'un écran placé à quelques centimètres en face de nos yeux, affichant une image stéréoscopique d'un environnement virtuel. Il dispose ainsi de divers capteurs permettant de détecter les mouvements de tête de l'utilisateur, ce qui permet d'adapter en temps réel l'image projetée sur l'écran, afin de produire l'illusion d'une immersion dans la scène restituée. L'utilisation d'un tel dispositif présente un autre moyen afin de caractériser la vision synthétique, puisqu'il augmente virtuellement notre champ visuel et permet un contrôle "naturel" du point de vue.

6.3 Gestion de la difficulté par les outils d'interface matérielle

Dans le chapitre 5, nous avons décrit nos démarches afin d'évaluer les effets de la répartition de l'attention entre différents loci, sur les performances des utilisateurs. Connaître les nouvelles contraintes imposées par le système dans ce type de situation, permet de fournir

des informations utiles afin d'améliorer ce type d'expérience.

Nous avons prouvé que la multiplication des écrans peut entraîner des variations de la performances de l'utilisateur. Nous avons pu définir une relation entre la difficulté d'une même tâche sur un dispositif mono-écran d'une part, et d'autre part, un dispositif multi-écrans. Le fait d'utiliser plusieurs écrans (vs. un seul) augmente la difficulté de la tâche demandée. Cette originalité peut être donc prise en compte dans la conception d'expériences interactives afin de moduler la difficulté d'une application en utilisant uniquement les outils d'interface matérielle.

Nous avons aussi trouvé que le temps de réaction dans une tâche d'un niveau élevé de difficulté, utilisant un ordonnancement alterné des stimuli, correspond à celui dans une tâche d'un niveau bas de difficulté, utilisant un ordonnancement aléatoire. Nous avons pu donc constater que le temps de réaction devient plus court d'être lorsque l'ordonnancement est connu avant l'exécution de la tâche

L'expérience variée de nos participants permettra à nos résultats se généralisé à une large population. Nous devons maintenant analyser nos résultats en séparant nos sujets selon leur profils. Nous devons étudier leurs actions dans plusieurs catégories (joueurs, non joueurs, adultes et jeunes).

De plus, nous devons étudier d'autres facteurs qui peuvent influencer la performance d'utilisateur dans ce type d'expérience, comme la taille d'écrans (petite, modérée, large), leur orientation les uns par rapport aux autres (verticale ou horizontale) et la distance entre eux (la trajectoire à parcourir afin d'acquérir les informations affichées).

Enfin, nous devons aussi questionner le grand public sur leurs manières de manipuler plusieurs écrans en parallèle (répondre au SMS et regarder un film par exemple), afin d'avoir une vue plus précise sur les nouveaux usages. Cela permettra d'intégrer plus efficacement ce type d'expérience, en fonction des performances des utilisateurs.

6.4 Conclusion

Dans cette thèse, nous avons étudié la problématique de l'évaluation d'attention ainsi que son implication dans la gestion de la difficulté dans le contexte du jeu vidéo.

Comme il est montré au cours de ce travail, l'analyse de l'attention peut jouer un rôle essentiel lors de la conception des applications interactives. Nous avons besoin de connaître pourquoi et quand la perte d'attention se produit, afin de caractériser les performances des utilisateurs.

Dans ce travail, nous avons montré comment nous pouvons utiliser les paramètres de l'attention afin d'ajuster la difficulté dans les jeux vidéo. La performance d'utilisateur est donc en relation directe avec l'interface qu'il/elle utilise. Plus la configuration d'une interface est complexe, plus nous avons besoin de se concentrer afin de récupérer les informations qui nous intéressent. Nos réactions face aux tâches demandées sont aussi affectées par les interfaces que nous manipulons. Nous pouvons affirmer que l'attention peut être augmentée/diminuée selon la configuration de l'interface qui l'opère, ce qui en conséquence influe sur son performance afin de produire ses actions. Les interfaces interactives doivent s'adapter aux besoins d'utilisateur afin de garantir leur utilisation, ce qui emploie un usage confortable, efficace et approprié.

Les modèles non-intrusifs exploités dans cette thèse, peuvent être généralisés et utilisés dans plusieurs contextes (la rééducation cognitive, la lecture numérique ou la navigation virtuelle) comme références pour analyser le comportement d'utilisateur. Nos modèles peuvent donc aider à la construction et la validation d'expériences interactives futures.

Pour finir, une suite naturelle de nos recherches consiste à étendre nos modèles pour permettre l'adaptation dynamique de l'interface au niveau attentionnel. Ce travail de thèse pourrait être complémentaire des travaux menés par Guillaume Levieux [Gui11], en adaptant dynamiquement la difficulté dans les jeux vidéo à partir de nos outils.

6.4. CONCLUSION

Bibliographie

- [ACL⁺06] Leila Cristina Vasconcelos de Andrade, Luis Alfredo Vidal Carvalho, Cabral Lima, Adriano Cruz, Paulo Mattos, Carlos Franco, Adriana Soares, and Bruno Grieco. Supermarket game : An adaptive intelligent computer game for attention deficit/hyperactivity disorder diagnosis. In *Proceedings of the Fifth Mexican International Conference on Artificial Intelligence*, pages 359–368, Washington, DC, USA, 2006. 35
- [And05] John R. Anderson. *Cognitive psychology and its implications*. Worth Publishers, 6 edition, 2005. 27, 57
- [Ass13] Entertainment Software Association. Essantial facts about the computer and video game industry. Online, 2013. 34
- [AWM⁺10] Ashmeed Ali, Edwin Wong, Gateley Meeker, David Gill, Paul Cushman, Brian Morel, Ira Amilhussin, Lauren Weinberg, and Jake Moskowitz. Mobile shopping framework. the role of mobile devices in the shopping process. Technical report, Yahoo! and The Nielsen Foundation, 2010. 56
- [Bad97] Alan Baddeley. *Human memory. Theory and practice*. (Rev.ed). Hove : Psychology Press, 1997. 29
- [BBM⁺03] James R. Booth, Douglas D. Burman, Joel R. Meyer, Zhang Lei, Barbara L. Trommer, Nicholas D. Davenport, Wei Li, Todd B. Parrish, Darren R. Gitelman, and M. Marsel Mesulam. Neural development of selective attention and response inhibition. *Neuroimage*, 20(2) :737–751, 2003. 33
- [BBR⁺88] Karlene K. Ball, Bettina L. Beard, Daniel L. Roenker, Richard L. Miller, and David S. Griggs. Age and visual search : Expanding the useful field of view. *Journal of the Optical Society of America*, 5 :2210–2219, 1988. 91

- [BEJ⁺14] Andy Brown, Michael Evans, Caroline Jay, Maxine Glancy, Rhianne Jones, and Simon Harper. Hci over multiple screens. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, CHI '14, pages 665–674, 2014. 59
- [BGH⁺11] D. Bavelier, C. S. Green, D. H. Han, P. F. Renshaw, M. M. Merzenich, and D. A. Gentile. Brains on video games. *Nature Reviews Neuroscience*, 12 :763–768, 2011. 35
- [BKS⁺08] Walter R. Boot, Arthur F. Kramer, Daniel J. Simons, Monica Fabiani, and Gabriele Gratton. The effects of video game playing on attention, memory, and executive control. *Acta Psychologica*, 129 :387–398, 2008. 35
- [BLF⁺98] R. Barbarotto, M. Laiacona, R. Frosio, M. Vecchio, A. Farinato, and E. Capitani. A normative study on visual reaction times and two stroop colour-word tests. *The Italian Journal of Neurological Sciences*, 19(3) :161–170, 1998. 33
- [BM13] Max Birk and Regan L. Mandryk. Control your game-self : Effects of controller type on enjoyment, motivation, and personality in game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 685–694, 2013. 66
- [BMKZ05] Markus Buehner, Marija Mangels, Stefan Krumm, and Matthias Ziegler. Are working memory and attention related constructs ? *Journal of Individual Differences*, 26(3) :121–131, 2005. 33
- [Bra11] Eric Braverman. Cognitive decline of aging : Important neuroendocrinological predictors of early cognitive decline in a clinical setting, 2011. 37
- [Bri66] Rolf Brickenkamp. *Le Test d2 d'attention concentrée*. Paris : Editest, 1966. 30
- [BSD⁺14] Jan Bobeth, Johann Schrammel, Stephanie Deutsch, Michael Klein, Mario Drobics, Christina Hochleitner, and Manfred Tscheligi. Tablet, gestures, remote control ? : influence of age on performance and user experience with itv applications. In *Proceedings of the 2014 ACM international conference*

- on Interactive experiences for TV and online video*, TVX '14, pages 139–146, 2014. 60
- [Bur13] Paolo Burelli. Adapting virtual camera behaviour. In *In proceedings of the 8th International Conference on the Foundations of Digital Games*, 2013. 53, 88
- [BWA⁺10] Nikolaus Bee, Johannes Wagner, Elisabeth Andre, Fred Charles, David Pizzi, and Marc Cavazza. Interacting with a gaze aware virtual character. In *Proceedings of the 2010 workshop on Eye gaze in intelligent human machine interaction*, EGIHMI 10, pages 71–77, New York, NY, USA, 2010. ACM. 52
- [Cad98] Bernard Cadet. *Psychologie cognitive*. In Press Editions, 1998. 27
- [CAT11] Burcu Cinaz, Bert Arnrich, and Gerhard Tröster. Monitoring of cognitive functioning by measuring reaction times with wearable devices. In *Proc. of the 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, pages 514–517, 2011. 41
- [CEY04] Charles E. Connor, Howard E. Egeth, and Steven Yantis. Visual attention : Bottom-up versus top-down. *Current Biology*, 14(19) :R850 – R852, 2004. 27, 28
- [Cha78] W.G. Chase. *Elementary information processes* :. W.K. Estes (ed.), Handbook of Learning and Cognitive Processes. Lawrence Erlbaum, 1978. 47
- [Che07] Jenova Chen. Flow in games (and everything else). *Commun. ACM*, 50(4) :31–34, April 2007. 73
- [CLR87] J E Clark, A K Lanphear, and C C Riddick. The effects of videogame playing on the response selection processing of elderly adults. *Journal of Gerontology*, 42 :82–85, 1987. 35, 37
- [CVdL02] Fabienne Collette and Martial Van der Linden. *La neuropsychologie de l'attention*, chapter Processus attentionnels et maladie d'Alzheimer, pages 157–173. 2002. <http://hdl.handle.net/2268/18058>. 43

- [CWC⁺10] Cho-Yan Chen, Chi Jen Wang, E Liang Chen, Chi Keng Wu, Yen Kuang Yang, Jeen-Shing Wang, and Pau Choo Chung. Detecting sustained attention during cognitive work using heart rate variability. In *Proc. of the Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 372–375, 2010. 39
- [CWJB12] Dale Cox, Justin Wolford, Carlos Jensen, and Dedrie Beardsley. An evaluation of game controllers and tablets as controllers for interactive tv applications. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction, ICMI '12*, pages 181–188, 2012. 66
- [DES⁺92] R. E. Dustman, R. Y. Emmerson, L. A. Steinhaus, D. E. Shearer, and T. J. Dustman. The effects of videogame playing on neuropsychological performance of elderly individuals. *Journal of Gerontology*, 47 :168–171, 1992. 37
- [ELM14] Katharina Emmerich, Stefan Liszio, and Maic Masuch. Defining second screen gaming : Exploration of new design patterns. In *roceedings of the 11th Conference on Advances in Computer Entertainment Technology, ACE 14*. ACM, 2014. 61
- [ENY06] Magy Seif El-Nasr and Su Yan. Visual attention in 3d video games. In *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology, ACE 06*, New York, NY, USA, 2006. ACM. 52
- [FBS99] J. K. Foster, M. Behrmann, and D. T. Stuss. Visual attention deficits in alzheimer’s disease : Simple versus conjoined feature search. *Neuropsychology*, 13 :223–245, 1999. 43
- [Fra68] Paul Fraisse. Le temps de reaction à un stimulus intercale dans une suite rythmique. *L annee psychologique*, 68(1) :11–22, 1968. 40
- [FRBK89] Paul M. Fischer, John W. Richards, Earl J. Berman, and Dean M. Krugman. Recall and eye tracking study of adolescents viewing tobacco advertisements. *Journal of the American Medical Association*, 261 (1) :90–94, 1989. 50

- [FSWB06] Clifton Forlines, Chia Shen, Daniel Wigdor, and Ravin Balakrishnan. Exploring the effects of group size and display configuration on visual search. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, CSCW '06*, pages 11–20, 2006. 59
- [GB03] C. Shawn Green and Daphne Bavelier. Action video game modifies visual selective attention. *Nature Publishing Group*, 423 :534–537, 2003. 34, 35, 71
- [GB09] William J. Gibbs and Ronan S. Bernas. Visual attention in newspaper versus tv oriented news websites. *Journal of Usability Studies*, 4 :147–165, 2009. 52, 53, 90
- [GDJ89] L. Gauthier, F. Deahut, and Y. Joannette. The bells test : a qualitative and quantitative test for visual neglect. *International Journal of Clinical Neuropsychology*, 11(2) :49–54, 1989. 29
- [GG89] Rudolf Groner and Marina T Groner. Attention and eye movement control : An overview. *European archives of psychiatry and neurological sciences*, 239(1) :9–16, 1989. 49
- [GLE14] Isabela Granic, Adam Lobel, and Rutger C. M. E. Engels. The benefits of playing video games. *American Psychologist*, 69(1) :66–78, 2014. 34
- [GLV⁺02] L. Gueye, E. Legallet, F. Viallet, E. Trouche, and G. Farnarier. Orientation spatiale de l'attention : étude des temps de réaction lors du mouvement de pointage spatial orienting of attention : a study of reaction time during pointing movement. *Neurophysiologie clinique*, 32 :361–368, 2002. 41
- [GM04] Victor M. González and Gloria Mark. "constant, constant, multi-tasking craziness" : Managing multiple working spheres. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04*, pages 113–120, 2004. 57
- [GNS⁺12] E. Guardiola, S. Natkin, D. Soriano, E. Loaser, and P. Vrignaud. Du jeu utile au jeu sérieux. In *Hermes*, volume 62, pages 87–93, 2012. 47
- [Gre97] R. L. Gregory. *Eye and brain. The psychology of seeing*. New York : Oxford University Press, 1997. 51

- [Gro14] NPD Group. The npd group reports 34 million core gamers spend an average of 22 hours per week playing video games. <https://www.npd.com/wps/portal/npd/us/news/press-releases/the-mpd-group-reports-34-million-core-gamers-spend-an-average-of-22-hours-per-week-playing-video-games/>, 2014. Last access 04/2015. 34
- [Gru01] Jonathan Grudin. Partitioning digital worlds : Focal and peripheral awareness in multiple monitor use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 458–465, 2001. 57
- [GSL⁺02] J. H. Goldberg, M. J. Stimson, M. Lewenstein, N. Scottand, and A. M. Wichansky. Eye tracking in web search tasks : design implications. In *Proceedings of the 2002 symposium on Eye tracking research and applications*, ETRA 02, pages 51–58, 2002. 52
- [Gui11] Guil. *Mesure de la difficulté dans les jeux vidéo*. PhD thesis, Conservatoire National des Arts et Métiers, 2011. 131
- [GWR⁺11] Ellen Greimel, Sina Wanderer, Aribert Rothenberger, Beate Herpertz-Dahlmann, Kerstin Konrad, and Veit Roessner. Attentional performance in children and adolescents with tic disorder and co-occurring attention-deficit/hyperactivity disorder : New insights from a 2*2 factorial design study. *Abnorm Child Psychol*, 39 :819–828, 2011. 33
- [HE98] Ignace Th.C. Hooge and Casper J. Erkelens. Adjustment of fixation duration in visual search. *Vision Research*, 38(9) :1295–1302, 1998. 50
- [HJC12] Michael E. Holmes, Sheree Josephson, and Ryan E. Carney. Visual attention to television programs with a second-screen application. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pages 397–400, 2012. 59
- [HLBC09] Sebastien Hillaire, Anatole Lecuyer, Gaspard Breton, and Tony Regia Corte. Gaze behavior and visual attention model when turning in virtual environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, VRST 09, pages 43–50, New York, NY, USA, 2009. ACM. 54, 56

- [HLRC⁺10] Sébastien Hillaire, Anatole Lécuyer, Tony Regia-Corte, Rémi Cozot, Jérôme Royan, and Gaspard Breton. A real-time visual attention model for predicting gaze point during first-person exploration of virtual environments. In *the 17th ACM Symposium on Virtual Reality Software and Technology, VRST '10*, pages 191–198, New York, NY, USA, 2010. ACM. 55
- [HPS08] Juan Pablo Hourcade, Keith B. Perry, and Aditya Sharma. Pointassist : helping four year olds point with ease. In *Proceedings of the 7th international conference on Interaction design and children, IDC 08*, pages 202–209, New York, NY, USA, 2008. ACM. 44
- [Hug06] Stephen B. Hughes. Coordination strategies for assisted viewpoint interaction. In *Proceedings of the 3D User Interfaces, 3DUI 06*, pages 17–24, Washington, DC, USA, 2006. IEEE Computer Society. 54
- [Hyr06] Aulikki Hyrskykari. *Eyes in Attentive Interfaces Experiences from Creating iDict, a Gaze Aware Reading Aid*. PhD thesis, Department of Computer Sciences, University of Tampere, 2006. 57
- [IKN98] 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(11) :1254–1259, 1998. 27
- [Jam81] William James. *The principles of psychology*, volume 1. London : Macmillan, 1981. 26
- [JC76] Marcel Adam Just and Patricia A. Carpenter. Eye fixations and cognitive processes. *Cognitive Psychology*, 8 :441–480, 1976. 51
- [JK03] R. J. K. Jacob and K. S. Karn. Eye tracking in human computer interaction and usability research : Ready to deliver the promises. *The minds eye : Cognitive and applied aspects of eye movement research*, pages 573–603, 2003. 50, 51, 90
- [JPMP04] Holly Jimison, Misha Pavel, James McKanna, and Jesse Pavel. Unobtrusive monitoring of computer interactions to detect cognitive status in elders. *IEEE Transactions on Information Technology in Biomedicine*, 8(3) :248–252, SEPTEMBER 2004. 37

- [Juu03] Jesper Juul. The game, the player, the world : Looking for a heart of gameness. In *Digital Games Research Conference*, pages 30–45. Utrecht : Utrecht University, 2003. 48
- [KF89] R. Klein and M. Farrell. Search performance without eye movements. *Percept Psychophys*, 46(5) :476–482, 1989. 49
- [KKD⁺03] Alan Kenny, Hendrik Koesling, Declan Delaney, Seamus Mcloone, and Tomas Ward. A preliminary investigation into eye gaze data in a first person shooter game. In *Proceedings of the European Conference Modelling and Simulation*, volume 5, pages 146–152, 2003. 52
- [KKS07] Jeongseok Ki, Yong-Moo Kwon, and Kwanghoon Sohn. 3d gaze tracking and analysis for attentive human computer interaction. In *Proceedings of the 2007 Frontiers in the Convergence of Bioscience and Information Technologies*, pages 617–621, Washington, DC, USA, 2007. IEEE Computer Society. 50
- [KM06] Chris Klochek and I. Scott MacKenzie. Performance measures of game controllers in a three-dimensional environment. In *Proceedings of Graphics Interface 2006, GI '06*, pages 73–79, 2006. 67
- [KMFK05] Azam Khan, Justin Matejka, George Fitzmaurice, and Gordon Kurtenbach. Spotlight : Directing users' attention on large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05*, pages 791–798, 2005. 58
- [KMS10] Dagmar Kern, Paul Marshall, and Albrecht Schmidt. Gazemarks : gaze-based visual placeholders to ease attention switching. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 10*, pages 2093–2102, 2010. 57
- [KYA10] Puspa Inayat Khalid, Jasmy Yunus, and Robiah Adnan. Extraction of dynamic features from hand drawn data for the identification of children with handwriting difficulty. *Research in Developmental Disabilities*, 31 :256–262, 2010. 45

BIBLIOGRAPHIE

- [Lik31] Rensis Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 22(140) :1–55, 1931. 106
- [LJM08] Amanda Lenhart, Sydney Jones, and Alexandra Rankin Macgill. Adults and video games. pew internet and american life project. available : www.pewinternet.org. Online, December 2008. 34
- [LKC07] Sungkil Lee, Gerard Jounghyun Kim, and Seungmoon Choi. Real-time tracking of visually attended objects in interactive virtual environments. In *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, VRST 07, pages 29–38, New York, NY, USA, 2007. ACM. 52, 55, 92, 129
- [LL94] M. F. Land and D. N. Lee. Where we look when we steer. *Nature*, 369 :742–744, 1994. 50
- [LLB79] Roy Lachman, Janet L. Lachman, and Earl C. Butterfield. *Cognitive psychology and information processing : An introduction*. Lawrence Erlbaum, 1979. 40
- [LM04] J. Leyba and J. (Malcolm. Eye tracking as an aiming device in a computer game. <http://cite-seerx.ist.psu.edu/viewdoc/download?doi=10.1.1.85.2175&rep=rep1&type=pdf>, 2004. Course work, Clemson University. 52
- [LMG⁺14] Monika Lewandowska, Rafal Milner, Malgorzata Ganc, Elzbieta Wlodarczy, and Henryk Skarzynski. Attention dysfunction subtypes of developmental dyslexia. *Medical Science Monitor*, 20 :2256–2268, 2014. 33
- [LP07] M. Leclercq and J. P. Peters. *BAWL Batterie Attention William Lennox*. Centre Neurologique William Lennox, B.1340 Ottignies Louvain la Neuve, 2007. 29, 32, 74
- [MD02] G. Marmitt and A. T. Duchowski. Modeling visual attention in vr : Measuring the accuracy of predicted scanpaths. In *Proc. of the Eurographics Workshop on Rendering Techniques*, pages 217–226, 2002. 50

BIBLIOGRAPHIE

- [MGB⁺84] G. Mulder, A. B. M. Gloerich, K. A. Brookhuis, H. J. van DeUen, and L. J. M. Mulder. Stage analysis of the reaction process using brain-evoked potentials and reaction time. *Psychological Research*, 46 :15–32, 1984. 47
- [Mik01] Tassos A. Mikropoulos. Brain activity on navigation in virtual environments. *Journal of Educational Computing Research*, 24 :1–12, 2001. 39
- [MNL12] Stephanie Mader, Stephane Natkin, and Guillaume Levieux. How to analyse therapeutic games : The player / game / therapy model. In *Entertainment Computing*, volume 7522 of *ICEC*, pages 193–206, 2012. 71
- [MOIY88] David E. Meyer, Allen M. Osman, David E. Irwin, and Steven Yantis. Modern mental chronometry. *Biol Psychol*, 26 :3–67, 1988. 40
- [MOT⁺12] T. Morooka, T. Ogino, A. Takeuchi, K. Hanafusa, M. Oka, and Y. Ohtsuka. Relationships between the color-word matching stroop task and the go/nogo task :toward multifaceted assessment of attention and inhibition abilities of children. *Acta Med Okayama*, 66(5) :377–386, 2012. 33
- [MR05] L. J. Muir and I.E. Richardson. Perception of sign language and its application to visual communications for deaf people. *The Journal of Deaf Studies and Deaf Education*, 10(4) :390–401, 2005. 50
- [MS04] Thierry Meulemans and Xavier Seron. *L’Examen neuropsychologique dans le cadre de l’expertise médico-légale*. 2004. 22, 27
- [MS09] Karen Murphy and Amy Spencer. Playing video games does not make for better visual attention skills. *Journal of Articles in Support of the Null Hypothesis. JASNH [Online]*, 6 (1) :1–20, 2009. 91
- [NB89] RD Nebes and CB Brady. Focused and divided attention in alzheimer’s disease. *Cortex*, 25(2) :305–315, 1989. 43
- [NCM09] Daniel Natapov, Steven J. Castellucci, and I. Scott MacKenzie. Iso 9241-9 evaluation of video game controllers. In *Proceedings of Graphics Interface 2009*, GI ’09, pages 223–230, 2009. 67

BIBLIOGRAPHIE

- [OI12] Masaki Oshita and Hirotaka Ishikawa. Gamepad vs. touchscreen : A comparison of action selection interfaces in computer games. In *Proceedings of the Workshop at SIGGRAPH Asia, WASA '12*, pages 27–31, 2012. 66
- [Oue03] N. Ouerhani. *Visual Attention : From Bio Inspired Modeling to Real Time Implementation*. PhD thesis, Université de Neuchatel Faculté des Sciences, 2003. 27
- [Pac73] Robert G. Pachella. The interpretation of reaction time in information processing research. Technical Report 45, Human Performance Center, Department of psychology, The university of Michigan, 1973. Prepared for : Air Force Office of Scientific Research. 40
- [Pla91] Reajean Plamondon. On the origin of asymmetric bell-shaped velocity profiles in rapid-aimed movements. In Jean Requin and GeorgeE. Stelmach, editors, *Tutorials in Motor Neuroscience*, volume 62 of *NATO ASI Series*, pages 283–295. Springer Netherlands, Kluwer Academic Publishers, 1991. 45
- [PM82] M. I. Posner and P. McLeod. Information processing models - in search of elementary operations. *Annual Review of Psychology*, 33 :477–514, 1982. 40
- [Pos78] Michael I. Posner. *Chronometric explorations of mind*. L. Erlbaum Associates, 1978. 40
- [PRBL14] Clément Pillias, Raphaël Robert-Bouchard, and Guillaume Levieux. Designing tangible video games : Lessons learned from the sifteo cubes. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems, CHI '14*, pages 3163–3166, 2014. 67
- [PT12] Tan Phat Pham and Yin-Leng Theng. Game controllers for older adults : Experimental study on gameplay experiences and preferences. In *Proceedings of the International Conference on the Foundations of Digital Games, FDG '12*, pages 284–285, 2012. 67
- [PWS08] David Pinelle, Nelson Wong, and Tadeusz Stach. Heuristic evaluation for games : Usability principles for video game design. In *Proceedings of the*

- SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1453–1462. ACM, 2008. 53
- [Ray95] Keith Rayner. Eye movements and cognitive processes in reading, visual search, and scene perception. *Eye Movement Research - Mechanisms, Processes, and Applications*, 6 :3–22, 1995. 50
- [Ray98] Keith Rayner. Eye movements in reading and information processing 20 years of research. *Psychological Bulletin*, 124(3) :372–422, 1998. 57
- [RCB⁺05] George Robertson, Mary Czerwinski, Patrick Baudisch, Brian Meyers, Daniel Robbins, Greg Smith, and Desney Tan. The large-display user experience. *IEEE Comput. Graph. Appl.*, 25 :44–51, 2005. 56
- [Rem80] R. W. Remington. Attention and saccadic eye movements. *J Exp Psychol Hum Percept Perform*, 6(4) :726–744., 1980. 49
- [RFR11a] Ney Renau-Ferrer and Celine Remi. Analyse procedurale : un nouveau paradigme pour leanalyse de traces manuscrits. In *Proc. of the 23rd French Speaking Conference on Human-Computer Interaction*, IHM 11, pages 16 :1–16 :4, New York, NY, USA, 2011. ACM. 46
- [RFR11b] Ney Renau-Ferrer and Celine Remi. A generic approach for recognition and structural modelling of drawersn sketching gestures. In *Proc. of the International Conference on Image Processing, Computer Vision, and Pattern Recognition*, volume 1, pages 99–105, 2011. 46
- [RFR12] Ney Renau-Ferrer and Celine Remi. Procedural analysis of a sketching activity : principles and applications. In *Proc. of the International Conference on Frontiers in Handwriting Recognition*, pages 461–466, 2012. 46
- [RNQ12] Umar Rashid, Miguel A. Nacenta, and Aaron Quigley. The cost of display switching : a comparison of mobile, large display and hybrid ui configurations. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, pages 99–106, 2012. 59
- [SBB96] David Schretlen, Julie Hoida Bobholz, and Jason Brandt. Development and psychometric properties of the brief test of attention. *The Clinical Neuropsychologist*, 10(1) :80–89, 1996. 29

- [SBPS95] Sergio Della Sala, A Baddeley, S Papagno, and H Spinnler. Dual task paradigm : a means to examine the central executive. In *J.Grafam K J, Holyoak & Boller (Eds). Structure and functions of the human prefrontal cortex. Special issue of Annals of the New-York Academy of Sciences*, 769 :161–172, 1995. 29
- [SCGTDT14] Pejman Sajjadi, Edgar Omar Cebolledo Gutierrez, Sandra Trullemans, and Olga De Troyer. Maze commander : A collaborative asynchronous game using the oculus rift & #38 ; the sifteo cubes. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-human Interaction in Play, CHI PLAY '14*, pages 227–236, 2014. 67
- [SDL⁺05] V. Sundstedt, K. Debattista, P. Longhurst, A. Chalmers, and T. Troscianko. Visual attention for efficient high-fidelity graphics. In *Proceedings of the 21st Spring Conference on Computer Graphics, SCCG '05*, pages 169–175. ACM, 2005. 27
- [SFM⁺07] Kurt P. Schulz, Jin Fan, Olga Magidina, David J. Marks, Bella Hahnc, and Jeffrey M. Halperin. Does the emotional go/no-go task really measure behavioral inhibition ? convergence with measures on a non-emotional analog. *Arch Clin Neuropsychol*, 22(2) :151–160, 2007. 33
- [SG06] J. David Smith and T. C. Nicholas Graham. Use of eye movements for video game control. In *Proceedings of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, ACE '06*. ACM, 2006. 52
- [Smi68] E. E. Smith. Choice reaction time : An analysis of the major theoretical positions. *Psychological Bulletin*, 69 :77–110, 1968. 40
- [SND10] Sophie Stellmach, Lennart Nacke, and Raimund Dachselt. Advanced gaze visualizations for three-dimensional virtual environments. In *Proceedings of the 2010 Symposium on Eye Tracking Research Applications, ETRA 10*, pages 109–112, New York, NY, USA, 2010. ACM. 54, 56
- [SSWR08] Veronica Sundstedt, Efstathios Stavrakis, Michael Wimmer, and Erik Reinhard. A psychophysical study of fixation behavior in a computer game. In

- Proceedings of the 5th symposium on Applied perception in graphics and visualization*, APGV 08, pages 43–50, New York, NY, USA, 2008. ACM. 50, 52
- [Str35] Jr Stroop. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18 :643–662, 1935. 29, 30
- [SW65] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52 :591–611, Dec. 1965. 76, 94
- [SZN⁺04] Ben Seltzer, Parvaneh Zolnouni, Margarita Nunez, Robert Goldman, Dinesh Kumar, John Ieni, and Sharon Richardson. Efficacy of donepezil in early-stage alzheimer disease : a randomized placebo controlled trial,. *Arch Neurol*, 61 :1852–1856, 2004. 37
- [Tay76] D.A. Taylor. Stage analysis of reaction time. *Psychological Bulletin*, 83 :161–191, 1976. 40
- [TG80] A. Treisman and G. Gelade. A feature-integration theory of attention. *Cognit Psychol*, 12(1) :97–136, 1980. 30
- [vdMR13] Rob van der Meulen and Janessa Rivera. Gartner says worldwide video game market to total 93 billion dollars in 2013. <http://www.gartner.com/newsroom/id/2614915>, 2013. Last access 04/2015. 34
- [vdSvOWA10] Erik D. van der Spek, Herre van Oostendorp, Pieter Wouters, and Laurens Aarnoudse. Attentional cueing in serious games. In *Proc. of the Second International Conference on Games and Virtual Worlds for Serious Applications*, pages 119–125. IEEE, 2010. 36
- [Viv91] P. Viviani. Eye movements in visual search : cognitive, perceptual and motor control aspects. *Reviews of oculomotor control research*, 4 :353–393, 1991. 50
- [VM14] Radu-Daniel Vatavu and Matei Mancas. Visual attention measures for multi-screen tv. In *Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video*, TVX 14, pages 111–118, 2014. 61

BIBLIOGRAPHIE

- [VT82] P. Viviani and C. Terzuolo. Trajectory determines movement dynamics. *Neuroscience*, 7(2) :431–437, 1982. 44
- [vZB94] Adriaan H. van Zomeren and Wiebo H. Brouwer. *Clinical Neuropsychology of Attention*, volume 1. Oxford University Press, 1994. 26, 29
- [Wan95] Brian A. Wandell. *Foundations of Vision*. Sinauer Associates Inc, 1995. 51
- [War00] C. Ware. *Information Visualization : Perception for Design*. San Francisco : Morgan Kaufman, 2000. 51
- [WGCB11] Jonathan Wonner, Jerome Grosjean, Antonio Capobianco, and Dominique Bechman. Speed : Prédiction de cibles. In *Proc. of the 23rd French Speaking Conference on Human Computer Interaction*, IHM 11, October 24-27 2011. 45
- [Wol98] J. M. Wolfe. Visual search. *H. Pashler*, pages 13–73, 1998. 31
- [Wol01] M. J. Wolf. *The Medium of the Video Game*. University of Texas Press, 2001. <http://utpress.utexas.edu/index.php/books/wolmed>. 34
- [Woo38] Robert Sessions Woodworth. *Experimental psychology*. H. Holt, 1938. 40
- [Yar67] A. Yarbus. *Eye movements and vision*. Plenum Press, 1967. 28
- [Yat66] Frances A. Yates. *The Art of Memory*. The University of Chicago Press, 1966. 25
- [YMJ10] Georgios N. Yannakakis, Héctor P. Martinez, and Arnav Jhala. Towards affective camera control in games. *User Modeling and User-Adapted Interaction*, 20(4) :313–340, 2010. 54
- [YS05] J. Yin and Z. Sun. An online multi-stroke sketch recognition method integrated with stroke segmentation. *Lecture Notes in Computer Science*, 3784 :803, 2005. 46
- [YWK06] Kenji Yokoi, Katsumi Watanabe, and Takashi Kawai. Dynamic evaluation of distribution of visual attention during playing video game. In *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*, ACE 06, New York, NY, USA, 2006. ACM. 52

[ZF94] P. Zimmermann and B. Fimm. Tests d'évaluation de l'attention (tea), 1994.
version 2.3. 29, 32, 108

BIBLIOGRAPHIE

Résumé :

Dans ce travail de thèse, nous proposons d'étudier les méthodes d'évaluation de l'attention ainsi que son implication dans la gestion de la difficulté dans le contexte du jeu vidéo. La prise en compte des processus attentionnels joue un rôle important dans la construction des interfaces homme-machine. La perturbation de ces processus peut, symétriquement, servir de variable d'ajustement des difficultés rencontrées dans les challenges proposés par les game designers. Notre première démarche a consisté à mesurer des temps de réaction de l'utilisateur dans un contexte de jeu sérieux. L'algorithme développé a permis, par une analyse des trajectoires du pointeur contrôlé par le joueur, de détecter des valeurs anormales de temps de réaction. Cela nous a amené à étudier dans quelle mesure la quantité d'information présentée à l'écran pouvait altérer les performances d'un joueur. Notre deuxième approche a consisté donc à mesurer les effets produits par une variation du champ de vision choisi sur la caméra virtuelle dans les jeux à la première personne. Dans la suite de ce travail, nous avons étudié dans quelle mesure la distribution de l'attention d'un utilisateur sur plusieurs dispositifs pour réaliser une simple tâche pouvait modifier les performances comparativement à un dispositif unique. Cette étude a montré que l'utilisation de deux dispositifs en parallèle provoque une augmentation significative de la difficulté.

Mots clés :

Attention, difficulté, performance, jeux vidéo, pointage, caméra virtuelle, écrans multiples.

Abstract :

In this thesis, we study attention evaluation methods and their application to the management of difficulty within the context of video games. The consideration of attentional processes plays an important role in the development of human-computer interfaces. The disruption of these attentional processes can be used symmetrically as a variable for adjusting game difficulty. Our first approach was to measure user reaction time, which we did in the context of a serious game. By analyzing the trajectories of a mouse cursor controlled by the player, our approach allows us to detect abnormal reaction time values. This led us to consider how the amount of information displayed on screen could alter player performance. Our second approach, therefore, was to measure the effects produced by variations in the field of view of the virtual camera in a 3D first-person game. Finally, we studied how the distribution of user attention across multiple devices for the execution of a simple task could affect performance relative to the use of a single device. This study showed that the use of two devices in parallel causes a significant increase in difficulty.

Keywords :

Attention, difficulty, performance, video games, pointing, virtual camera, multiple screens.