HAL open science

Interaction avec les interfaces visuelles dynamiques complexes des jeux vidéo : effets des arrière-plans sur la performance et le comportement du regard

Loïc Caroux

▶ **To cite this version:**

Loïc Caroux. Interaction avec les interfaces visuelles dynamiques complexes des jeux vidéo : effets des arrière-plans sur la performance et le comportement du regard. Psychologie. Université de Poitiers (France), 2012. Français. NNT : . tel-01470826

HAL Id: tel-01470826

https://hal.science/tel-01470826

Submitted on 17 Feb 2017

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

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

Université de Poitiers

UFR Sciences Humaines et Arts

Centre de Recherches sur la Cognition et l'Apprentissage – UMR 7295

INTERACTION AVEC LES INTERFACES VISUELLES DYNAMIQUES COMPLEXES DES JEUX VIDEO

EFFETS DES ARRIERE-PLANS SUR LA PERFORMANCE ET LE COMPORTEMENT DU REGARD

Thèse de doctorat de l'Université de Poitiers

Mention Psychologie

Présentée par

Loïc CAROUX

Sous la direction de

Nicolas VIBERT et **Ludovic LE BIGOT**

Membres du jury

Thierry BACCINO, Professeur, Université Paris 8 (rapporteur)

Christian BASTIEN, Professeur, Université de Lorraine (rapporteur)

Didier BAZALGETTE, Docteur, Direction Générale de l'Armement

Mireille BETRANCOURT, Professeur, Université de Genève

Ludovic LE BIGOT, Professeur, Université de Poitiers (directeur)

Stéphane NATKIN, Professeur, Conservatoire National des Arts et Métiers

Nicolas VIBERT, Directeur de Recherche, Centre National de la Recherche Scientifique (directeur)

20 novembre 2012

Remerciements

Mes premiers remerciements vont à mes directeurs de thèse, Nicolas et Ludovic, qui m'accompagnent depuis le début de mon aventure dans la recherche. Sans eux, je n'aurais pas pu présenter ce travail. Ils m'ont fait partager leur passion sans faille, m'ont quasiment tout appris sur la recherche et ont su m'inculquer le gout de la rigueur intellectuelle nécessaire bien sûr à une science de qualité, mais également à la vie d'un homme. J'espère que je continuerai longtemps à travailler avec eux d'une manière ou d'une autre.

Merci également à Jean-François Rouet qui était là aussi au début de l'aventure et continue depuis à m'apporter son soutien.

Merci à Mme Mireille Bétrancourt, MM. Thierry Baccino, Christian Bastien, Didier Bazalgette et Stéphane Natkin d'avoir accepté d'être membres du jury et d'évaluer cette thèse.

Merci à la Direction Générale de l'Armement d'avoir financé cette thèse, et de m'avoir permis de travailler sereinement et de consacrer toute mon énergie mentale à l'avancée de ce travail.

Merci à mes collègues et amis doctorants du CeRCA qui ont partagé le même bateau que moi. Je pense bien sûr d'abord à mon ex-colocataire du R.09 qui a longtemps cru qu'avec mes superpouvoirs je finirai avant elle. Et non Margot, tu vois que les choses se sont déroulées dans l'ordre des choses. Je pense également à tous ceux qui sont encore dans le bateau, notamment mes compagnons du R.09 (Dom et Nadia) et du R.08 (Julien et Solen), à ceux qui sont se sont échappés avant moi, mais aussi à mes deux condisciples du Master Recherche Alice et Geoffroy. Ils ont tous contribués à leur façon à la réussite de ce travail. Enfin, je pense aux « anciens » qui n'étaient déjà plus doctorants quand je suis arrivé (Antonine, Vanessa, Manu), mais qui étaient là pour guider le nouveau que j'étais.

Merci aussi aux nombreux collègues doctorants et chercheurs que j'ai pu croiser lors de mes nombreux voyages en France et ailleurs, avec qui j'ai pu discuter de jeux vidéo, d'ergonomie et de plein d'autres choses. Ils ont rendu mon aventure encore plus passionnante.

Merci aux membres du CeRCA, et en particulier au personnel technique et administratif, qui ont fait que le plan s'est déroulé sans accroc.

Merci aux 302 participants des expériences réalisées pendant la thèse (j'ai recompté plusieurs fois) pour leur contribution à ce travail et à la science, ils sont si souvent et si rapidement oubliés.

Enfin merci aux personnes qui comptent le plus pour moi. Merci à mes amis « de ch'nord », ma sœur, mon frère, qui ont su accepter que je m'expatrie si loin et tout ce temps juste pour des jeux vidéo (et ça ne va pas s'arranger), et qui ne m'en tiennent pas rigueur lorsqu'on se retrouve autour d'une bonne bière. Merci à mes parents pour leur soutien pendant ces longues années d'étude (13 ans de carte d'étudiant quand même) qui n'auraient pas existées sans eux.

Merci à mes petits bouts, Théo et Yanis, de supporter tant bien que mal que Papa « y va travailler à son travail ». Et puis bien sûr, Pauline, merci tout court.

Table des matières

INTRODUCTION **9**

1. Les jeux vidéo aujourd'hui **9**

2. Vers une utilisabilité spécifique des jeux vidéo ? **12**
2.1. De l'utilisabilité à l'expérience utilisateur 12
2.2. Modèles d'utilisabilité adaptés aux jeux vidéo 14

3. Présentation de la thèse **16**

CHAPITRE 1 : LES INTERACTIONS JOUEUR-JEU VIDEO **19**

1. Les jeux vidéo comme environnements virtuels **20**
1.1. Caractéristiques des interfaces des jeux vidéo 20
1.2. Le point de vue de l'utilisateur 24

2. Les jeux vidéo comme environnements dynamiques complexes **28**
2.1. Interaction avec des environnements dynamiques complexes 28
2.2. Partage attentionnel lors de l'interaction avec des environnements dynamiques complexes 31

3. Synthèse **37**

CHAPITRE 2 : PERCEPTION DES ENVIRONNEMENTS VISUELS DYNAMIQUES COMPLEXES **39**

1. La perception visuelle de scènes complexes **40**
1.1. Le guidage de l'attention visuelle 40
1.2. Les mouvements oculaires en tant qu'indicateurs de l'attention visuelle 43
1.3. La recherche visuelle 44

2. Perception et interaction personne-système dans des scènes visuelles complexes et/ou dynamiques **48**
2.1. Scènes visuelles dynamiques 48
2.2. Scènes visuelles complexes 54

3. Synthèse **58**

CHAPITRE 3 : SYNTHESE ET HYPOTHESES GENERALES 61

1. Synthèse des principaux éléments théoriques 61

2. Raisonnement et hypothèses générales 63

CHAPITRE 4 : MAXIMISER L'ANTICIPATION VISUELLE DES JOUEURS EN APPLIQUANT LE PRINCIPE DE PROXIMITE-COMPATIBILITE A LA CONCEPTION DES JEUX VIDEO 67

1. Introduction 67

2. Expérience 1 : Score en haut ou en bas de l'écran de jeu 70
2.1. Méthode 70
2.2. Résultats 75
2.3. Discussion 77

3. Expérience 2 : Score en haut, en bas, au milieu ou en dessous de l'écran de jeu 80
3.1. Méthode 80
3.2. Résultats 82
3.3. Discussion 84

4. Discussion générale 86

CHAPITRE 5 : LA PERFORMANCE A UN JEU VIDEO EST PLUS DEGRADEE PAR LA COMPLEXITE STRUCTURALE DE L'ARRIERE-PLAN QUE PAR SON MOUVEMENT 89

1. Introduction 89

2. Expérience 3a : Tâche de recherche visuelle 92
2.1. Méthode 94
2.2. Résultats 97
2.3. Discussion 99

3. Expérience 3b : Tâche de recherche visuelle et structure radiale 101
3.1. Méthode 101
3.2. Résultats 102
3.3. Discussion 104
3.4. Discussion générale des expériences 3a et 3b 105

4. Expérience 4 : Tâche de tir 106
4.1. Méthode 107
4.2. Résultats 110
4.3. Discussion 115

5. Discussion générale 118

CHAPITRE 6 : COMPLEXITE VISUELLE DE L'ARRIERE-PLAN ET DIFFICULTE DE LA TACHE INTERAGISSENT AVEC LE MOUVEMENT DE L'ARRIERE-PLAN DANS UN JEU VIDEO 121

1. Introduction 121

2. Expérience 5 : Complexité et mouvement de l'arrière-plan 123
2.1. Méthode 124
2.2. Résultats 128
2.3. Discussion 131
2.4. Conclusion 134

3. Expérience 6 : Difficulté de la tâche et mouvement de l'arrière-plan 135
3.1. Méthode 136
3.2. Résultats 138
3.3. Discussion 141

4. Discussion générale 142

CHAPITRE 7 : DISCUSSION GENERALE 145

1. Partage attentionnel dans les interfaces visuelles de jeux vidéo 146
1.1. Modèles et théories existants 146
1.2. Modèle de partage attentionnel entre deux sources visuelles d'information superposées 148
1.3. Eléments non pris en compte dans le modèle 150

2. Recommandations pour la conception des jeux vidéo 152
2.1. Anticipation visuelle du mouvement et positionnement des informations contextuelles 153
2.2. Mouvement de l'arrière-plan, structure visuelle, et complexité de la tâche 154

3. Perspectives 156

CONCLUSION 161

BIBLIOGRAPHIE 165

ANNEXES 181

Introduction

1. Les jeux vidéo aujourd'hui

Les jeux vidéo occupent aujourd'hui une place importante dans notre société. Le jeu vidéo est désormais la première industrie culturelle en France et dans le monde, devant le cinéma notamment (Syndicat National du Jeu Vidéo, 2011). Le chiffre d'affaires du jeu vidéo représentait en 2011, 2,7 milliards d'euros en France (52 milliards dans le monde). Depuis le premier succès commercial du jeu « Pong » en 1972, les jeux vidéo se sont développés sans cesse et de manière très diversifiée. Les jeux vidéo sont pratiqués par différentes catégories de joueurs et sur différents types de support. 63 % des Français de 10 ans et plus (âge moyen 35 ans) ont déclaré avoir joué à des jeux vidéo en 2011, dont 52% de femmes et 48% d'hommes (Syndicat National du Jeu Vidéo, 2011). Actuellement, 40% du temps consacré au jeu vidéo se passe sur ordinateur, 24% sur des consoles de salon, 15% sur des consoles portables, 13% sur des téléphones mobiles et 7% sur des tablettes (Institut de sondage GfK ISL, 2012).

A l'origine, le seul but d'un jeu vidéo était de divertir le joueur. Les finalités sont aujourd'hui multiples. Elles incluent par exemple l'éducation (Egenfeldt-Nielsen, 2006 ; Orvis, Horn, & Belanich, 2008), la simulation d'aviation militaire (Gopher, Weil, & Bareket, 1994), l'entrainement à des techniques chirurgicales (E. Boyle, Kennedy, Traynor, & Hill, 2011) ou encore l'expérimentation

scientifique (Frey, Hartig, Ketzel, Zinkernagel, & Moosbrugger, 2007 ; Washburn, 2003). Ces jeux sont appelés « jeux sérieux » (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012), jeux éducatifs (Virvou & Katsionis, 2008) ou jeux d'exercice physique (Jin, 2009). Cependant, malgré ces objectifs différents, un même cadre scientifique peut être utilisé quel que soit le type de jeu pour étudier l'interaction entre le joueur et le système.

L'étude scientifique des jeux vidéo reste difficile car la définition d'un jeu vidéo est encore aujourd'hui imprécise (Pagulayan, Keeker, Wixon, Romero, & Fuller, 2008). Le dénominateur commun aux jeux vidéo est la possibilité pour le joueur d'interagir avec un environnement virtuel, généralement à l'aide d'un artefact (e.g., manette de jeu, capteur de reconnaissance de mouvement), environnement dont l'interface principale est restituée sous une forme audio-visuelle par un terminal spécifique aux jeux vidéo (e.g., console de jeux portable) ou non (e.g., téléviseur, téléphone mobile de dernière génération). Cet environnement réagit aux ordres de l'individu et renvoie alors de nouvelles informations par le biais de l'interface. Contrairement à d'autres médias qui véhiculent des informations audio-visuelles comme le cinéma, les jeux vidéo permettent l'interaction avec ces informations. Cette interaction joueur-jeux vidéo est certainement le point central qui différencie les jeux des autres types de médias culturels.

Les jeux vidéo sont considérés par certains comme avant tout des produits artistiques (Gee, 2006 ; Murphy, 2010). Cette idée est confortée par plusieurs éléments. La création d'un jeu vidéo requiert notamment un scénario, linéaire ou non, des éléments de graphisme généralement animés ou encore des éléments sonores et musicaux. Plusieurs créateurs ou artistes aux compétences différentes travaillent ainsi ensemble pour aboutir à un projet commun. Dans le cadre de cette création artistique, le but des concepteurs est souvent de suivre la vision artistique voulue par le ou les créateurs du jeu sans nécessairement prendre en compte l'interaction finalisée joueur-jeu vidéo.

La démocratisation des jeux vidéo a fait apparaitre dans la littérature scientifique de nombreuses publications liées à cette interaction joueur-jeu vidéo. Elles restent toutefois encore majoritairement

dédiées à l'impact de l'utilisation du jeu sur le comportement extérieur du joueur. Des études ont par exemple examiné l'influence négative des jeux vidéo violents sur le comportement (e.g., Anderson et al., 2010). Ces études ont montré que la pratique de jeux vidéo violents et asociaux pouvait faire naitre des pensées et des comportements agressifs, aussi bien à court- qu'à long-terme. A l'opposé de ces études, d'autres études ont mis en avant les effets positifs de l'entrainement cognitif observés lors de l'interaction avec des jeux vidéo d'action (e.g., Bavelier et al., 2011). Ces études ont notamment montré que les joueurs de jeux vidéo d'action avaient une attention spatiale plus développée et une acuité visuelle plus fine que les non-joueurs. Ces améliorations s'obtenaient par un simple entraînement.

Plus récemment, une autre communauté de chercheurs a publié certains résultats non pas sur l'effet des jeux vidéo sur le joueur, mais sur la nature même de l'interaction entre le joueur et le jeu vidéo, en s'attachant aux motivations des joueurs (e.g., recherche de plaisir, de divertissement, de challenge, d'émotions) (Barr, Noble, & Biddle, 2007). Les jeux vidéo représentent aujourd'hui une part importante des expériences digitales accessibles au grand public, notamment les jeux vidéo de dernière génération qui proposent des scènes visuelles en relief, ou ceux qui utilisent la reconnaissance de mouvements. Etudier les jeux vidéo et comprendre l'influence, par exemple, de leurs caractéristiques visuelles sur le comportement des joueurs s'avère très utile pour optimiser leurs interfaces. Ces études aident à minimiser les perturbations dans l'interaction entre joueurs et jeux vidéo de sorte que l'expérience vécue par les joueurs soit la meilleure possible. Les études sur l'interaction joueur-jeu vidéo centrées sur le joueur sont d'une importance capitale, en particulier parce qu'elles permettent aux concepteurs de créer des jeux vidéo qui s'ajustent aux attentes et aux capacités spécifiques des joueurs (Pagulayan et al., 2008).

Le marché des jeux vidéo est un secteur concurrentiel. Les jeux vidéo qui ont une viabilité commerciale sont ceux qui bénéficient d'une conception qui prend en considération le joueur. L'idée selon laquelle une approche centrée-utilisateur serait un coût supplémentaire pour la conception

d'un jeu vidéo semble perdurer chez les concepteurs. En réalité, une conception centrée-utilisateur est rentabilisée économiquement par une plus grande réussite commerciale (Nørgaard & Sørensen, 2008). Plusieurs entreprises de jeu vidéo communiquent régulièrement sur la conception centrée sur le joueur, comme par exemple Microsoft Studios (Etats-Unis) qui possède un département de « recherche utilisateur », le « Microsoft Studios' user research group » (e.g., Jun et al., 2008 ; Pagulayan et al., 2008).

2. Vers une utilisabilité spécifique des jeux vidéo ?

Pour concevoir les systèmes interactifs classiques du point de vue de l'utilisateur, la principale notion utilisée est l'utilisabilité. Shackel (2009) définissait l'utilisabilité d'un système interactif comme sa capacité à être utilisé facilement et efficacement par des utilisateurs particuliers, ayant un niveau d'expertise précis, pour une tâche et dans un environnement spécifiques. Toutefois, l'utilisabilité reste mal adaptée à la conception des jeux vidéo, qui sont des systèmes de divertissement. Les systèmes pour lesquels l'utilisabilité est définie sont généralement des systèmes où ce qui doit être maximisé est la productivité. La notion « d'expérience utilisateur » est préférée lorsque les sensations de l'utilisateur pendant l'interaction doivent être prises en compte. Finalement, des adaptations de l'utilisabilité ont été proposées pour qualifier l'interaction joueur-jeu vidéo de manière plus spécifique.

2.1. De l'utilisabilité à l'expérience utilisateur

Dans le domaine des applications informatiques, le critère d'utilisabilité est aujourd'hui prépondérant dans la conception des interactions personne-système (Alonso-Ríos, Vázquez-García, Mosqueira-Rey, & Moret-Bonillo, 2010 ; Dumas & Salzman, 2006 ; Hertzum, 2010 ; Hornbæk, 2006 ; Shackel, 2009). Il est particulièrement pertinent pour mesurer la qualité d'une interaction entre un système et son utilisateur. Classiquement, l'utilisabilité d'un système répond à trois critères

principaux (Norme ISO 9241-11, 1998). Elle permet à l'utilisateur d'atteindre son but (efficacité), de réaliser sa tâche avec un minimum de moyens et de ressources (efficience), et enfin d'avoir un sentiment positif pendant la réalisation de la tâche (satisfaction).

Les jeux vidéo sont cependant des systèmes interactifs particuliers, car ce sont des systèmes de divertissement où les buts des joueurs sont définis par un concepteur et se situent à l'intérieur d'un monde virtuel (Pagulayan et al., 2008). Pagulayan et al. soulignent que l'utilisabilité des jeux vidéo, au sens classique du terme, risque de ne pas suffire pour prendre en compte toutes les spécificités de l'interaction joueur-jeu vidéo. Les études traitant de l'utilisabilité des jeux vidéo se heurtent notamment aux limites du concept proposé à l'origine pour des systèmes liés au travail et non au loisir ou au divertissement (voir aussi Zaphiris & Ang, 2007). L'utilisabilité pour un jeu vidéo consiste en général à adapter d'une manière empirique des critères utilisés pour les interfaces de logiciels plus classiques (Isbister & Schaffer, 2008). Pourtant, des auteurs tels que Johnson et Wiles (2003) ont argumenté que plusieurs règles d'utilisabilité des interactions personne-système devaient être violées pour que le joueur puisse ressentir du plaisir en jouant. Par exemple, la tâche ne devrait pas être simplifiée au maximum pour faciliter l'interaction, mais adaptée pour maintenir élevée la motivation du joueur (Malone & Lepper, 1987).

La problématique de l'adaptation de l'utilisabilité à des systèmes non productifs s'est aussi posée pour d'autres systèmes que les jeux vidéo. Plus généralement, les systèmes et produits de la vie quotidienne font aussi l'objet de critères d'utilisabilité trop restrictifs (Norman, 2004). Depuis plusieurs années, certaines études ont montré notamment que l'esthétique était importante pour l'utilisation optimale d'un système (Hassenzahl & Monk, 2010 ; Lavie, Oron-Gilad, & Meyer, 2011 ; Moshagen, Musch, & Göritz, 2009 ; Sonderegger & Sauer, 2010 ; Sonderegger, Zbinden, Uebelbacher, & Sauer, 2012). Les performances d'utilisateurs d'appareils courants comme les téléphones portables (Sonderegger & Sauer, 2010) ou les systèmes d'aide à la navigation (Lavie et al., 2011) sont meilleures avec les appareils dont les interfaces sont jugées les plus esthétiques.

Un concept différent de l'utilisabilité est de plus en plus utilisé aujourd'hui pour l'étude des systèmes qui ne sont pas tournés exclusivement vers la productivité : « l'expérience utilisateur » (Hassenzahl, Diefenbach, & Göritz, 2010 ; Hassenzahl & Tractinsky, 2006). L'expérience utilisateur se concentre sur les perceptions et les réponses qui résultent de l'utilisation ou de l'anticipation de l'utilisation d'un produit, d'un système ou d'un service (Norme ISO 9241-210, 2010). De manière schématique, ce concept prend en compte les critères de l'utilisabilité (efficacité, efficience et satisfaction) tout en mettant en avant les aspects sociocognitifs et affectifs de l'interaction entre l'utilisateur et le système (Barcenilla & Bastien, 2009 ; Law & van Schaik, 2010). Toutefois, et contrairement à l'utilisabilité, il n'existe pas de consensus dans la communauté scientifique sur la définition de l'expérience utilisateur, ni d'ailleurs sur son positionnement par rapport à l'utilisabilité (Law & van Schaik, 2010). Par exemple, la question de savoir si l'expérience utilisateur est une subdivision de l'utilisabilité ou l'inverse reste posée. Ou encore, les critères importants qui définissent une expérience utilisateur optimale, comme la qualité pragmatique (utilisabilité perçue par l'utilisateur), la qualité hédonique (qualités du produit générant du plaisir), la beauté (esthétique) ou encore la qualité globale du produit, ne sont pas encore bien définis. Malgré une définition encore incertaine de l'expérience utilisateur, les études sur l'interaction joueur-jeu vidéo semblent désormais privilégier ce concept à celui d'utilisabilité (Bernhaupt, 2010). Pourtant, plusieurs travaux ont proposé des approches alternatives aux concepts d'utilisabilité et d'expérience utilisateur dans le cadre de la conception des jeux vidéo.

2.2. Modèles d'utilisabilité adaptés aux jeux vidéo

La notion d'expérience utilisateur permet d'étudier n'importe quelle interaction personne-système. Mais parallèlement, d'autres modèles d'étude de l'interaction joueur-jeu vidéo sont proposés (Barr et al., 2007 ; Calvillo Gámez, Cairns, & Cox, 2010 ; Fabricatore, Nussbaum, & Rosas, 2002 ; González Sánchez, Zea, & Gutiérrez, 2009 ; Wood, Griffiths, Chappell, & Davies, 2004). L'objectif de ces modèles, présentés dans les paragraphes suivants, est de prendre en compte le plus possible et le

plus précisément possible les spécificités de l'interaction joueur-jeu vidéo, comme le plaisir du jeu ou le challenge.

Une première approche consiste à garder les critères d'utilisabilité des interfaces classiques, mais à en changer la signification. González Sánchez et al. (2009) ont proposé une évolution du concept d'utilisabilité vers le concept de « jouabilité ». La jouabilité d'un jeu vidéo serait le point jusqu'auquel des utilisateurs spécifiques réalisent des buts spécifiques avec efficacité, efficience, et surtout satisfaction et amusement dans un contexte ludique. En plus du critère d'efficacité, le critère d'efficience serait divisé en 2 critères distincts (apprenabilité et immersion), et le critère de satisfaction en 4 critères (satisfaction, motivation, émotion et socialisation). Cette définition se rapproche de la notion d'expérience utilisateur utilisée pour les systèmes interactifs en général.

Une deuxième approche, complémentaire de la conception centrée-utilisateur des jeux vidéo, consiste à développer des modèles explicatifs de l'utilisation des jeux par les joueurs. Basés sur l'observation de joueurs en situation de jeu, ces modèles décrivent les différents éléments constitutifs d'une interaction joueur-jeu vidéo. Fabricatore et al. (2002) ont proposé un modèle de « jouabilité » regroupant des aspects-clés de la conception d'un jeu basés sur les préférences des joueurs. Ces auteurs ont étudié qualitativement des sessions de jeu sur une sélection de jeux vidéo d'action bien notés par les critiques du domaine. Ils ont pu déterminer un modèle hiérarchique de paramètres qui influencent la jouabilité, regroupés en 3 catégories : les paramètres de l'interface, ceux du scénario, et ceux liés à la hiérarchisation des buts. Chaque catégorie est elle-même divisée en sous-catégories. Les paramètres de l'interface regroupent les éléments liés à la présentation physique du jeu, notamment visuelle, et particulièrement les objets et/ou personnages avec lesquels le joueur interagit. Le scénario regroupe les éléments concernant le contexte et l'histoire dans lesquels évoluent les personnages et les objets, qui peuvent par exemple déterminer les points de vue de caméra. Enfin, la hiérarchisation des buts détermine la manière dont se déroule le jeu dans le temps. Pour chaque catégorie de paramètres, les auteurs ont proposé des recommandations aux

concepteurs de jeux. Par exemple, pour les paramètres de l'interface, ils ont recommandé aux concepteurs de permettre au joueur d'avoir une perception précise du statut d'énergie de son avatar (objet ou entité qui représente le joueur dans le monde virtuel) et de ses changements.

Barr et al. (2007) ont quant à eux proposé de qualifier l'interaction joueur-jeu vidéo en termes de « valeurs » qui correspondraient aux objectifs spécifiques que poursuivent les joueurs en fonction des types de jeu vidéo. Par exemple, dans les jeux de type « jeu de rôle », (a) l'accomplissement de quêtes, (b) l'exploration du monde, et (c) l'amélioration du niveau de l'avatar sont trois valeurs qui font que les joueurs ont l'impression de progresser dans le jeu. D'après ces auteurs, c'est la satisfaction de ces valeurs qu'il est nécessaire de prendre en compte et de quantifier pour évaluer l'utilisabilité, ou plus généralement l'attrait d'un jeu vidéo. Toutefois, la notion de valeur reste difficile à maîtriser puisque chaque type de jeu posséderait ses valeurs propres.

Tous ces modèles proposent des visions alternatives de l'utilisabilité des jeux vidéo. Il reste difficile d'évaluer ces modèles les uns par rapport aux autres. Chacun prône des approches et utilise des notions assez différentes. Finalement, la question de savoir si les jeux vidéo nécessitent une redéfinition de l'utilisabilité est posée. Les besoins de redéfinition de l'utilisabilité semblent moins présents depuis que la notion d'expérience utilisateur est utilisée pour qualifier les interactions personne-système.

3. Présentation de la thèse

La thèse défendue ici se situe dans le cadre des recherches centrées sur la compréhension et l'optimisation des interactions joueur-jeu vidéo pour une meilleure expérience du joueur. L'objectif est de comprendre l'influence de la conception des interfaces visuelles, en fonction de leurs caractéristiques dynamiques et de leur complexité, sur la performance et le comportement du regard des joueurs.

Ce document de thèse se structure en une partie théorique (chapitres 1 à 3), une partie expérimentale (chapitres 4 à 6) et une discussion générale (chapitre 7). Le chapitre 1 aborde les interactions entre le joueur et le jeu vidéo du point de vue du joueur, en soulignant plus généralement les similitudes avec les interactions entre environnements virtuels et utilisateurs. Ce chapitre précise aussi que les jeux vidéo sont des environnements dynamiques complexes, et que les théories de l'attention partagée élaborées à la base pour des activités multitâches peuvent être appliquées à leur conception. Le chapitre 2 montre qu'une conception précise des interfaces visuelles des jeux vidéo nécessite aussi de prendre en compte les théories de la perception de scènes visuelles dynamiques et complexes. Le raisonnement de la thèse et les hypothèses sont ensuite résumées dans le chapitre 3.

Les trois chapitres de la partie expérimentale s'intéressent aux caractéristiques dynamiques et à la complexité des interfaces (objets avec lesquels le joueur doit interagir et arrière-plans), et à leur impact sur la performance et le comportement du regard des joueurs. L'objectif du chapitre 4 est de comprendre l'influence de la position d'éléments contextuels d'un jeu vidéo, comme le score, dans un jeu ou les éléments de l'arrière-plan sont en mouvement. L'objectif du chapitre 5 est de comprendre l'influence des mouvements d'ensemble de l'arrière-plan et de la complexité des structures visuelles qui permettent de générer ces mouvements. L'objectif du chapitre 6 est de comprendre de manière plus générale les influences croisées des différents mouvements d'ensemble de l'arrière-plan et du niveau de complexité visuelle de cet arrière-plan. L'influence du type de tâche à réaliser ou de la difficulté de cette tâche sont aussi prises en compte dans les chapitres 5 et 6.

Enfin, le chapitre 7 est consacré à la discussion de l'ensemble des résultats obtenus. Cette section propose un cadre théorique actualisé des interactions avec les environnements dynamiques complexes en général, et avec les jeux vidéo en particulier, et en tire des implications pratiques.

Chapitre 1 :
Les interactions joueur-jeu vidéo

L'objectif de ce chapitre est de montrer que du point de vue ergonomique, les jeux vidéo doivent être considérés avant tout comme des systèmes avec lesquels un individu souhaite interagir dans le cadre d'une activité finalisée. La première section du chapitre défend l'idée que les jeux vidéo sont assimilables à des environnements virtuels. Contrairement avec ce qui se passe dans les systèmes informatiques classiques, des aspects subjectifs tels que « l'engagement » du joueur ou les sensations ressenties pendant l'utilisation du jeu font également partie de l'interaction. La seconde section défend l'idée que l'interaction avec les jeux vidéo est dynamique et complexe. Le joueur gère des informations complexes qui évoluent de manière permanente. De ces aspects complexes et dynamiques découle une spécificité importante, la nécessité du partage du temps et des ressources attentionnelles du joueur entre plusieurs tâches simultanées. Certaines théories, comme la théorie des ressources multiples de Wickens (1984, 2002, 2008), peuvent aider à la conception de ces systèmes en prenant en compte les limites et les capacités des joueurs.

1. Les jeux vidéo comme environnements virtuels

Les environnements virtuels sont des environnements qui simulent la présence physique de l'utilisateur dans un lieu qui imite le monde réel, ou représente un monde imaginaire (Stanney, Mourant, & Kennedy, 1998). Ces environnements se caractérisent d'une part par des particularités de l'interface du système, et d'autre part par la manière dont l'utilisateur aborde et utilise cette interface (Stanney, Mollaghasemi, Reeves, Breaux, & Graeber, 2003). Leur interface présente deux caractéristiques principales : (a) l'utilisateur peut interagir avec les éléments de l'environnement et (b) les informations sont présentées sous des formes visuelles, auditives, haptiques, ou diverses combinaisons de deux ou trois de ces modalités. Du point de vue de l'utilisateur, la manière dont est perçue l'interface dépend du degré d'engagement de l'utilisateur dans l'activité et des sensations qu'il ressent pendant et après l'utilisation de l'environnement virtuel.

1.1. Caractéristiques des interfaces des jeux vidéo

Deux éléments sont à prendre en compte pour décrire l'interface d'un environnement virtuel tel qu'un jeu vidéo : d'une part la manière dont l'individu peut interagir avec l'environnement, et d'autre part les informations données par l'interface qui permettent cette interaction à l'utilisateur (Stanney et al., 2003). Les paragraphes suivants se concentrent essentiellement sur l'impact du mode de présentation des informations nécessaires à l'interaction sur le comportement du joueur. Ces informations peuvent être fournies selon des modalités sensorielles différentes, mais la modalité visuelle est celle qui reste généralement privilégiée (Stanney et al., 2003).

1.1.1. Les informations visuelles

Les interfaces visuelles des jeux vidéo sont généralement constituées d'une scène d'action principale composée d'objets avec lesquels le joueur doit interagir (e.g., avatars, ennemis ou cibles) et d'un arrière-plan complexe et en mouvement (e.g., décors, paysages). Un « affichage tête-haute » est souvent superposé à la scène d'action principale. Cet affichage fournit des informations

contextuelles liées à la situation en cours (Brooksby, 2008), sous forme de mots, listes de mots, chiffres ou symboles. Généralement, les informations relativement permanentes sont disposées dans les zones périphériques de l'écran (e.g., score, points de vie, carte du monde virtuel). Les informations non-permanentes sont par contre affichées sur la zone centrale (e.g., messages d'alertes) (voir exemples en figure 1). Les affichages tête-haute sont conçus en fonction du contexte. Certains augmentent la perception de l'environnement réel par les utilisateurs en superposant des informations contextuelles à la scène réelle (réalité augmentée). Les affichages tête-haute utilisés dans les jeux vidéo sont assimilables à ceux utilisés, par exemple, pour aider au pilotage d'avions (Alexander, Wickens, & Hardy, 2005 ; Fadden, Ververs, & Wickens, 2001 ; Martin-Emerson & Wickens, 1997 ; Wickens & Long, 1995). Cependant, la plupart des éléments des affichages tête-haute des jeux vidéo sont opaques et n'ont pas de lien particulier avec la partie de l'écran à laquelle ils sont superposés. Ils peuvent alors masquer des éléments de la scène principale du jeu (figure 1).

L'influence d'éléments particuliers de ces interfaces visuelles sur la performance et/ou l'expérience du joueur a fait l'objet de différentes études. Par exemple, Wolfson et Case (2000) ont étudié les effets de la couleur du fond de l'écran sur la performance des joueurs. L'hypothèse des auteurs était que la performance et l'éveil mental (mesuré par le rythme cardiaque) des joueurs au cours du jeu étaient influencés par la couleur, rouge ou bleu, du fond d'écran. Les principaux résultats ont montré que le score et le rythme cardiaque augmentaient régulièrement au cours des parties lorsque le fond d'écran était bleu. Lorsque le fond était rouge, le score et le rythme cardiaque augmentaient plus rapidement que lorsqu'il était bleu, mais diminuaient dans la deuxième moitié des parties. La couleur rouge augmentait plus rapidement l'éveil, mais un effet d'habituation en provoquait ensuite la diminution. Quelle que soit la couleur du fond, la performance était liée au niveau d'éveil mental. Hou, Nam, Peng, et Lee (2012) ont quant à eux étudié l'influence de la taille de l'écran de jeu sur l'impression « d'immersion » dans le jeu et les sensations subjectives des joueurs. La notion d'immersion est détaillée plus loin dans ce chapitre mais brièvement, le niveau d'immersion ressentie est lié au niveau d'implication du joueur dans le monde virtuel proposé par le jeu. Un

même jeu vidéo était présenté sur un écran de petite taille (type téléviseur) ou de grande taille (type vidéo-projection). Les résultats ont montré que jouer sur un écran de grande taille augmentait l'impression d'immersion, mais n'influençait pas l'évaluation subjective du jeu en termes d'intérêt, d'originalité, de qualité graphique, et d'intention de recommander le jeu à d'autres.

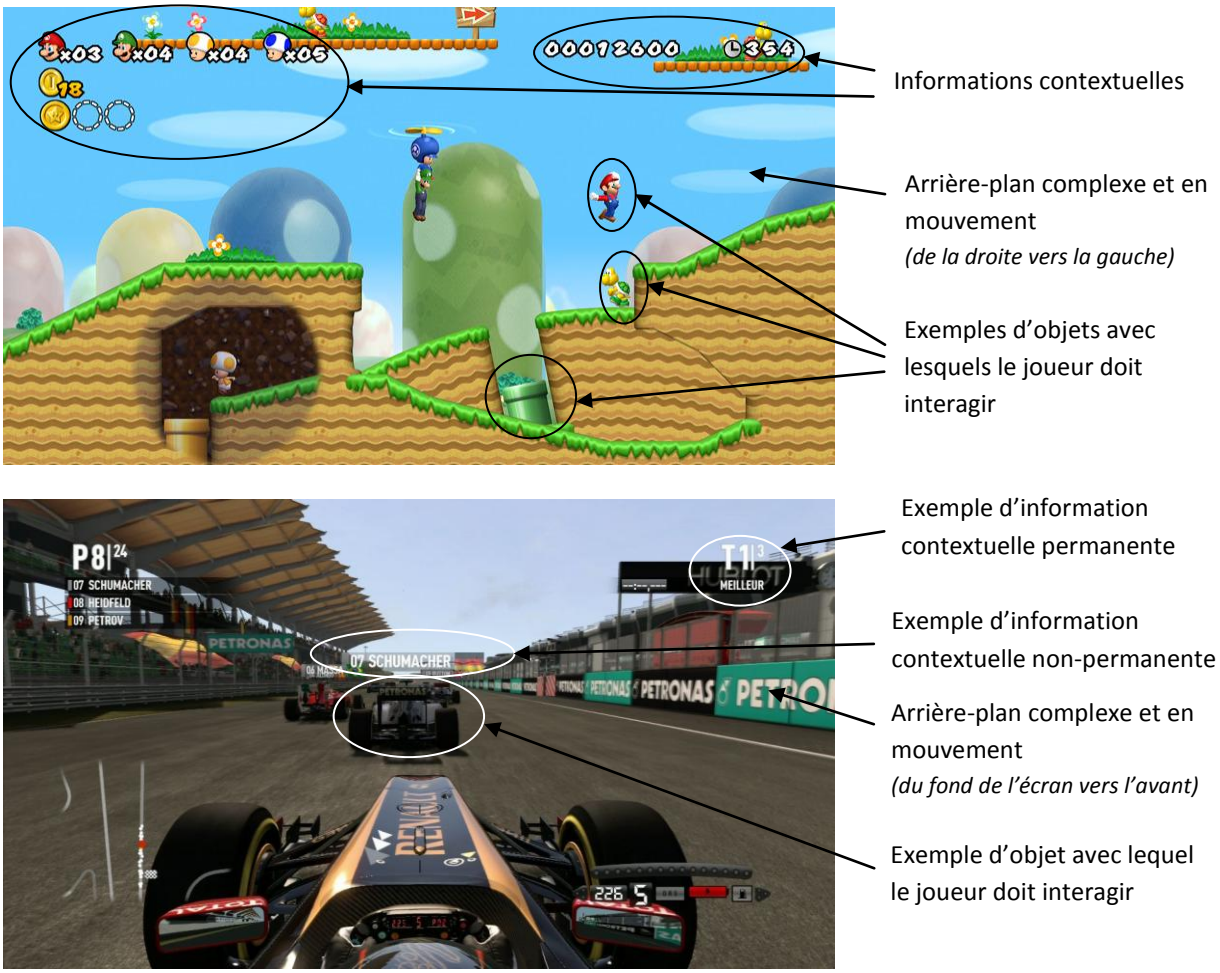


Figure 1. Captures d'écran des interfaces visuelles de deux jeux vidéo commerciaux récents.
Elles montrent les trois éléments principaux d'une interface visuelle classique d'un jeu vidéo : les objets avec lesquels le joueur doit interagir, un arrière-plan complexe et en mouvement (ces deux types d'éléments composent la scène d'action principale) et les informations contextuelles (qui composent l'affichage tête-haute). Les informations contextuelles disposées à la périphérie de l'écran représentent des informations permanentes et celles disposées dans la partie centrale représentent des informations non-permanentes (exemple dans la capture d'écran du bas). Les informations contextuelles peuvent masquer des objets de la scène principale d'action (exemple dans la capture d'écran du haut). Les types de mouvement de l'arrière-plan et leurs conséquences sont détaillés dans le chapitre 5. La capture d'écran du haut est tirée du jeu *New Super Mario Bros Wii* (Nintendo, 2009). La capture d'écran du bas est tirée du jeu *F1 2011* (Codemasters, 2011).

Les effets de la disposition de différents éléments constitutifs des interfaces visuelles des jeux vidéo ont également été étudiés. Par exemple, Sabri, Ball, Fabian, Bhatia, et North (2007) ont montré que sur des interfaces de jeu vidéo en haute résolution comprenant plusieurs moniteurs, les informations

contextuelles importantes les plus regardées ou utilisées devaient être les plus proches possibles du curseur contrôlé par le joueur. Sur la configuration à plusieurs moniteurs, les joueurs étaient plus performants quand les informations contextuelles étaient affichées sur le moniteur effectivement utilisé.

Ces études ont montré que la présentation des informations visuelles a une influence importante sur la performance des joueurs et l'expérience qu'ils ressentent. La section suivante s'intéresse aux autres types d'informations, auditives et haptiques, qui font également partie de l'expérience de jeu.

1.1.2. Les informations auditives et haptiques

Les informations auditives et haptiques sont utilisées pour renforcer ou faciliter l'interaction joueur-jeu. Généralement, le rôle des informations auditives est d'augmenter la saillance des indices visuels en facilitant leur localisation ou leur détection par « sonification » (Stanney et al., 2003). La sonification se réfère à la présentation de certaines informations simultanément sous forme visuelle et sonore. Le son peut servir à annoncer l'apparition ou la modification d'une information visuelle, mais peut aussi aider à sa localisation pour que le joueur repère mieux les indices dans un espace visuel en trois dimensions. Ces éléments sonores améliorent la perception des indices visuels et augmentent la performance du joueur (Nacke, Grimshaw, & Lindley, 2010 ; Wolfson & Case, 2000), mais peu d'études expérimentales se sont intéressées à cet effet ou à l'impact des informations auditives sur l'expérience subjective. Aucune étude ne s'est apparemment intéressée à l'influence d'informations verbales entendues par le joueur, ou de l'intégration d'informations auditives et visuelles sur son comportement global.

Enfin, certains jeux proposent des interactions haptiques aux joueurs, matérialisées notamment par des rétroactions kinesthésiques ou tactiles (Stanney et al., 2003). Certains artefacts de contrôle de jeu prennent en compte le retour de force : le système envoie au joueur des signaux qui simulent les perceptions proprioceptives qui seraient ressenties si son action se déroulait dans le monde réel.

Une manière minimaliste de fournir ces informations au joueur est de les présenter sous forme de vibrations dont la durée et l'intensité varient en fonction de l'intensité du retour. Le peu d'études consacrées aux informations auditives et haptiques est certainement lié à leur rôle mineur dans l'interaction utilisateur-environnement virtuel par rapport aux informations visuelles (Stanney et al., 2003), auxquelles se consacrent en priorité la majorité des études sur les interfaces des jeux vidéo.

Du point de vue de l'interface du système, les jeux vidéo sont assimilables aux environnements virtuels. Cependant, peu d'études ont été consacrées à l'influence des différents éléments constitutifs des interfaces sur le comportement du joueur. Par exemple, des caractéristiques essentielles des interfaces visuelles comme les caractéristiques des arrière-plans ou des informations contextuelles en termes de complexité et de mouvement, n'ont pas encore été abordées de manière scientifique. Les environnements virtuels se caractérisent aussi par la manière dont l'utilisateur aborde et utilise ce type d'interface, et cette question, abordée dans la section suivante, a en revanche été plus étudiée dans le cas des jeux vidéo.

1.2. Le point de vue de l'utilisateur

Si les paramètres de l'interface du système concernent des aspects techniques de l'interaction joueur-jeu vidéo, le point de vue de l'utilisateur sur l'interface est directement lié à l'expérience que le joueur recherche. Les deux types d'éléments liés à l'utilisateur sont « l'engagement » du joueur dans l'interaction, et les sensations perçues par le joueur (Stanney et al., 2003). L'engagement, généralement agréable, car recherché par le joueur, inclut principalement les aspects « d'immersion » et de « présence ». Les sensations perçues peuvent par contre être désagréables, car elles ne sont pas forcément recherchées par le joueur.

1.2.1. Engagement, immersion, présence

L'engagement est en relation avec le niveau de motivation de l'utilisateur dans l'interaction avec l'environnement virtuel, et se compose principalement de l'immersion et de la présence (Stanney et

al., 2003). L'immersion est définie comme « un état psychologique caractérisé par la perception d'être enveloppé par, inclus dans, et en interaction avec un environnement qui fournit un flux continu de stimuli et d'expériences » (Stanney & Salvendy, 1998 ; Witmer & Singer, 1998). La présence, quant à elle, est définie comme « la perception d'une expérience dans un environnement généré par une machine plutôt que dans l'environnement physique réel » (Stanney & Salvendy, 1998). La frontière entre les concepts d'immersion et de présence reste floue dans la plupart des études sur les environnements virtuels, car les termes sont généralement utilisés de manière interchangeable, bien que la distinction entre les deux notions soit réelle (Stanney et al., 2003). L'immersion est directement liée aux caractéristiques du système (e.g., taille de l'écran, volume sonore), alors que la présence est liée à la volonté de l'utilisateur « d'être dans l'environnement ». Cette différence reste toutefois peu explicitée dans la littérature sur les jeux vidéo (E. A. Boyle, Connolly, Hainey, & Boyle, 2012). De manière générale, le terme d'engagement regroupe tous les aspects liés à l'implication du joueur dans le jeu, qu'elle soit volontaire ou non (Brockmyer et al., 2009). Selon les études, l'engagement inclut les expériences subjectives vécues lors de l'utilisation d'un jeu vidéo, comme l'immersion ou la présence (Jennett et al., 2008), mais aussi les aspects qui motivent un individu à jouer à un jeu vidéo et/ou à continuer à jouer (Bostan, 2009).

Brown et Cairns (2004) définissent l'immersion dans un jeu vidéo selon une échelle à 3 niveaux en fonction d'une combinaison de facteurs humains, de facteurs liés au système et à son interface, et de facteurs contextuels. Le premier niveau, le plus faible, est la simple implication dans le jeu. Elle correspond au fait que le joueur accepte d'investir du temps, de l'effort et de l'attention pour apprendre les mécanismes et le contrôle du jeu. Le deuxième niveau est « l'absorption ». Ce niveau est atteint lorsque les émotions du joueur sont directement affectées et que le contrôle du jeu devient inconscient. A ce niveau, le joueur est moins conscient de son environnement réel. « L'immersion totale » est le troisième niveau d'immersion (le plus élevé). Le joueur est inattentif à l'environnement réel et se situe uniquement dans l'environnement virtuel. L'immersion totale implique toutes les ressources attentionnelles. Selon Brown et Cairns, l'immersion totale est une

expérience rare et plutôt passagère dans une session de jeu, tandis que l'implication ou l'absorption sont plus fréquentes. D'autres concepts proches de l'immersion ont été étudiés dans la littérature liée aux jeux vidéo. Par exemple, le « flux » proposé par Csikszentmihalyi (1990) est une expérience d'engagement dans une activité tellement intensive que plus rien d'autre ne semble important pour le joueur. Cette expérience est décrite comme positive et plus extrême que l'immersion. L'absorption cognitive proposée par Agarwal et Karahanna (2000) est un état d'implication profonde dans un logiciel. Selon ces auteurs, l'absorption cognitive est similaire au flux, mais plutôt applicable dans le cadre de l'utilisation de technologies informatiques.

La présence est le concept le plus étudié dans le cadre des environnements virtuels. La présence est un concept particulièrement bien adapté à ces environnements, puisqu'il s'agit d'un état d'illusion perceptive où la barrière entre mondes réel et virtuel tombe (Draper, Kaber, & Usher, 1998 ; K. M. Lee, 2004 ; Stanney & Salvendy, 1998). L'expérience de présence apparait quand l'individu a l'impression « d'être » dans l'environnement virtuel, ce qui la rapproche de l'immersion totale.

Deux types de méthode existent pour mesurer l'immersion et la présence du joueur. Le moyen le plus simple est de questionner le joueur après la session de jeu (Park, Lee, Jin, & Kang, 2010 ; Qin, Rau, & Salvendy, 2009 ; Ravaja et al., 2006 ; Weibel, Wissmath, Habegger, Steiner, & Groner, 2008). Les limites de cette méthode sont d'une part le caractère subjectif du questionnaire, d'autre part la mesure différée de l'immersion. Des méthodes objectives et en temps réel ont été validées comme l'enregistrement des mouvements du regard, qui pourraient refléter le niveau d'immersion du joueur (Jennett et al., 2008).

L'immersion et la présence sont des sensations ressenties pendant l'interaction. Ce sont généralement des sensations recherchées par le joueur. Parallèlement, d'autres sensations plus ou moins agréables peuvent également être ressenties.

1.2.2. Les sensations ou effets ressentis

L'utilisation d'environnements virtuels peut provoquer des effets secondaires plus ou moins agréables, voire indésirables, sur le confort physique de l'utilisateur. Ces effets secondaires sont principalement liés aux éléments visuels des interfaces. Stanney et al. (2003) les ont regroupés en trois catégories : l'inconfort, la cinétose et les « effets de répercussion ». L'inconfort est ressenti lorsque les conditions de l'interaction ne sont pas optimales (trop faible luminosité de l'écran, artefact trop lourd, etc.). En cas de très grand inconfort, des douleurs peuvent survenir. La cinétose est une sensation ressentie en cas de conflit entre le système perceptif visuel et le système vestibulaire de détection des mouvements de la tête (Chang, Pan, Tseng, & Stoffregen, 2012 ; Merhi, Faugloire, Flanagan, & Stoffregen, 2007 ; Stoffregen, Faugloire, Yoshida, Flanagan, & Merhi, 2008). Les effets de répercussion sont observés une fois que l'utilisateur est « retourné » dans le monde réel, comme l'effet de répercussion du mouvement observé à l'écran (Dyson, 2010).

Les effets indésirables dus à l'utilisation des jeux vidéo ont été peu étudiés. Ceci tient probablement au fait que dans des conditions normales, l'environnement de l'utilisateur est moins envahi par le jeu que lors de l'utilisation d'autres environnements virtuels. L'environnement visuel est généralement affiché sur un écran de taille modérée (e.g., téléviseur, moniteur), alors que certains autres environnements virtuels peuvent remplir l'intégralité du champ visuel de l'utilisateur.

Ces données montrent que les jeux vidéo s'apparentent à des environnements virtuels. Leurs caractéristiques sont semblables en tous points à celles de ces environnements (Stanney et al., 2003). Cependant, peu d'études se sont réellement intéressées à la manière dont l'interface des jeux vidéo doit être conçue. Les jeux vidéo, comme tous les environnements virtuels, sont des environnements dynamiques complexes dont les caractéristiques sont principalement déterminées par la complexité et les mouvements dans l'interface visuelle. L'influence de ces caractéristiques sur le comportement des joueurs reste à étudier précisément. Le but de la section suivante est de

montrer que les théories ergonomiques de l'interaction avec des environnements dynamiques complexes peuvent éclairer les interactions avec les jeux vidéo.

2. Les jeux vidéo comme environnements dynamiques complexes

Interagir avec des systèmes dynamiques complexes comme les jeux vidéo implique des processus cognitifs variés (Boot, Kramer, Simons, Fabiani, & Gratton, 2008), parmi lesquels l'attention joue un rôle prédominant (Dye, Green, & Bavelier, 2009 ; Green & Bavelier, 2003, 2006, 2007). Le joueur doit porter son attention sur plusieurs sources d'informations auditives et visuelles tout en contrôlant (manuellement ou vocalement) différents éléments du jeu. La multiplicité des informations encodées et produites crée un environnement multitâche dans lequel les capacités attentionnelles du joueur sont facilement surchargées.

L'interaction avec des environnements dynamiques complexes a surtout été étudiée en environnement réel (e.g., conduite de centrale nucléaire, pilotage d'avion). Après une définition des caractéristiques de ces environnements, cette partie expose les théories particulièrement adaptées pour étudier le comportement d'un opérateur dans ce type de situation, comme la théorie des ressources multiples (Wickens, 1984, 2002, 2008).

2.1. Interaction avec des environnements dynamiques complexes

Un environnement dynamique complexe est dynamique car l'opérateur interagit avec différentes variables qui évoluent en temps réel. Ces évolutions peuvent être autonomes, ou résulter des actions de l'opérateur. La complexité des environnements dynamiques complexes dépend quant à elle de la structure des systèmes impliqués, qui peut intégrer une infinie variété de combinaisons de relations linéaires, non-linéaires ou bruitées entre les entrées et les sorties d'information (Osman, 2010).

L'étude de l'interaction avec des environnements dynamiques complexes s'est principalement faite sur les situations de travail, qui sont des sources potentielles d'accidents à grande échelle en termes de santé et de sécurité (Hoc & Ceillier, 2001). Par exemple, la conduite de centrale nucléaire consiste à exécuter plusieurs tâches simultanément, comme surveiller des indicateurs de la salle de commande concernant le comportement du cœur du réacteur nucléaire ou la répartition des températures, détecter les anomalies et modifier en conséquence les paramètres (Vicente, Roth, & Mumaw, 2001). Progressivement, l'étude de ces situations s'est généralisée à tous les types d'environnements dynamiques complexes qui ne sont pas forcément liés à des situations à risques humains, comme par exemple l'activité d'un opérateur de marché financier. Entre autres, cet opérateur doit gérer simultanément la réception permanente d'informations concernant des fluctuations de marché ou des rapports d'actualités et l'achat-vente d'actions en fonction des profits éventuels qu'il peut réaliser (Brandouy, 2001). Par ailleurs, l'interaction avec des environnements dynamiques complexes a également été étudiée dans la vie quotidienne. La conduite automobile en est l'exemple le plus courant, notamment parce que la technologie embarquée y est de plus en plus présente avec le développement des systèmes de régulation de vitesse (Stanton & Young, 2005) ou d'aide à la navigation (Horrey & Wickens, 2004).

Plusieurs théories encadrent la gestion des environnements dynamiques complexes. Elles sont toutes basées sur la manière dont un opérateur contrôle l'incertitude et gère les risques (Osman, 2010). Osman souligne que deux activités prédominent dans la gestion des environnements dynamiques complexes : le contrôle et la surveillance. Les comportements de contrôle sont liés aux buts de l'opérateur, et impliquent la création et la réalisation d'actions conçues pour générer un évènement particulier dans le futur. La surveillance consiste pour l'opérateur d'une part à comprendre la tâche elle-même, et d'autre part à suivre et évaluer l'efficacité des plans d'action qui ont été élaborés pour atteindre le but désiré.

Le modèle « Compétences-Règles-Connaissances » (modèle SKR) de Rasmussen (1983) postule que le contrôle d'un environnement dynamique complexe par un opérateur s'effectue selon 3 niveaux de traitement. Par défaut, l'opérateur adopte un contrôle basé sur ses habitudes, sous une forme d'automatisation (niveau 1). Quand celles-ci sont insuffisantes pour contrôler la situation, l'opérateur utilise les indices à sa disposition et applique les règles du système (niveau 2). Le troisième niveau de décision, plus performant mais aussi plus coûteux en temps et en ressources, consiste pour l'opérateur à utiliser un plus haut niveau d'abstraction pour prendre une décision, dans le cas où les règles du système s'avèrent insuffisantes. Selon Rasmussen le niveau de traitement se modifie de manière sérielle, c'est-à-dire que pour une même tâche un seul niveau de contrôle est activé à un moment donné. Hoc et Amalberti (1994) ont toutefois montré qu'une activation parallèle des différents niveaux de traitement pour un même processus était également possible, même si le coût en temps et en ressources cognitives était alors plus élevé.

Du point de vue de l'opérateur, les aspects complexe et dynamique de l'activité dans un environnement dynamique complexe ont souvent pour conséquence la nécessité de réaliser simultanément plusieurs tâches. La gestion de ces environnements se caractérise aussi le plus souvent par un partage du temps entre les tâches (Hoc & Ceillier, 2001). Toutefois, les modèles de gestion des environnements dynamiques complexes ne renseignent pas très bien sur l'utilisation des ressources attentionnelles dans l'activité, et plus particulièrement sur le partage de ces ressources entre les tâches. Certains modèles de l'attention permettent toutefois d'expliquer et de prédire au moins en partie le partage de l'attention entre les différentes tâches d'un opérateur en activité.

2.2. Partage attentionnel lors de l'interaction avec des environnements dynamiques complexes

2.2.1. L'attention dans des situations multitâches

L'attention est une fonction cognitive qui permet de percevoir et traiter spécifiquement une ou plusieurs informations présentes parmi une multitude d'autres. Classiquement, un modèle simple d'attention se compose de 2 éléments (Wickens & McCarley, 2008). Un filtre permet de sélectionner les informations présentes dans l'environnement qui sont utiles à la tâche. Puis un « réservoir » unique attribue à chaque information les ressources nécessaires pour son traitement (sa perception, son intégration, et la réponse qu'elle provoque éventuellement). Toutes les tâches que le joueur doit exécuter requièrent des ressources attentionnelles. Lors d'une activité multitâche, l'accumulation des demandes attentionnelles peut générer des conflits dans l'allocation de ces ressources, et par conséquent détériorer la performance du joueur. Les modèles de l'attention prennent généralement en compte un réservoir unique de ressources qui peut être partagé de manière flexible entre les différentes modalités des informations et les différents processus.

Selon les modèles de ressources attentionnelles les plus anciens (e.g., Broadbent, 1958), le traitement de l'information s'effectue de manière sérielle à travers un canal mental unique. Le concept de goulot d'étranglement était alors utilisé pour expliquer les baisses de performance d'un opérateur dans une situation de double-tâche. Le concept de traitement parallèle des informations a été ensuite défini à partir d'un modèle de ressources limitées qui supposait un réservoir de ressources unique, mais faisait aussi intervenir un processeur central qui affectait à chaque demande une partie des ressources disponibles en fonction de sa difficulté (e.g., Kahneman, 1973). Toutefois, ces modèles étaient insuffisants pour rendre compte des résultats qui montraient une variabilité des interférences entre tâches selon leur nature, mais indépendamment de leur difficulté. Ces modèles restaient donc limités pour expliquer et prédire les interférences entre tâches dans des

environnements multitâches. Le modèle des ressources multiples proposé par Wickens (1984, 2002, 2008) offre un meilleur cadre pour expliquer et prédire la manière dont les ressources attentionnelles sont allouées lorsque plusieurs tâches doivent être réalisées simultanément par un opérateur. Wickens a proposé qu'un processeur central puisse partager les ressources attentionnelles (limitées) en fonction à la fois des caractéristiques qualitatives des tâches engagées et de leur difficulté.

2.2.2. La théorie des ressources multiples

Le modèle des ressources multiples de Wickens (1984, 2002, 2008) suppose que quatre dimensions dichotomiques catégorisent les ressources attentionnelles nécessaires à une tâche (voir figure 2 pour une représentation graphique). Selon les tâches, l'attention mobilise un seul des deux types de ressources de chaque dimension. Deux tâches qui font appel aux mêmes types de ressources d'une dimension donnée interférent entre elles plus que deux tâches qui font appel aux deux types différents de ressources de cette dimension. Chacune des quatre dimensions est liée à des substrats neurophysiologiques précis, et a été validée par des études expérimentales. Ces dimensions correspondent par ailleurs à divers choix de conception possibles pour les environnements multimodaux (Wickens & McCarley, 2008).

Dans une première dimension, les demandes attentionnelles sont caractérisées par le niveau de traitement que nécessite la tâche (Isreal, Chesney, Wickens, & Donchin, 1980). Le premier niveau de traitement inclut l'encodage (perception) et le traitement central (cognition) de l'information, tandis que le second est celui de l'action de l'opérateur humain (réponse). Si une tâche requiert la perception d'une information et une autre concurrente un traitement central, par exemple en mémoire de travail, elles vont probablement interférer. De même, une interférence forte est attendue si les ressources attentionnelles doivent être partagées entre la sélection d'une réponse et l'exécution d'une autre. Au contraire, une tâche perceptive interfère de manière plus faible avec une tâche motrice concurrente, et les deux tâches se réalisent ensemble plus facilement.

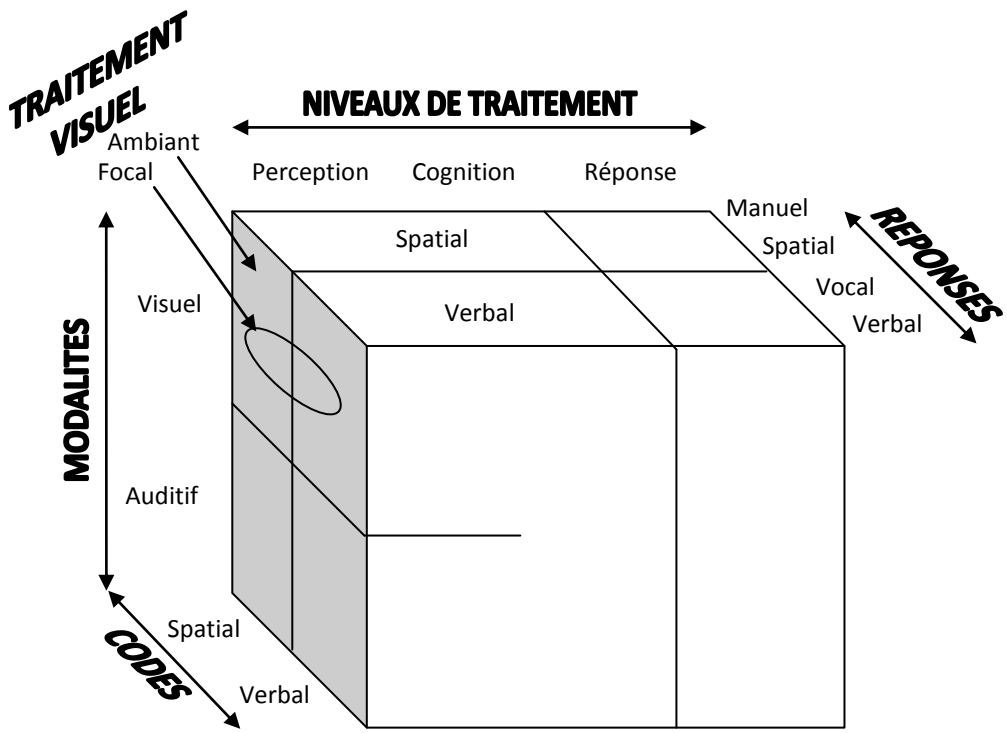


Figure 2. Représentation graphique (adaptée de l'anglais) des quatre dimensions de la théorie des ressources multiples selon Wickens (2002, 2008).

La deuxième dimension utilisée pour catégoriser les demandes attentionnelles des tâches est la modalité perceptive. Les modalités auditive et visuelle impliqueraient des ressources attentionnelles distinctes, mais seulement dans le cadre de la perception. Cette dichotomie repose sur le fait qu'un partage de ressources intra-modal est plus difficile qu'un partage intermodal (Pierno, Caria, Glover, & Castiello, 2005 ; Wickens, Goh, Helleberg, Horrey, & Talleur, 2003). En effet, la proximité spatiale ou temporelle de deux sources d'informations visuelles ou auditives génère des effets de confusion ou des effets de masquage chez l'opérateur (Wickens, Alexander, Ambinder, & Martens, 2004). Par ailleurs, le suivi simultané de deux sources visuelles éloignées augmente le coût de la tâche à cause de la charge supplémentaire liée au balayage visuel.

La troisième dimension des demandes attentionnelles se réfère à la distinction entre les codes de traitement spatial (analogique) et verbal (symbolique). Au niveau moteur, la distinction s'effectue entre les réponses manuelles qui sont généralement de nature plus spatiale, et les réponses vocales qui utilisent souvent des ressources verbales. Par exemple, Wickens et Liu (1988) ont montré par une

méthode de double tâche que la performance à une tâche spatiale continue (poursuite manuelle d'une cible) est détériorée par une tâche secondaire de décision spatiale, mais pas par une tâche secondaire verbale.

La quatrième et dernière dimension a été introduite plus récemment (Wickens, 2002). Dans la modalité visuelle, il semble que les traitements visuels focalisés et non focalisés (i.e., globaux, ou périphériques) utilisent des ressources différentes (Weinstein & Wickens, 1992). Les ressources du « canal focal », qui implique surtout la vision fovéale, sont allouées aux tâches attentionnelles les plus fines telles que la reconnaissance d'objets. Les ressources du « canal ambiant », qui implique surtout la vision périphérique, sont utilisées pour des tâches plus générales telles que la perception du déplacement de soi et de l'orientation, ou pour détecter des changements dans le champ visuel périphérique.

Les ressources attentionnelles nécessaires pour une tâche peuvent être caractérisées selon chacune de ces quatre dimensions. Dans des situations multitâches, plus les tâches utilisent des ressources de niveaux communs dans une ou plusieurs dimensions, plus le partage de ces ressources provoque des interférences. A l'inverse, il y a moins d'interférences entre les demandes attentionnelles qui exploitent des ressources de niveaux distincts dans chaque dimension. Par conséquent, le modèle peut prédire l'intensité des interférences entre des tâches réalisées simultanément lors d'une activité impliquant des environnements dynamiques complexes.

Deux exemples d'application du modèle de Wickens ont été présentés par Wickens et al. (2003) et Horrey, Wickens, et Consalus (2006). Dans la première étude (Wickens et al., 2003), un appareil d'aide à la navigation d'un avion présentait des informations de manière verbale auditive et/ou visuelle à un pilote pendant qu'il dirigeait son avion. Les détections de cibles et le suivi du trajet étaient meilleurs lorsque la modalité auditive seule était utilisée par rapport à quand la modalité visuelle seule, ou combinée avec la modalité auditive, était utilisée. Un problème de mémorisation survenait tout de même si le message présenté était trop long. Dans la deuxième étude (Horrey et

al., 2006), l'interférence visuelle entre l'activité de conduite automobile et le suivi d'informations visuelles sur un écran embarqué a également été démontrée. Plus l'information à traiter sur l'écran embarqué était importante, plus le suivi du trajet et la détection de cibles étaient défaillants. Les auteurs ont néanmoins supposé que la vision périphérique pouvait être une aide non-négligeable à la détection de cibles lorsque l'attention était posée sur l'écran embarqué.

Dans une activité multitâche, l'attention de l'opérateur doit être partagée. Cependant le partage de l'attention est aussi nécessaire lors de l'utilisation de plusieurs sources d'information pour une même tâche.

2.2.3. Intégration de plusieurs sources d'information

Jouer à un jeu vidéo requiert la réalisation simultanée de plusieurs tâches (Ang, Zaphiris, & Mahmood, 2007), pour lesquelles le joueur doit partager son attention entre les différentes sources d'information auditives ou visuelles disponibles, acquérir les informations nécessaires et les intégrer. Plusieurs études qui se sont intéressées à la conception optimale des interfaces multimodales complexes suggèrent qu'il faudrait minimiser le coût d'intégration des informations (Wickens & McCarley, 2008). Les interfaces devraient faciliter la combinaison mentale des informations issues de sources séparées, en particulier quand l'attention doit être partagée entre ces sources.

Principe de proximité-compatibilité. Le principe de proximité-compatibilité peut être utilisé pour expliquer et optimiser l'intégration d'informations. Wickens et Carswell (1995) ont démontré que deux sources d'information qui requièrent de l'attention partagée devaient être placées à proximité l'une de l'autre dans l'interface pour être intégrées en vue d'une tâche ou d'une opération mentale commune. A l'inverse, les informations utilisées en attention focalisée (de manière isolée) doivent être éloignées les unes des autres. En effet, maximiser la proximité spatiale et/ou temporelle de deux informations réduit l'amplitude des mouvements attentionnels nécessaires dans les situations d'attention partagée. Renshaw, Finlay, Tyfa, et Ward (2004) ont étudié une tâche de recherche

d'information dans un graphique qui requérait l'intégration de sources multiples d'informations. Ils ont montré que les participants répondaient plus rapidement quand les sources d'informations, telles que les noms des courbes et les données numériques, étaient disposées à l'intérieur du graphique plutôt qu'à l'extérieur comme légendes.

Cependant, une forte proximité spatiale entre les signaux visuels ne permet pas toujours une meilleure intégration des informations. Premièrement, la proximité peut conduire à la confusion et au masquage, par exemple quand un affichage tête-haute est superposé à la scène principale. Ceci peut avoir un effet négatif ou réduire le bénéfice pour la performance de l'opérateur (Wickens et al., 2004). Deuxièmement, le continuum du principe de proximité-compatibilité indique que moins des sources d'information sont utilisées simultanément dans une même tâche, plus leur proximité spatiale peut conduire à des interférences (Wickens & Carswell, 1995). Elles doivent donc être tenues plus éloignées les unes des autres. Dans les cas les plus extrêmes, le continuum s'étend à des tâches où une intégration complète des différentes informations est nécessaire, et au contraire à d'autres où les sources d'information ne sont jamais utilisées simultanément. Entre ces deux extrêmes, il existe des cas où seules quelques caractéristiques des deux informations doivent être combinées, et d'autres où les deux informations doivent être utilisées simultanément, mais sans être ni combinées, ni intégrées.

Lors de l'intégration d'informations, une des sources peut fournir l'information de manière dynamique. Dans ce cas, des mécanismes d'anticipation peuvent également être mis en jeu.

Anticipation visuelle des informations dans les jeux vidéo. Dans les jeux vidéo, anticiper l'apparition ou la modification des informations apparait essentiel. Les objets visuels impliqués dans l'action et visibles à l'arrière-plan de la scène sont généralement en mouvement. De plus, les informations données par les affichages tête-haute sont souvent modifiées en temps réel. La performance du joueur est optimale lorsqu'il anticipe le plus possible les mouvements et qu'il réussit à détecter les modifications d'information le plus rapidement possible.

Quand l'arrière-plan de la scène se déplace dans une direction, par exemple vers le haut, pour donner l'illusion d'un déplacement de soi, le joueur doit continuellement porter son attention sur les nouvelles informations qui apparaissent en bas de l'écran. Il doit détecter aussi vite que possible des obstacles potentiels, des alliés ou des ennemis. Dans le même temps, il doit continuer à exécuter la tâche principale et regarder (ou écouter) les informations données par le système.

Bien que l'anticipation visuelle des joueurs de jeu vidéo n'ait pas été empiriquement étudiée, des situations similaires sont souvent rencontrées par des opérateurs humains dans d'autres environnements virtuels ou réels. Les meilleurs exemples sont actuellement le pilotage d'avions ou la conduite automobile. Conducteurs et pilotes doivent porter leur attention à de multiples sources d'information tout en anticipant ce qui peut arriver en avant de leur engin ou véhicule, vers où ils se dirigent. Les enregistrements des mouvements oculaires ont été utilisés pour observer les mouvements du regard et évaluer le degré d'anticipation des opérateurs (Owsley & McGwin Jr, 2010 ; Underwood, 2007 ; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003 ; Wickens et al., 2003). Underwood et al. ont étudié les mouvements oculaires de conducteurs novices et experts. Dans des situations de conduite naturelle, les conducteurs experts accordaient plus d'attention aux régions les plus lointaines de la route qu'aux autres zones de l'environnement visuel. Cette stratégie reflète le fait qu'ils passent la majorité de leur temps à anticiper de possibles accidents. Par contre, les conducteurs novices utilisaient moins souvent ce type de stratégie. Ces résultats suggèrent qu'une plus grande part des ressources attentionnelles est dévolue à l'anticipation quand la conduite devient plus automatique.

3. Synthèse

Les jeux vidéo sont assimilables à des environnements virtuels. Leurs caractéristiques en termes d'interface et de point de vue de l'utilisateur sont très proches. L'interaction avec les jeux vidéo a toutefois été peu étudiée avec cette perspective. Une manière d'étudier les interactions avec les jeux

vidéo de manière spécifique par rapport aux interactions avec les systèmes informatiques classiques est donc de faire appel aux théories des interactions avec des environnements dynamiques complexes. La gestion de ce genre d'environnements, où les erreurs peuvent avoir des conséquences graves, a été très étudiée. Plusieurs cadres théoriques sont utilisés pour comprendre et optimiser ces situations.

Parmi celles-ci, la théorie des ressources multiples permet d'expliquer et de prédire le partage des ressources attentionnelles dans des activités multitâches. Cependant, cette théorie ne peut pas prédire le partage de ressources lorsque deux tâches requièrent les mêmes ressources (Wickens, 2008). Concrètement, cette théorie peut souvent très bien expliquer la manière dont un joueur peut partager son attention entre plusieurs tâches. Elle peut également expliquer la manière dont les joueurs intègrent différentes informations visuelles et sonores, ou visuelles fovéales et périphériques pour suivre, par exemple, l'action principale du jeu au centre de l'écran et l'évolution des informations contextuelles en périphérie. Par contre, cette théorie ne peut pas expliquer la manière dont plusieurs informations vont être traitées lorsqu'elles sont en concurrence dans le même canal perceptif visuel, spatial focal ou ambiant. Par exemple, l'influence des caractéristiques visuelles des arrière-plans de l'écran de jeu sur le traitement attentionnel des éléments visuels de l'action principale superposés ne peut pas être précisément décrite à l'aide de cette théorie. D'autres connaissances liées à la perception visuelle des environnements visuels dynamiques complexes, exposées dans le chapitre suivant, sont nécessaires.

Chapitre 2 :
Perception des environnements visuels dynamiques complexes

Que l'environnement soit interactif ou non, la présence de plus d'une source d'information dans l'interface visuelle entraine un partage de l'attention de l'individu. En effet, les ressources attentionnelles disponibles pour le traitement de ces différentes informations sont limitées. Dans le cadre de la théorie des ressources multiples (Wickens, 1984, 2002, 2008), des informations issues de modalités sensorielles différentes (e.g., visuelle ou auditive) ou faisant appel à des modes de codage différents (e.g., verbal ou spatial), peuvent être traitées simultanément. En revanche, lorsque les informations font appel à la même modalité sensorielle ou partagent le même code, l'utilisation des mêmes ressources attentionnelles se fait aux dépens des informations les moins importantes.

Les jeux vidéo, comme tous les environnements dynamiques complexes, sont construits de telle manière que leurs interfaces visuelles requièrent une grande part de l'attention du joueur. L'objectif de ce chapitre est de montrer que des connaissances spécifiques sur la perception visuelle humaine sont nécessaires pour comprendre et optimiser l'interaction entre un joueur et ce type d'interface. La première section de ce chapitre donne les bases de la perception visuelle de scènes complexes. Le guidage de l'attention visuelle obéit à des règles bien étudiées dans le cadre de l'exploration libre ou

contrainte de scènes. La seconde section de ce chapitre montre que ces règles sont modulées dans le cadre de l'exploration de scènes visuelles dynamiques et/ou complexes. La présence de ce type de scène visuelle dans une activité peut influencer la performance de l'individu.

1. La perception visuelle de scènes complexes

L'étude des interfaces visuelles des jeux vidéo suppose la prise en compte des connaissances sur la perception des scènes visuelles statiques. Le guidage de l'attention visuelle lors de l'exploration d'une scène est aujourd'hui assez bien connu, et ses déplacements sont généralement reflétés par les mouvements du regard (Carrasco, 2011). Ainsi, les paradigmes de recherche visuelle sont intéressants pour aborder le guidage de l'attention dans la majorité des tâches à réaliser dans un jeu vidéo.

1.1. Le guidage de l'attention visuelle

Lors de l'exploration de scènes visuelles statiques, l'attention de l'observateur ne se déplace pas de manière aléatoire. Deux types de guidage attentionnel coexistent (Henderson, 2007). Le guidage « ascendant », plutôt réactif, est régi par les propriétés visuelles (dites « caractéristiques de bas niveau ») de la scène perçue. Par exemple, les endroits visuellement saillants (par leur luminosité, leur couleur, etc.) ou les frontières entre les différents éléments présents dans la scène attirent l'attention (Itti & Koch, 2000 ; Tatler, Hayhoe, Land, & Ballard, 2011). Le guidage « descendant » dépend quant à lui des buts et des connaissances de chaque individu, qui impliquent notamment la mémoire à court-terme et la mémoire épisodique (Folk, Remington, & Johnston, 1992). Généralement, ces deux modes de guidage sont activés simultanément et interagissent pendant la perception d'une scène (Torralba, Oliva, Castelhano, & Henderson, 2006).

Une des théories les plus utilisées pour expliquer le fonctionnement de l'attention visuelle dans ces situations est la « théorie d'intégration des traits » (Treisman & Gelade, 1980). Cette théorie permet

de modéliser le guidage de l'attention lors de la perception d'une scène visuelle. Elle suppose que l'attention de l'observateur est guidée en deux étapes. Dans la première étape dite « pré-attentive », toutes les informations de la scène sont traitées simultanément. Les caractéristiques de bas-niveau sont utilisées pour décomposer tous les objets de la scène en leurs traits visuels distinctifs. Cette première étape permet de préparer le déploiement de l'attention à l'étape suivante. Dans la seconde étape, plus sélective, l'attention est portée de manière séquentielle aux différents objets de la scène. L'attention est guidée d'une part de manière ascendante à partir des éléments extraits à l'étape précédente, et d'autre part de manière descendante en fonction des buts et connaissances de l'observateur. Ce modèle a été repris et adapté par Wolfe et ses collaborateurs dans le modèle de « recherche guidée » (Wolfe, 1994, 2003 ; Wolfe, Cave, & Franzel, 1989). Contrairement au modèle de Treisman et Gelade (1980), ce modèle suggère que le traitement des objets est à la fois séquentiel et parallèle dans la phase d'attention sélective. Les objets s'engagent un par un dans le traitement attentionnel, mais plusieurs objets peuvent être traités simultanément une fois leur traitement entamé. Les deux types de guidage (ascendant et descendant) sont détaillés dans les paragraphes qui suivent.

1.1.1. Guidage ascendant

Les propriétés visuelles de bas niveau qui guident l'attention incluent principalement la couleur, le contraste, la luminance, mais également le mouvement dans les scènes visuelles dynamiques (Itti, 2005 ; Le Meur, Le Callet, & Barba, 2007). Les travaux ont permis d'élaborer des modèles computationnels de prédiction des déplacements de l'attention. Le Meur et al. ont proposé un modèle spatio-temporel de déploiement de l'attention visuelle basé sur les caractéristiques visuelles de bas niveau de scènes visuelles dynamiques. Ce modèle est fondé sur une décomposition de l'image en plusieurs cartes de saillance spatiales. Chacune correspond à une propriété des objets telle que la couleur, la luminance, le contraste ou les fréquences spatiales. Une carte supplémentaire, celle du mouvement, est basée sur la décomposition temporelle de la séquence

d'images qui forment une animation décrite, pour chaque image, par un vecteur de mouvement général et des vecteurs de mouvement local. Ce modèle prédit que lors de l'exploration libre d'une scène, le guidage attentionnel est fortement influencé par les caractéristiques de bas niveau de la scène. La confrontation avec les données expérimentales (visionnage libre de clips vidéo) a montré que le modèle était fortement prédictif. Dans le cas d'une scène statique, le guidage ascendant est maximal juste après la présentation de la scène, mais diminue ensuite au cours du temps. Par contre, dans le cas d'une scène en mouvement, le guidage reste majoritairement ascendant tout au long de l'observation. Dans le cadre de ce modèle en effet, le guidage de l'attention visuelle est avant tout prédit par les mouvements des différents objets de la scène.

1.1.2. Guidage descendant

L'activité de jeu implique aussi une exploration de l'environnement visuel guidée par des intentions, liées principalement au but du jeu (Bernhard, Stavrakis, & Wimmer, 2010). Les joueurs ne se trouvent que rarement en situation de simple exploration de la scène. Toutefois, certains jeux (e.g., jeux de course, jeux musicaux) contraignent l'activité du joueur dans le temps. Dans ce cas le joueur n'a pas le choix du rythme de son activité. Lorsqu'un rythme rapide est imposé par le jeu, le guidage de l'attention visuelle peut être géré uniquement de manière réactive.

Plusieurs auteurs ont observé et modélisé les déplacements de l'attention dans les scènes visuelles des jeux vidéo en fonction des propriétés de la scène, mais également des intentions du joueur (Bernhard et al., 2010 ; Peters & Itti, 2008). Peters et Itti ont montré que la prise en compte des caractéristiques visuelles de la scène dans un jeu vidéo aidait à prédire l'orientation de l'attention, et donc du regard. Cependant, les résultats dépendaient du type de jeu utilisé. Dans un jeu vidéo qui nécessite de l'exploration visuelle (e.g., un jeu d'aventure), où l'attention du joueur est libre, les propriétés de la scène visuelle sont prédominantes. Dans un jeu vidéo où l'attention est contrainte par le temps (e.g., un jeu de course), la prédiction par les caractéristiques visuelles de la scène est plus difficile. De plus, Bernhard et al. (2010) ont montré que dans un jeu de tir en vision subjective, la

prédiction du déplacement de l'attention était biaisée à cause de la présence fréquente d'un réticule de visée au centre de l'écran. Cet élément, pas nécessairement très saillant, est utilisé pour réaliser les tâches qui imposent de viser et de tirer sur des cibles.

Ces différentes études ont montré qu'il était difficile de généraliser le comportement de l'attention visuelle à tous les types de jeu vidéo. Comme l'activité est accomplie dans un but différent selon le jeu, l'attention est orientée selon des contraintes différentes, et les propriétés de la scène visuelle du jeu sont généralement insuffisantes pour prédire les mouvements du regard du joueur. Il est donc important de prendre en compte les règles du jeu et les instructions données au joueur pour concevoir les différents éléments d'aide ou de challenge des interfaces visuelles d'un jeu vidéo.

1.2. Les mouvements oculaires en tant qu'indicateurs de l'attention visuelle

L'étude de la perception et des déplacements attentionnels peut être menée relativement facilement grâce aux techniques d'enregistrement des mouvements oculaires, déjà éprouvées dans plusieurs domaines. Une des raisons du succès de ces techniques est que les mouvements oculaires semblent effectivement refléter les traitements attentionnels en cours (Hayhoe & Ballard, 2005 ; Kowler, 2011 ; Liversedge & Findlay, 2000 ; Rayner, 1998, 2009 ; Schütz, Braun, & Gegenfurtner, 2011 ; Yarbus, 1967). Plus spécifiquement, les mouvements oculaires reflètent des déplacements « explicites » et visibles de l'attention visuelle (Zelinsky, 2008). L'attention visuelle peut en effet se déplacer soit de manière masquée (« covert attention »), soit de manière explicite (« overt attention ») (Posner, 1980). Un déplacement explicite de l'attention implique un mouvement physique de l'œil, de la tête ou du corps pour aligner le plus possible les capteurs sensoriels de l'individu avec l'objet d'intérêt. A l'inverse, le déplacement masqué de l'attention est un changement du point d'attention sans mouvement physique. Ce dernier type de déplacement ne peut évidemment pas être détecté par les mouvements du regard.

Une scène visuelle, quelle que soit sa nature, est balayée par le regard, qui passe de point de fixation en point de fixation par des mouvements successifs appelés saccades oculaires. L'alternance des fixations et des saccades permet de traiter l'ensemble des informations de la scène pendant les fixations (Henderson, 2003). Seule la région centrale de faible diamètre de la rétine, la fovéa, permet de percevoir les informations avec l'acuité maximale. Autour de la fovéa, dans les régions parafovéale puis périphérique, l'information n'est pas traitée précisément, mais suffisamment pour permettre de programmer la prochaine saccade (pendant laquelle l'information n'est pas ou quasiment pas traitée) qui déplacera le regard vers le point de fixation suivant.

Plusieurs études ont utilisé le suivi des mouvements oculaires pour comprendre précisément les mécanismes de la perception des scènes visuelles dynamiques des jeux vidéo. Certains auteurs ont utilisé les mouvements oculaires comme outil d'évaluation de la sensation d'immersion du joueur (Jennett et al., 2008), ou de la qualité globale du jeu vidéo (Léger, Fouquereau, & Tijus, 2009). D'autres auteurs ont observé et modélisé les mouvements du regard dans les interfaces visuelles des jeux vidéo en fonction des propriétés de la scène et des intentions du joueur (Bernhard et al., 2010 ; Jie & Clark, 2008 ; Peters & Itti, 2008).

Le suivi des mouvements oculaires permet d'étudier le guidage de l'attention visuelle lors de la perception de scènes. En psychologie cognitive, une autre méthode d'étude des mécanismes attentionnels est le paradigme de recherche visuelle, exposé dans la section suivante.

1.3. La recherche visuelle

Un des paradigmes les plus utilisés dans la littérature pour étudier l'attention visuelle est la recherche visuelle d'un objet connu à l'avance parmi d'autres objets (voir Eckstein, 2011 pour une revue). En effet, ce paradigme reflète une tâche courante dans la vie quotidienne, mais également dans l'utilisation d'applications informatiques (Fleetwood & Byrne, 2006 ; Léger, Tijus, & Baccino, 2005 ; Ojanpää, Näsänen, & Kojo, 2002 ; Pearson & van Schaik, 2003) et d'environnements virtuels

de type jeu vidéo (Castel, Pratt, & Drummond, 2005 ; Hubert-Wallander, Green, Sugarman, & Bavelier, 2011). Les principes de la recherche visuelle peuvent être appliqués aux environnements multimodaux tels que les jeux vidéo.

1.3.1. Principes

Dans les paradigmes classiques de recherche visuelle, les participants doivent dire le plus rapidement possible si une cible particulière est présente ou non parmi un ensemble d'objets distracteurs plus ou moins similaires à la cible (Thornton & Gilden, 2007 ; Wolfe, 1998). La performance des participants est habituellement évaluée par le taux d'erreur ou le temps de réaction, qui diminuent au fur et à mesure que la performance augmente. Une des caractéristiques particulières de la recherche visuelle est de faire rechercher la cible dans des ensembles de taille variable. De nombreux travaux ont montré que la cible pouvait être identifiée de deux manières différentes en fonction des caractéristiques visuelles de base des différents distracteurs, comme la couleur, la taille, la luminance ou le mouvement. Selon la théorie de Wolfe (1994), la cible peut être détectée très rapidement, sans attention volontaire, si elle diffère de tous les autres objets présents par au moins une de ses caractéristiques visuelles de base (e.g., une barre rouge parmi des barres vertes). Avec cet effet de « pop-out », la cible est repérée instantanément et sans effort, quel que soit le nombre de distracteurs présents. Si la cible n'est pas immédiatement détectable, notamment parce que toutes ses caractéristiques visuelles sont partagées avec au moins une partie des distracteurs (e.g., une barre rouge horizontale parmi des barres rouges verticales et des barres vertes horizontales), un deuxième processus, sériel, est engagé (Gilden, Thornton, & Marusich, 2010). Dans ce cas, l'attention visuelle du participant se déplace d'objet en objet ou de groupe d'objets en groupe d'objets de manière séquentielle jusqu'à la découverte de la cible. Ainsi, plus le nombre d'objets présents dans la scène est grand, plus le temps de recherche s'allonge.

Plusieurs auteurs ont étudié le comportement du regard pendant des tâches de recherche visuelle (voir Rayner, 2009 pour une revue). Le suivi des mouvements oculaires est particulièrement

pertinent lorsqu'un processus de recherche sérielle est engagé, puisque le déplacement du regard est principalement calqué sur le déplacement de l'attention. Par exemple, Najemnik et Geisler (2005) ont modélisé la stratégie optimale d'exploration oculaire en recherche visuelle. Le modèle prend en compte les deux phases importantes qui alternent pendant la recherche sérielle, la phase d'intégration des informations (fixation) et celle de sélection des points de fixation suivants (programmation des saccades). Ce modèle intègre également un principe d'inhibition de retour qui empêche le regard de balayer plusieurs fois les mêmes zones de la scène. Les auteurs ont montré que ce modèle est représentatif de l'observateur humain, sauf pour l'inhibition de retour qui n'est pas appliquée de manière vraiment performante par l'humain.

1.3.2. Recherche visuelle et multimodalité

La recherche visuelle n'est pas toujours purement visuelle. Certains auteurs ont montré qu'une intégration des informations visuelles avec des informations auditives, verbales ou non, modifie le comportement de l'individu.

Plusieurs études ont abordé la recherche visuelle d'objets accompagnée d'instructions sonores verbales (Dahan & Tanenhaus, 2005 ; Gibson, Eberhard, & Bryant, 2005 ; Koelewijn, Bronkhorst, & Theeuwes, 2009 ; Mosimann, Felblinger, Colloby, & Müri, 2004 ; Olds & Fockler, 2004 ; Reali, Spivey, Tyler, & Terranova, 2006 ; Spivey, Tyler, Eberhard, & Tanenhaus, 2001). Leurs auteurs cherchaient notamment à comprendre comment le langage pouvait gêner ou améliorer la perception visuelle d'une scène. Par exemple, Spivey et al. ont présenté verbalement les caractéristiques de la cible dans le cadre d'une recherche visuelle. Les participants devaient trouver une cible (une barre) en entendant une voix indiquer sa couleur (rouge ou verte) et son orientation (verticale ou horizontale). Les auteurs ont manipulé le moment de présentation de l'indication et le nombre d'objets présents dans la scène. L'indication pouvait être donnée avant que le champ de recherche n'apparaisse, ou au même moment. Les résultats ont montré que les temps de réaction étaient plus longs lorsque l'indication était présentée en même temps plutôt qu'avant les objets, mais étaient en revanche

moins dégradés par l'augmentation du nombre de distracteurs. Ces données suggèrent que les systèmes perceptifs spécialisés dans le langage et la vision interagissent efficacement. De manière concrète, quand l'indication était présentée en même temps que le champ de recherche, dès que la première caractéristique (couleur ou orientation) était énoncée, les objets correspondants devenaient plus saillants pour l'observateur. Une partie des distracteurs pouvait alors être directement ignorée. La recherche était ensuite facilitée dès que la deuxième caractéristique était donnée. Ces résultats ont été répliqués avec des recherches impliquant une triple conjonction de caractéristiques (couleur, orientation et taille) (Reali et al., 2006). Gibson et al. (2005) ont tempéré ces résultats en montrant que lorsque l'indication verbale était énoncée à une vitesse d'élocution plus proche de celle utilisée dans la vie courante, l'atténuation de la dégradation en fonction de la taille du champ de recherche était plus faible.

D'autres auteurs se sont intéressés à l'interaction entre informations visuelles et auditives dans la recherche visuelle, mais en utilisant des sons plutôt que des indications verbales (Iordanescu, Grabowecky, Franconeri, Theeuwes, & Suzuki, 2010 ; Iordanescu, Guzman-Martinez, Grabowecky, & Suzuki, 2008 ; Pierno et al., 2005). Par exemple, Iordanescu et al. (2008) ont montré que lorsqu'un son habituellement associé à la cible à trouver est présenté en même temps que les objets, la cible est détectée plus rapidement que lorsque le son n'est pas congruent. D'après ces auteurs, ces résultats suggèrent que le traitement sensoriel des informations est fondamentalement multimodal.

La perception de scènes visuelles statiques a été très étudiée. Toutefois, les études classiques du domaine n'insistent pas sur le caractère dynamique et complexe de certaines scènes visuelles. Ce sont donc les influences de deux caractéristiques spécifiques des scènes visuelles des jeux vidéo, la complexité et le mouvement des arrière-plans, qui sont plus précisément mises en avant dans la section suivante. L'accent est mis sur la compréhension de l'influence de ces caractéristiques sur l'interaction avec les systèmes, notamment en termes de performance des individus.

2. Perception et interaction personne-système dans des scènes visuelles complexes et/ou dynamiques

Dans les environnements virtuels tels que les jeux vidéo, le mouvement et la complexité sont deux caractéristiques importantes des arrière-plans. Plusieurs études ont montré que le mouvement et la complexité de l'arrière-plan pouvaient gêner le traitement des informations superposées.

2.1. Scènes visuelles dynamiques

Généralement, le caractère dynamique des scènes visuelles se traduit par des mouvements des objets qui constituent la scène (Itti, 2005 ; Le Meur et al., 2007). Dans les jeux vidéo, ce sont non seulement les objets avec lesquels le joueur doit interagir qui sont en mouvement, mais aussi souvent les arrière-plans.

2.1.1. Perception d'objets en mouvement

Les jeux vidéo qui présentent des scènes visuelles en mouvement nécessitent une anticipation visuelle du mouvement d'éléments qu'un avatar doit éviter (e.g., les jeux vidéo de course) ou toucher (e.g., les jeux vidéo de tir) pour réussir une tâche. En présence de plusieurs éléments en mouvement, le joueur doit être capable de porter son attention sur l'ensemble de ces éléments en parallèle. Il doit donc être capable d'adopter une stratégie visuelle optimale en fonction du but du jeu. Plusieurs études sur l'anticipation visuelle dans des situations dynamiques existent, notamment dans les domaines de la conduite automobile (Underwood et al., 2003) et du pilotage d'avions (Wickens et al., 2003). Les résultats de ces études sont détaillés dans le précédent chapitre.

Le suivi d'objets multiples tous en mouvement a quant à lui été peu étudié dans des conditions naturelles (Cavanagh & Alvarez, 2005 ; Jans, Peters, & De Weerd, 2010). Des études en laboratoire ont toutefois permis d'observer des stratégies différentes selon la tâche demandée (Fehd & Seiffert, 2010 ; McIntire, Havig, Watamaniuk, & Gilkey, 2010 ; Rothkopf, Ballard, & Hayhoe, 2007). Le suivi

d'objets multiples en mouvement se caractérise par des séquences alternées de saccades et de poursuites oculaires des objets (Fehd & Seiffert, 2010). Rothkopf et al. (2007) ont montré que dans un même environnement dynamique, les individus adoptaient des stratégies de suivi visuel différentes selon la tâche. Quand la tâche était de toucher des objets en mouvement, les fixations oculaires étaient localisées majoritairement sur les centres des objets, tandis que quand la tâche était d'éviter ces mêmes objets, les fixations oculaires se portaient majoritairement sur leurs extrémités. La tâche « éviter les objets » induisait un comportement du regard caractérisé plutôt par des séquences de saccades et de fixations oculaires de balayage de l'ensemble de l'environnement, tandis que la tâche « toucher les objets » induisait un comportement de suivi successif des différents objets (Voir Caroux, Le Bigot, & Vibert, 2011 pour un exemple d'application, annexe 1).

Le comportement observé lors du suivi d'objets en mouvement est un comportement actif. Dans les jeux vidéo, ce suivi peut être lié à la tâche principale. Cependant, le mouvement sur l'écran d'un jeu vidéo concerne aussi les arrière-plans des scènes. Le comportement du joueur face à ce type d'élément visuel, qui n'apporte pas toujours d'information utile à la tâche principale, peut distraire l'attention portée aux objets présentés en superposition à cet arrière-plan dans la scène visuelle.

2.1.2. Les effets d'arrière-plans en mouvement

La vection. La plupart des environnements virtuels affichent des scènes visuelles complexes et dynamiques qui incluent un arrière-plan visuel structuré en mouvement. Les conséquences de ce genre d'arrière-plans sur le comportement de l'observateur sont bien documentées. Un des effets les mieux connus des scènes visuelles en mouvement sur des observateurs immobiles est la vection (Riecke, Schulte-Pelkum, Avraamides, Von Der Heyde, & Bülthoff, 2006 ; Trutoiu, Mohler, Schulte-Pelkum, & Bülthoff, 2009). La vection est une perception illusoire d'un déplacement de soi qui est souvent ressentie par un observateur immobile exposé à un mouvement global de son environnement visuel (Lepecq et al., 2006). La vection est induite quand le déplacement visuel correspond à ce que l'observateur verrait pendant son propre déplacement, c'est-à-dire quand le flux

optique sur la rétine reproduit le flux optique induit par un déplacement propre. Cependant, le mouvement illusoire n'est jamais permanent. Les phases de déplacement illusoire alternent toujours avec des phases pendant lesquelles l'observateur réalise que c'est l'arrière-plan qui bouge, et pas lui (Giannopulu & Lepecq, 1998 ; Lepecq et al., 2006 ; Trutoiu et al., 2009). La durée et la fréquence de la vection varient en fonction des individus, mais les flux optiques qui induisent le mieux la vection sont ceux qui reproduisent des vitesses de déplacement propre vraisemblables.

Différents flux optiques induisent différents types de vection. Les mouvements d'ensemble de l'environnement visuel de la gauche vers la droite ou de la droite vers la gauche créent des sensations de mouvement propre dans la direction opposée (vection latérale). Les rotations de l'environnement visuel induisent des sensations de rotation (vection circulaire). Enfin, des flux optiques en expansion radiale, comme lors d'un déplacement dans un couloir ou un tube, créent des sensations de vection vers l'avant (Bubka & Bonato, 2010 ; Lepecq et al., 2006 ; Trutoiu et al., 2009). Dans les environnements virtuels, et plus spécifiquement dans les jeux vidéo, deux types de mouvements de l'arrière-plan sont souvent utilisés : des flux optiques radiaux simulent des mouvements vers l'avant ou vers l'arrière, tandis que des flux optiques latéraux simulent des mouvements latéraux (le plus souvent de la gauche vers la droite).

Le réflexe optocinétique. Un des phénomènes principaux associés à la perception de son déplacement, que ce soit dans des environnements réels ou virtuels, est le nystagmus optocinétique (OKN) déclenché par les mouvements d'ensemble de l'environnement visuel (Kim & Palmisano, 2010 ; Riecke et al., 2006). Contrairement à la vection, l'OKN est un réflexe qui persiste aussi longtemps que l'arrière-plan visuel est en mouvement. L'OKN est un mouvement conjugué des deux yeux dans lequel deux phases alternent : la phase lente emmène les yeux dans la direction du mouvement de l'arrière-plan, idéalement à la même vitesse, tandis que la phase rapide ramène régulièrement les yeux dans la direction opposée (Ilg, 1997 ; Waespe & Schwarz, 1987). La phase lente de l'OKN est un mouvement compensatoire de l'œil qui permet à l'observateur de garder

automatiquement l'information visuelle stable sur la rétine. Enfin, pendant l'OKN, l'orientation moyenne du regard est généralement déviée vers l'origine du mouvement de la scène.

Les humains ne peuvent supprimer volontairement leur OKN que si un élément immobile est superposé sur l'arrière-plan en mouvement (Ilg, 1997 ; Pola, Wyatt, & Lustgarten, 1992 ; Pola, Wyatt, & Lustgarten, 1995 ; Rubinstein & Abel, 2011 ; Waespe & Schwarz, 1987 ; Williams, Mulhall, Mattingley, Lueck, & Abel, 2006 ; Wyatt & Pola, 1984 ; Wyatt, Pola, & Lustgarten, 1988). Dans ce cas, l'individu peut fixer cet élément et supprimer l'OKN de manière complète et rapide dans un délai de 100 à 200 ms (Pola et al., 1992 ; Wyatt & Pola, 1984). Le niveau de suppression de l'OKN (en termes d'amplitude ou de fréquence du déplacement oscillatoire des yeux) peut être variable, et dépend notamment des caractéristiques des éléments que peut fixer l'individu. Selon la localisation de l'élément immobile dans le champ de vision, la suppression de l'OKN est plus ou moins complète (Pola et al., 1995 ; Rubinstein & Abel, 2011). Un élément situé dans le champ visuel central permet de mieux supprimer l'OKN qu'un élément situé dans le champ périphérique. Un élément que peut fixer l'individu, même s'il n'est pas immobile, peut aussi atténuer l'OKN. Rubinstein et Abel (2011) et Williams et al. (2006) ont montré que lorsque des individus doivent rechercher une cible en mouvement, faisant partie d'une scène en mouvement latéral global, tout en gardant leur regard sur un point statique en face d'eux (situation de déplacement masqué de l'attention), l'OKN était atténué par rapport à une exploration libre de la scène. L'atténuation est toutefois moins forte que lorsque l'individu peut fixer son attention sur un élément visuel immobile.

Influence des arrière-plans en mouvement sur la performance de l'individu. L'OKN a des implications pour les activités réalisées sur des arrière-plans en mouvement. Il a des effets négatifs sur la performance des observateurs dans des tâches perceptives rapides. Kaminiarz, Krekelberg, et Bremmer (2007) et Tozzi, Morrone, et Burr (2007) ont étudié l'impact de l'OKN sur les mouvements oculaires de participants et sur leur performance dans une tâche de localisation de cible visuelle. Les participants devaient localiser une cible affichée brièvement sur un arrière-plan structuré en

mouvement. Quand la cible apparaissait pendant la phase lente de l'OKN, une erreur de localisation était faite dans la direction du mouvement du regard. Quand la cible apparaissait juste avant la phase rapide, l'erreur de localisation était faite dans la direction de la phase rapide, alors que l'erreur était dans la direction opposée si la cible apparaissait juste après la phase rapide. L'amplitude de l'erreur variait aussi en fonction du moment de l'apparition de la cible. Les plus petites erreurs étaient obtenues pour les cibles qui apparaissaient juste avant la phase rapide, et les plus grandes pour les cibles qui apparaissaient juste après. Harrison, Thompson, et Sanderson (2010) ont également montré qu'un arrière-plan en mouvement dégradait la performance dans une tâche d'intégration auditive et visuelle impliquant des cibles mobiles. Les participants devaient compter en silence le nombre d'évènements dans lesquels son et image n'étaient pas cohérents. Les objets visuels étaient affichés par le biais d'un visiocasque et étaient superposés au monde réel. Les auteurs ont montré que lorsque les participants marchaient ou lorsque les murs autour du participant bougeaient, la performance était plus faible que lorsque les participants étaient assis ou que les murs étaient immobiles.

Cependant, selon d'autres auteurs, les arrière-plans en mouvement n'ont pas toujours des effets négatifs sur la performance. Par exemple, Menozzi et Koga (2004) ont étudié la lecture d'un texte affiché sur un arrière-plan en mouvement latéral ou vertical. Le déplacement de l'arrière-plan n'avait pas d'effet sur les temps de lecture et le comportement du regard, à savoir le nombre de saccades réalisées pendant la lecture. A la différence des tâches utilisées dans les études précédentes, la lecture est une activité qui requiert du temps et beaucoup de ressources attentionnelles. Les lecteurs devaient en permanence fixer les mots, qui étaient immobiles et restaient en permanence visibles. L'OKN et les effets négatifs qu'il pouvait avoir sur la tâche étaient vraisemblablement annulés. De plus, Menozzi et Koga ont montré que la perception de déplacement de soi était inhibée, probablement parce que les participants devaient se concentrer sur la tâche de lecture pour discriminer le texte de l'arrière-plan structuré. De la même manière, Seno, Ito, et Sunaga (2011) ont montré que l'augmentation de la charge attentionnelle de l'observateur lors d'une tâche visuelle

complexe (e.g., imposant le suivi d'objets multiples) réalisée sur un arrière-plan en mouvement diminuait la force de la perception du déplacement de soi. Ainsi, selon la nature de la tâche réalisée, l'OKN peut ou non dégrader la performance.

L'utilisation d'environnements virtuels se fait généralement dans un but finalisé. Dans une perspective ergonomique, les sensations ressenties par l'utilisateur sont, comme ses compétences, des déterminants de la performance. Lorsqu'un individu est face à un environnement visuel en mouvement, des sensations désagréables comme la cinétose peuvent être ressenties (Diels, Ukai, & Howarth, 2007 ; Golding, 1998, 2006 ; Kennedy, Drexler, & Kennedy, 2010 ; Kuze & Ukai, 2008 ; Webb & Griffin, 2003). Certaines études ont mesuré la satisfaction de l'utilisateur face à ce type de scène. Par exemple, Trutoiu et al. (2009) ont demandé à des participants de se tenir immobile face à un environnement virtuel où une scène visuelle en mouvement était présentée. Ces auteurs ont manipulé le type de déplacement du monde visuel utilisé et ont mesuré l'inconfort induit par les différents mouvements. L'inconfort ressenti par les participants était plus élevé quand la vection était latérale (vers la gauche ou vers la droite) que pour une vection vers l'avant. Menozzi et Koga (2004) ont montré que pendant la lecture d'un texte affiché sur un arrière-plan structuré, l'inconfort était plus élevé quand l'arrière-plan était en mouvement latéral ou vertical que quand il était immobile. Par ailleurs, l'inconfort était plus élevé quand l'arrière-plan bougeait dans les directions de la lecture (de la gauche vers la droite ou de haut en bas) que dans les directions opposées. Prises ensemble, ces deux études suggèrent qu'un déplacement latéral de l'arrière-plan place les individus dans des situations plus inconfortables que lorsque le déplacement suggère un mouvement vers l'avant ou que lorsqu'il n'y a pas de mouvement du tout.

En résumé, l'influence du mouvement de l'arrière-plan est généralement négative sur la performance et l'expérience subjective ressentie par l'utilisateur. Le mouvement peut être sans effet visible si la tâche requiert beaucoup d'attention, plus particulièrement si de nombreux objets statiques sont disposés sur l'arrière-plan et utilisés de manière soutenue (Pola et al., 1992 ; Wyatt &

Pola, 1984). Dans ce cas les effets perturbateurs des sensations de vection et des mouvements réflexes compensatoires du regard (OKN) peuvent être annulés.

En plus de leur caractère dynamique, les interfaces visuelles des jeux vidéo sont souvent très complexes. Leur structuration visuelle et son degré de complexité peuvent également être des sources de baisse de performance.

2.2. Scènes visuelles complexes

2.2.1. Principes

La complexité d'une scène visuelle est caractérisée par la densité des informations visuelles qu'elle présente. Plusieurs manières de définir la complexité existent dans la littérature. Du point de vue de l'activité, la notion d'encombrement visuel est généralement mise en avant. Une scène visuelle est encombrée quand il y a un excès d'information dans celle-ci, et que la « variabilité » de l'information est élevée (Beck, Lohrenz, & Trafton, 2010). La littérature considère qu'il existe deux types d'encombrement qui peuvent détériorer la performance : l'encombrement local et l'encombrement global. L'encombrement global est une mesure de la quantité totale d'information visuelle contenue dans la scène. L'encombrement local est lié à la densité d'information visuelle entourant immédiatement l'objet utilisé par l'individu ou l'endroit de la scène sur lequel il se concentre. Selon les situations, les effets d'encombrement global et local peuvent être additifs ou interactifs.

D'autres auteurs ont utilisé la notion d'entropie, qui relève du même type de raisonnement (Jie & Clark, 2008 ; Raj, Geisler, Frazor, & Bovik, 2005). Ils distinguent également deux types d'entropie similaires aux types d'encombrement, l'entropie globale et l'entropie locale. La différence avec l'encombrement réside principalement dans le fait que l'entropie est, dans la pratique, considérée comme une mesure de la quantité d'information de l'arrière-plan d'une scène, sans considérer les cibles nécessaires à la tâche superposées à cet arrière-plan. L'encombrement est quant à lui utilisé

pour mesurer la quantité d'information de l'ensemble de la scène, dont les cibles à utiliser font alors partie intégrante. Cette différenciation dépend du type de tâche demandé aux individus et s'observe dans les situations concrètes développées dans la section suivante.

Dans tous les cas, la mesure de quantité d'information se base uniquement sur les propriétés de bas niveau de la scène visuelle. Une scène visuelle encombrée ou avec une entropie forte est une scène qui comporte un grand nombre de différences de couleurs, de contraste ou de luminance. Autrement dit, l'encombrement ou l'entropie d'une scène visuelle est intimement lié à la quantité d'éléments visuels saillants différents qu'elle contient. D'un point de vue perceptif, encombrement et entropie peuvent être considérés comme identiques.

2.2.2. Influence de la complexité sur la performance d'un individu

La complexité d'une scène visuelle a généralement une influence négative sur différentes tâches, comme la recherche visuelle (Beck et al., 2010 ; Neider & Zelinsky, 2011 ; Wolfe, Oliva, Horowitz, Butcher, & Bompas, 2002) ou l'utilisation d'un jeu vidéo (Jie & Clark, 2008), et impacte également le comportement du regard (Henderson, Chanceaux, & Smith, 2009), comme le détaillent les paragraphes qui suivent.

La plupart des études des effets de la complexité visuelle d'une scène sur la performance à une tâche ont utilisé des situations de la vie quotidienne. Parmi celles-ci, la plus courante est probablement la recherche d'un objet disposé dans un environnement en désordre. Wolfe et al. (2002) ont étudié la performance de participants lors de la recherche visuelle d'un objet spécifique dans des environnements visuels encombrés. Par exemple, ils devaient chercher un aimant de réfrigérateur sur un bureau encombré. Ces auteurs ont montré que le temps de détection de la cible augmentait en fonction de l'encombrement. Wolfe et al. ont expliqué ces résultats à l'aide de la théorie de la « recherche guidée » (Wolfe, 1994). Comme la séparation entre l'arrière-plan et les objets qui sont des cibles potentielles s'effectue dans la phase pré-attentive, Wolfe et al. (2002) indiquent que plus

la scène est encombrée, plus cette séparation est imparfaite, ce qui allonge le temps de traitement des objets.

D'autres études se sont intéressées à la recherche d'une information dans une carte géographique, un environnement visuel généralement très chargé. Par exemple, Beck et al. (2010) ont étudié le temps de réponse de participants lors de la recherche d'un symbole particulier dans plusieurs cartes différentes. Les auteurs ont fait varier l'encombrement global, l'encombrement local et le nombre de distracteurs dans ces cartes. Ils ont observé une interaction entre les impacts des facteurs d'encombrement global et local sur la performance de recherche. Les temps de réponse augmentaient quand l'encombrement global était élevé, et l'augmentation était encore plus forte lorsque l'encombrement local était également élevé. Par contre, quand l'encombrement local était faible, l'effet de l'encombrement global était minimisé. Par ailleurs, les effets d'encombrement étaient tellement importants qu'ils masquaient l'effet classique du nombre de distracteurs présents dans la scène. Neider et Zelinsky (2011) ont étudié la recherche d'un objet dans des scènes de paysages urbains virtuels, semblables à ceux proposés dans des jeux vidéo de gestion immobilière. Chez eux aussi, le niveau de performance était lié au niveau d'encombrement global. Plus l'encombrement était élevé, plus la performance était faible. Ces auteurs ont aussi montré que mesurer l'encombrement d'une scène visuelle en fonction uniquement de ses caractéristiques visuelles n'était pas la seule méthode possible. Une mesure de jugement subjectif d'encombrement donnait les mêmes résultats en termes d'effets de l'encombrement sur la performance.

Le temps de traitement d'une scène visuelle encombrée pendant une recherche visuelle peut aussi être appréhendé en analysant les mouvements oculaires des observateurs (Rayner, 2009). Par exemple, la durée des fixations oculaires est souvent considérée comme un marqueur des ressources attentionnelles nécessaires au traitement des objets disposés dans l'environnement. En cohérence avec les résultats de Wolfe et al. (2002), Henderson et al. (2009) ont montré que plus une scène réelle est encombrée, plus la durée moyenne des fixations est élevée. Ceci reflète la difficulté pour

les observateurs d'extraire l'information utile parmi le « bruit » visuel local. La durée moyenne des fixations oculaires est donc un indicateur pertinent de l'efficience de la recherche dans une scène encombrée.

Ces études de recherche visuelle montraient des scènes où l'objet à rechercher faisait partie intégrante de l'arrière-plan. Toutefois, dans certains jeux vidéo, les objets avec lesquels le joueur doit interagir peuvent être simplement superposés à l'arrière-plan. L'influence des caractéristiques visuelles de l'arrière-plan sur la performance des joueurs dans un jeu vidéo a été abordée par Jie et Clark (2008), qui ont étudié l'effet de l'entropie de l'arrière-plan visuel sur la difficulté à percevoir des éléments disposés sur l'écran de jeu. Les participants devaient jouer à un jeu de tir, où leur tâche était de repérer des cibles statiques ou mouvantes qui apparaissaient aléatoirement sur l'écran, et de tirer dessus le plus rapidement possible. L'arrière-plan était une scène statique composée de zones d'entropie locale plus ou moins élevée. Les temps de réponse étaient plus lents lorsque les ennemis apparaissaient dans des zones à forte entropie que lorsqu'ils apparaissaient dans des zones à faible entropie. L'enregistrement des mouvements oculaires a également montré que le nombre de saccades effectuées par le joueur était d'autant plus élevé que l'entropie locale autour de la cible à atteindre était importante.

Enfin, d'autres auteurs se sont intéressés aux moyens d'aider les utilisateurs d'environnements visuels encombrés. Du point de vue de l'ergonomie, il est en effet important de concevoir des moyens de contrecarrer la baisse de performance de recherche d'un objet dans les scènes de complexité visuelle élevée. Comme l'encombrement est lié aux caractéristiques visuelles de la scène, une manière de le réduire est de manipuler ces caractéristiques. Wickens et al. (2004) ont utilisé une tâche de localisation d'éléments symboliques sur une carte picturale fortement chargée. Ils ont montré qu'une différence de luminosité entre deux catégories d'éléments de la carte (e.g., immeubles et véhicules) diminuait significativement le temps de recherche d'un de ces éléments. Elle favoriserait la discrimination et permettrait une ségrégation plus aisée des objets. Yeh et Wickens

(2001) ont eux montré qu'un codage par des couleurs améliorait plus le temps de recherche qu'un codage par intensités lumineuses dans une carte picturale et topographique très chargée.

En résumé, le mouvement de l'arrière-plan dégrade généralement la performance de l'individu dans une tâche. Lorsque la tâche de l'individu est suffisamment complexe ou implique des objets fixes superposés à l'arrière-plan en mouvement auxquels doit être portée une attention soutenue, cette dégradation ne serait pas observée. Par contre, la littérature a montré que la « complexité » d'une scène visuelle dégrade toujours la performance. Le niveau de dégradation est lié au degré de complexité de la scène. Des solutions ont toutefois été proposées pour atténuer, voire supprimer cette dégradation.

3. Synthèse

Les interfaces visuelles des jeux vidéo sont dynamiques et complexes. Les théories de l'attention telles que la théorie des ressources multiples de Wickens (2002, 2008) sont encore insuffisantes pour complètement expliquer et prédire le traitement des informations présentées dans ce type d'interfaces.

Les mécanismes de la perception visuelle chez l'humain permettent de mieux comprendre le guidage attentionnel dans des scènes de ce type. Les travaux sur l'exploration de scènes visuelles ont montré que les propriétés visuelles de bas niveau, notamment le mouvement, jouent un rôle important dans ce guidage. Les intentions de l'observateur ont aussi un rôle important, mais ont moins d'influence dans des tâches à fortes contraintes temporelles comme celles des jeux vidéo.

Les travaux sur les scènes dynamiques ou complexes ont montré que des processus particuliers étaient mis en œuvre lors de leur observation. Ceux-ci sont souvent reflétés par les mouvements du regard des observateurs. Par exemple, une conséquence importante du mouvement de l'arrière-plan d'une scène est le déclenchement d'un OKN. Par ailleurs, face à un arrière-plan visuel statique de

structure complexe, les temps de traitement des informations superposées, reflétés par la durée moyenne des fixations oculaires, sont allongés. Différentes études ont enfin montré que la complexité d'une scène ou son caractère dynamique peuvent avoir des conséquences négatives sur la performance et/ou le ressenti de l'individu.

Cependant, contrairement à la théorie des ressources multiples qui modélise le traitement attentionnel dans le cadre de l'activité, la plupart des études présentées dans ce chapitre ne s'intéressent qu'à la perception visuelle dans des situations simples d'observation, et non dans le cadre de l'activité. La plupart des résultats de ces études n'ont donc pas été reproduits avec des individus en interaction avec un système, et a fortiori avec un système complexe.

Chapitre 3 :
Synthèse et hypothèses générales

1. Synthèse des principaux éléments théoriques

Les jeux vidéo sont pour la plupart des applications informatiques qui peuvent être considérées comme des environnements virtuels. L'interaction avec les jeux vidéo est souvent multimodale, puisque les informations peuvent être présentées de manière visuelle, mais aussi sonore ou haptique. De plus, les interfaces des jeux vidéo sont généralement dynamiques et complexes. Leurs interfaces visuelles, notamment, fournissent trois grandes catégories d'informations qui évoluent en temps réel. La source d'information principale correspond aux éléments principaux de la scène d'action avec lesquels le joueur interagit. Une deuxième source d'information est l'arrière-plan sur lequel ces éléments évoluent. Enfin, la troisième source d'information concerne les informations contextuelles qui renseignent le joueur sur l'état du jeu en cours. De fait, l'interaction joueur-jeu vidéo est comparable à l'interaction avec n'importe quel autre environnement dynamique complexe.

Certains modèles théoriques de l'attention sont particulièrement adaptés pour l'étude de ce type d'interaction. De nombreuses études sur les interactions avec des systèmes analogues aux jeux vidéo tels que les systèmes d'aide au pilotage d'avion s'appuient sur la théorie des ressources attentionnelles multiples (Wickens, 2002, 2008). Cette théorie considère que lorsque plusieurs types

d'informations sont traités simultanément, des réservoirs de ressources différents sont activés en fonction de la nature des traitements à réaliser. Un système est optimal lorsque les différents types de ressources attentionnelles impliquées dans le traitement des informations et l'exécution des tâches proviennent de réservoirs différents. Selon cette théorie, les informations de la scène principale d'action du jeu placées en vision fovéale et les informations contextuelles placées en vision périphérique sont traitées par des ressources attentionnelles différentes. Cependant, le partage de l'attention n'est pas expliqué lorsque plusieurs informations doivent être traitées par un même réservoir de ressources attentionnelles. Dans un jeu vidéo par exemple, la théorie des ressources multiples ne permet pas de prédire ou d'expliquer le partage de l'attention entre plusieurs types d'informations présentées en vision fovéale, comme les éléments avec lesquels le joueur interagit et leur arrière-plan. Un principe de conception des interfaces en lien avec la théorie des ressources multiples est le principe de proximité-compatibilité (Wickens & Carswell, 1995). Ce principe indique que pour une tâche nécessitant le traitement de plusieurs sources d'informations, la proximité spatiale et/ou temporelle entre ces sources dans l'interface (visuelle et/ou auditive) doit être proportionnelle à leur niveau d'utilisation conjointe dans la tâche à réaliser.

La conception des interfaces des jeux vidéo nécessite de bien connaitre le fonctionnement de la perception visuelle humaine. La perception de scènes visuelles statiques ou contenant des éléments en mouvement a été largement étudiée. L'attention de l'observateur peut être guidée par les propriétés visuelles de bas niveau de la scène ou par ses intentions. Dans certains cas, le suivi des mouvements oculaires est un indicateur pertinent du guidage attentionnel. Un élément moins documenté, pertinent dans le cas des interfaces visuelles des jeux vidéo, est l'influence des caractéristiques de l'arrière-plan de la scène d'action sur la perception globale de l'interface et sur la performance du joueur. Les variations de mouvement ou de complexité visuelle de l'arrière-plan ont pourtant une influence sur l'attention portée aux objets superposés avec lesquels le joueur doit interagir. En effet, le mouvement d'ensemble d'une scène visuelle non uniforme génère notamment des mouvements compensatoires du regard tels que le réflexe optocinétique. Ce mouvement

oscillatoire du regard, constitué d'une phase lente dans le sens du mouvement de la scène et d'une phase rapide dans le sens inverse, a aussi pour particularité de déporter le regard vers l'origine du mouvement. Un aspect important du réflexe optocinétique est son annulation lorsque l'individu fixe de manière soutenue des éléments immobiles par rapport à l'arrière-plan en mouvement. Le mouvement de l'arrière-plan provoque généralement une dégradation de la performance dans les tâches visuelles les plus simples et rapides, telles que repérer une cible ponctuelle affichée de manière transitoire. Toutefois, dans des tâches où l'individu porte une attention soutenue à des éléments fixes, la dégradation est atténuée, voire supprimée. Enfin, la complexité d'une scène visuelle statique, caractérisée par la densité des informations visuelles qu'elle présente, peut influencer le comportement du regard de l'observateur. La complexité d'une scène visuelle contribue également à dégrader la performance d'un observateur dans diverses tâches de perception ou de recherche d'information.

2. Raisonnement et hypothèses générales

L'objectif de la thèse était de comprendre l'influence des caractéristiques des interfaces visuelles des jeux vidéo sur la performance et les mouvements du regard des joueurs. Précisément, les expériences ont porté sur l'impact des caractéristiques et des mouvements des arrière-plans et des éléments avec lesquels le joueur doit interagir. Les mouvements oculaires des joueurs ont été systématiquement analysés pour comprendre plus en détail l'influence de ces caractéristiques sur leur comportement.

La thèse défendue dans ce document est qu'un mouvement constant de l'arrière-plan ou des éléments constitutifs de l'arrière-plan modifie normalement les mouvements du regard et dégrade la performance du joueur dans les environnements virtuels de type jeux vidéo. Cette modification et cette dégradation sont cependant atténuées ou supprimées si les informations contextuelles sont affichées de telle manière que l'anticipation visuelle est maximisée (chapitre 4), si le mouvement

apparent de l'arrière-plan implique une structuration visuelle qui contient des points fixes saillants (chapitre 5 et chapitre 6), ou si la tâche demandée implique une attention soutenue sur des objets fixes superposés à l'arrière-plan (chapitre 5 et chapitre 6).

L'objectif du chapitre 4 était de montrer que les théories du partage attentionnel telles que le principe de proximité-compatibilité (Wickens & Carswell, 1995) ou la théorie des ressources multiples (Wickens, 2002, 2008) peuvent être utilisées pour guider le positionnement des informations d'une interface visuelle de jeux vidéo. Deux expériences ont permis d'étudier l'influence des informations contextuelles sur la performance et les mouvements du regard des joueurs dans un jeu où les éléments avec lesquels le joueur doit interagir sont en mouvement. Un jeu vidéo simple mais « écologique », où des obstacles qui constituaient l'arrière-plan étaient en mouvement vertical, a été utilisé. La tâche du joueur était d'éviter qu'une boule, qui pouvait être manœuvrée sur un axe horizontal, ne heurte ces obstacles en mouvement du bas vers le haut de l'écran. La localisation du score, information contextuelle typique, a été manipulée. Dans l'expérience 1, le score était localisé soit en haut, soit en bas de l'écran de jeu. Dans l'expérience 2, deux positions supplémentaires ont été testées, au milieu et juste en dessous de l'écran de jeu. L'hypothèse générale était que selon le principe de proximité-compatibilité, la performance est meilleure lorsque l'information contextuelle est au plus près de la zone principale de jeu, en l'occurrence la zone d'anticipation visuelle du mouvement des obstacles.

L'objectif du chapitre 5 était de comprendre dans quelle mesure des théories liées aux caractères dynamiques (Ilg, 1997 ; Pola et al., 1992 ; Wyatt & Pola, 1984) et complexes (Wolfe et al., 2002) de scènes visuelles sont nécessaires pour l'interprétation des variations de performance des joueurs lorsque deux sources d'information visuelles sont superposées. Trois expériences se sont centrées sur l'influence du type de mouvement de l'arrière-plan et de la complexité de sa structure sur la performance et les mouvements du regard des joueurs dans une tâche de recherche visuelle (expériences 3a et 3b) et une tâche de tir simple (expérience 4). La nature du mouvement de

l'arrière-plan a été manipulée. Ce mouvement était soit latéral (de la droite vers la gauche) et généré par une structure visuelle de complexité réduite, soit radial (du fond de l'écran vers l'avant) et généré par une structure visuelle plus complexe. Contrairement à la précédente, cette structure radiale comportait des éléments saillants qui restaient fixes même pendant le mouvement, sur lesquels le joueur pouvait donc fixer son regard et potentiellement annuler le réflexe optocinétique induit par le mouvement. L'hypothèse générale était que le mouvement de type latéral dégrade plus la performance que le mouvement radial, car ce dernier implique une structure visuelle qui contient des points fixes saillants. Dans le cas d'une tâche où une attention soutenue est portée aux objets fixes superposés à l'arrière-plan en mouvement, l'hypothèse était que seule la complexité de la structure dégrade la performance.

L'objectif principal du chapitre 6 était de répliquer les résultats du chapitre 5 dans des environnements et avec des tâches plus proches de ceux rencontrés dans les jeux vidéo, à l'aide de deux expériences. Précisément, l'objectif de l'expérience 5 était de montrer que la complexité visuelle d'un arrière-plan en mouvement détériore toujours la performance à une tâche de jeu vidéo simple, même lorsque cette complexité est intrinsèque à l'arrière-plan et non directement conçue pour générer le mouvement. Trois niveaux de complexité visuelle de l'arrière-plan ont été comparés. La performance et les mouvements du regard ont été mesurés en présence et en l'absence d'un mouvement latéral de ces arrière-plans. L'objectif de l'expérience 6 était de montrer que, comme dans tout autre cadre, le mouvement d'un arrière-plan complexe détériore la performance à une tâche de jeu vidéo uniquement quand celle-ci demande peu de ressources attentionnelles. Le niveau de difficulté de la tâche de jeu vidéo a aussi été manipulé. La tâche difficile impliquait une attention plus soutenue sur les objets fixes superposés à l'arrière-plan que la tâche facile. Comme dans l'expérience 5, la performance et les mouvements du regard ont été mesurés en présence ou en l'absence de mouvement des arrière-plans. L'hypothèse générale était que le mouvement de l'arrière-plan dégrade la performance du joueur à une tâche de jeu vidéo uniquement quand celle-ci est facile et ne nécessite pas une attention trop soutenue. Dans ce type de tâche, l'augmentation de

la complexité de l'arrière-plan viendrait potentialiser la dégradation de la performance induite par le mouvement.

Enfin, l'objectif de la discussion générale était double. Premièrement, elle a permis de synthétiser les résultats obtenus dans les chapitres expérimentaux pour répondre à l'objectif principal de la thèse. Le premier objectif était de proposer un nouveau cadre théorique pour la modélisation du partage attentionnel dans les interfaces visuelles de jeux vidéo, notamment quand les sources d'information visuelles sont superposées. Ce nouveau cadre est basé sur les théories existantes du partage attentionnel dans des situations multitâches, et de la perception visuelle des scènes dynamiques complexes. Plus précisément, il permet d'optimiser le modèle des ressources multiples de Wickens (2002, 2008) pour ces situations spécifiques. Le second objectif était de proposer des solutions pratiques aux concepteurs de jeux vidéo, mais aussi plus généralement aux concepteurs d'environnements virtuels, voire d'environnements dynamiques complexes. Des recommandations de conception des interfaces visuelles ont été élaborées sur la base des résultats expérimentaux.

Chapitre 4 :
Maximiser l'anticipation visuelle des joueurs en appliquant le principe de proximité-compatibilité à la conception des jeux vidéo[1]

1. Introduction

L'interface visuelle des jeux vidéo est généralement constituée d'une scène d'action principale, où les objets avec lesquels le joueur interagit et l'arrière-plan peuvent tous être en mouvement. Des informations contextuelles regroupées dans un affichage tête-haute sont superposées à cette scène d'action, en vision périphérique ou centrale selon la nature de l'information.

Pour concevoir ce type d'interface, les concepteurs de jeux vidéo font généralement appel à des heuristiques ou des recommandations issues de travaux concernant les applications informatiques à interfaces statiques (e.g., Nielsen, 1993). Cependant, les caractéristiques dynamiques des jeux vidéo nécessitent une meilleure spécification de ces heuristiques. En particulier, les concepteurs de jeux

[1] Ce chapitre est une adaptation de l'article publié dans la revue *Human Factors* : Caroux, L., Le Bigot, L., & Vibert, N. (2011). Maximizing players' anticipation by applying the proximity-compatibility principle to the design of video games. *Human Factors, 53*, 103-117.

vidéo pourraient adapter les recommandations issues des recherches sur les interactions entre personnes et systèmes dynamiques (e.g., simulateurs de vols ou cockpits d'avions), comme le principe de proximité-compatibilité (Wickens & Carswell, 1995) exposé dans le chapitre 1, ou encore le modèle « Saillance-Effort-Attente-Valeur » (modèle SEEV) (Wickens et al., 2003 ; Wickens & McCarley, 2008). Ce dernier modèle permet de prédire les mouvements oculaires des opérateurs sur les différentes zones d'intérêt des interfaces visuelles dynamiques. Il postule que la probabilité que le regard se pose sur une zone particulière résulte de 4 facteurs. La saillance visuelle de la zone et l'effort nécessaire pour accéder à l'information, c'est-à-dire la distance entre cette zone et les autres sources d'information, dépendent directement des caractéristiques physiques de la scène. Les deux autres facteurs sont l'attente qu'a l'opérateur de trouver de l'information pertinente dans cette zone, qui est liée à la fréquence de changement de l'information présentée, et la valeur de cette information pour l'opérateur. Par exemple, Horrey et al. (2006) ont prédit les mouvements oculaires de conducteurs d'automobiles entre l'environnement extérieur et les indications fournies à l'intérieur du véhicule en fonction de la priorité des tâches à réaliser dans chaque zone et du taux de renouvellement de l'information à l'intérieur des deux zones.

Le premier objectif du chapitre est de caractériser les mouvements oculaires des joueurs de jeux vidéo qui doivent anticiper leurs actions en regardant notamment vers l'origine du mouvement de l'arrière-plan. Les joueurs doivent aussi intégrer cette information continuellement renouvelée avec d'autres informations disponibles sur l'écran, comme par exemple le score à atteindre. Par conséquent, le deuxième objectif est d'évaluer l'influence de la proximité spatiale entre la zone d'anticipation (partie de la zone principale du jeu où de nouveaux éléments en mouvement, comme par exemple des obstacles, apparaissent) et le lieu d'affichage des informations contextuelles sur le comportement d'anticipation et la performance du joueur. En effet, une plus grande attention des joueurs à la zone d'anticipation devrait induire une meilleure performance au jeu.

Dans les deux expériences, la localisation de l'affichage tête-haute sur l'écran de jeu (i.e., le score) était manipulée. Le jeu vidéo utilisé avait des règles simples et une conception graphique minimaliste en deux dimensions. Il était composé d'une balle en mouvement sur un axe horizontal, contrôlée par le joueur, et d'obstacles de forme carrée qui apparaissaient continuellement en bas de l'écran et montaient vers le haut où ils disparaissaient. L'affichage tête-haute était composé d'un seul élément : le score. Le joueur devait anticiper visuellement les mouvements des obstacles afin d'éviter de les toucher avec la balle, tout en suivant l'évolution de l'affichage du score. Le score augmentait automatiquement avec le temps tant que la balle ne touchait aucun obstacle, mais la partie était perdue dès que la balle heurtait un obstacle.

Une première hypothèse, préliminaire, était que tous les éléments de l'affichage (la balle, les obstacles en mouvement et l'affichage du score) sont pertinents pour le jeu et que par conséquent, les participants les examinent pendant qu'ils jouent. Conformément aux précédentes études, présentées dans le chapitre 1 (e.g., Underwood et al., 2003), qui ont montré que les fixations des conducteurs automobiles étaient le plus souvent observées sur les zones d'anticipation de la route à suivre, la deuxième hypothèse était que les joueurs font plus de fixations sur les zones situées en dessous de l'axe du mouvement de la balle que sur celles situées au-dessus pour anticiper le mouvement des obstacles. La troisième hypothèse était que localiser l'affichage du score en bas de l'écran plutôt qu'en haut réduit la distance entre l'affichage du score et la zone de jeu qui doit être balayée pour anticiper les obstacles (zone d'anticipation), et par conséquent augmente l'anticipation des mouvements des obstacles. En accord avec le principe de proximité-compatibilité (Wickens & Carswell, 1995), la quatrième hypothèse était que les joueurs réalisent une meilleure performance quand le score est affiché en bas plutôt qu'en haut de l'écran en raison d'une amplitude réduite du balayage visuel, et d'une augmentation de l'anticipation. Le tableau 1 montre les caractéristiques principales des affichages du score utilisés dans les conditions Haut et Bas de l'expérience 1.

Tableau 1. *Résumé des caractéristiques principales des quatre conditions de position du score des expériences 1 et 2.*

Condition	Haut (Expériences 1 et 2)	Bas (Expériences 1 et 2)	Milieu (Expérience 2)	Dessous (Expérience 2)
Position du score	En haut de l'écran de jeu	En bas de l'écran de jeu	Au milieu de l'écran de jeu, juste au-dessus de l'axe du mouvement de la balle	Juste en dessous de l'écran de jeu
Distance entre le score et la zone d'anticipation	Très au-dessus (faible proximité)	A l'intérieur de la zone principale de jeu (forte proximité)	Juste au-dessus (proximité intermédiaire)	Juste en dessous (proximité intermédiaire)
Encombrement au niveau de la position du score	Intermédiaire: certains obstacles sont masqués, mais en dehors de la zone d'anticipation	Elevé: le sore masque certains obstacles à l'intérieur de la zone d'anticipation	Intermédiaire: certains obstacles sont masqués, mais en dehors de la zone d'anticipation	Faible: aucun obstacle n'est masqué

2. Expérience 1 : Score en haut ou en bas de l'écran de jeu

2.1. Méthode

Participants. Trente participants (10 femmes, 20 hommes) volontaires droitiers ont pris part à l'expérience. Leur âge moyen était de 22,9 ans (*ET* = 3,2) et leur durée moyenne d'études de 15,8 ans (*ET* = 2,4). Huit des participants n'avaient jamais joué à des jeux vidéo, dix participants jouaient ou avaient joué occasionnellement (moins de 2 heures par semaine) et douze étaient des joueurs réguliers (plus de 2 heures par semaine) (voir le questionnaire pré-expérimental en annexe 2).

Appareillage. Un oculomètre Tobii 1750, non-intrusif, qui ressemble à un écran d'ordinateur de 17'' (résolution de 1024x768 pixels) et permet aux participants de garder leur liberté de mouvement a été utilisé pour expérimenter dans des conditions naturelles. L'oculomètre était contrôlé par un ordinateur qui recueillait les données et exécutait le programme du jeu vidéo. Les positions du regard étaient échantillonnées à une fréquence de 50 Hz avec une précision de 0,5 degré d'angle

visuel. Comme la distance moyenne entre l'écran de l'oculomètre et les yeux des participants était de 660 mm (*ET* = 66), 0,5 degré d'angle visuel correspondait en moyenne à 5,75 mm et 17 pixels sur l'écran. Les mouvements oculaires ont été analysés avec le logiciel « ClearView 2.7.0 ». Les fixations oculaires ont été définies comme toute période où le regard restait au moins 60 ms (trois points successifs à 50 Hz) dans une zone de 30 pixels (en moyenne 0,89 degré d'angle visuel) de diamètre.

Matériel. Le jeu a été créé avec le logiciel « Game Maker 7.0 ». La règle du jeu était d'éviter les obstacles en mouvement à l'aide d'une balle. Les obstacles étaient des carrés noir et gris foncé d'une taille de 32x32 pixels (11x11 mm) qui montaient vers le haut de la fenêtre de jeu à une vitesse de 150 pixels (51 mm) par seconde sur un arrière-plan gris clair. La fenêtre de jeu ne couvrait pas la totalité de l'écran, mais mesurait 480 pixels de large sur 672 pixels de haut et couvrait une zone de 162x227 mm de l'écran de l'oculomètre. Environ cinquante obstacles étaient visibles en même temps. Deux ou trois obstacles apparaissaient toutes les 210 millisecondes en bas de la fenêtre de jeu, à des positions aléatoires sur la même ligne horizontale. Une balle verte positionnée au milieu de la fenêtre pouvait être déplacée par le joueur, à l'aide d'un clavier, sur l'axe horizontal seulement. La balle se déplaçait à une vitesse constante de 210 pixels (environ 6,2 degrés d'angle visuel) par seconde vers la droite ou vers la gauche quand la touche « gauche » ou « droite » du clavier était pressée. La distance verticale entre l'extrémité basse de la balle et le bas de l'écran de jeu était de 320 pixels (108 mm, 9,4 degrés). Par conséquent, la balle pouvait être déplacée sur la largeur complète de l'écran de jeu en 2210 ms, tandis que les obstacles prenaient 2130 ms pour se déplacer du bas de l'écran à l'axe du mouvement de la balle. La partie était perdue dès que la balle touchait un obstacle.

Un score était affiché dans un rectangle bleu dans l'une ou l'autre de deux positions, à savoir les coins en haut à droite et en bas à droite de la fenêtre de jeu (conditions Haut et Bas, figure 3). Au début de chaque essai, le score était de 0 et augmentait de 1 point par ligne d'obstacles évitée.

Autrement dit, le score augmentait de 1 point toutes les 210 ms tant que la balle ne touchait aucun carré.

Figure 3. Captures d'écran des interfaces du jeu dans les quatre conditions de positions du score des expériences 1 et 2.
Seules les conditions Haut et Bas étaient manipulées en inter-participants dans l'expérience 1. Les quatre conditions étaient manipulées en intra-participant dans l'expérience 2. Le trait en pointillés représente l'axe horizontal de déplacement de la balle, mais n'était pas affichée sur les interfaces réelles.

Plan expérimental et procédure. La position de l'affichage du score était manipulée en inter-participants. Les participants ont été répartis aléatoirement entre les deux conditions. Il n'y avait pas de différence significative d'expertise des jeux vidéo entre les groupes. Cinq non joueurs, trois joueurs occasionnels et sept joueurs réguliers ont joué dans la condition Bas, tandis que trois non joueurs, sept joueurs occasionnels et cinq joueurs réguliers ont joué dans la condition Haut.

Le but initial de chaque participant était d'atteindre un score seuil de 100 points en un essai, mais aucune instruction sur le suivi de l'affichage du score ne leur a été donnée. Ils pouvaient essayer d'atteindre le but autant de fois qu'ils voulaient, mais le temps total de jeu effectif (à savoir la durée totale de tous les essais) était limité à dix minutes, même si le score seuil de 100 points n'avait pas été atteint.

Le participant commençait chaque essai en appuyant sur la touche « entrée ». La balle, les obstacles en mouvement et le score apparaissaient et le participant pouvait contrôler la balle à l'aide des touches « gauche » et « droite » pour éviter les obstacles en mouvement vers le haut. Le score augmentait avec le temps tant que les obstacles étaient évités, mais l'essai s'arrêtait dès que la balle touchait un des carrés, et tous les éléments du jeu disparaissaient. Le score était enregistré et remis ensuite à 0 pour l'essai suivant. Quand un participant atteignait le score seuil de 100 points en un essai, le jeu continuait jusqu'à ce que la boule ait touché un obstacle. Un nouveau score seuil majoré de 50 points (i.e., 150 points) était proposé au participant pour les essais suivants. Si le participant atteignait ce nouveau score seuil avant le temps de jeu effectif de dix minutes, le score seuil était augmenté à nouveau de 50 points (soit 200 points), et ainsi de suite jusqu'à ce que les dix minutes soient atteintes. En moyenne, chaque participant a fait *M* = 83,4 essais (*ET* = 17,2), d'une durée moyenne de *M* = 7,2 secondes (*ET* = 4,9) et dont le score moyen était de *M* = 32,3 points (*ET* = 7,4).

Mesures dépendantes. La première mesure était le score moyen obtenu par chaque participant pour chaque essai. Les autres mesures étaient calculées à partir des données de mouvements oculaires. La fenêtre de jeu était divisée en huit zones d'intérêt (AOI, figure 4). Les AOI 1 à 6 étaient horizontales, tandis que les deux autres (AOI « ScoreHaut » et « ScoreBas », de taille 123x89 pixels) couvraient les positions d'affichage du score utilisées dans les conditions Haut et Bas. L'AOI ScoreHaut était extraite de l'AOI 1 (taille 480x310 pixels), qui incluait tout l'espace au-dessus de l'axe de déplacement de la balle. Le reste de l'écran de jeu était divisée en cinq AOI horizontales de tailles quasiment égales numérotées, de haut en bas, de 2 à 6. La ligne médiane de l'AOI 2 (480x72 pixels) était l'axe horizontal de déplacement de la balle. Les tailles des AOI 3 à 6, où les participants pouvaient anticiper le mouvement des obstacles, étaient les mêmes que celle de l'AOI 2, excepté pour l'AOI 6 (480x74 pixels). L'AOI ScoreBas était extraite des AOI 5 et 6. Les deux premières variables dépendantes étaient le nombre total de fixations oculaires réalisées pendant une partie et le nombre de fixations dans chaque AOI.

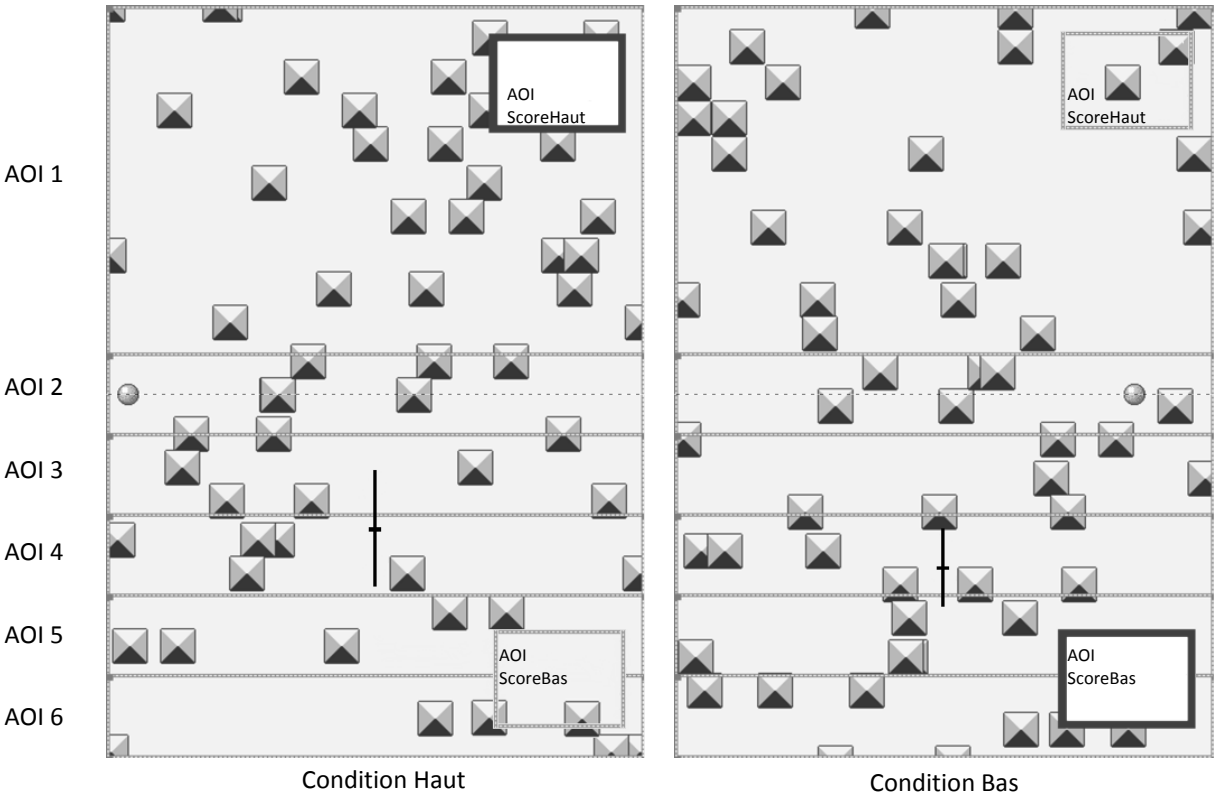


Figure 4. Captures d'écran du jeu vidéo utilisé dans l'expérience 1 dans les conditions Haut et Bas.
Les lignes horizontales en gris foncé marquent les limites des AOI horizontales. La plus haute est l'AOI 1 et la plus basse l'AOI 6. Les rectangles en trait plein noir marquent les limites des AOI où le score était positionné dans chaque condition. Les rectangles gris indiquent la limite des AOI du score vide. Les barres horizontales et verticales représentent la position et la dispersion moyennes des fixations oculaires dans chaque condition.

Deux mesures supplémentaires ont été calculées à partir des coordonnées verticales (axe des ordonnées Y) de la position de chaque fixation oculaire pour évaluer l'anticipation par les participants du mouvement des obstacles et l'amplitude du balayage visuel. L'anticipation du mouvement des obstacles a été évaluée par la position verticale moyenne du regard, mesurée en pixels du haut (pixel 1) vers le bas (pixel 672) de la fenêtre de jeu. Une position moyenne basse du regard devait en effet permettre aux participants de détecter l'apparition des obstacles en bas de l'écran et d'anticiper leur mouvement plus rapidement. L'amplitude du balayage visuel pendant le jeu était évaluée par l'écart-type de la position verticale moyenne du regard, qui donnait une idée de la dispersion verticale des fixations. Les écarts-types de la position verticale moyenne du regard les plus grands indiquaient que les participants balayaient régulièrement une plus grande partie de l'écran de jeu.

2.2. Résultats

Pour optimiser l'analyse des données, les essais ont été filtrés en utilisant deux critères. Les vingt premiers essais de chaque participant, approximativement les 2 premières minutes de jeu effectif, ont été exclus de l'analyse. Ils correspondaient à une phase d'entrainement durant laquelle les participants découvraient le jeu et son contrôle à l'aide du clavier. Parmi les essais restants (environ 60 par participant), seuls ceux dans lesquels un score de 20 ou plus a été atteint ont été retenus pour l'analyse. En effet, laisser le jeu s'exécuter seul, sans contrôle humain de la balle, pendant 250 essais, donnait un score moyen de 14,6 (*ET* = 3,6). Par conséquent, les scores en dessous de 20 étaient liés au hasard et ne donnait pas d'informations sur les capacités réelles des joueurs. Le nombre moyen d'essais exclus en raison de scores trop faibles est de *M* = 21,3 (*ET* = 10,5) pour la condition Bas et de *M* = 21,5 (*ET* = 13,9) pour la condition Haut. Au final, l'analyse de données a été réalisée sur *M* = 43,1 essais (*ET* = 7,0) par participant dans la condition Bas et *M* = 40,9 essais (*ET* = 7,4) dans la condition Haut. Leur durée était de *M* = 9,3 secondes (*ET* = 1,8) pour la condition Bas et de *M* = 9,5 secondes (*ET* = 2,1) pour la condition Haut.

Nombre de fixations oculaires. Le nombre total de fixations faites par les participants dans la fenêtre de jeu (ensemble des AOI) ne variait pas significativement entre les conditions Haut (*M* = 1052, *ET* = 302) et Bas (*M* = 1246, *ET* = 292), $t(28) = 1{,}79$, $p = 0{,}08$. La distribution des fixations oculaires entre les huit AOI dans chaque condition est donnée dans le tableau 2. Quelle que soit la condition, environ 40% des fixations étaient localisées dans l'AOI 3, 20% dans l'AOI 2 (autour de l'axe de déplacement de la balle) et 20% dans l'AOI 4. Plus de 70% des fixations étaient localisées dans les AOI 3 à 6. Par conséquent, les participants regardaient plus les AOI en dessous de l'axe horizontal de déplacement de la balle, là où ils pouvaient anticiper le mouvement des obstacles.

Des tests t pour échantillons indépendants ont été utilisés pour comparer les proportions de fixations dans chaque AOI entre les deux conditions. La proportion de fixations était plus grande dans l'AOI ScoreHaut dans la condition Haut que dans la condition Bas, $t(28) = -8{,}53$, $p < 0{,}001$, *d* de Cohen =

3,12. Par contre, la proportion de fixations était plus grande dans l'AOI ScoreBas dans la condition Bas que dans la condition Haut, $t(28) = 6{,}30$, $p < 0{,}001$, $d = 4{,}59$. Pour les autres AOI, les proportions de fixations n'étaient pas significativement différentes entre les deux conditions.

Tableau 2. *Nombre et proportion moyens de fixations oculaires faites par les participants dans chaque AOI et dans chaque condition de l'expérience 1.*

	Condition Haut		**Condition Bas**	
AOI	Nombre de fixations	Proportion	Nombre de fixations	Proportion
1	43 (61)	4.1%	16 (38)	1.3%
2	206 (131)	19.6%	251 (225)	20.1%
3	418 (118)	39.7%	504 (152)	40.4%
4	245 (92)	23.2%	284 (109)	22.8%
5	76 (62)	7.3%	117 (74)	9.4%
6	16 (38)	1.5%	21 (40)	1.6%
ScoreHaut	43 (20)	4.1%	0 (0)	0.0%
ScoreBas	5 (6)	0.1%	54 (30)	4.3%

Note. Les nombres entre parenthèses correspondent aux écart-types.

Position verticale et dispersion des fixations. Comme l'indique la figure 4, la position verticale moyenne des fixations oculaires était plus basse dans la condition Bas ($M = 501{,}9$, $ET = 31{,}7$) que dans la condition Haut ($M = 467{,}1$, $ET = 29{,}2$), $t(28) = 3{,}13$, $p < 0{,}01$, $d = 1{,}14$. De plus, la dispersion verticale moyenne des fixations était plus faible dans la condition Bas ($M = 70{,}2$, $ET = 12{,}1$) que dans la condition Haut ($M = 103{,}7$, $ET = 16{,}5$), $t(28) = 6{,}35$, $p < 0{,}001$, $d = 2{,}32$.

Score de performance. Le score de performance ne différait pas significativement entre les conditions Haut ($M = 40{,}7$, $ET = 7{,}7$) et Bas ($M = 39{,}3$, $ET = 8{,}9$), $t(28) = 0{,}39$, $p = 0{,}70$. Le score maximal moyen était de 96,0 ($ET = 26{,}0$). Cinq (33%) et huit (53%) participants sur 15 ont atteint le score seuil de 100 points respectivement dans les conditions Bas et Haut. Un seul participant jouant dans la condition Bas a atteint le seuil suivant de 150 points.

2.3. Discussion

2.3.1. Hypothèse préliminaire: balayage visuel de l'interface visuelle du jeu

L'hypothèse préliminaire a été confirmée (tableau 2). Les enregistrements des mouvements oculaires ont montré que quelle que soit la position de l'affichage du score, plus de 95% des fixations des participants se faisaient soit dans la zone de la balle (AOI 2, environ 20% des fixations), soit dans la zone d'anticipation du mouvement des obstacles (AOI 3 à 6, entre 70 et 75% des fixations), soit dans l'AOI du score (entre 4 et 5% des fixations). Le nombre de fixations dans les AOI ScoreHaut et ScoreBas variait selon la position de l'affichage du score. Les participants ne fixaient les AOI ScoreHaut ou ScoreBas que quand le score y était affiché, c'est-à-dire respectivement dans les conditions Haut et Bas. Par conséquent, l'affichage du score était pertinent pour la tâche de jeu et était bien suivi par les participants pendant qu'ils jouaient. Selon le modèle SEEV (Wickens et al., 2003), le grand nombre de fixations à l'intérieur des zones d'anticipation de l'écran peut résulter des attentes du joueur, qui sait que les nouvelles informations viennent de ces AOI. L'affichage du score serait aussi une position soumise à une forte attente, qui plus est de valeur élevée pour les joueurs, et pourrait aussi être plus saillant que les autres éléments de l'interface.

2.3.2. Amplitude réduite du balayage visuel dans la condition Bas

La deuxième hypothèse a été confirmée. Quelle que soit la position de l'affichage du score, plus de 70% de fixations étaient faites dans la partie de la fenêtre de jeu où le mouvement des obstacles pouvait être anticipé. La troisième hypothèse a également été vérifiée. Le comportement visuel des participants était modifié quand le score était positionné en bas plutôt qu'en haut de la fenêtre de jeu. La position verticale moyenne des fixations était plus basse et leur dispersion verticale plus faible dans la condition Bas que dans la condition Haut. Par conséquent, positionner le score en bas de la fenêtre de jeu orientait le regard des participants vers le point d'origine des obstacles en réduisant l'amplitude du balayage visuel et en incitant à l'inspection du bas de l'écran.

2.3.3. Influence d'une amplitude réduite du balayage visuel sur la performance

Selon le principe de proximité-compatibilité (Wickens & Carswell, 1995), une plus grande proximité spatiale entre l'affichage du score et la zone d'anticipation devait faciliter l'anticipation du mouvement des obstacles et augmenter la performance du joueur. Cependant, cette quatrième hypothèse n'a pas été confirmée. Minimiser l'amplitude du balayage visuel en affichant le score en bas de la fenêtre n'a pas provoqué d'augmentation significative de la performance des joueurs en comparaison avec la condition Haut, malgré le fait qu'une position verticale moyenne plus basse des fixations devait les aider à anticiper le mouvement des obstacles.

Trois facteurs peuvent expliquer pourquoi déplacer l'affichage du score vers la zone d'anticipation n'a pas augmenté la performance. Premièrement, comme l'a précisé la description du principe de proximité-compatibilité (chapitre 1), une trop grande proximité spatiale entre les signaux visuels qui doivent être suivis simultanément peut provoquer une interférence intramodale, et générer des effets de confusion ou de masquage qui perturbent la perception et l'intégration des informations (Wickens et al., 2004). Puisque dans la condition Bas le score était à l'intérieur de la zone d'anticipation du jeu, ce chevauchement pourrait avoir gêné l'anticipation du mouvement des obstacles.

Deuxièmement, les participants pourraient partager leurs ressources attentionnelles visuelles entre les canaux focal et ambiant pour optimiser l'anticipation du mouvement des obstacles dans la condition Haut, la moins favorable. Les données ont montré que bien que la position verticale des fixations était plus basse dans la condition Bas que dans la condition Haut, la performance n'était pas significativement supérieure. Ceci suggère que les participants pouvaient utiliser efficacement la vision parafovéale pour observer la zone d'anticipation même dans la condition Haut.

Troisièmement, le score seuil de 100 points était difficile à atteindre pour beaucoup de participants. Dix-sept participants ont joué 10 minutes sans atteindre le but fixé. Les jeux vidéo sont des objets de divertissement qui doivent être agréables à manipuler, et le challenge doit être ajusté aux capacités

des joueurs pour être accepté. Ici, les participants qui ont eu des difficultés à atteindre le score seuil ont pu estimer que le jeu était trop difficile et être par conséquent peu motivés. Un effet plancher pourrait dans ce cas expliquer pourquoi la manipulation de l'affichage du score n'a pas influencé significativement la performance.

A la suite de ces résultats, une nouvelle expérience a été conçue pour préciser comment réduire l'amplitude du balayage visuel tout en obtenant de meilleures performances dans les jeux vidéo, en appliquant le principe de proximité-compatibilité.

2.3.4. Raisonnement de l'expérience 2

L'expérience 1 a été modifiée de deux façons. Premièrement, il fallait réduire l'interférence visuelle provoquée par l'affichage du score dans la zone d'anticipation du jeu en condition Bas. Deux nouvelles positions d'affichage du score ont été testées (figure 3). Le score pouvait être positionné sur la droite de la fenêtre de jeu, juste au-dessus de l'axe de déplacement de la balle (condition Milieu), ou en bas à droite de l'écran de l'ordinateur sous la fenêtre de jeu (condition Dessous) pour inciter les participants à regarder l'endroit d'où viennent les obstacles, comme dans la condition Bas. L'idée était de placer l'affichage du score plus proche de la zone d'anticipation que dans la condition Haut, mais pas à l'intérieur de cette zone. Deuxièmement, le score seuil initial a été divisé par 2,5 (de 100 à 40) pour rendre le jeu plus facile, augmenter la motivation des joueurs les plus novices et éviter les éventuels effets planchers.

Les hypothèses ont également été modifiées. Une plus grande proximité entre les signaux visuels à suivre simultanément n'amènerait une meilleure performance qu'en l'absence d'interférence visuelle ou de chevauchement entre les signaux. Dans des environnements dynamiques, la solution optimale serait de minimiser l'amplitude du balayage visuel entre les zones pertinentes du jeu tout en évitant leur chevauchement.

La première hypothèse de l'expérience 2 était donc que l'anticipation, évaluée par la position verticale moyenne des fixations, dépend de la position d'affichage du score par rapport à la zone d'anticipation : l'anticipation est plus forte dans les conditions Bas et Dessous que dans les conditions Milieu et Haut. La deuxième hypothèse était que l'amplitude du balayage visuel de l'interface, évaluée par la dispersion verticale des fixations, est minimale dans la condition Bas, intermédiaire (et optimale) dans les conditions Milieu et Dessous et maximale dans la condition Haut. L'hypothèse 3a stipulait que la performance est la meilleure dans la condition Dessous, qui associe la meilleure anticipation avec une amplitude optimale du balayage visuel. Selon l'hypothèse 3b, la deuxième meilleure performance était attendue pour la condition Milieu, qui combine la même amplitude optimale de balayage visuel que la condition Dessous avec des possibilités plus faibles d'anticipation. Dans les deux cas, la performance devait être plus forte que dans la condition Haut, qui n'incite pas à l'anticipation, et la condition Bas, qui crée de l'interférence du fait du chevauchement avec la zone d'anticipation. Le tableau 1 montre les caractéristiques principales des affichages du score utilisées dans les quatre conditions de l'expérience 2.

3. Expérience 2 : Score en haut, en bas, au milieu ou en dessous de l'écran de jeu

3.1. Méthode

Participants. Vingt-et-un participants (12 femmes, 9 hommes) volontaires droitiers ont pris part à l'expérience. Leur âge moyen était de 26,2 ans (*ET* = 7,1) et leur durée moyenne d'études de 16,1 ans (*ET* = 2,8). Neuf d'entre eux n'avaient jamais joué à des jeux vidéo, neuf étaient des joueurs occasionnels et trois des joueurs réguliers.

Matériel et appareillage. Le jeu vidéo utilisé était le même que celui de l'expérience 1, mais deux nouvelles positions ont été utilisées pour le score. Le score était affiché à droite juste au-dessus de l'axe de déplacement de la balle (dans l'AOI 1) dans la condition Milieu, et à droite en dessous de la fenêtre de jeu dans la condition Dessous (figure 3). L'appareillage était le même que dans la première expérience. La distance moyenne entre l'écran de l'oculomètre et les yeux des participants était de 651 mm (*ET* = 71), et donc 0,5 degré d'angle visuel correspondait à environ 5,67 mm à l'écran.

Plan expérimental et procédure. La position de l'affichage du score était manipulée en intra-participant. La durée totale de jeu effectif était de 20 minutes. Le score était placé aléatoirement dans une des quatre positions testées (Haut, Bas, Milieu, Dessous) au début de chaque essai. Le but des participants était le même que dans l'expérience 1, mais le score seuil initial était de 40 points au lieu de 100, et ce score seuil augmentait de 40 points (au lieu de 50) chaque fois que les participants l'atteignaient. L'ajustement du score seuil était valable pour tous les essais suivants, indépendamment de la condition d'affichage du score dans laquelle le score seuil avait été atteint. Les participants étaient invités à prendre une courte pause approximativement toutes les cinq minutes de jeu effectif. Les participants ont réalisé en moyenne 176,7 essais (*ET* = 44,2), qui duraient en moyenne 7,4 secondes (*ET* = 5,3) et dont le score moyen était de *M* = 33,8 points (*ET* = 10,3).

Mesures dépendantes. L'expérience 2 utilisait les mêmes mesures dépendantes que l'expérience 1, excepté le nombre de fixations par AOI. En effet, l'analyse de l'expérience 1 a seulement révélé que le regard des participants allait dans toutes les parties de l'écran qui affichaient de l'information, quelle que soit la condition. Les mesures de mouvements oculaires plus informatives, à savoir la position verticale moyenne et la dispersion verticale des fixations, apparaissaient suffisantes pour estimer l'anticipation visuelle des participants et leur amplitude de balayage visuel.

3.2. Résultats

Comme dans l'expérience 1, les vingt premiers essais de chaque participant et tous les essais dont les scores étaient inférieurs à 20 ont été exclus de l'analyse. Sur l'ensemble des participants, le nombre d'essais exclus pour un score trop faible allait de *M*= 12,2 (*ET* = 7,1) pour la condition Haut à *M* = 14,8 (*ET* = 8,2) pour la condition Dessous. L'analyse des données a été réalisée sur *M* = 26,0 essais (*ET* = 6,5) par participant pour la condition Haut, *M* = 27,2 essais (*ET* = 6,1) pour la condition Bas, *M* = 25,2 essais (*ET* = 6,0) pour la condition Milieu et *M* = 25,8 essais (*ET* = 6,4) pour la condition Dessous. Leur durée allait de *M* = 9,1 secondes (*ET* = 2,1) pour la condition Haut à *M* = 10,2 secondes (*ET* = 2,9) pour la condition Dessous.

Les données ont été analysées en utilisant des ANOVA avec la position du score en facteur intra-participant. Quand le test de sphéricité de Mauchley était significatif, une correction de Greenhouse-Geisser a été appliquée.

Position verticale et dispersion des fixations. La position d'affichage du score avait un effet sur la position verticale des fixations, $F(3, 60) = 17{,}65$, $p < 0{,}001$, $\eta^2_p = 0{,}469$ (figure 5 et figure 6). Des comparaisons planifiées ont montré que la position verticale des fixations était marginalement inférieure quand le score était affichée en bas (*M* = 467,6, *ET* = 45,5) plutôt qu'en dessous de la fenêtre de jeu (*M* = 464,7, *ET* = 43,5), $F(1, 20) = 4{,}02$, $p = 0{,}059$. La position verticale des fixations était aussi plus basse dans les conditions Bas et Dessous que dans les conditions Milieu (*M* = 453,8, *ET* = 44,1) ou Haut (*M* = 453,3, *ET* = 46,2), $Fs(1, 20) > 12{,}16$, $p < 0{,}01$, mais n'était pas significativement différente entre ces deux dernières conditions, $F(1, 20) = 0{,}04$, $p = 0{,}84$.

La dispersion verticale des fixations était aussi influencée par la position d'affichage du score, $F(3, 60) = 23{,}80$, $p < 0{,}001$, $\eta^2_p = 0{,}543$ (figure 5 et figure 6). La dispersion verticale des fixations était plus grande dans la condition Haut (*M* = 87,8, *ET* = 14,4) que dans les trois autres, $Fs(1, 20) > 16{,}62$, $p < 0{,}001$. La dispersion verticale des fixations était plus grande dans la condition Dessous (*M* = 77,9, *ET* = 9,9) que dans les conditions Bas (*M* = 73,1, *ET* = 8,3) et Milieu (*M* = 74,9, *ET* = 10,1), $Fs(1, 20) > 7{,}12$, *p*

< 0,05, mais n'était pas significativement différente entre ces deux dernières conditions, $F(1, 20) = 2,34$, $p = 0,14$.

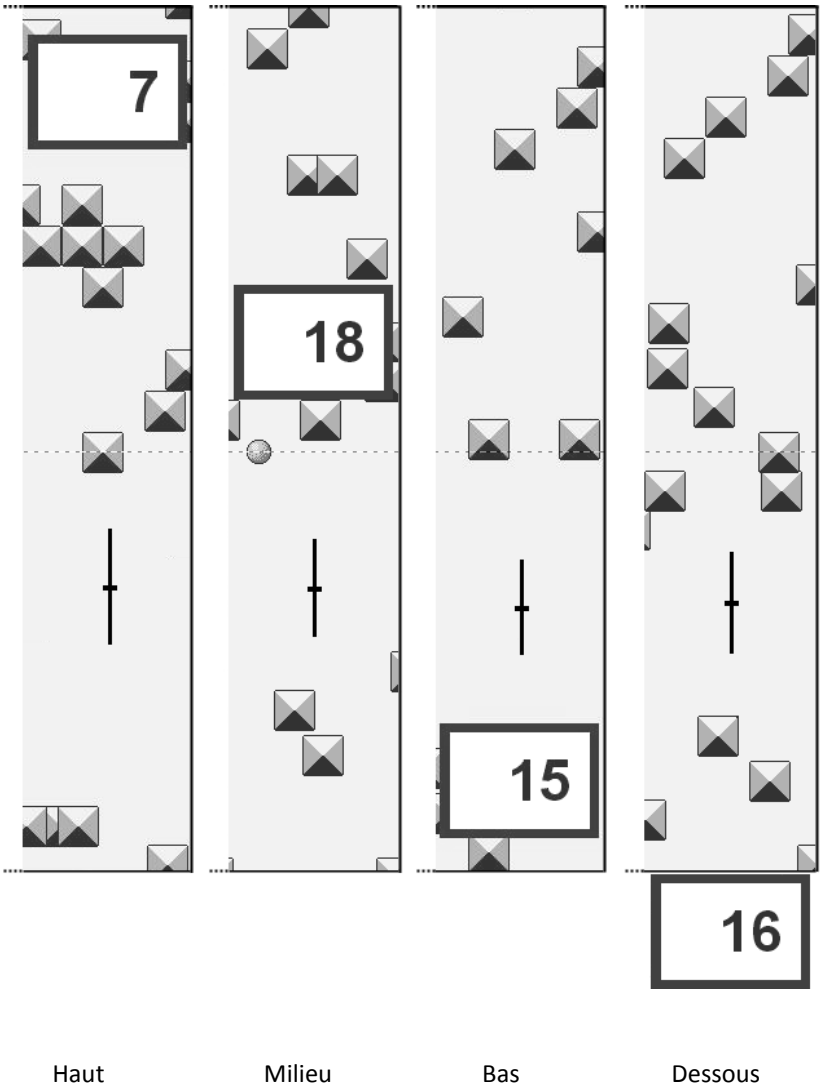


Figure 5. Côtés droits des captures d'écran du jeu vidéo utilisé dans l'expérience 2 dans les quatre conditions de position du score.
Les barres noires horizontale et verticale représentent respectivement la position et la dispersion verticales moyennes des fixations oculaires dans chaque condition.

Score de performance. La figure 6 indique que la position de l'affichage du score avait un effet sur la performance, $F(3, 60) = 2.90$, $p < 0,05$, $\eta^2_p = 0,127$. Des scores meilleurs ou marginalement meilleurs étaient obtenus dans la condition Dessous ($M = 43,3$, $ET = 12,5$) que dans les conditions Haut ($M = 38,7$, $ET = 8,7$), $F(1, 20) = 6,83$, $p < 0,05$, et Bas ($M = 40,2$, $ET = 9,4$), $F(1, 20) = 4,04$, $p = 0,058$. Le score de performance n'était pas significativement différent entre les conditions Dessous et Milieu (M =

42,3, *ET* = 14,6), *F*(1, 20) = 0,26, *p* = 0,62. Aucune autre comparaison n'a atteint la significativité statistique. Le score maximal moyen était de 123,9 (*ET* = 45,4). Tous les participants ont atteint le premier score seuil de 40 points au cours des 20 minutes de jeu, neuf d'entre eux (43%) ont atteint le troisième score seuil de 120 points, et deux d'entre eux (10%) le cinquième score seuil de 200 points.

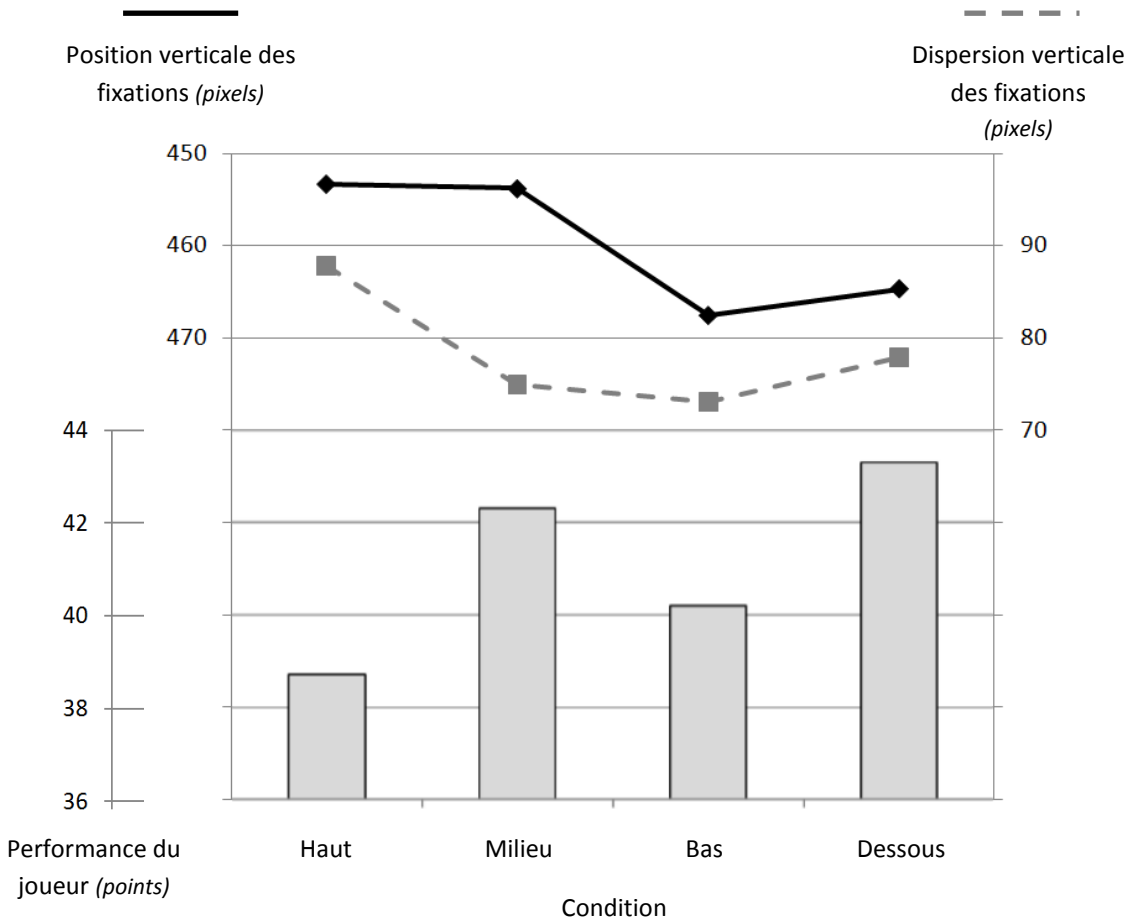


Figure 6. Mesures moyennes obtenues pour les trois variables dépendantes dans chaque condition de l'expérience 2.
Les barres représentent la performance moyenne des joueurs.

3.3. Discussion

Comme attendu, la position verticale moyenne des fixations était plus basse dans les conditions Dessous et Bas que dans les conditions Milieu et Haut. Elle n'était pas significativement différente entre ces deux dernières conditions, mais était marginalement plus basse dans la condition Bas que dans la condition Dessous. Ceci pourrait être du à une utilisation plus importante par les participants de leur champ visuel parafovéal pour traiter les informations visuelles dans la zone d'anticipation la

plus basse (AOI 6) et suivre l'affichage du score dans la condition Dessous. En effet, un plus grand effort était nécessaire pour aller de la zone principale de jeu au score dans cette condition.

La deuxième hypothèse, à savoir que la dispersion verticale des fixations est minimale dans la condition Bas, intermédiaire dans les conditions Milieu et Dessous et maximale dans la condition Haut, n'a pas été vérifiée. Comme attendu, la dispersion verticale des fixations était la plus grande dans la condition Haut et était plus grande dans la condition Dessous que dans la condition Bas. Cependant, la dispersion verticale était plus faible dans la condition Milieu que dans la condition Bas. Une explication serait que dans l'expérience 2, le score était plus proche de la position verticale moyenne des fixations dans les conditions Milieu et Bas que dans la condition Dessous (figure 5 et figure 6). En effet, la position verticale moyenne des fixations n'a pas varié autant que dans l'expérience 1 entre les conditions, et est toujours restée à la limite ou juste en dessous des AOI 3 et 4.

Plus important, l'hypothèse 3a a été vérifiée. Les scores des participants étaient plus élevés dans la condition Dessous que dans les deux conditions de l'expérience 1. L'absence de différence significative entre les scores obtenus dans les conditions Haut et Bas réplique les résultats de l'expérience 1. En revanche, l'hypothèse 3b n'a pas été vérifiée du fait de l'absence de différence significative entre les conditions Dessous et Milieu. Dans cette dernière condition, les scores étaient intermédiaires entre ceux de la condition Dessous et ceux des conditions Haut et Bas. L'avantage attendu de la condition Dessous sur la condition Milieu a pu être réduit du fait que même si la position verticale moyenne des fixations était plus faible dans la condition Dessous, l'amplitude du balayage visuel restait plus faible dans la condition Milieu.

L'hypothèse théorique principale qui a guidé la conception de l'expérience 2 a été essentiellement vérifiée. Dans le contexte dynamique des jeux vidéo, une plus grande proximité entre des informations visuelles suivies simultanément semble favoriser leur intégration seulement en l'absence de chevauchement entre les zones visuelles pertinentes de l'écran de jeu. L'amplitude

optimale du balayage visuel correspondrait à la présentation des différentes sources d'information côte à côte plutôt qu'en superposition, même partielle.

4. Discussion générale

Le principe de proximité-compatibilité recommande aux concepteurs d'interface de placer le plus proche possible les sources d'information qui partagent des critères d'utilisation communs, c'est-à-dire qui contribuent à la réalisation d'une même tâche ou d'une même opération mentale. Cette règle de conception aboutirait à minimiser l'effort nécessaire pour déplacer l'attention entre les positions des deux sources.

Les deux expériences de cette étude étaient basées sur un jeu vidéo où la zone principale d'attention était la zone d'anticipation visuelle placée dans la partie basse de la fenêtre de jeu. Dans l'expérience 1, l'affichage du score était positionné soit en haut de la fenêtre, comme dans la plupart des jeux de ce type, soit en bas. Puisque la position haute était éloignée de la zone d'anticipation visuelle, le principe de proximité-compatibilité recommanderait plutôt de positionner l'affichage du score en bas de la fenêtre de jeu, pour minimiser les déplacements attentionnels entre cette zone et le score. Les résultats ont montré des fixations en moyenne plus basses et une amplitude de balayage visuel réduite dans la condition Bas mais, contrairement à ce qui était attendu, la performance n'a pas augmenté significativement.

Dans l'expérience 2, deux positions supplémentaires d'affichage du score ont été testées, à savoir les positions Milieu et Dessous. Les deux ont été choisies pour être proches de la zone d'anticipation visuelle comme le recommande le principe de proximité-compatibilité, mais pas dans la zone d'anticipation, pour tenir compte des résultats de l'expérience 1. L'expérience 2 a démontré que quand le score était affiché juste en dessous de la fenêtre de jeu, l'anticipation visuelle des joueurs et la performance augmentaient significativement par rapport à la condition Haut, malgré une

augmentation significative de l'amplitude du balayage visuel par rapport à la condition Bas. Au final, même si suivre le score faisait partie du jeu, une meilleure performance était obtenue en l'absence de chevauchement entre la position du score et la zone d'anticipation visuelle.

Deux interprétations peuvent être avancées. La première suppose que le principe de proximité-compatibilité s'applique aux jeux vidéo dynamiques, mais seulement dans des contextes d'absence de chevauchement visuel entre la zone d'anticipation et les informations qui doivent être suivies simultanément. La proximité spatiale entre ces sources d'informations faciliterait l'attention divisée seulement en l'absence d'interférence provoquée par le chevauchement. En cas de chevauchement, l'incapacité à focaliser l'attention sur une seule des deux sources deviendrait préjudiciable à la performance. Les données expérimentales ont montré que les joueurs pouvaient plus anticiper les mouvements des obstacles quand le score était placé en bas ou juste en dessous de la fenêtre de jeu. Le fait que la performance soit meilleure dans la condition Dessous que dans la condition Bas, bien que l'amplitude du balayage visuel soit plus grande, suggère que minimiser le chevauchement visuel des sources d'information est plus bénéfique pour la performance des joueurs que réduire à tout prix l'amplitude du balayage visuel.

Dans le modèle des ressources multiples (Wickens, 2002, 2008), la modalité visuelle inclut deux canaux attentionnels différents, à savoir les canaux focal et ambiant. Dans la condition Dessous, le canal focal peut être utilisé pour anticiper les mouvements des obstacles tandis que le canal ambiant suivrait les modifications de l'affichage du score. Cette interprétation est quelque peu contre-intuitive, car le canal focal est considéré comme plus adéquat pour la reconnaissance d'objets que le canal ambiant. Cependant, l'anticipation précise du mouvement des obstacles pourrait nécessiter plus d'attention que le suivi par intermittence de l'affichage d'un score écrit en gros caractères. La grande taille et la familiarité des chiffres composant le score devraient rendre l'identification du score en vision périphérique plus facile. Plusieurs auteurs ont déjà démontré que des mots pouvaient

être lus et reconnus dans le champ visuel périphérique par des déplacements implicites de l'attention (H.-W. Lee, Legge, & Ortiz, 2003 ; Pelli et al., 2007).

Une interprétation alternative est possible en considérant le « continuum » du principe de proximité-compatibilité (Wickens & Carswell, 1995). Ce continuum souligne que plus les sources d'information sont utilisées simultanément pour la même tâche, plus elles doivent être placées proches l'une de l'autre. Ici, les deux sources d'information (le mouvement des obstacles et l'affichage du score) pourraient n'être pas suffisamment utilisées ensemble pendant la tâche pour justifier la forte proximité de la condition Bas. Les joueurs devaient suivre le score pendant qu'ils jouaient pour voir quand ils avaient atteint le score seuil, mais connaitre le score ne les aidait pas à éviter les obstacles, et la nécessité d'utiliser conjointement les deux sources d'information était assez faible. Par conséquent, placer les deux sources d'information très proches l'une de l'autre ne serait pas nécessaire pour une performance optimale. Les placer modérément proches l'une de l'autre (comme dans les conditions Dessous et Milieu) plutôt qu'à l'extrémité haute de l'écran (condition Haut) serait suffisant. Quelle que soit l'explication, le principe de proximité-compatibilité et le modèle des ressources multiples fournissent un cadre suffisant pour prédire et interpréter les données.

Ces deux expériences avaient pour but de comprendre le partage de l'attention entre les éléments d'une scène d'action d'un jeu vidéo et les informations contextuelles. Les résultats ont montré que la performance du joueur est meilleure si ces deux informations sont disposées de telle manière qu'elles puissent être traitées par deux canaux attentionnels différents. La question reste posée dans le cas où deux types d'informations, comme les éléments avec lesquels le joueur doit interagir et l'arrière-plan, sont forcément superposés. Malheureusement, ni le modèle des ressources multiples, ni le principe de proximité-compatibilité n'apportent de réponse à cette question, posée dans le chapitre suivant.

Chapitre 5 :
La performance à un jeu vidéo est plus dégradée par la complexité structurale de l'arrière-plan que par son mouvement

1. Introduction

Les expériences du chapitre précédent ont montré que le positionnement des informations contextuelles avait une influence sur la performance des joueurs dans le cadre d'un jeu dont les éléments sont en mouvement. Plusieurs théories, comme la théorie des ressources multiples (Wickens, 2002, 2008) aident à la conception de ce type d'interfaces visuelles. Cependant, dans les expériences précédentes, les ressources attentionnelles étaient partagées entre des sources d'informations spatialement séparées. La scène d'action pouvait être traitée en vision fovéale, et l'information contextuelle en vision périphérique.

L'objectif de ce chapitre est de comprendre le fonctionnement du partage attentionnel entre deux ou plusieurs stimuli visuels traités simultanément par la vision fovéale, comme des éléments avec

lesquels le joueur doit interagir et leur arrière-plan, en fonction des caractéristiques de cet arrière-plan.

Plus précisément, le but des expériences était d'étudier l'influence de deux mouvements typiques des arrière-plans d'environnements virtuels sur la performance des joueurs dans un jeu vidéo. Dans les jeux vidéo, l'information de mouvement est souvent perçue par les joueurs à travers une caméra virtuelle (Yannakakis, Martínez, & Jhala, 2010). Le mouvement de cette caméra peut être soit directement contrôlé par le joueur, soit être une conséquence du mouvement de son avatar. Dans tous les cas, ces mouvements de caméra impliquent un mouvement de l'ensemble de la scène visuelle, et plus particulièrement de l'arrière-plan. En général, deux types de mouvements se rencontrent dans les jeux. Les mouvements « en profondeur » vers l'avant ou vers l'arrière sont simulés par une expansion, ou inversement une rétraction radiale de l'ensemble de l'arrière-plan. Les mouvements latéraux (le plus souvent orientés de la droite vers la gauche) sont simulés par un mouvement d'ensemble de l'arrière-plan dans la direction opposée[2].

A chaque type de mouvement est associée une structuration visuelle particulière de l'arrière-plan, dont le mouvement donne l'illusion d'un mouvement cohérent du joueur, ou de son avatar, dans l'environnement virtuel. La structure visuelle des scènes en mouvement latéral est plutôt homogène. Tous les éléments de l'arrière-plan apparaissent à la droite de l'écran, par exemple, pour aller vers la gauche. La structure visuelle des scènes en mouvement radial est plus hétérogène. Les objets de l'arrière-plan apparaissent (ou disparaissent) au centre de l'écran (point de fuite) pour aller vers (ou en venant de) l'un de ses quatre côtés. Ce mouvement s'appuie sur des lignes de fuite qui séparent visuellement les objets selon le côté de l'écran où ils vont aboutir, ou d'où ils viennent. Par conséquent, la structure visuelle qui génère un mouvement radial comporte de fait des éléments

[2] Exemples de jeux vidéo récents à succès affichant un mouvement radial: jeux de courses (e.g., *F1 2011*, Codemasters, 2011) ou jeux de tir à la première personne (e.g., *Call of Duty: Modern Warfare 3*, Activision, 2011). Exemples de jeux vidéo récents à succès affichant un mouvement latéral: jeux de plateforme (e.g., *New Super Mario Bros Wii*, Nintendo, 2009) ou de « shoot them up » (e.g., *Deathsmiles*, Cave, 2007). Voir aussi la Figure 1.

fixes saillants qui s'ajoutent aux éléments qui composent par ailleurs l'arrière-plan. Ces éléments augmentent la complexité visuelle de l'arrière-plan de structure radiale par rapport à l'arrière-plan associé à un mouvement latéral, qui n'affiche pas de lignes ou de points de fuite saillants additionnels.

Plusieurs études, décrites dans le chapitre 2, ont montré des effets du mouvement et de la complexité des arrière-plans sur la performance ou les mouvements du regard dans diverses tâches. Les mouvements d'ensemble de la scène visuelle induisent des mouvements oculaires réflexes, comme l'OKN (Ilg, 1997), qui perturbent le guidage attentionnel vers les objets superposés à la scène en mouvement et diminuent les performances d'utilisation de ces objets (Harrison et al., 2010 ; Kaminiarz et al., 2007 ; Tozzi et al., 2007). Cependant, l'OKN peut être volontairement annulé quand des éléments fixes superposés sont présents (Ilg, 1997 ; Pola et al., 1992 ; Rubinstein & Abel, 2011 ; Wyatt & Pola, 1984). La variation de la densité d'informations visuelles présente dans l'arrière-plan, qui est une des causes de complexité visuelle, est aussi un facteur de perturbation du guidage attentionnel vers les objets de l'interface. La présence de points de saillance, comme des différences de couleur, de contraste ou de luminance, dans l'arrière-plan d'une scène rend le guidage vers les éléments fixes superposés moins efficace et diminue la performance (Jie & Clark, 2008 ; Wolfe et al., 2002). Une plus grande densité des informations visuelles provoquerait aussi une augmentation de la durée moyenne des fixations oculaires, du fait de la plus grande quantité d'informations à traiter à chaque fixation (Henderson et al., 2009). Cependant, l'interaction entre la complexité de l'arrière-plan et le fait qu'il soit ou non en mouvement n'a pas encore été étudiée à notre connaissance.

Les expériences 3a et 3b avaient pour but d'évaluer l'impact de mouvements latéral et radial des arrière-plans, et des structurations visuelles associées à ces mouvements sur la performance et les mouvements du regard des joueurs dans une tâche de recherche visuelle. L'influence de ces mouvements et de la structuration visuelle a aussi été étudiée pendant une tâche plus simple (en termes de demande attentionnelle) de jeu vidéo (expérience 4). Les deux types de mouvements de

l'arrière-plan ont été générés en utilisant un même motif visuel simple, relativement dépourvu d'information, pour évaluer le mieux possible les effets de la structuration de l'arrière-plan qui provoque l'illusion de chaque type de mouvement pour le participant. Les arrière-plans ont été conçus pour déclencher un OKN quand ils sont en mouvement, et potentiellement générer une perception de déplacement de soi de la gauche vers la droite (vection latérale) ou vers l'avant (vection radiale).

2. Expérience 3a : Tâche de recherche visuelle

Initialement, l'expérience 3a avait uniquement pour but d'étudier l'influence de mouvements apparents latéral et radial de l'arrière-plan sur la performance d'individus dans une tâche typique en psychologie : la recherche visuelle (décrite dans le chapitre 2). Même si la recherche visuelle d'une cible immobile parmi des objets immobiles présentés sur un arrière-plan en mouvement n'a pas encore été étudiée, plusieurs expériences de recherche visuelle d'une cible immobile ou en mouvement parmi des distracteurs en mouvement ont apporté des données intéressantes (Royden, Wolfe, & Klempen, 2001 ; Van Loon, Hooge, & Van den Berg, 2003). Van Loon et al. ont comparé la recherche visuelle dans des structures de points en mouvement d'ensemble radial (qui simulent un déplacement de soi vers l'avant) et dans des structures de traits immobiles visuellement similaires, qui reproduisaient les déplacements des points en expansion radiale. Dans le premier cas, les participants devaient trouver un point cible qui ne suivait pas le mouvement d'expansion radial cohérent des autres points de la scène. Dans le second cas, ils devaient trouver un trait cible immobile qui ne suivait pas l'orientation radiale cohérente des autres lignes. Les auteurs n'ont pas trouvé de différence significative dans les taux de bonnes réponses, mais ont décrit des stratégies de recherche différentes dans les deux scènes. Dans la scène en mouvement, les participants faisaient moins de saccades et des premières fixations plus longues que dans la scène statique. Royden et al. (2001) ont quant à eux étudié la recherche visuelle d'un rond immobile parmi des ronds en

mouvement. Ils ont montré que la recherche était plus efficace quand le mouvement d'ensemble des ronds était homogène (simulant une expansion radiale) que quand leur mouvement était aléatoire. Dans ces deux études, le repérage de la cible était facilité par sa grande dissimilitude avec les distracteurs.

L'expérience 3a a été construite pour comparer les effets sur la performance de recherche visuelle d'arrière-plans en mouvement apparent latéral ou radial et d'un arrière-plan statique. Cet arrière-plan statique utilisé comme référence portait le motif visuel le plus simple, celui dont le mouvement donnait l'illusion d'un déplacement latéral du participant. Les participants devaient détecter le plus rapidement possible sur ces différents arrière-plans la présence ou l'absence d'une cible parmi plusieurs distracteurs. D'un point de vue théorique, le chapitre 2 suggère que les mouvements d'ensemble des arrière-plans dégradent les performances de perception visuelle pour des tâches très simples, telles que repérer une cible ponctuelle affichée de manière transitoire (Kaminiarz et al., 2007 ; Tozzi et al., 2007), mais pas pour une tâche de lecture demandant une attention soutenue et où le texte est en permanence surimposé sur l'arrière-plan en mouvement (Menozzi & Koga, 2004). Dans la recherche visuelle, les individus doivent prêter attention à de multiples éléments d'une même scène pour décider si la cible est présente ou absente (Thornton & Gilden, 2007 ; Wolfe, 1998), en passant très rapidement de l'un à l'autre. Bien que la littérature ne donne que peu d'indications précises sur le niveau d'attention que requiert une tâche de recherche visuelle simple comparativement aux tâches citées plus haut, il est raisonnable de penser qu'il est intermédiaire entre celui requis pour localiser le plus rapidement possible une cible ponctuelle et celui requis par la lecture. Par ailleurs, les résultats obtenus en recherche visuelle d'une cible parmi des distracteurs en mouvement suggèrent que la présence de mouvements dans la scène visuelle diminue l'efficacité de la recherche (Royden et al., 2001), ou au moins modifie les mouvements du regard associés (Van Loon et al., 2003). L'hypothèse principale de l'expérience était donc que la performance est plus faible sur les deux arrière-plans en mouvement que sur l'arrière-plan statique, du fait notamment de

l'OKN que provoquent normalement ces mouvements. Cet OKN se manifeste notamment sous la forme d'une déviation continue du regard vers l'origine du mouvement (Ilg, 1997).

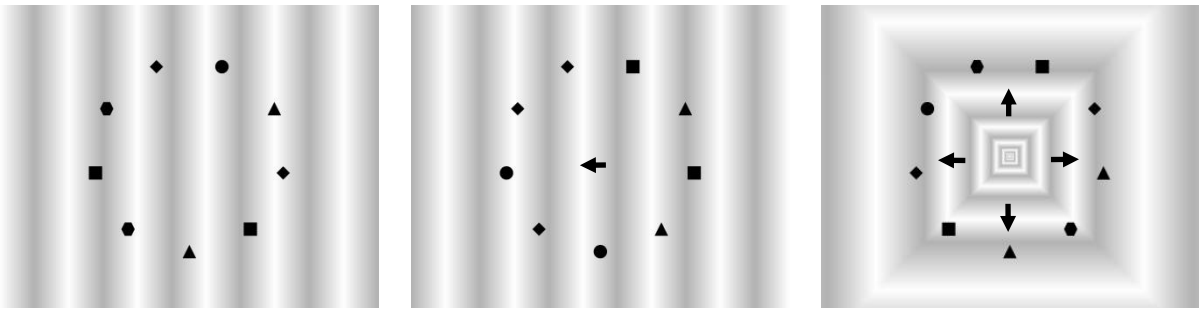
2.1. Méthode

Participants. Vingt-huit volontaires (23 femmes, 5 hommes) ont pris part à l'expérience. Leur âge moyen était de 19,4 ans (*ET* = 1,1) et leur durée d'études moyenne de 13,0 ans (*ET*=0,2). Tous les participants avaient une vision normale ou corrigée. Vingt-trois d'entre eux étaient droitiers, quatre étaient gauchers et un ambidextre. Quinze participants avaient pour œil dominant l'œil droit et treize l'œil gauche. Onze participants étaient des joueurs réguliers, qui avaient joué au moins une fois par semaine lors des 12 derniers mois, quatre étaient d'anciens joueurs réguliers, qui avaient joué moins d'une fois par semaine lors des 12 derniers mois mais au moins une fois par semaine auparavant, et treize étaient des joueurs occasionnels, qui avaient joué moins d'une fois par semaine tout au long de leur vie (voir le questionnaire pré-expérimental en annexe 3).

Appareillage. Un oculomètre Tobii T120 laissant la tête libre et non invasif, qui ressemblait à un écran d'ordinateur TFT 17" (résolution 1280x1024 pixels), a été utilisé pour imiter le plus possible une interaction naturelle avec l'environnement virtuel. L'oculomètre était contrôlé par un ordinateur qui recueillait les données et contrôlait l'exécution de la tâche. Les positions du regard étaient échantillonnées à une fréquence de 120 Hz avec une précision de 0,5 degré d'angle visuel (environ 5 mm sur l'écran). Les mouvements oculaires étaient analysés avec le logiciel « Tobii Studio 2.0.6 ». Les fixations oculaires étaient définies comme toute période où le regard restait pendant 58 ms (7 points successifs à 120 Hz) ou plus dans une zone de 30 pixels de diamètre (8 mm ou 0,8 degré d'angle visuel).

Matériel. Le matériel a été créé avec le logiciel « Adobe Director 11 ». Trois arrière-plans différents ont été construits (figure 7). Ils étaient basés sur un motif déjà utilisé par Tozzi et al. (2007), et constitué de rayures grises (RVB: 179, 179, 179) et blanches (RVB: 255, 255, 255) alternant

progressivement selon un profil sinusoïdal de 0,17 cycle/deg (0,17 cycle/cm). Le motif visuel utilisé pour générer la perception d'un mouvement latéral était composé uniquement de rayures verticales. Le motif visuel utilisé pour générer le mouvement radial affichait ce même motif de largeur croissante dans les quatre quadrants allant du centre à chacun des côtés de l'écran. Les rayures étaient horizontales en haut et en bas de l'écran, mais verticales à gauche et à droite. Du fait de cette disposition du motif, il y avait plus de points fixes saillants (point et lignes de fuite du mouvement apparent) dans l'arrière-plan de structure radiale que dans celui de structure latérale. L'arrière-plan de structure radiale était donc plus complexe que celui de structure latérale. Dans l'arrière-plan latéral en mouvement, les rayures se déplaçaient vers la gauche à la vitesse de 121 mm/s (environ 12 degrés/s). Dans l'arrière-plan radial, les mêmes rayures se déplaçaient dans quatre directions divergentes à partir du centre de l'écran, de manière à simuler un déplacement dans un couloir dans lequel les murs, le plafond et le sol afficheraient le motif à rayures. La vitesse de déplacement des rayures était de 121 mm/s dans chaque direction. Le troisième arrière-plan reproduisait le motif visuel de l'arrière-plan de structure latérale, mais restait statique.



Latéral / Statique Latéral / En mouvement Radial / En mouvement

Figure 7. Les trois arrière-plans utilisés dans l'expérience 3a avec des exemples de champs de recherche. La structure radiale a été réutilisée dans l'expérience 3b, où elle pouvait être soit statique, soit en mouvement. Les flèches noires indiquent les directions du mouvement de l'arrière-plan, mais n'étaient pas visibles des participants.

Dix formes géométriques simples ont été utilisées comme cibles et distracteurs dans la tâche de recherche visuelle (annexe 4). Il s'agissait d'un carré, d'un rond, d'un triangle, d'un hexagone, et d'un losange de couleur noire qui étaient soit pleins, soit vides. Leur taille était de 12 mm (environ 1,2 degré d'angle visuel). Le champ de recherche visuelle était toujours composé de neuf formes

disposées en un cercle virtuel de 170 mm (environ 17 degrés d'angle visuel) de diamètre (figure 7). Les formes étaient disposées de manière régulière sur le cercle, à une distance de 61 mm (environ 6 degrés d'angle visuel) les unes des autres.

Plan expérimental et procédure. Les participants devaient détecter si une forme géométrique particulière était présente ou absente parmi les neuf formes présentées sur l'écran du Tobii T120. Le type d'arrière-plan et la présence ou l'absence de la cible étaient manipulés en intra-participant. Au début de chaque bloc d'essais, une forme cible était présentée au centre de l'écran pendant 5 secondes. Un arrière-plan de structure latérale statique ou en mouvement ou un arrière-plan de structure radiale en mouvement était ensuite affiché seul pendant 30 secondes (pour assurer une activation maximale de l'OKN). Un cadre bleu de 74 mm de côté (environ 7,4 degrés d'angle visuel) était ensuite affiché au centre de l'écran pendant 3 secondes pour orienter le regard des participants vers le centre de l'écran sans annuler l'OKN. Les participants devaient garder leur regard le plus possible à l'intérieur du cadre. Le cadre disparaissait et l'arrière-plan était à nouveau présenté seul pendant 1 seconde. Les neuf formes apparaissaient alors, et les participants devaient dire si la cible était présente ou absente aussi vite que possible en appuyant sur une des deux touches « présent » ou « absent ». Ces touches correspondaient aux touches « 7 » et « 9 » d'un pavé numérique pour la moitié des participants, et aux touches « 9 » et « 7 » pour l'autre moitié.

Trente blocs de quatre essais étaient présentés à chaque participant, avec une courte pause entre chaque bloc. Les quatre essais d'un même bloc étaient réalisés sur le même arrière-plan et avec la même forme cible. Quand l'arrière-plan était en mouvement, le mouvement était continu et persistait pendant l'ensemble du bloc. La position de la forme cible à l'intérieur du champ de recherche était aléatoire et différente à chaque essai. Si la cible était une forme vide, toutes les formes distractrices étaient également vides, si la cible était pleine, tous les distracteurs étaient pleins. Quand la cible (e.g., un carré) était présente, elle était présentée parmi deux exemplaires de chacune des quatre autres formes (e.g., deux losanges, deux ronds, deux triangles et deux

hexagones). Quand la cible était absente, une des quatre autres formes était répétée trois fois dans le champ de recherche. L'ordre de présentation des différents arrière-plans et des dix formes cibles possibles était aléatoire, mais chacune des dix formes cibles était recherchée quatre fois (i.e., pendant un bloc d'essais) sur chaque arrière-plan par chaque participant.

Mesures dépendantes. Le temps écoulé entre l'apparition des formes et la réponse du participant a été utilisé pour évaluer la performance de recherche. Pour tenter d'expliquer la performance, les mouvements oculaires des participants ont été analysés en utilisant deux mesures : la durée moyenne des fixations oculaires, et la position (définie par les coordonnées horizontale et verticale X, Y) de la fixation initiale enregistrée sur l'écran de l'oculomètre pendant ou juste avant l'apparition du champ de recherche. Cette dernière mesure avait pour but d'évaluer l'influence de chaque arrière-plan sur la déviation initiale moyenne du regard (Ilg, 1997). Toutes ces variables ont été analysées en utilisant des ANOVA pour mesures répétées avec le type d'arrière-plan et la présence/absence de la cible dans le champ de recherche en tant que facteurs intra-participant. Quand le test de sphéricité de Mauchley était significatif, une correction de Greenhouse-Geisser a été appliquée.

2.2. Résultats

Les réponses incorrectes, prématurées (temps de réponse inférieurs à 100 ms) ou trop tardives (temps de réponse supérieurs à 3000 ms) ont été exclues de l'analyse. Le nombre d'essais exclus représentait en moyenne 5,8% du nombre total d'essais par participant (min = 0.8%, max = 13.3%). Les données de temps de réponse ont été transformées logarithmiquement avant que les ANOVA ne soient réalisées. Pour la clarté de la présentation, ce sont les moyennes non-transformées qui sont indiquées dans le texte et le tableau.

Temps de réponse. Comme indiqué dans le tableau 3, le type d'arrière-plan avait une influence sur les temps de réponse, $F(2, 54) = 8{,}80$, $p < 0{,}001$, $\eta^2_p = 0{,}25$. Les comparaisons planifiées ont montré que les temps de réponse étaient plus longs sur l'arrière-plan radial en mouvement que sur l'arrière-

plan latéral, qu'il soit statique, $F(1, 27) = 8{,}72$, $p < 0{,}01$, ou en mouvement, $F(1, 27) = 15{,}77$, $p < 0{,}001$. Les temps de réponse n'étaient pas significativement différents entre les arrière-plans latéraux statique et en mouvement, $F(1, 27) = 1{,}29$, $p = 0{,}27$. La présence/absence de la cible avait aussi un effet sur les temps de réponse, $F(1, 27) = 447{,}13$, $p < 0{,}001$, $\eta^2_p = 0{,}94$. Les temps de réponse étaient plus longs quand la cible était absente (*M* = 1591 ms, *ET* = 291) que quand elle était présente (*M* = 1149 ms, *ET* = 195). L'interaction entre ces deux facteurs n'était pas significative, $F(2, 54) = 2{,}46$, $p = 0{,}09$.

Tableau 3. *Temps de réponse moyens (en ms) dans les expériences 3a et 3b.*

Arrière-plan	**Latéral / Statique**	**Latéral / En mouvement**	**Radial / Statique**	**Radial / En mouvement**
Expérience 3a	1361 (348)	1336 (337)	-	1413 (311)
Expérience 3b	-	-	1445 (276)	1486 (291)

Note. Les écarts-type correspondants sont donnés entre parenthèses.

Durée moyenne des fixations oculaires. Ni le type d'arrière-plan, $F(2, 54) < 1$, ni la présence/absence de la cible, $F(1, 27) = 3{,}58$, $p = 0{,}07$, n'avaient d'effet significatif sur la durée moyenne des fixations oculaires. L'interaction entre ces deux facteurs n'était pas non plus significative, $F(2, 54) = 1{,}70$, $p = 0{,}19$.

Fixation initiale. Comme l'indique la figure 8, le type d'arrière-plan avait un effet sur la coordonnée horizontale (X) de la fixation initiale, $F(2, 54) = 51{,}68$, $p < 0{,}001$, $\eta^2_p = 0{,}66$. Comme attendu, les comparaisons planifiées ont montré que la fixation initiale était positionnée plus vers la droite sur l'arrière-plan latéral en mouvement (*M* = 678 pixels, *ET* = 43) que sur l'arrière-plan latéral statique (*M* = 620 pixels, *ET* = 31), $F(1, 27) = 72{,}02$, $p < 0{,}001$, et sur l'arrière-plan radial en mouvement (*M* = 626 pixels, *ET* = 22), $F(1, 27) = 58{,}26$, $p < 0{,}001$. Le type d'arrière-plan avait aussi une influence sur la coordonnée verticale (Y) de cette fixation, $F(2, 54) = 31{,}41$, $p < 0{,}001$, $\eta^2_p = 0{,}54$. La fixation était plus basse sur l'arrière-plan radial en mouvement (*M* = 489 pixels, *ET* = 49) que sur les arrière-plans latéraux statique (*M* = 451 pixels, *ET* = 62), $F(1, 27) = 32{,}51$, $p < 0{,}001$, ou en mouvement (*M* = 445 pixels, *ET* = 61), $F(1, 27) = 41{,}17$, $p < 0{,}001$. Aucune autre comparaison n'était significative.

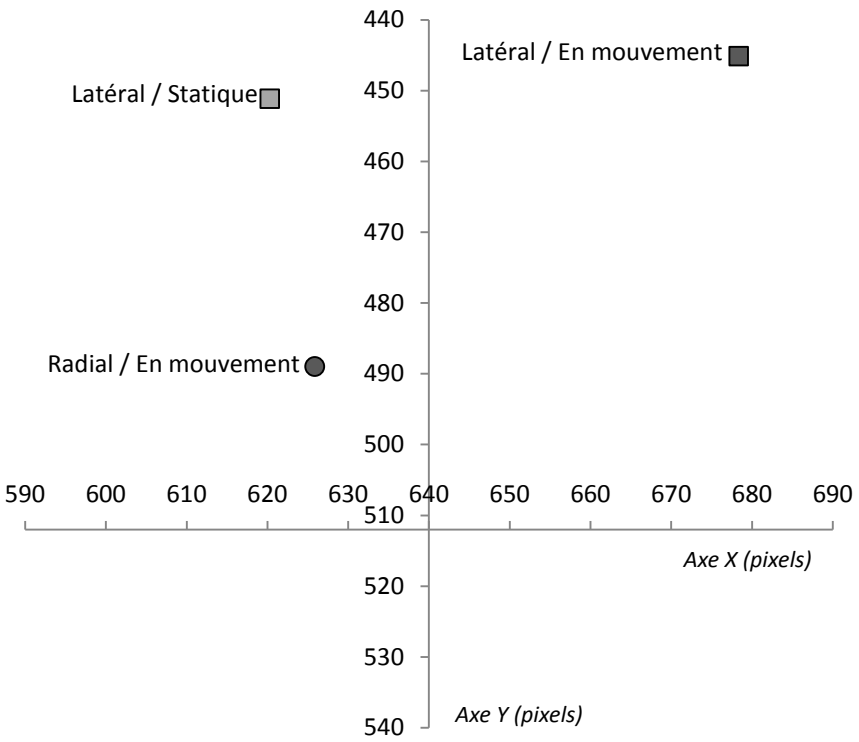


Figure 8. Coordonnées moyennes de la fixation initiale (en pixels) selon le type d'arrière-plan utilisé dans l'expérience 3a.
Le graphique montre un zoom sur le centre de l'écran de l'oculomètre. Les axes correspondent aux axes centraux de l'écran (coordonnées de l'origine X = 640, Y = 512).

2.3. Discussion

L'hypothèse principale n'a pas été vérifiée. Quand l'arrière-plan était en mouvement, la performance n'était plus faible que si le mouvement et la structure de l'arrière-plan étaient de type radial. Cette absence d'effet du mouvement de l'arrière-plan de structure latérale sur la recherche visuelle semble contredire les résultats de Harrison et al. (2010), Kaminiarz et al. (2007) et Tozzi et al. (2007), qui ont montré qu'un arrière-plan latéral en mouvement diminuait la performance dans diverses tâches simples de perception visuelle, mais sont en accord avec ceux de Menozzi et Koga (2004) qui ont montré que ce mouvement ne diminuait pas la performance de lecture d'un texte superposé. Ce résultat suggère donc que le niveau d'attention aux éléments fixes requis par la tâche de recherche visuelle permet d'éviter la dégradation des performances induite par un mouvement apparent latéral de l'arrière-plan.

Pourtant, les deux arrière-plans en mouvement semblaient bien induire un OKN, visible à l'observation des tracés des mouvements oculaires, même s'il était difficile à quantifier du fait de la faible résolution temporelle de l'oculomètre. Cet OKN provoquait, avant l'apparition des formes géométriques, une déviation de la position moyenne initiale du regard des participants vers la droite de l'écran sur l'arrière-plan en mouvement latéral, et vers le centre de l'écran sur l'arrière-plan en mouvement radial. Autrement dit, le regard des participants était dévié vers l'origine du mouvement de l'arrière-plan. Ces données confirment celles de précédentes études (Ilg, 1997 ; Kaminiarz et al., 2007 ; Tozzi et al., 2007), qui ont montré qu'un arrière-plan structuré en mouvement déclenche un OKN pendant lequel les individus ont tendance à regarder vers le lieu apparent de renouvellement de l'arrière-plan. Pendant la tâche de recherche visuelle, l'OKN déclenché par le mouvement latéral de l'arrière-plan aurait une influence trop faible pour être observée, peut-être du fait des possibilités de suppression de l'OKN offertes par les formes géométriques que doit fixer l'observateur (Ilg, 1997 ; Pola et al., 1992 ; Rubinstein & Abel, 2011 ; Wyatt & Pola, 1984).

Les résultats ont montré que seul l'arrière-plan de structure radiale en mouvement, qui suggère un mouvement de l'observateur vers l'avant, induisait une augmentation du temps moyen mis pour détecter la présence/ absence de la cible par rapport à l'arrière-plan statique de structure latérale. Une première interprétation de ce résultat serait qu'un mouvement radial de l'arrière-plan dégrade plus la performance de recherche visuelle qu'un mouvement latéral.

Toutefois, une interprétation alternative possible serait que ce n'est pas le mouvement radial lui-même qui provoque la dégradation, mais plutôt la structure visuelle radiale de l'arrière-plan utilisé pour donner l'illusion du mouvement. En effet, la présence dans l'arrière-plan de structure radiale de lignes et d'un point de fuite saillants rend cet arrière-plan plus complexe que l'arrière-plan composé uniquement de rayures verticales utilisé pour le mouvement latéral. Or, comme l'indique la partie théorique de la thèse (chapitre 2), plusieurs auteurs ont montré que les performances de recherche visuelle dans des tâches typiques des jeux vidéo sont plus faibles quand la complexité de l'arrière-

plan, même statique, augmente (Jie & Clark, 2008 ; Wolfe et al., 2002). Il est donc possible que ce soit la complexité supérieure de l'arrière-plan radial par rapport à l'arrière-plan latéral qui provoque la dégradation de la performance, indépendamment du fait que l'arrière-plan soit ou non en mouvement. Si c'est le cas cependant, cette augmentation de complexité n'est apparemment pas suffisante pour induire l'augmentation significative de la durée des fixations oculaires sur l'arrière-plan de structure radiale qui aurait pu être observée d'après Henderson et al. (2009). En effet, le type d'arrière-plan utilisé n'avait aucun effet significatif sur ce paramètre. Il est possible également que le fait que l'arrière-plan radial soit en mouvement rende invisible cette augmentation de durée dans le contexte de la suppression de l'OKN.

Un des moyens de trancher entre ces deux interprétations était de déterminer si la dégradation des performances de recherche visuelle observée sur l'arrière-plan de structure radiale en mouvement par rapport à l'arrière-plan de structure latérale (statique ou en mouvement) persistait ou non lorsque l'arrière-plan de structure radiale était utilisé en version statique. L'expérience 3b a été construite dans ce but, pour comparer les effets des arrière-plans de structure radiale statique et en mouvement sur la performance de recherche visuelle. L'hypothèse était que la performance est plus faible sur l'arrière-plan radial en mouvement que sur l'arrière-plan radial statique, du fait notamment de l'OKN généré par le mouvement (Ilg, 1997).

3. Expérience 3b : Tâche de recherche visuelle et structure radiale

3.1. Méthode

Participants. Quinze volontaires (13 femmes, 2 hommes) ont pris part à l'expérience. Leur âge moyen était de 18,9 ans (*ET* = 1,3) et leur durée moyenne d'études de 13,0 ans (*ET* = 0,0). Tous les

participants avaient une vision normale ou corrigée et étaient droitiers. Dix participants avaient l'œil dominant droit et cinq l'œil dominant gauche. Sept participants étaient des joueurs réguliers et huit étaient des joueurs occasionnels.

Appareillage et matériel. L'appareillage et la tâche de recherche visuelle étaient les mêmes que ceux de l'expérience 3a. Deux types d'arrière-plans ont été construits à partir du motif à rayures de structure radiale : une version en mouvement (la même que dans l'expérience 3a) et une version statique (figure 7).

Plan expérimental et procédure. Le mouvement de l'arrière-plan et la présence/absence de la cible étaient manipulés en intra-participant. Vingt blocs de quatre essais étaient présentés à chaque participant, avec une courte pause entre chaque bloc. Les autres détails de procédure étaient identiques à ceux de l'expérience 3a.

Mesures dépendantes. Les mêmes mesures dépendantes que celles de l'expérience 3a ont été utilisées. Toutes ont été analysées avec des ANOVA pour mesures répétées avec le mouvement de l'arrière-plan et la présence/absence de la cible comme facteurs intra-participants.

3.2. Résultats

Comme pour l'expérience 3a, les réponses incorrectes, prématurées et tardives ont été exclues de l'analyse. Deux participants ont été exclus de l'analyse parce que plus de 15% de leurs essais étaient exclus avec ces critères. Pour les treize autres participants, le nombre d'essais exclus représentaient en moyenne 4,6% du nombre total d'essais par participant (min = 2.8%, max = 10,0%). Les données de temps de réponse ont été transformées logarithmiquement avant que les ANOVA ne soient réalisées. Pour la clarté de la présentation, ce sont les moyennes non-transformées qui sont indiquées dans le texte et le tableau.

Temps de réponse. Comme l'indique le tableau 3, le mouvement de l'arrière-plan n'avait pas d'influence significative sur les temps de réponse, $F(1, 12) = 1{,}90$, $p = 0{,}19$. La présence/absence de la cible avait un effet sur les temps de réponse, $F(1, 12) = 102{,}8$, $p < 0{,}001$, $\eta^2_p = 0{,}90$, qui étaient plus longs quand la cible était absente (*M* = 1696 ms, *ET* = 178) que quand elle était présente (*M* = 1236 ms, *ET* = 143). L'interaction entre ces deux facteurs n'était pas significative, $F(1, 12) < 1$.

Durée moyenne des fixations oculaires. La présence/absence de la cible avait un effet sur la durée moyenne des fixations, $F(1, 12) = 27{,}8$, $p < 0{,}001$, $\eta^2_p = 0{,}70$, qui étaient plus longues quand la cible était présente (*M* = 163 ms, *ET* = 33) que quand elle était absente (*M* = 150 ms, *ET* = 28). Le mouvement de l'arrière-plan n'avait pas d'effet significatif sur la durée des fixations oculaires, $F(1, 12) < 1$, et l'interaction entre ces deux facteurs n'était pas significative, $F(1, 12) = 1{,}03$, $p = 0{,}33$.

Fixation initiale. Comme le montre la figure 9, le mouvement de l'arrière-plan n'avait pas d'influence significative sur les coordonnées horizontale (X), $F(1, 12) < 1$, et verticale (Y), $F(1, 12) < 1$, de la fixation initiale. Aucune autre comparaison n'était significative.

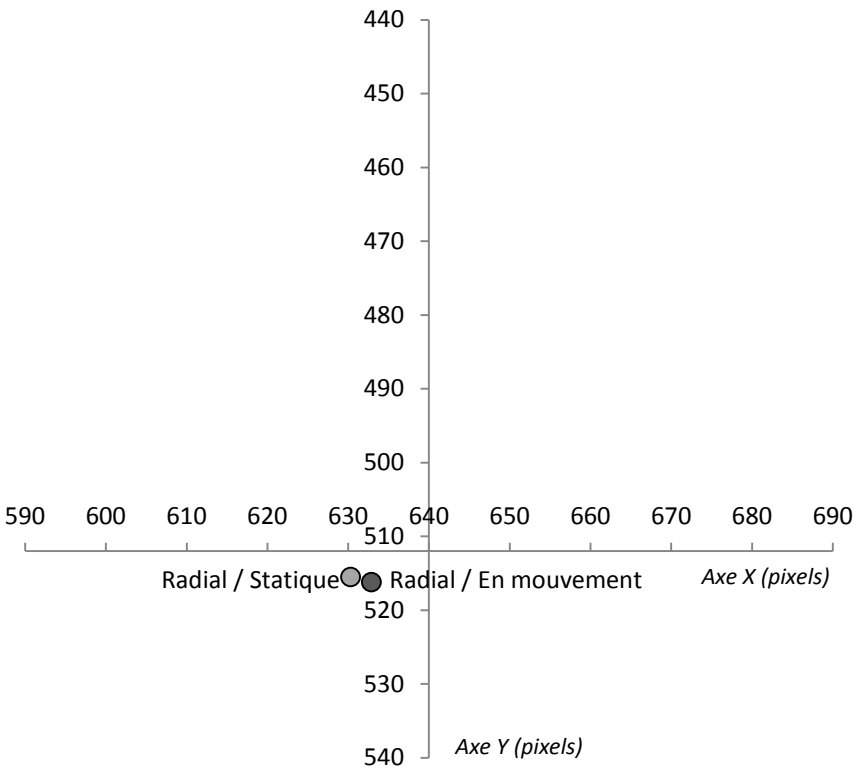


Figure 9. Coordonnées moyennes de la fixation initiale (en pixels) selon le type d'arrière-plan utilisé dans l'expérience 3b.
Le graphique montre un zoom sur le centre de l'écran de l'oculomètre. Les axes correspondent aux axes centraux de l'écran (coordonnées de l'origine X = 640, Y = 512).

3.3. Discussion

L'hypothèse n'a pas été vérifiée. Les résultats n'ont pas montré de différence de performance de recherche visuelle selon que l'arrière-plan de structure radiale était statique ou en mouvement. Ces données indiquent que l'effet de l'arrière-plan de structure radiale en mouvement obtenu dans l'expérience 3a était probablement lié à la plus grande complexité de l'arrière-plan utilisé par rapport à l'arrière-plan utilisé pour générer la sensation de mouvement latéral, et non au mouvement de l'arrière-plan. Ce résultat est similaire à celui de Van Loon et al. (2003), qui n'ont pas montré de différence significative de taux de bonnes réponses entre la recherche d'un point en mouvement isolé dans un ensemble de points en expansion radiale cohérente et la recherche d'un trait correspondant dans un ensemble de traits immobiles qui imitait le pattern de points en expansion. Malheureusement, les temps de réponse n'ont pas été mesurés dans cette étude antérieure. Par

ailleurs, les résultats de l'expérience 3b montrent également une absence d'effet significatif du mouvement de l'arrière-plan radial sur la position moyenne de la fixation initiale. La déviation vers le centre de l'écran de la fixation initiale observée dans l'expérience 3a était donc due à une attirance vers le point saillant central de la structure visuelle utilisée, et non vers l'origine centrale du mouvement de l'arrière-plan.

3.4. Discussion générale des expériences 3a et 3b

3.4.1. La tâche de recherche visuelle

Comme dans toutes les tâches de recherche visuelle (Thornton & Gilden, 2007 ; Wolfe, 1998), les temps de réponse étaient plus longs quand la cible était absente que quand elle était présente. En effet, les participants doivent balayer l'ensemble des objets avant de pouvoir conclure sur l'absence de la cible, alors qu'ils peuvent dire si la cible est présente dès qu'ils l'ont détectée. De plus, les fixations oculaires étaient plus longues quand la cible était présente dans l'expérience 3b, et tendaient à l'être dans l'expérience 3a. Ceci est probablement dû à la fixation finale sur la cible, qui est généralement plus longue que les autres. En résumé, les résultats de la tâche de recherche visuelle utilisée dans cette étude étaient comparables à ceux présentés dans la littérature.

3.4.2. Influence de la complexité structurale de l'arrière-plan sur la recherche visuelle

Les résultats des deux expériences suggèrent que la performance à la tâche de recherche visuelle est plus faible quand la complexité structurale de l'arrière-plan augmente. Par contre, la performance ne serait pas significativement influencée par la présence ou l'absence de mouvement de l'arrière-plan, quelle que soit la complexité de sa structure visuelle. Seule la complexité structurale de l'arrière-plan aurait un effet significatif sur les performances de recherche visuelle des participants. Les résultats montrent que même si un arrière-plan en mouvement latéral influence le déplacement du regard, la

performance n'est pas affectée si la tâche implique suffisamment d'objets fixes auxquels l'observateur doit porter une attention soutenue. En recherche visuelle, comme dans d'autres tâches complexes comme la lecture (Menozzi & Koga, 2004), les participants doivent en permanence observer ces objets, ce qui aboutit probablement à annuler l'OKN et/ou la vection (Ilg, 1997 ; Pola et al., 1992 ; Rubinstein & Abel, 2011 ; Wyatt & Pola, 1984) et contrecarrer l'influence négative qu'il peut avoir sur la tâche en cours (Menozzi & Koga, 2004 ; Seno et al., 2011).

Pour valider ce raisonnement, il était intéressant de vérifier l'impact des deux types de mouvement de l'arrière-plan et des structures visuelles associées sur une tâche plus simple en termes de demandes attentionnelles. C'est cette question qui a fait l'objet de l'expérience 4.

4. Expérience 4 : Tâche de tir

L'expérience 4 avait pour but d'évaluer l'influence du mouvement et de la complexité structurale de l'arrière-plan sur la performance de joueurs à un jeu vidéo de tir, plus simple en termes de demandes attentionnelles que la recherche visuelle. Les participants devaient tirer le plus rapidement possible sur des cibles pour obtenir le meilleur score sur plusieurs parties. Toutes les modalités manipulées dans les expériences 3a et 3b ont été regroupées en une seule expérience.

Les deux premières hypothèses étaient basées sur les résultats des deux expériences précédentes. Comme la tâche de jeu vidéo utilisée dans l'expérience présente était a priori plus simple que la tâche de recherche visuelle, l'OKN ne devait pas pouvoir être aussi bien annulé pendant la tâche que dans l'expérience 3a. Aussi, la première hypothèse était que la performance est plus faible sur un arrière-plan en mouvement que sur un arrière-plan statique, mais seulement lorsque la structure de l'arrière-plan ne contient pas de points fixes saillants qui peuvent servir de support pour annuler l'OKN. En d'autres termes, le mouvement n'aurait un effet significatif que pour l'arrière-plan de structure latérale, qui permet de donner l'illusion d'un mouvement de droite à gauche. La deuxième

hypothèse était que la performance est plus faible quand la structure visuelle de l'arrière-plan est complexe et contient plusieurs points ou lignes fixes saillants (i.e., radiale) que quand elle est moins complexe (i.e., latérale), et ce indépendamment du caractère statique ou en mouvement de l'arrière-plan. Cette dégradation serait due à un traitement plus lent de l'information visuelle, reflété par une durée moyenne des fixations oculaires plus élevée (Henderson et al., 2009).

En complément de l'étude des effets de la complexité de la structure visuelle de l'arrière-plan sur la performance, l'effet de la présence ou non d'une structure visuelle en arrière-plan a également été testé. La troisième hypothèse était que la performance est plus faible quand un arrière-plan visuellement structuré, quel qu'il soit, est utilisé que quand l'arrière-plan est uni.

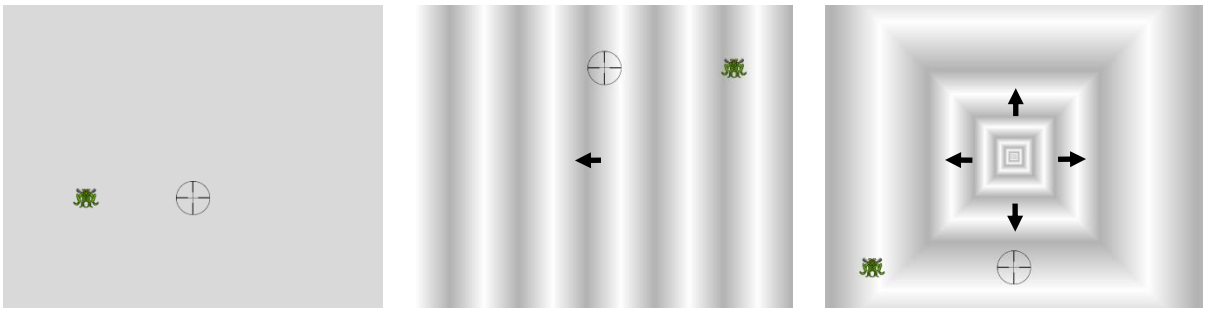
Enfin, un signal auditif était parfois présenté pendant le jeu pour tenter de réduire ou annuler ces effets de détérioration de la performance. Avant l'apparition de chaque cible à viser, ce signal indiquait approximativement sa position au participant. Wickens (2002, 2008) a montré que l'intégration de deux sources d'information mène à une meilleure performance lorsqu'elles peuvent être traitées par deux canaux attentionnels différents. La quatrième hypothèse était donc que ce signal auditif réduit la baisse de performance induite par les arrière-plans visuellement complexes et/ou en mouvement.

4.1. Méthode

Participants. Trente-huit volontaires (24 femmes, 14 hommes) ont pris part à l'expérience. Leur âge moyen était de 19,9 ans (*ET* = 2,8) et leur durée d'études moyenne de 13,6 ans (*ET* = 1,1). Tous les participants étaient de langue maternelle française et avaient une vision normale ou corrigée. Trente-six d'entre eux étaient droitiers, un était gaucher, et un était ambidextre. Vingt-deux participants avaient l'œil droit dominant et seize l'œil gauche. Dix-sept participants étaient des joueurs réguliers, qui avaient joué au moins une fois par semaine lors des 12 derniers mois, treize étaient d'anciens joueurs réguliers, qui avaient joué moins d'une fois par semaine lors des 12

derniers mois mais au moins une fois par semaine auparavant, et huit étaient des joueurs occasionnels, qui avaient joué moins d'une fois par semaine tout au long de leur vie.

Appareillage et matériel. L'appareillage était similaire à celui utilisé dans les expériences 3a et 3b. Le matériel a été créé avec le logiciel « Adobe Director 11 ». Cinq arrière-plans différents ont été utilisés (figure 10). Les deux arrière-plans visuellement structurés (latéral et radial) étaient les mêmes que ceux utilisés dans les expériences 3a et 3b. Ils pouvaient être statiques ou en mouvement, donnant l'illusion d'un mouvement du participant de droite à gauche ou vers l'avant. Le cinquième arrière-plan était un arrière-plan contrôle de couleur gris uni (RVB: 217, 217, 217). La cible du jeu de tir était un dessin de créature verte de 85x66 pixels (22x17 mm, 2,2 degrés). Le curseur était une représentation d'un viseur d'arme de 114 pixels de diamètre (30 mm, 3 degrés) (annexe 5). Enfin, aucune information contextuelle n'était affichée à l'écran. Deux indices sonores ont été construits à partir d'une voix de synthèse masculine. Ces indices prononçaient « gauche » ou « droite » sur une durée de 500 ms. Ils étaient diffusés par un haut-parleur disposé face au participant (sans biais de localisation spatiale).



Uni | Latéral / Statique ou En mouvement | Radial / Statique ou En mouvement

Figure 10. Les cinq arrière-plans utilisés dans l'expérience 4 avec des exemples de position initiale de cible et de curseur.
Les flèches noires indiquent les directions de mouvement de l'arrière-plan, mais n'étaient pas visibles sur les interfaces réelles.

Plan expérimental et procédure. La structure visuelle de l'arrière-plan (latéral ou radial) était manipulée en inter-participants. Les participants étaient assignés à un des deux types d'arrière-plan, en s'assurant de l'absence de différence significative dans la distribution de l'expertise du jeu vidéo

entre les groupes. La nature de l'arrière-plan était manipulée en intra-participant, en présentant soit l'arrière-plan structuré statique, soit l'arrière-plan structuré en mouvement, soit l'arrière-plan uni. L'arrière-plan uni était identique pour les deux conditions de structuration visuelle de l'arrière-plan. La présence/absence de l'indice sonore était aussi manipulée en intra-participant.

Une tâche de tir était utilisée. Au début de chaque essai, l'arrière-plan était affiché et un cadre bleu de 74 mm de côté (environ 7,4 degrés) était superposé au centre de l'écran pendant 1 seconde pour orienter le regard du participant vers le centre de l'écran sans annuler les effets du mouvement éventuel de l'arrière-plan sur les mouvements oculaires (OKN). Les participants devaient garder leur regard le plus possible à l'intérieur du cadre. L'indice sonore, « gauche » ou « droite », était présenté en même temps que le cadre. La cible et le curseur apparaissaient ensuite à l'écran sur une même ligne horizontale dont la coordonnée était déterminée aléatoirement. Le curseur apparaissait toujours sur l'axe vertical central de l'écran, et la cible aléatoirement sur la droite ou sur la gauche. La coordonnée horizontale d'apparition de la cible (entre le curseur et le côté de l'écran) était déterminée aléatoirement. Les participants devaient placer le curseur sur la cible à l'aide des touches « flèche gauche » et « flèche droite » du clavier, puis tirer sur la cible en appuyant sur la touche « espace ». Le participant pouvait déplacer le curseur et tirer autant de fois qu'il le souhaitait tant que la cible n'était pas touchée, autrement dit tant que le curseur et la cible n'étaient pas superposés au moment du tir. Quand la cible était touchée, un autre essai débutait. L'expérimentateur présentait cette procédure au participant comme un jeu vidéo dans lequel le participant devait toucher le plus d'ennemis possibles en un temps limité de 60 secondes. Chaque bloc d'essais de 60 secondes était présenté au participant comme une partie. Le score global de performance était donné à la fin de la passation.

Vingt-quatre parties étaient proposées à chaque participant avec une pause entre chaque. Tous les essais d'une partie étaient réalisés sur le même arrière-plan, qui restait affiché entre chaque essai. Quand l'arrière-plan était en mouvement, ce mouvement continu persistait pendant toute la partie.

De même, l'indice était présent ou absent sur l'ensemble de la partie. L'ordre de présentation des 24 parties (4 parties x [3x2] conditions différentes d'arrière-plans et d'indices sonores) à chaque participant était pseudo-aléatoire.

Mesures dépendantes. Les mesures dépendantes utilisées pour évaluer la performance au jeu étaient le nombre de cibles touchées (ou d'essais réalisés) par partie et le temps moyen mis pour toucher chaque cible (ou pour réaliser un essai). Les mouvements oculaires des participants étaient analysés en utilisant les mêmes mesures dépendantes que dans les expériences 3a et 3b, à savoir la durée moyenne des fixations oculaires et les coordonnées des fixations initiales avant l'apparition de la cible et du curseur. Toutes ces variables ont été analysées en utilisant des ANOVA pour mesures répétées avec la structure de l'arrière-plan en facteur inter-participants, et la nature de l'arrière-plan et la présence/absence de l'indice en facteurs intra-participants.

4.2. Résultats

Le nombre de cibles touchées par partie et le temps moyen par cible ont été transformés logarithmiquement avant que les ANOVA ne soient réalisées. Pour la clarté de la présentation, ce sont les moyennes non-transformées qui sont indiquées dans le texte et les graphiques.

Nombre de cibles touchées par partie. La nature de l'arrière-plan avait une influence sur le nombre de cibles touchées par partie, $F(2, 72) = 35{,}16$, $p < 0{,}001$, $\eta^2_p = 0{,}49$. Comme le montre la figure 11, cet effet interagissait avec la structure visuelle de l'arrière-plan, $F(2, 72) = 17{,}19$, $p < 0{,}001$, $\eta^2_p = 0{,}32$. Les comparaisons planifiées ont démontré qu'avec la structure radiale, le nombre de cibles touchées était plus faible quand l'arrière-plan était statique, $F(1, 36) = 51{,}83$, $p < 0{,}001$, ou en mouvement, $F(1, 36) = 78{,}59$, $p < 0{,}001$, que quand il était uni, mais n'était pas significativement différent entre les arrière-plans structurés statique et en mouvement, $F(1, 36) = 2{,}04$, $p = 0{,}16$. Avec la structure latérale, le nombre de cibles touchées était plus faible quand l'arrière-plan était en mouvement que quand il était statique, $F(1, 36) = 10{,}86$, $p < 0{,}01$, ou uni, $F(1, 36) = 7{,}64$, $p < 0{,}01$, mais n'était pas

significativement différent entre les arrière-plans structuré statique et uni, $F(1, 36) < 1$. L'effet principal de la structure de l'arrière-plan n'était pas significatif, $F(1, 36) = 1{,}52$, $p = 0{,}23$. L'indice sonore avait aussi un effet sur le nombre de cibles touchées par partie, $F(1, 36) = 118.26$, $p < 0{,}001$, $\eta^2_p = 0{,}77$. Le nombre de cibles touchées était supérieur quand l'indice était présenté ($M = 29{,}82$, $ET = 3{,}07$) qu'en son absence ($M = 27{,}57$, $ET = 2{,}88$). L'effet de l'indice sonore n'interagissait par ailleurs significativement ni avec la structure, $F(1, 36) < 1$, ni avec la nature de l'arrière-plan, $F(2, 72) < 1$. Aucune autre comparaison n'était significative.

Temps moyen pour toucher chaque cible. Les temps moyens par cible étaient liés aux nombres de cibles touchées par partie. Comme le montre la figure 12, la nature de l'arrière-plan avait une influence sur le temps moyen utilisé par cible, $F(2, 72) = 26{,}85$, $p < 0{,}001$, $\eta^2_p = 0{,}43$. Cependant, cet effet interagissait avec la structure visuelle de l'arrière-plan, $F(2, 72) = 15{,}29$, $p < 0{,}001$, $\eta^2_p = 0{,}30$. Les comparaisons planifiées ont montré qu'avec la structure radiale, ce temps moyen était plus long quand l'arrière-plan était statique, $F(1, 36) = 39{,}36$, $p < 0{,}001$, ou en mouvement, $F(1, 36) = 79{,}56$, $p < 0{,}001$, que quand il était uni, sans différence significative entre les arrière-plans structurés statique et en mouvement, $F(1, 36) < 1$. Avec la structure latérale, le temps moyen utilisé par cible était plus long quand l'arrière-plan était en mouvement que quand il était statique, $F(1, 36) = 9{,}11$, $p < 0{,}01$, ou uni, $F(1, 36) = 7{,}74$, $p < 0{,}01$, sans différence significative entre les arrière-plans structuré statique et uni, $F(1, 36) < 1$. L'effet principal de la structure de l'arrière-plan n'était pas significatif, $F(1, 36) < 1$. L'indice sonore avait aussi un effet sur le temps moyen mis pour toucher chaque cible, $F(1, 36) = 153{,}49$, $p < 0{,}001$, $\eta^2_p = 0{,}81$. Ce temps moyen était plus faible quand l'indice était présenté ($M = 1025$ ms, $ET = 225$) qu'en son absence ($M = 1186$ ms, $ET = 285$). L'effet de l'indice sonore n'interagissait significativement ni avec la structure, $F(1, 36) < 1$, ni avec la nature de l'arrière-plan, $F(2, 72) < 1$. Aucune autre comparaison n'était significative.

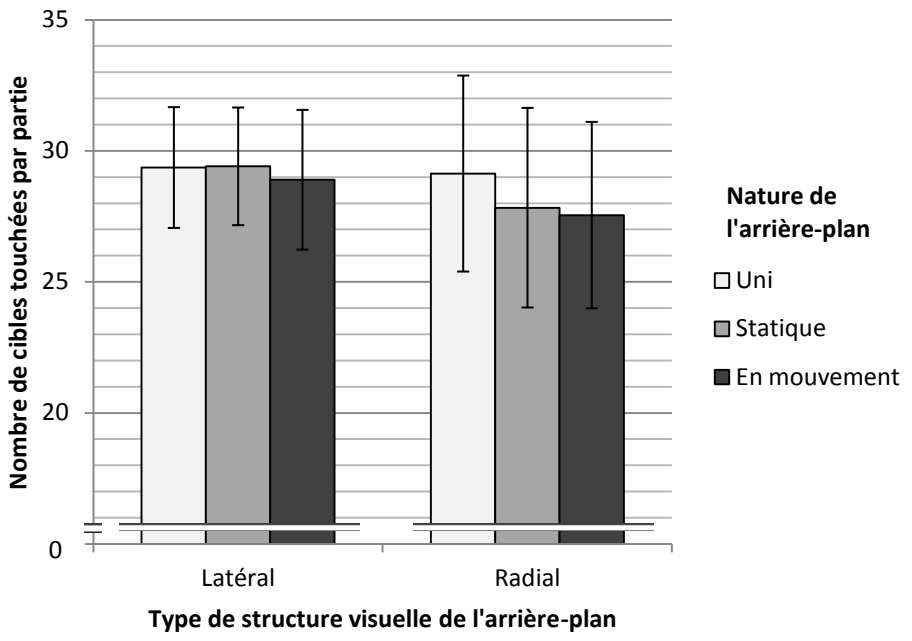


Figure 11. Nombre d'ennemis touchés par partie dans l'expérience 4 dans chaque condition (structure visuelle x nature de l'arrière-plan).
Les barres d'erreur représentent ±1 écart-type de la moyenne.

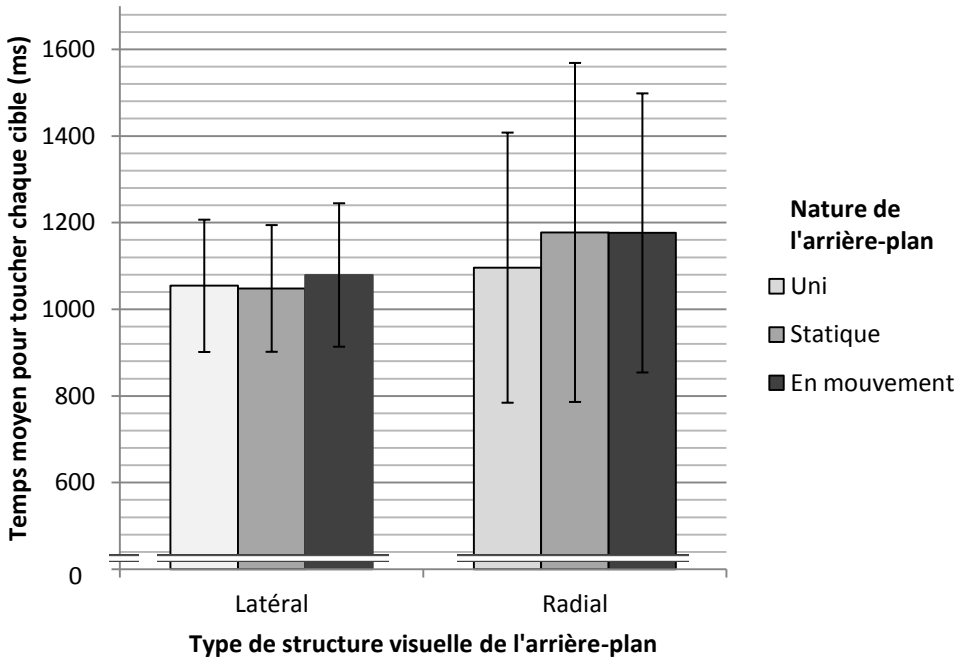


Figure 12. Temps moyen (en ms) pour toucher chaque cible dans l'expérience 4 dans chaque condition (structure visuelle x nature de l'arrière-plan).
Les barres d'erreur représentent ±1 écart-type de la moyenne.

Durée moyenne des fixations oculaires. La structure visuelle de l'arrière-plan avait une influence sur la durée moyenne des fixations oculaires, $F(1, 36) = 4{,}16$, $p < 0{,}05$, $\eta^2_p = 0{,}10$. Les fixations étaient plus longues quand la structure était radiale (M = 189 ms, ET = 46) que quand elle était latérale (M = 162 ms, ET = 40). La nature de l'arrière-plan avait aussi une influence sur la durée moyenne des fixations, $F(2, 72) = 4{,}05$, $p < 0{,}05$, $\eta^2_p = 0{,}10$. Elles étaient plus longues quand un arrière-plan structuré statique (M = 179 ms, ET = 48), $F(1, 36) = 5{,}62$, $p < 0{,}05$, ou en mouvement (M = 178 ms, ET = 48), $F(1, 36) = 7{,}17$, $p < 0{,}05$, était affiché plutôt qu'un arrière-plan uni (M = 171 ms, ET = 40). L'interaction entre la structure visuelle et la nature de l'arrière-plan n'était pas significative, $F(2, 72) = 1{,}02$, $p = 0{,}37$. La présence/absence de l'indice sonore avait un effet sur la durée moyenne des fixations oculaires, $F(1, 36) = 13{,}75$, $p < 0{,}001$, $\eta^2_p = 0{,}28$. De plus, cet effet interagissait avec la structure de l'arrière-plan, $F(2, 72) = 4{,}89$, $p < 0{,}05$, $\eta^2_p = 0{,}12$. Comme le montre la figure 13, avec la structure radiale, les fixations étaient plus courtes quand l'indice était présent que quand il était absent, $F(1, 36) = 17{,}51$, $p < 0{,}001$. Avec la structure latérale, la durée des fixations n'était pas significativement différente que l'indice soit présent ou absent, $F(1, 36) = 1{,}12$, $p = 0{,}30$. Aucune autre comparaison n'était significative.

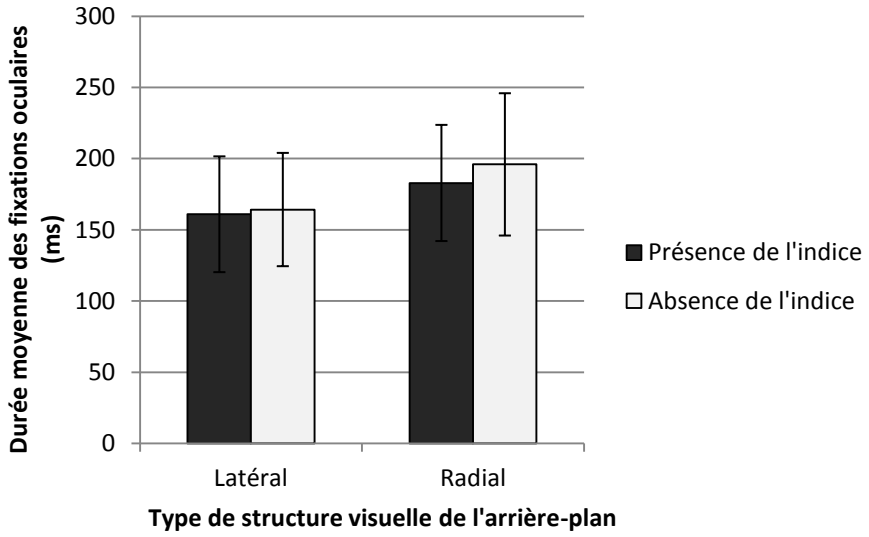


Figure 13. Durée des moyennes des fixations oculaires en fonction du type de structure de l'arrière-plan et de la présence ou de l'absence de l'indice sonore dans l'expérience 4.
Les barres d'erreur représentent ±1 écart-type de la moyenne.

Fixation initiale. La figure 14 montre que la nature de l'arrière-plan avait une influence sur la coordonnée verticale (Y) de la fixation initiale, $F(2, 72) = 10{,}64$, $p < 0{,}001$, $\eta^2_p = 0{,}23$. Cet effet interagissait avec la structure de l'arrière-plan, $F(2, 72) = 14{,}39$, $p < 0{,}001$, $\eta^2_p = 0{,}29$. Les comparaisons planifiées ont montré qu'avec la structure radiale, la fixation initiale était positionnée plus vers le bas de l'écran quand l'arrière-plan était statique (*M* = 486 pixels, *ET* = 34), $F(1, 36) = 56{,}79$, $p < 0{,}001$, ou en mouvement (*M* = 494 pixels, *ET* = 30), $F(1, 36) = 40{,}69$, $p < 0{,}001$, que quand il était uni (*M* = 464 pixels, *ET* = 33), sans différence significative entre les arrière-plans structurés statique et en mouvement, $F(1, 36) = 1{,}98$, $p = 0{,}17$. Avec la structure latérale, il n'y avait pas de différence significative entre les arrière-plans statique (*M* = 476 pixels, *ET* = 54) et en mouvement (*M* = 467 pixels, *ET* = 39), $F(1, 36) = 2{,}47$, $p = 0{,}12$, statique et uni (*M* = 471 pixels, *ET* = 49), $F(1, 36) = 2{,}08$, $p = 0{,}16$, et en mouvement et uni, $F(1, 36) < 1$. L'indice sonore avait aussi un effet sur la coordonnée verticale (Y) de la fixation initiale, $F(1, 36) = 5{,}68$, $p < 0{,}05$, $\eta^2_p = 0{,}14$. La fixation initiale était positionnée plus vers le haut de l'écran quand l'indice était présent (*M* = 474 pixels, *ET* = 41) que quand il était absent (*M* = 479 pixels, *ET* = 42). Aucune autre comparaison n'était significative.

Ni la structure visuelle, $F(1, 36) <1$, ni la nature de l'arrière-plan, $F(2, 72) = 2{,}51$, $p = 0{,}09$, ni la présence de l'indice sonore, $F(1, 36) < 1$, n'avaient d'effet significatif sur la coordonnée horizontale (X) de la fixation initiale. Les interactions entre ces facteurs n'étaient pas non plus significatives.

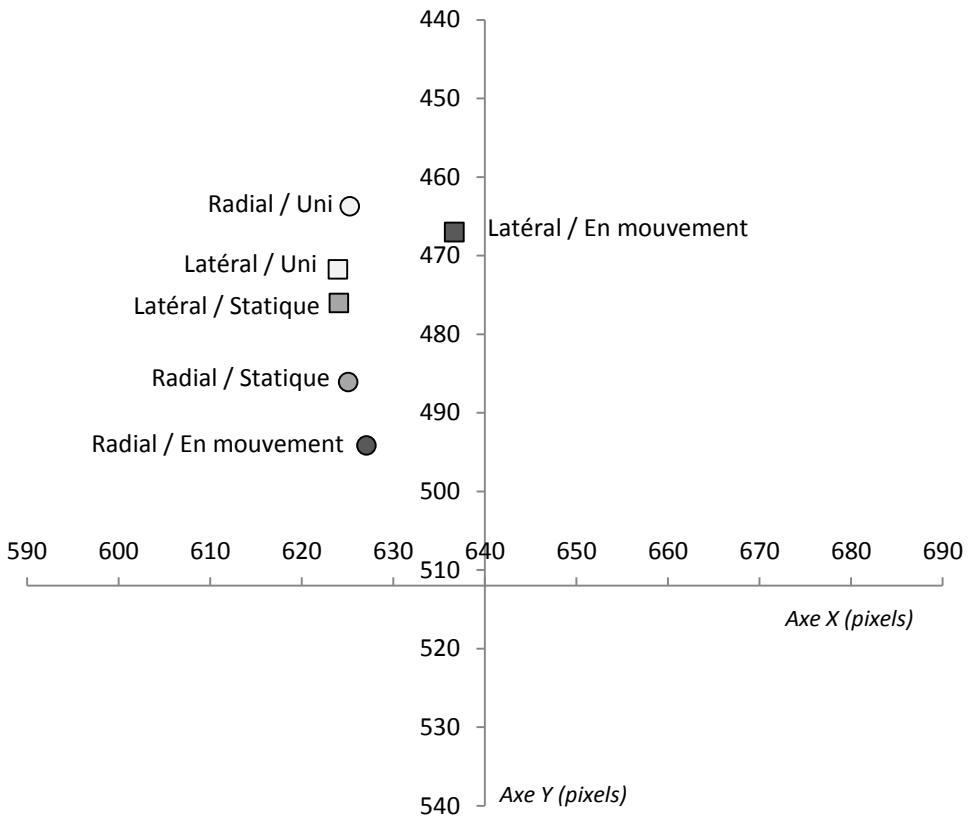


Figure 14. Coordonnées moyennes de la fixation initiale (en pixels) selon la structure et la nature de l'arrière-plan dans l'expérience 4.
Le graphique montre un zoom sur le centre de l'écran. Les axes correspondent aux axes centraux de l'écran (coordonnées de l'origine X = 640, Y = 512).

4.3. Discussion

La première hypothèse a été vérifiée. La performance est plus faible quand l'arrière-plan est en mouvement que quand il est statique, mais seulement quand sa structure visuelle est latérale et ne contient pas de points saillants de fixation grâce auxquels l'OKN peut être annulé (Ilg, 1997 ; Pola et al., 1992 ; Rubinstein & Abel, 2011 ; Wyatt & Pola, 1984). En accord avec l'expérience 3b, l'analyse des performances n'a pas montré d'effet significatif du mouvement apparent vers l'avant sur la performance quand la structure de l'arrière-plan était radiale. Contrairement à la structure latérale, la structure radiale contient des points de fixation statiques, comme le point et les lignes de fuite, qui pourraient faciliter l'annulation de l'OKN. Ces résultats sont en accord avec ceux de Kaminiarz et al. (2007) et Tozzi et al. (2007) qui ont montré que l'OKN déclenché par un arrière-plan en mouvement

latéral diminuait la précision de la localisation de cibles visuelles. Ils sont aussi cohérents avec ceux de Harrison et al. (2010), qui ont montré qu'un arrière-plan réel en mouvement latéral diminue la performance dans une tâche d'intégration auditive et visuelle. Au contraire, ils contredisent en apparence ceux de Menozzi et Koga (2004) qui avaient montré que le mouvement latéral n'affectait pas la performance de lecture. Dans une tâche simple, comme le jeu de tir utilisé dans cette expérience, le mouvement de l'arrière-plan provoquerait une diminution de la performance lorsque sa structure visuelle est simple, et ne contient pas de ligne ou de point fixe saillants. Cependant, les coordonnées de la fixation initiale n'expliquent qu'imparfaitement ces résultats. En effet, contrairement à ce qu'Ilg (1997) a décrit dans un contexte où l'OKN était déclenché, le regard des participants n'était pas initialement dévié significativement vers la droite quand l'arrière-plan de structure latérale était en mouvement. Néanmoins, la tendance était présente. En revanche, et comme dans l'expérience 3b, la structure visuelle radiale, qu'elle soit statique ou en mouvement, provoquait une déviation de la fixation initiale vers le centre de l'écran et le point fixe central de la structure. En effet, la coordonnée verticale n'était pas significativement différente entre les arrière-plans de structure visuelle radiale statique et en mouvement.

La deuxième hypothèse n'a pas été vérifiée. La performance ne dépendait pas significativement de la complexité de la structure visuelle de l'arrière-plan, plus élevée pour la structure radiale que pour la structure latérale, malgré une tendance dans ce sens. Cependant, en accord avec Henderson et al. (2009), les enregistrements des mouvements oculaires ont montré que la durée moyenne des fixations était plus longue avec la structure radiale qu'avec la structure latérale. L'impact sur la performance était peut-être trop faible pour être significatif.

La troisième hypothèse a été en partie vérifiée. Elle stipulait que la performance est plus faible sur un arrière-plan visuellement structuré que sur un arrière-plan uni. Les résultats ont montré que la structure visuelle radiale, en mouvement ou non, diminuait le nombre de cibles touchées par partie et augmentait le temps moyen mis pour toucher chaque cible par rapport à l'arrière-plan uni. En

revanche, la structure latérale avait ce même effet uniquement lorsqu'elle était en mouvement. Malgré cela, la durée moyenne des fixations était plus longue sur un arrière-plan structuré (latéral ou radial) que sur l'arrière-plan uni, en accord avec Henderson et al. (2009).

Dans leur ensemble, ces résultats suggèrent que les durées moyennes de fixations sont directement influencées par la complexité visuelle de l'arrière-plan, indépendamment du fait qu'il soit ou non en mouvement. En effet, la durée moyenne des fixations est plus longue sur un arrière-plan de structure latérale que sur l'arrière-plan uni, et plus longue sur un arrière-plan de structure radiale que sur un arrière-plan de structure latérale. En revanche, contrairement à ce qu'avaient suggéré de précédentes études (Jie & Clark, 2008 ; Wolfe et al., 2002), la complexité de l'arrière-plan ne dégrade pas toujours la performance à la tâche de manière significative. Il est possible que l'attention spécifique portée aux éléments visuels fixes superposés utilisés pour la tâche permette d'atténuer les effets de la complexité visuelle de l'arrière-plan, ou que les différences de complexité entre les différents arrière-plans utilisés dans cette étude ne soient pas suffisamment importantes.

Enfin, l'hypothèse 4 n'a pas été vérifiée. La présence d'un indice sonore ne compensait pas significativement la baisse de performance induite par les arrière-plans en mouvement ou de structure complexe. La présence de l'indice améliorait logiquement la performance, mais sans interagir avec les deux autres facteurs de complexité et de nature de l'arrière-plan. Toutefois, la présence de l'indice diminuait la durée moyenne des fixations sur l'arrière-plan de structure visuelle radiale le plus complexe. De nouveau, la durée des fixations oculaires semble être un indicateur plus sensible de la difficulté « potentielle » de la tâche que la performance elle-même.

En résumé, ces résultats ont montré que dans une situation de jeu vidéo simple, le type de mouvement du participant suggéré par les mouvements des arrière-plans se combine avec leur structure visuelle pour influencer la performance des joueurs. Globalement cependant, la complexité de la structure visuelle de l'arrière-plan semble être un facteur de perturbation plus puissant que le mouvement de l'arrière-plan. En effet, quand la structure visuelle est complexe (radiale), la

performance est dégradée, que l'arrière-plan soit en mouvement ou non, par rapport à un arrière-plan uni. Avec une structure moins complexe (latérale), la performance n'est dégradée que si l'arrière-plan est en mouvement.

5. Discussion générale

Les effets du mouvement et de la complexité des arrière-plans sur la performance des individus dans diverses tâches ont déjà été étudiés de manière isolée. Dans l'étude rapportée ici, trois expériences ont été menées pour évaluer les influences conjointes de ces caractéristiques des interfaces visuelles des jeux vidéo sur la performance des joueurs. Le mouvement et la complexité de la structure visuelle associée ont été manipulés ensemble dans une tâche typique de recherche visuelle et une tâche simple de jeu vidéo. Les résultats ont montré que le type de mouvement et le type de structure visuelle associé se combinent pour influencer la performance dans une tâche simple de visée de cible typique des jeux vidéo. Par contre, seule la complexité structurale de l'arrière-plan a une influence sur la performance de recherche visuelle, une tâche plus exigeante en termes de demandes attentionnelles. De plus, les enregistrements des mouvements oculaires ont montré que les propriétés de bas niveau de la scène visuelle avaient une influence sur le comportement du regard, qui pouvait être à l'origine des modifications de performance.

Selon de précédentes recherches (Jie & Clark, 2008 ; Wolfe et al., 2002), une scène visuelle complexe tend à diminuer la performance dans des tâches de jeu vidéo et de recherche visuelle. Les résultats des trois expériences ont confirmé ce point, même si avec les arrière-plans utilisés ici, la diminution n'était pas toujours suffisante pour être significative. Selon d'autres études (Harrison et al., 2010 ; Kaminiarz et al., 2007 ; Tozzi et al., 2007), le mouvement d'ensemble de l'arrière-plan diminue la performance dans des tâches de détection de cible quand la structure de l'arrière-plan est latérale. Les résultats de l'expérience 4 ont confirmé ce point. Dans une tâche simple de jeu vidéo, quand la complexité structurale est faible, la performance diminue sur un arrière-plan en mouvement par

rapport à un arrière-plan statique. Cet effet disparait cependant quand la complexité de la structure visuelle est plus élevée (expérience 4), ou quand la tâche est plus exigeante en termes de demandes attentionnelles (expériences 3a et 3b). La performance ne serait influencée par le mouvement de l'arrière-plan que quand la complexité de la structure visuelle de l'arrière-plan est faible et quand la tâche ne nécessite pas beaucoup de ressources attentionnelles, ou ne fournit que peu de points fixes qui peuvent servir à annuler l'OKN. Selon la quantité d'attention allouée aux objets superposés à l'arrière-plan et leur nombre, les effets négatifs de l'OKN et de la sensation de vection induits par le mouvement de l'arrière-plan pourraient être plus ou moins bien annulés ou contrôlés (Ilg, 1997 ; Menozzi & Koga, 2004 ; Pola et al., 1992 ; Rubinstein & Abel, 2011 ; Seno et al., 2011 ; Wyatt & Pola, 1984).

Même si la complexité structurale d'un arrière-plan a une influence sur la performance du joueur, les résultats ne permettent pas de comprendre clairement son rôle et l'importance du fait que la structure contienne ou non des points fixes saillants. Le chapitre suivant a pour objectif d'étudier l'influence de la complexité d'un arrière-plan indépendamment de celles des structures visuelles nécessaires pour générer les différentes sensations de mouvement propre du participant. Comme de plus, les motifs des arrière-plans et la tâche de recherche visuelle utilisés dans les expériences précédentes peuvent être considérés comme assez éloignés des environnements de jeux vidéo, les expériences du chapitre suivant ont été élaborées dans un environnement plus proche des jeux vidéo existants.

Chapitre 6 :
Complexité visuelle de l'arrière-plan et difficulté de la tâche interagissent avec le mouvement de l'arrière-plan dans un jeu vidéo

1. Introduction

Les expériences du précédent chapitre ont montré que le mouvement de l'arrière-plan pouvait dans certaines conditions détériorer la performance à une tâche qui utilise des objets superposés à cet arrière-plan. Les expériences 3a et 3b ont montré que le mouvement ne détériorait pas la performance dans une tâche de recherche visuelle qui impliquait une attention soutenue aux objets fixes affichés sur l'arrière-plan. L'expérience 4 a montré que le mouvement ne détériorait la performance que lorsque la tâche visuelle à réaliser était simple et n'impliquait que peu d'objets fixes, et que par ailleurs la structure visuelle de l'arrière-plan en mouvement ne présentait pas de lignes ou de points saillants fixes. En effet, contrairement à un arrière-plan de structure visuelle latérale, les particularités de l'arrière-plan de structure radiale permettraient à l'observateur d'annuler les effets négatifs du mouvement apparent radial, qui normalement génère un OKN et potentiellement une sensation de vection. Par ailleurs, dans les trois expériences, des effets de la

complexité de la structure visuelle de l'arrière-plan ont pu être mis en évidence. Son augmentation tend à détériorer la performance aux tâches visuelles demandées.

L'objectif des expériences 5 et 6 décrites dans ce chapitre est double. D'abord, les expériences avaient pour but de répliquer les résultats précédents dans le cadre d'environnements plus proches de ceux réellement rencontrés dans les jeux vidéo. Les arrière-plans basés sur des motifs visuels utilisés plutôt en psychophysique ont été remplacés par des arrière-plans basés sur des décors réels de jeux vidéo. De plus, comme la tâche de recherche visuelle reste relativement éloignée des tâches habituellement proposées dans les jeux vidéo, c'est une tâche de tir similaire à celle de l'expérience 4 qui a été déclinée en plusieurs versions dans les expériences 5 et 6. L'idée était de mettre en évidence directement l'influence de variations du niveau de difficulté d'une même tâche sur l'impact du mouvement de l'arrière-plan.

Plus précisément, l'objectif de l'expérience 5 était de mieux comprendre l'influence de la complexité visuelle de l'arrière-plan et de son interaction avec le mouvement de l'arrière-plan sur la performance du joueur. En effet, le matériel utilisé dans le chapitre précédent n'a pas permis de complètement expliquer les résultats observés. Le fait qu'une détérioration de la performance par le mouvement de l'arrière-plan ait été observée avec la structure latérale, mais jamais avec la structure radiale pourrait être la conséquence soit de la plus grande complexité de la structure radiale en général, soit de propriétés visuelles spécifiques de cette structure telles que l'existence d'un point et de lignes de fuite fixes. Dans l'expérience 5, la complexité visuelle d'une structure visuelle latérale a été manipulée pour vérifier que les propriétés visuelles de bas niveau de l'arrière-plan seules (i.e., sans lien avec le type d'illusion de mouvement généré par le mouvement de l'arrière-plan) peuvent être source de dégradation de la performance, que l'arrière-plan soit fixe ou en mouvement.

En résumé, le but de l'expérience 5 était de montrer que la complexité visuelle d'un arrière-plan détériore en elle-même la performance à une tâche simple de jeu vidéo, que l'arrière-plan soit fixe ou en mouvement. Le but de l'expérience 6 était de montrer que, comme le suggèrent les

expériences précédentes, le mouvement de l'arrière-plan ne détériore la performance à une tâche de tir inspirée des jeux vidéo que quand elle est peu complexe en termes de demandes attentionnelles.

Enfin, des indications sonores sur la cible ont également été utilisées dans ces deux expériences. Toutefois, leur rôle était différent de celui de l'expérience 4. La présence ou l'absence n'a pas été manipulée, puisque aucun avantage n'a été montré en cas d'arrière-plan en mouvement ou complexe en présence d'un indice de localisation de la cible. Dans les expériences 5 et 6, les indications sonores étaient toujours présentes. Elles ont été utilisées pour diversifier le type de cible à repérer (couleur, taille ou localisation) dans le cadre des jeux de tir proposé aux participants. Aucune interaction n'étant attendue avec les facteurs principaux manipulés dans ces expériences (mouvement et complexité de l'arrière-plan), aucune hypothèse particulière n'a été posée en fonction des variations de ces indications.

2. Expérience 5 : Complexité et mouvement de l'arrière-plan

L'expérience 5 avait pour but de préciser les influences conjointes du mouvement latéral et de la complexité visuelle d'un arrière-plan sur une tâche de jeu vidéo simple. Le jeu de tir utilisé était proche de celui de l'expérience 4. Le but du jeu était de viser et tirer le plus rapidement possible sur diverses cibles visuelles désignées par des indications sonores pour obtenir le meilleur score possible.

La première hypothèse était que la performance est plus faible quand l'arrière-plan est complexe que quand il est peu complexe (Jie & Clark, 2008 ; Wolfe et al., 2002). Le chapitre précédent, ainsi que les précédentes recherches, décrites dans le chapitre 2 (Harrison et al., 2010 ; Kaminiarz et al., 2007 ; Tozzi et al., 2007), ont montré qu'un arrière-plan en mouvement, parce qu'il modifie le comportement du regard (Ilg, 1997), dégrade la performance dans des tâches plutôt simples. Cette dégradation de la performance devrait interagir de manière multiplicative avec celle liée à la

complexité visuelle de l'arrière-plan. Aussi, la seconde hypothèse était que sur un arrière-plan de structure latérale complexe, la performance est plus dégradée par la complexité quand il est en mouvement que quand il est statique.

2.1. Méthode

Participants. Vingt-deux volontaires (16 femmes, 6 hommes) ont pris part à l'expérience. Leur âge moyen était de 19,9 ans (*ET* = 1,8) et leur durée moyenne d'études de 13,6 ans (*ET* = 1,4). Tous les participants étaient de langue maternelle française et avaient une vision normale ou corrigée. Dix-neuf d'entre eux étaient droitiers et trois étaient gauchers. Douze participants avaient l'œil droit dominant et dix l'œil gauche. Neuf participants étaient des joueurs réguliers, qui avaient joué au moins une fois par semaine lors des 12 derniers mois, trois étaient d'anciens joueurs réguliers, qui avaient joué moins d'une fois par semaine lors des 12 derniers mois mais au moins une fois par semaine auparavant, neuf étaient des joueurs occasionnels, qui avaient joué moins d'une fois par semaine tout au long de leur vie, et un dernier participant n'avait jamais joué à des jeux vidéo.

Appareillage et matériel. L'appareillage utilisé était le même que celui utilisé pour les expériences du chapitre précédent (oculomètre Tobii T120 et haut-parleur central).

Le matériel a été créé avec le logiciel « Adobe Director 11 ». Deux types différents d'arrière-plans ont été construits (figure 15). L'un représentait un décor naturel, et l'autre un décor non-naturel. Le décor naturel était composé d'un fond composé d'un motif neutre à dominante verte parsemé d'arbres, tirés du jeu vidéo « Sim City 4 » (Electronic Arts, 2003). Le décor non-naturel était composé d'un autre motif à dominante verte parsemé d'objets abstraits, tirés du jeu vidéo « Child of Eden » (Ubisoft, 2011). Chaque type d'arrière-plan a été décliné en 3 versions de complexité visuelle croissante, déterminée en fonction de la superficie du recouvrement du motif de base par les objets. La première version était composée du motif seul, la deuxième était composée du motif recouvert par les objets (arbres ou objets abstraits) sur 25 % de sa superficie et la troisième était composée du

motif recouvert par les objets sur 50 % de sa superficie. Lorsque l'arrière-plan était en mouvement latéral, il se déplaçait vers la gauche à une vitesse de 121 mm/s (environ 12 degrés par seconde). La cible du jeu de tir et le distracteur étaient représentés par des dessins d'une créature, similaire à celle de l'expérience 4. Quatre versions de la créature ont été utilisées (annexe 6). Elle pouvait être de couleur bleue ou rouge, et de petite (85x66 pixels, 22x17 mm, 2,2x1,7 degrés) ou grande taille (85x99 pixels, 22x26 mm, 2,2x2,6 degrés). Le curseur était une représentation d'un viseur d'arme de 114 pixels de diamètre (30 mm, 3 degrés). Six indices sonores ont été construits à partir d'une voix naturelle masculine enregistrée. Ces indices prononçaient « gauche », « droite », « bleu », « rouge », « grand » ou « petit » sur une durée de 500 ms.

Plan expérimental et procédure. Le mouvement (présence ou absence) et le taux de recouvrement (nul, faible ou élevé) des arrière-plans par les objets incorporés étaient manipulés en intra-participant. La nature de l'indication sonore (localisation, couleur ou taille) était également manipulée en intra-participant. Même si les participants ont été confrontés aux deux types de décor (naturel et non-naturel) des arrière-plans, l'analyse des résultats a été faite en les séparant du fait d'une impossibilité de les comparer de manière objective. Le type d'arrière-plan n'était donc pas considéré comme un facteur faisant partie du plan expérimental.

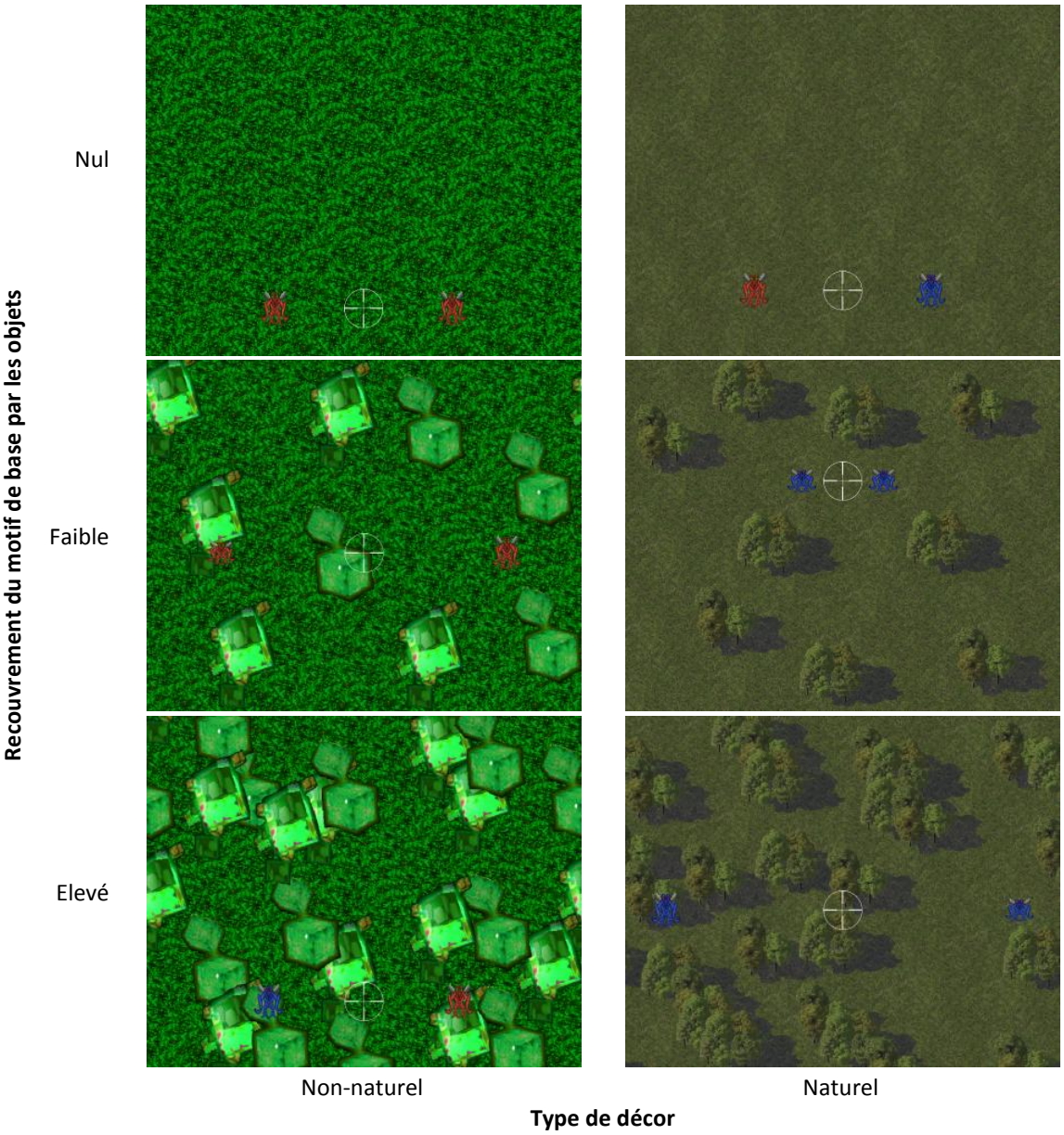


Figure 15. Les deux types de décor utilisés pour les arrière-plans dans l'expérience 5 avec différents exemples de dispositions d'une cible, d'un distracteur et du curseur.
Chaque arrière-plan était décliné en 3 versions correspondant à différents taux de recouvrement du motif de base par les objets (arbres ou objets abstraits).

Une tâche de tir était utilisée. Au début de chaque essai, l'arrière-plan était affiché et un cadre noir de 74 mm de côté (7,4 degrés) était affiché au centre de l'écran pendant 1 seconde pour orienter le regard du participant sans annuler les effets éventuels (OKN et vection) du mouvement de l'arrière-plan. Les participants devaient garder leur regard le plus possible à l'intérieur du cadre. L'indication sonore était présentée en même temps que le cadre. Dans le cas d'une indication de localisation, le participant entendait « gauche » ou « droite », dans le cas d'une indication de couleur, il entendait

« bleu » ou « rouge » et dans le cas d'une indication de taille, il entendait « grand » ou « petit ». La cible, le distracteur et le curseur apparaissaient à l'écran sur une même ligne horizontale dont la coordonnée était déterminée aléatoirement. Le curseur était toujours positionné sur l'axe vertical central de l'écran. La cible était affichée aléatoirement sur la droite ou sur la gauche, et sa coordonnée horizontale (entre le curseur et l'extrémité de l'écran) déterminée elle aussi aléatoirement. Le distracteur était positionné à l'exact opposé de la cible par rapport à l'axe vertical central de l'écran. La taille et la couleur du distracteur ne différait de la cible que sur le trait visuel correspondant à l'indication présentée. Par exemple, si l'indication était « rouge », la cible était obligatoirement de couleur rouge mais pouvait être soit grande, soit petite. Le distracteur était alors de couleur opposée (ici bleue), mais de taille identique à celle de la cible. Les participants devaient placer le viseur sur la cible à l'aide des touches « flèche gauche » et « flèche droite » du clavier et tirer sur celle-ci en appuyant sur la touche « espace ». Le participant pouvait déplacer le curseur et tirer autant de fois qu'il le souhaitait tant que la cible n'était pas touchée, autrement dit tant que le curseur et la cible n'étaient pas superposés. Quand la cible était touchée, un autre essai débutait.

L'expérimentateur présentait cette procédure au participant comme un jeu vidéo. Il devait toucher le plus rapidement possible chaque ennemi en négligeant le distracteur. Chaque bloc d'essais était présenté au participant comme une partie. Douze parties de soixante essais étaient présentées à chaque participant avec une pause entre chaque partie. Le score global de performance était donné à la fin de la passation. Tous les essais d'une partie étaient réalisés avec le même arrière-plan, qui ne disparaissait pas entre les essais. Quand l'arrière-plan était en mouvement, le mouvement était continu pendant toute la partie. L'ordre de présentation des 12 parties (2x2x3 conditions différentes de type d'arrière-plan, de mouvement et de taux de recouvrement de l'arrière-plan) était pseudo-aléatoire. Dans chaque partie, l'ordre de présentation des 3 types d'indications (20 essais par type) était également pseudo-aléatoire.

Mesures dépendantes. La mesure dépendante utilisée pour évaluer la performance au jeu était le temps de réponse moyen pour toucher chaque cible (i.e., pour réaliser un essai). Pour expliquer la performance, les mouvements oculaires des participants ont été analysés en utilisant deux mesures dépendantes : la durée moyenne des fixations oculaires, influencée par la complexité visuelle (Henderson et al., 2009), et la position de chaque fixation initiale enregistrée sur l'écran de l'oculomètre pendant ou juste avant l'apparition de la cible, définie par ses coordonnées horizontale et verticale X et Y. Cette dernière mesure servait à évaluer l'influence de chaque arrière-plan sur la position moyenne du regard (Ilg, 1997). Toutes ces variables ont été analysées en utilisant des ANOVA pour mesures répétées avec le mouvement de l'arrière-plan, le recouvrement de l'arrière-plan et la nature de l'indication en facteurs intra-participants. Quand le test de sphéricité de Mauchley était significatif, une correction de Greenhouse-Geisser a été appliquée.

2.2. Résultats

Deux analyses séparées ont été réalisées pour les deux types de décor (naturel et non-naturel). Les réponses tardives (temps de réponse supérieurs à 3000 ms) ont été exclues de l'analyse. Le nombre d'essais exclus représentait en moyenne 2,0% du nombre total d'essais par participant (min = 0,0%, max = 10,8%) pour le décor naturel, et 2,1% (min = 0,0%, max = 13,3%) pour le décor non-naturel. Les données de temps de réponse ont été transformées logarithmiquement avant que les ANOVA ne soient réalisées. Pour la clarté de la présentation, ce sont les moyennes non-transformées qui sont indiquées dans le texte et les graphiques.

2.2.1. Temps de réponse moyen

Décor non-naturel. Comme le montre la figure 16, le mouvement de l'arrière-plan avait une influence sur le temps de réponse, $F(1, 21) = 32{,}08$, $p < 0{,}001$, $\eta^2_p = 0{,}60$. Le temps de réponse était supérieur sur l'arrière-plan en mouvement (M = 1404 ms, ET = 244) par rapport à l'arrière-plan statique (M = 1324 ms, ET = 272). Le taux de recouvrement de l'arrière-plan avait également une

influence sur le temps de réponse, $F(2, 42) = 4{,}87$, $p < 0{,}05$, $\eta^2_p = 0{,}19$. Les comparaisons planifiées ont montré que le temps de réponse en condition de faible recouvrement (*M* = 1392 ms, *ET* = 264) était plus élevé que dans les conditions sans recouvrement (*M* = 1341 ms, *ET* = 253), $F(1, 21) = 7{,}65$, $p < 0{,}05$, ou de recouvrement élevé (*M* = 1359 ms, *ET* = 265), $F(1, 21) = 4{,}90$, $p < 0{,}05$. La différence entre ces deux dernières conditions n'était pas significative, $F(1, 21) = 1{,}02$, $p = 0{,}32$. L'interaction entre les facteurs mouvement et taux de recouvrement n'était pas significative, $F(2, 42) = 2{,}08$, $p = 0{,}14$. Enfin, la nature de l'indication avait aussi une influence sur le temps de réponse, $F(2, 42) = 13{,}54$, $p < 0{,}001$, $\eta^2_p = 0{,}39$. Les comparaisons planifiées ont révélé que le temps de réponse obtenu avec l'indication de taille (*M* = 1438 ms, *ET* = 245) était supérieur à ceux obtenus avec les indications de localisation (*M* = 1339 ms, *ET* = 288), $F(1, 21) = 14{,}27$, $p < 0{,}05$, ou de couleur (*M* = 1315 ms, *ET* = 232), $F(1, 21) = 75{,}74$, $p < 0{,}001$, sans différence significative entre ces deux dernières conditions, $F(1, 21) < 1$. Aucune autre comparaison n'était significative.

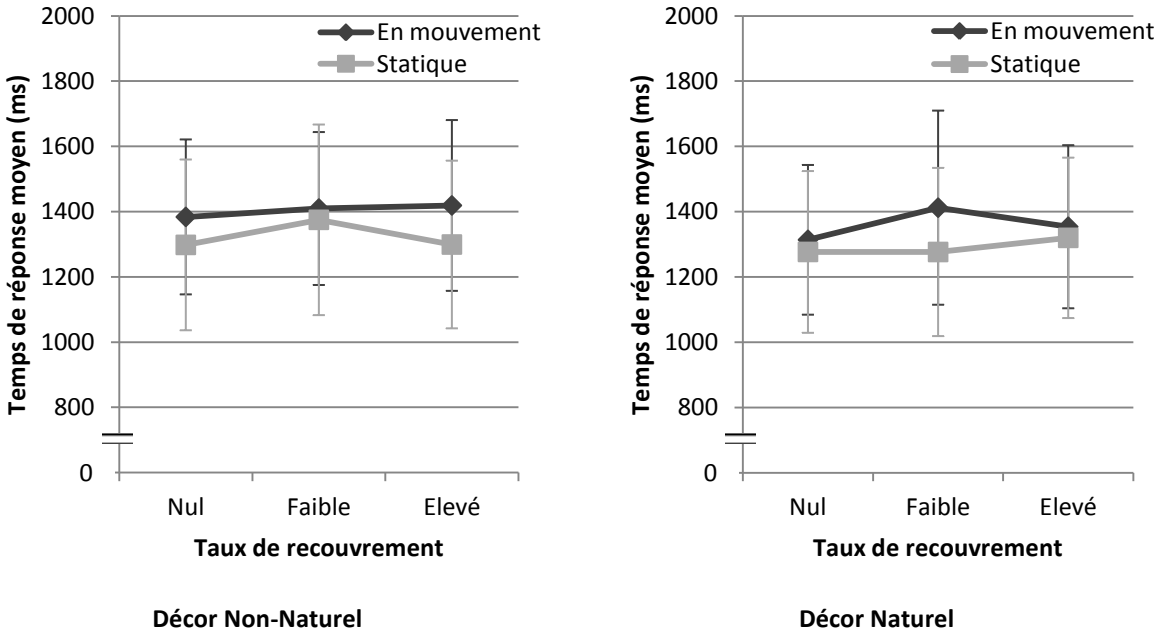


Figure 16. Temps de réponse moyen (en ms) pour toucher chaque cible dans l'expérience 5 dans chaque condition (taux de recouvrement X mouvement).
Les barres d'erreur représentent ±1 écart-type de la moyenne.

Décor naturel. Comme le montre la figure 16, le mouvement de l'arrière-plan avait une influence sur le temps de réponse, $F(1,21) = 6{,}71$, $p < 0{,}05$, $\eta^2_p = 0{,}24$. Le temps de réponse était plus long sur

l'arrière-plan en mouvement (*M* = 1360 ms, *ET* = 262) que sur l'arrière-plan statique (*M* = 1291 ms, *ET* = 250). Le taux de recouvrement de l'arrière-plan n'avait en revanche pas d'influence significative sur le temps de réponse, $F(2, 42) = 2{,}65$, $p = 0{,}08$. Par contre, l'interaction entre le mouvement et le taux de recouvrement de l'arrière-plan était significative, $F(2, 42) = 3{,}81$, $p < 0{,}05$, $\eta^2_p = 0{,}15$. Les comparaisons planifiées ont montré que quand l'arrière-plan était en mouvement, le temps de réponse de la condition de faible recouvrement était supérieur à celui de la condition sans recouvrement, $F(1, 21) = 6{,}96$, $p < 0{,}05$. Les temps de réponse n'était pas significativement différent entre les conditions sans recouvrement et de recouvrement élevé, $F(1, 21) = 1{,}88$, $p = 0{,}18$, et de recouvrement faible et élevé, $F(1, 21) = 3{,}07$, $p = 0{,}09$. Quand l'arrière-plan était statique, le temps de réponse n'était pas significativement différent entre les conditions de recouvrement nul et faible, $F(1, 21) < 1$, nul et élevé, $F(1, 21) = 3{,}56$, $p = 0{,}07$, et faible et élevé, $F(1, 21) = 2{,}87$, $p = 0{,}10$.

La nature de l'indication sonore avait une influence sur le temps de réponse moyen, $F(2, 42) = 12{,}32$, $p < 0{,}01$, $\eta^2_p = 0{,}37$. Les comparaisons planifiées ont montré que le temps de réponse obtenu avec l'indication de taille (*M* = 1392 ms, *ET* = 232) était supérieur à ceux obtenus avec les indications de localisation (*M* = 1307 ms, *ET* = 292), $F(1, 21) = 10{,}91$, $p < 0{,}01$, et de couleur (*M* = 1277 ms, *ET* = 234), $F(1, 21) = 120{,}76$, $p < 0{,}001$. Le temps de réponse n'était pas significativement différent entre ces deux dernières conditions, $F(1, 21) < 1$. De plus, il y avait une interaction entre la nature de l'indication et le mouvement de l'arrière-plan, $F(2, 42) = 3{,}81$, $p < 0{,}05$, $\eta^2_p = 0{,}15$. Les comparaisons planifiées ont montré que le temps de réponse sur l'arrière-plan en mouvement était plus élevé que sur l'arrière-plan statique après la présentation d'indications de localisation (respectivement *M* = 1358 ms, *ET* = 296, et *M* = 1257 ms, *ET* = 281), $F(1, 21) = 10{,}84$, $p < 0{,}01$, ou de taille (*M* = 1423 ms, *ET* = 241, et *M* = 1361 ms, *ET* = 220), $F(1, 21) = 5{,}79$, $p < 0{,}05$, mais pas après l'indication de couleur (*M* = 1299 ms, *ET* = 235, et *M* = 1255 ms, *ET* = 233), $F(1, 21) = 2{,}06$, $p = 0{,}17$. Aucune autre comparaison n'était significative.

2.2.2. Durée moyenne des fixations oculaires

Décor non-naturel. Ni le mouvement de l'arrière-plan, $F(1, 21) = 3{,}78$, $p = 0{,}07$, ni le taux de recouvrement de l'arrière-plan, $F(2, 42) < 1$, ni la nature de l'indication, $F(2, 42) = 1{,}52$, $p = 0{,}23$, n'avaient d'influence significative sur la durée moyenne des fixations oculaires. Les interactions entre ces facteurs n'étaient pas non plus significatives.

Décor naturel. Ni le mouvement de l'arrière-plan, $F(1, 21) = 1{,}99$, $p = 0{,}17$, ni le taux de recouvrement de l'arrière-plan, $F(2, 42) < 1$, ni la nature de l'indication, $F(2, 42) = 2{,}31$, $p = 0{,}11$, n'avaient d'influence significative sur la durée moyenne des fixations oculaires. Les interactions entre ces facteurs n'étaient pas non plus significatives.

2.2.3. Fixation initiale

Décor non-naturel. La fixation initiale était positionnée plus vers la droite sur l'arrière-plan en mouvement (M = 654 pixels, ET = 49) que sur l'arrière-plan statique (M = 630 pixels, ET = 49), $F(1,21) = 28{,}94$, $p < 0{,}001$, $\eta^2_p = 0{,}58$. Par contre, aucun des trois facteurs testés n'avait d'influence significative sur la coordonnée verticale (Y) de la fixation initiale. Aucune autre comparaison n'était significative.

Décor naturel. La fixation initiale était positionnée plus vers la droite sur l'arrière-plan en mouvement (M = 660 pixels, ET = 66) que sur l'arrière-plan statique (M = 634 pixels, ET = 68), $F(1,21) = 18{,}41$, $p < 0{,}001$, $\eta^2_p = 0{,}47$. Par contre, aucun des trois facteurs testés n'avait d'influence significative sur la coordonnée verticale (Y) de la fixation initiale. Aucune autre comparaison n'était significative.

2.3. Discussion

La première hypothèse était que la performance est plus faible quand l'arrière-plan est complexe (taux de recouvrement élevé) que quand il est peu complexe. Elle n'a été que partiellement vérifiée.

Le temps de réponse ne variait significativement en fonction du taux de recouvrement de l'arrière-plan que lorsque le décor était de type non-naturel. De plus, le temps de réponse en condition de recouvrement faible était plus élevé que dans les deux autres conditions de recouvrement, nul et élevé. Selon Beck et al. (2010) et Wolfe et al. (2002), la dégradation de la performance dans une tâche de recherche visuelle serait linéairement liée à la complexité visuelle de la scène, et le temps de réponse des participants dans la condition de complexité élevée aurait dû être supérieur à celui de la condition de complexité faible. Il semble donc que la manière dont la complexité visuelle des arrière-plans a été définie dans cette expérience n'était pas complètement pertinente. D'après les résultats, la présence d'objets superposés à un motif de base augmente bien la complexité visuelle de l'arrière-plan, mais un accroissement de la superficie recouverte ne semble pas nécessairement augmenter cette complexité, et pourrait même en fait la faire diminuer. Ceci vient peut-être du fait que la quantité de points saillants ne dépendrait pas directement de la proportion de la scène recouverte d'objets.

La seconde hypothèse était qu'avec un arrière-plan complexe, la performance est encore plus dégradée quand il est, en plus, en mouvement que quand il est statique. Elle n'a été que partiellement vérifiée. Les résultats ont effectivement montré que la performance était plus faible quand l'arrière-plan était en mouvement que quand il était statique, et ce quel que soit le type de décor. Le temps de réponse des participants était plus long lorsque l'arrière-plan était en mouvement. L'analyse des mouvements oculaires a montré que cette détérioration était vraisemblablement due aux effets de l'OKN généré par le mouvement. Les résultats ont en effet montré une déviation du regard vers l'origine du mouvement, quand il était présent (Ilg, 1997). Cependant l'interaction attendue entre les facteurs de mouvement et de taux de recouvrement de l'arrière-plan n'était significative que lorsque le décor était de type naturel. Les résultats ont montré que lorsque le taux de recouvrement de l'arrière-plan était nul ou élevé, il n'y avait pas de différence de temps de réponse entre les arrière-plans statique et en mouvement. Par contre, quand le taux de recouvrement était faible, le temps de réponse des participants en condition de mouvement était

supérieur à celui de la condition statique. Lorsque le décor était de type naturel, la présence d'objets superposés au motif de base de l'arrière-plan dégradait la performance au moins quand le taux de recouvrement n'était pas trop élevé, mais seulement lorsque l'arrière-plan était en mouvement.

En résumé, dans l'expérience 5, l'arrière-plan ne pouvait apparemment être considéré comme plus complexe que si les objets recouvraient 25 % de la superficie de l'écran. Lorsque le recouvrement était trop élevé, l'influence de la complexité visuelle de l'arrière-plan n'était pas différente de celle d'un arrière-plan composé seulement du motif de base. Des expériences complémentaires sont nécessaires pour déterminer dans quelle mesure le taux de recouvrement d'un motif de base par des objets peut être un indicateur de complexité visuelle de l'arrière-plan.

Par ailleurs, l'analyse des mouvements oculaires n'a pas montré de différence significative de durée moyenne des fixations entre les différentes conditions de recouvrement, quel que soit le type de décor utilisé. Ce résultat n'est pas en accord avec ceux de l'expérience 4 et celle de Henderson et al. (2009), et confirme qu'effectivement, le taux de recouvrement ne serait pas un bon indice de la complexité d'un arrière-plan construit comme dans cette expérience. L'influence du taux de recouvrement sur les durées de fixation était peut-être aussi trop faible pour pouvoir être observée.

Enfin, la nature de l'indication utilisée avait également une influence sur la performance des participants. D'après les résultats, le fait de repérer la cible par sa taille était plus difficile que de la repérer par sa localisation ou sa couleur, et ce quel que soit le type de décor. Cet effet est probablement dû au fait que repérer la cible par sa taille nécessitait une comparaison avec le distracteur, puisque la taille de la cible était définie comparativement à celle du distracteur, contrairement à sa couleur ou sa localisation. Pour le décor de type naturel, les résultats ont également montré qu'un arrière-plan en mouvement dégradait la performance avec une indication de localisation ou de taille mais pas avec une indication de couleur. Dans une tâche simple comme celle utilisée ici, repérer un objet par sa couleur se fait généralement de manière automatique (effet pop-out), ce qui atténuerait l'influence du mouvement sur la performance dans cette condition.

2.4. Conclusion

La performance à une tâche simple de jeu vidéo peut être dégradée par le mouvement de l'arrière-plan, comme l'expérience 4 l'a déjà montré. Par contre, à l'inverse de ce qui a été montré dans l'expérience 4, la complexité visuelle de l'arrière-plan ne permet pas forcément de réduire les effets négatifs du mouvement. Les résultats de l'expérience 5 ont montré que les effets de la complexité visuelle et du mouvement sur la performance étaient additifs pour un type de décor (décor non-naturel), voire même multiplicatifs pour un autre type de décor (décor naturel). Malheureusement, une limite de cette expérience est que des résultats spécifiques ont été obtenus pour chaque type de décor. Des études complémentaires sont nécessaires pour déterminer les raisons précises de cette différence, qui seraient liées aux caractéristiques visuelles des décors utilisés.

En tout état de cause, les résultats ont permis de montrer que la complexité visuelle d'un arrière-plan masque les effets négatifs du mouvement uniquement si les éléments à l'origine de cette complexité comprennent des points ou lignes qui restent fixes pendant le mouvement, comme dans l'expérience 4. Dans l'expérience présente, la complexité visuelle des arrière-plans variait indépendamment de la présence ou de l'absence de points saillants fixes, puisque les objets étaient en mouvement avec l'ensemble de l'arrière-plan. Du coup, les effets du mouvement n'étaient pas annulés par ceux de la complexité, et pouvaient même dégrader la performance de manière synergique.

Cette expérience avait pour but d'étudier les interactions entre le mouvement et la complexité de l'arrière-plan dans un environnement de jeu vidéo. L'objectif de l'expérience suivante était d'étudier dans le même type d'environnement l'interaction entre le mouvement de l'arrière-plan et la difficulté de la tâche.

3. Expérience 6 : Difficulté de la tâche et mouvement de l'arrière-plan

L'engagement dans un jeu vidéo est lié au challenge proposé par celui-ci (E. A. Boyle et al., 2012). Contrairement aux systèmes interactifs classiques, la tâche ne doit pas être simplifiée au maximum pour faciliter l'interaction, mais adaptée pour maintenir élevée la motivation du joueur (Malone & Lepper, 1987). Dans un jeu vidéo, le challenge est généralement déterminé par la difficulté de la tâche à réaliser (Aponte, Levieux, & Natkin, 2011 ; Chanel, Rebetez, Betrancourt, & Pun, 2011 ; C. Liu, Agrawal, Sarkar, & Chen, 2009 ; Orvis et al., 2008 ; Qin, Rau, & Salvendy, 2010).

L'expérience 6 avait pour but d'étudier simultanément l'influence d'un mouvement latéral de l'arrière-plan et de la difficulté de la tâche sur la performance des joueurs dans un environnement de jeu vidéo. Dans le chapitre précédent, les conclusions sur l'interaction entre le mouvement de l'arrière-plan et la difficulté de la tâche ont été faites à partir de deux tâches différentes, une tâche simple de jeu de tir et une tâche plus complexe de recherche visuelle. Dans l'expérience présente, une même tâche de jeu de tir avec des niveaux de difficultés différents a été utilisée, basée sur celle de l'expérience 5. Le but du jeu était de viser et de tirer le plus rapidement possible sur diverses cibles pour obtenir le meilleur score. La cible était toujours présentée parmi un nombre fixe (4) de distracteurs. La difficulté de la tâche variait en fonction des différences de caractéristiques visuelles (ici couleur et taille) entre la cible et les distracteurs. Les études sur la recherche visuelle de « conjonction de caractéristiques » (e.g., Reali et al., 2006 ; Thornton & Gilden, 2007) ont montré que le niveau de difficulté de la recherche est lié, entre autres, au nombre de caractéristiques visuelles partagées entre la cible et les distracteurs. Lorsque la cible ne partage aucune caractéristique avec les distracteurs, la recherche s'effectue en parallèle, de manière quasi-instantanée (effet pop-out). Lorsque la cible partage une ou plusieurs caractéristiques avec les distracteurs, la recherche devient sérielle et dure alors plus longtemps.

La première hypothèse était que la performance est plus faible quand le niveau de difficulté de la tâche est élevé (difficile) que quand il est faible (facile), conformément aux résultats des expériences de recherche visuelle de conjonction de caractéristiques. La seconde hypothèse était que la performance est plus faible quand l'arrière-plan est en mouvement seulement si la tâche est facile. En effet, le précédent chapitre a montré qu'une tâche impliquant une attention soutenue sur des éléments fixes superposés à l'arrière-plan annulait les effets négatifs du mouvement.

3.1. Méthode

Participants. Vingt-deux volontaires (15 femmes, 7 hommes) ont pris part à l'expérience. Leur âge moyen était de 20,1 ans (*ET* = 3,0) et leur durée moyenne d'études de 13,5 ans (*ET* = 1,2). Tous les participants étaient de langue maternelle française et avaient une vision normale ou corrigée. Dix-neuf d'entre eux étaient droitiers et trois étaient gauchers. Treize participants avaient l'œil droit dominant et neuf l'œil gauche. Onze participants étaient des joueurs réguliers, cinq étaient d'anciens joueurs réguliers, et six étaient des joueurs occasionnels.

Appareillage et matériel. L'appareillage était similaire à celui de l'expérience 5. Le matériel a été créé avec le logiciel « Adobe Director 11 ». Deux types différents d'arrière-plans ont été construits (figure 17). Le décor non-naturel de l'expérience 5 a été réutilisé et décliné en deux versions. La première était identique à la version de taux de recouvrement élevé de l'expérience 5. La seconde était une version noir et blanc de la précédente. Cette seconde version a été utilisée pour vérifier que les effets observés sont indépendants de la colorisation de l'arrière-plan. Lorsque l'arrière-plan était en mouvement latéral, il se déplaçait vers la gauche à une vitesse identique à celle de l'expérience 5. La cible du jeu de tir et les distracteurs étaient représentés par le même dessin de créature que dans l'expérience 5. Les quatre mêmes versions (couleur bleue ou rouge X petite ou grande taille) ont été utilisées. Quatre indices sonores ont été construits à partir d'une voix naturelle masculine enregistrée. Ces indices prononçaient « bleu », « rouge », « grand » ou « petit » sur une durée de 500 ms. Ils étaient identiques à ceux utilisés dans l'expérience 5.

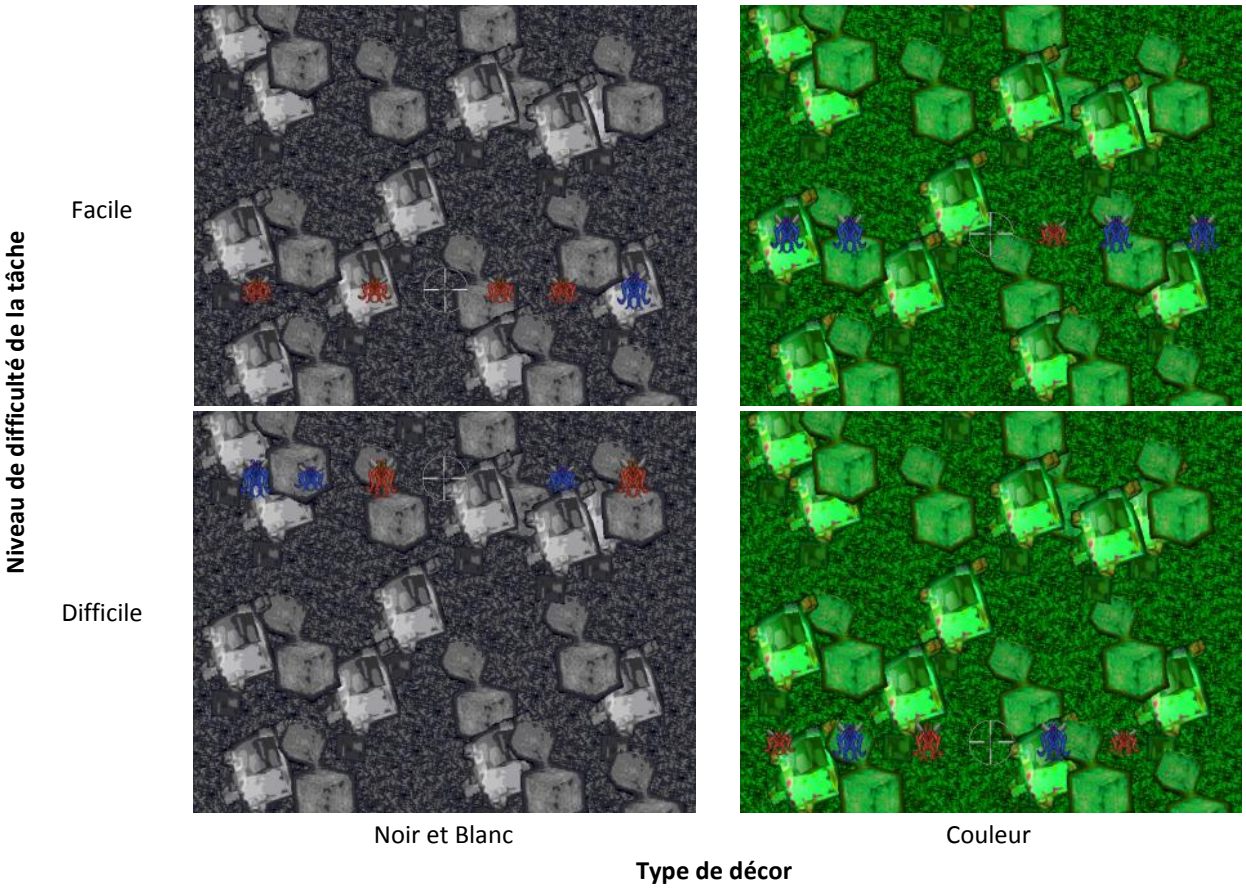


Figure 17. Les deux types de décor utilisés pour les arrière-plans dans l'expérience 6 et des exemples de dispositions de la cible, des distracteurs et du curseur en fonction du niveau de difficulté de la tâche.

Plan expérimental et procédure. Le mouvement (présence ou absence) des arrière-plans et le niveau de difficulté (facile ou difficile) de la tâche étaient manipulés en intra-participant. Même si les participants ont été confrontés aux deux types de décor (en couleur et en noir et blanc) des arrière-plans, l'analyse des résultats a été faite en les séparant du fait d'une impossibilité de les comparer de manière objective. Le type d'arrière-plan n'était donc pas considéré comme un facteur faisant partie du plan expérimental.

La tâche de tir était identique à celle de l'expérience 5. La procédure était également la même excepté sur les deux points suivants. Premièrement, deux indications étaient présentées successivement pendant l'affichage du cadre central au lieu d'un seul. Ils renseignaient sur la taille et la couleur de la cible. L'ordre de présentation des deux indications était contrebalancé. Deuxièmement, quatre distracteurs étaient présentés au lieu d'un seul. Ils étaient disposés de

manière aléatoire sur la même ligne horizontale que la cible, en veillant à ce que deux d'entre eux soient positionnés à gauche de l'axe central vertical de l'écran et deux à droite. Quand la tâche était facile, les distracteurs ne partageaient aucune caractéristique de couleur ou de taille avec la cible. Quand la tâche était difficile, deux distracteurs au hasard étaient de même taille que la cible mais de couleur différente, et les deux autres étaient de même couleur mais de taille différente (exemples en figure 17). Huit parties de soixante-quatre essais chacune étaient présentées à chaque participant selon les mêmes contraintes que dans l'expérience 5.

Mesures dépendantes. Les mesures dépendantes étaient similaires à celles de l'expérience 5. Même si la complexité visuelle de l'arrière-plan n'a pas été manipulée dans cette expérience, la durée moyenne des fixations oculaires a été enregistrée car, selon Rayner (2009), elle peut refléter la difficulté du processus de recherche d'une cible dans une scène visuelle. Ces variables ont été analysées en utilisant des ANOVA pour mesures répétées avec le mouvement de l'arrière-plan et le niveau de difficulté de la tâche en facteurs intra-participants.

3.2. Résultats

Des analyses séparées ont été réalisées pour chacun des deux types de décor (couleur et noir et blanc) utilisés pour les arrière-plans.

Les réponses tardives ont été exclues de l'analyse comme pour l'expérience 5. Un participant a été exclu de l'analyse parce qu'il avait plus de 15% d'essais exclus selon ce critère. Pour les 21 autres participants, le nombre d'essais exclus représentaient en moyenne 2,5% du nombre total d'essais (min = 0,0%, max = 9,8%) pour le décor en couleur, et 2,8% (min = 0,0%, max = 9,8%) pour le décor en noir et blanc. Les temps de réponse ont été transformés logarithmiquement avant que les ANOVA ne soient réalisées. Pour la clarté de la présentation, ce sont les moyennes non-transformées qui sont rapportées dans le texte et les graphiques.

3.2.1. Temps de réponse moyen

Décor en noir et blanc. Comme le montre la figure 18, la difficulté de la tâche avait une influence sur le temps de réponse, $F(1, 20) = 199{,}13$, $p < 0{,}001$, $\eta^2_p = 0{,}91$. Le temps de réponse en condition difficile (M = 1586 ms, ET = 240) était plus élevé qu'en condition facile (M = 1381 ms, ET = 239). Le mouvement de l'arrière-plan avait également une influence sur le temps de réponse, $F(1, 20) = 8{,}16$, $p < 0{,}01$, $\eta^2_p = 0{,}29$. De plus, l'interaction entre les facteurs mouvement de l'arrière-plan et difficulté de la tâche était significative, $F(1, 20) = 6{,}90$, $p < 0{,}05$, $\eta^2_p = 0{,}26$. Les comparaisons planifiées ont montré que le mouvement de l'arrière-plan n'avait un effet que lorsque la tâche était facile. Dans ce cas, le temps de réponse en condition de mouvement de l'arrière-plan était plus long qu'en condition statique, $F(1, 20) = 14{,}80$, $p < 0{,}05$. En revanche, quand la tâche était difficile, les temps de réponse obtenus en présence ou en absence de mouvement de l'arrière-plan n'étaient pas significativement différents, $F(1, 20) = 1{,}71$, $p = 0{,}21$.

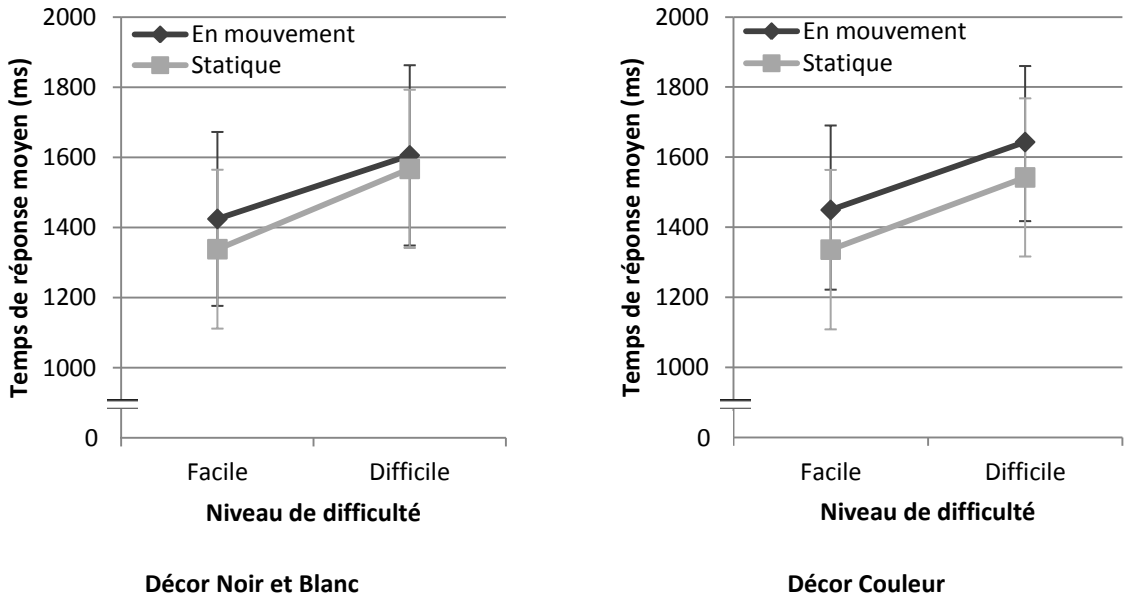


Figure 18. Temps de réponse moyen (en ms) pour toucher chaque cible dans l'expérience 6 dans chaque condition (difficulté de la tâche X mouvement).
Les barres d'erreur représentent ±1 écart-type de la moyenne.

Décor en couleur. Comme le montre la figure 18, la difficulté de la tâche avait également une influence sur le temps de réponse, $F(1, 20) = 188{,}00$, $p < 0{,}001$, $\eta^2_p = 0{,}90$. Le temps de réponse en

condition difficile (*M* = 1592 ms, *ET* = 225) était plus long qu'en condition facile (*M* = 1392 ms, *ET* = 239). Le mouvement de l'arrière-plan avait également une influence sur le temps de réponse, $F(1, 20) = 35{,}50$, $p < 0{,}001$, $\eta^2_p = 0{,}64$. Le temps de réponse en condition de mouvement (*M* = 1546 ms, *ET* = 247) était plus long qu'en condition statique (*M* = 1439 ms, *ET* = 247). L'interaction entre les facteurs mouvement de l'arrière-plan et difficulté de la tâche n'était cependant pas significative, $F(1, 20) < 1$.

3.2.2. Durée moyenne des fixations oculaires

Décor en noir et blanc. La difficulté de la tâche avait une influence sur la durée moyenne des fixations, $F(1, 20) = 5{,}55$, $p < 0{,}05$. La durée des fixations était plus longue dans la condition difficile (*M* = 195 ms, *ET* = 56) que dans la condition facile (*M* = 187 ms, *ET* = 47). Le mouvement de l'arrière-plan n'avait par contre pas d'effet significatif, $F(1, 20) < 1$. L'interaction entre ces deux facteurs n'était pas non plus significative.

Décor en couleur. La durée des fixations n'était pas significativement plus longue dans la condition difficile (*M* = 201 ms, *ET* = 55) que dans la condition facile (*M* = 197 ms, *ET* = 64), $F(1, 20) = 1{,}26$, $p = 0{,}27$. Le mouvement de l'arrière-plan n'avait pas d'influence significative sur la durée des fixations, $F(1, 20) < 1$. L'interaction entre ces deux facteurs n'était pas significative.

3.2.3. Fixation initiale

Décor en noir et blanc. Ni le mouvement de l'arrière-plan, $F(1, 20) < 1$, ni la difficulté de la tâche, $F(1, 20) = 1{,}03$, $p = 0{,}32$, n'avaient d'influence sur la coordonnée horizontale (X) de la fixation initiale. L'interaction entre ces deux facteurs n'était pas non plus significative.

De même, ni le mouvement de l'arrière-plan, $F(1, 20) < 1$, ni la difficulté de la tâche, $F(1, 20) < 1$, n'avaient d'influence sur la coordonnée verticale (Y) de la fixation initiale. L'interaction entre ces deux facteurs n'était pas non plus significative.

Décor en couleur. Ni le mouvement de l'arrière-plan, $F(1, 20) = 1{,}51$, $p = 0{,}24$, ni la difficulté de la tâche, $F(1, 20) < 1$, n'avaient d'influence sur la coordonnée horizontale (X) de la fixation initiale. L'interaction entre ces deux facteurs n'était pas non plus significative.

De même, ni le mouvement de l'arrière-plan, $F(1, 20) < 1$, ni la difficulté de la tâche, $F(1, 20) = 4{,}01$, $p = 0{,}06$, n'avaient d'influence sur la coordonnée verticale (Y) de la fixation initiale. L'interaction entre ces deux facteurs n'était pas non plus significative.

3.3. Discussion

La première hypothèse a été vérifiée. La performance était plus faible quand le niveau de difficulté de la tâche était plus élevé. En accord avec les précédentes études sur la recherche visuelle de « conjonction de caractéristiques » (Reali et al., 2006 ; Thornton & Gilden, 2007), les temps de réponse pour toucher une cible partageant une caractéristique visuelle de taille ou de couleur avec les distracteurs étaient plus longs que pour une cible ne partageant aucune caractéristique avec les distracteurs, et ce quel que soit le type de décor. L'analyse des mouvements oculaires a montré que la durée moyenne des fixations était plus longue lorsque le niveau de difficulté était plus élevé, conformément à la littérature (Rayner, 2009), mais uniquement pour l'arrière-plan en noir et blanc. Sur l'arrière-plan en couleur, la différence n'était qu'une tendance.

La seconde hypothèse était que la performance est plus faible quand l'arrière-plan est en mouvement uniquement si la tâche est facile. Elle a été vérifiée avec le décor en noir et blanc. Le temps de réponse était plus long quand l'arrière-plan était en mouvement que quand il était statique uniquement quand la tâche était facile. Il n'y avait pas de différence significative de temps de réponse quand la tâche était difficile. L'hypothèse n'a en revanche pas été vérifiée avec le décor en couleur. Seul un effet simple de ralentissement de la réponse en cas de mouvement de l'arrière-plan a été montré, indépendamment du niveau de difficulté de la tâche. Cependant, pour les deux types de décor, l'analyse des mouvements oculaires a confirmé la présence probable d'un OKN avec les

arrière-plans en mouvement, en tout cas de la déviation du regard vers l'origine du mouvement normalement associée à l'OKN.

Les expériences du précédent chapitre avaient montré qu'une tâche complexe en termes de demandes attentionnelles pouvait annuler les effets négatifs du mouvement de l'arrière-plan si elle impliquait des objets visuels fixes superposés à l'arrière-plan. L'expérience présente a en partie répliqué ce résultat (décor en noir et blanc) dans un environnement et une tâche spécifiques des jeux vidéo. Cependant, l'interaction n'a pas été retrouvée sur le décor en couleur. Le niveau de difficulté de la tâche difficile n'était peut-être pas assez suffisant pour que l'interaction entre le mouvement de l'arrière-plan et la difficulté de la tâche soit observée quelles que soit les caractéristiques visuelles de l'arrière-plan utilisé.

4. Discussion générale

Les expériences présentées dans ce chapitre ont confirmé ou affiné deux éléments importants abordés dans le chapitre précédent concernant l'influence du mouvement de l'arrière-plan sur la performance des joueurs. Elles ont permis d'étudier plus précisément les effets conjoints du mouvement et de la complexité de l'arrière-plan d'une part, et du mouvement de l'arrière-plan et de la difficulté de la tâche d'autre part, dans des environnements plus spécifiques des jeux vidéo.

Premièrement, l'expérience 5 a montré que dans un environnement de jeu vidéo, un mouvement latéral de l'arrière-plan dégrade la performance du joueur quelle que soit la complexité visuelle de celui-ci, comme l'ont déjà montré de précédentes recherches (Harrison et al., 2010 ; Kaminiarz et al., 2007 ; Tozzi et al., 2007). Toutefois, cette dégradation se produit seulement si cette complexité est intrinsèque à l'arrière-plan et n'augmente pas le nombre d'éléments saillants fixes de la scène visuelle. Dans certains cas, la complexité pouvait même accentuer la dégradation de la performance par le mouvement. Autrement dit, les expériences 4 et 5 indiquent que les points saillants fixes d'un

arrière-plan par ailleurs en mouvement peuvent être utilisés comme point de fixation par le regard du joueur et permettre d'annuler l'OKN et ses effets négatifs sur la performance (Ilg, 1997 ; Pola et al., 1992 ; Rubinstein & Abel, 2011 ; Wyatt & Pola, 1984). Si les points saillants qui créent la complexité sont en mouvement avec l'ensemble de l'arrière-plan, le joueur ne peut pas annuler l'OKN et les effets négatifs sur la performance demeurent.

Deuxièmement, l'expérience 6 a confirmé que dans un environnement et avec une tâche spécifiques des jeux vidéo, le mouvement de l'arrière-plan ne dégrade systématiquement la performance du joueur que si la tâche est simple en termes de demandes attentionnelles. Ceci confirme les résultats des expériences 3a et 3b qui ont montré que lorsque la tâche implique une attention soutenue sur des éléments visuels fixes disposés sur un arrière-plan en mouvement (tâche de recherche visuelle), l'individu peut annuler l'OKN et ses effets négatifs sur la performance, comme cela a déjà été montré dans une tâche de lecture (Menozzi & Koga, 2004).

Une limite importante peut toutefois nuancer les résultats obtenus dans les expériences de ce chapitre. L'influence sur la performance des joueurs de la complexité de l'arrière-plan d'une part (expérience 5) et de la difficulté de la tâche d'autre part (expérience 6), n'a pas été montrée quel que soit le type de décor utilisé. Dans l'expérience 5, quand le décor était non-naturel (composé d'objets abstraits), l'effet simple de la complexité de l'arrière-plan était significatif. Quand le décor était naturel (composé d'arbres), l'effet simple n'était pas significatif, mais l'interaction avec le mouvement de l'arrière-plan l'était. Dans l'expérience 6, quand le décor était en couleur (à dominante verte), l'effet simple de la difficulté de la tâche était significatif. Quand le décor était en noir et blanc, l'interaction avec le mouvement de l'arrière-plan était significative en plus de l'effet simple. Ces différences de résultats ne sont dues qu'à la nature des décors utilisés puisque tous les autres paramètres des expériences étaient strictement identiques. Parmi les raisons possibles de cette différence, on peut citer pour l'expérience 5 les différentes dispositions des objets (arbres ou cubes) sur le motif de base qui ont pu créer des contrastes différents, ou les teintes différentes des

motifs de base et/ou des objets superposés. Cette différence de contrastes a pu faire varier la difficulté de la tâche de perception. Pour l'expérience 6, la colorisation est la seule raison possible pour expliquer la différence de résultats. Pourtant, la couleur verte a été choisie pour sa faible capacité à faire varier les états d'éveil mental ou émotionnels de l'observateur, par rapport à celle des couleurs bleue ou rouge par exemple (Bonnardel, Piolat, & Le Bigot, 2011 ; Valdez & Mehrabian, 1994 ; Wolfson & Case, 2000). Dans le cas présent, il semble que la couleur verte ait tout de même eu une influence significative sur les résultats. Des expériences supplémentaires sont nécessaires pour déterminer les raisons de cette non-réplication de résultats. Pour comprendre l'influence de la nature des décors sur les résultats, il faudrait manipuler les propriétés visuelles de bas niveau évoquées ci-dessus pour chacune des expériences.

Chapitre 7 :
Discussion générale

La thèse défendue dans ce document est qu'un mouvement constant de l'arrière-plan, ou des éléments constitutifs de l'arrière-plan, modifie les mouvements du regard et dégrade la performance du joueur dans un environnement virtuel de type jeu vidéo. Les résultats des expériences présentées ont confirmé, en accord avec les recherches précédentes (e.g., Harrison et al., 2010 ; Ilg, 1997 ; Kaminiarz et al., 2007 ; Tozzi et al., 2007), qu'un arrière-plan en mouvement modifie le comportement du regard du joueur et dégrade par conséquent la performance dans une tâche de jeu vidéo simple. L'originalité de ce travail de thèse est de montrer que plusieurs éléments déterminants des interfaces visuelles des jeux vidéo, comme le positionnement des informations contextuelles, la structuration et la complexité de l'arrière-plan, ou la difficulté de la tâche, peuvent moduler, atténuer ou supprimer ces effets du mouvement de l'arrière-plan sur le comportement du regard et la dégradation des performances.

La première section de la discussion reprend les principaux éléments de la partie expérimentale qui répondent aux questions posées dans la thèse. La discussion de ces résultats permet de proposer un nouveau cadre théorique pour la modélisation du partage attentionnel dans des interfaces visuelles de jeux vidéo. Quelques limites notables sont également indiquées. La deuxième section propose des

recommandations ergonomiques pour la conception des jeux vidéo, mais aussi des environnements virtuels en général, en fonction des résultats des expériences. Enfin, des perspectives de recherche sont avancées dans la troisième section.

1. Partage attentionnel dans les interfaces visuelles de jeux vidéo

Dans cette section, est discutée la question théorique de la thèse qui est de savoir comment prédire et expliquer le partage de l'attention d'un opérateur entre deux sources d'information présentes dans des interfaces visuelles d'environnements virtuels de type jeux vidéo. La partie théorique a montré que les théories existantes ne permettaient de répondre que partiellement à cette question. Les résultats des expériences de la thèse permettent d'apporter une réponse plus complète. Une proposition de modèle du partage attentionnel lors de l'interaction avec une interface visuelle de jeux vidéo, qui complète le modèle des ressources multiples de Wickens (2002, 2008), est ensuite exposée. Enfin, les éléments non intégrés dans ce modèle sont discutés dans le but d'indiquer les travaux encore nécessaires pour pouvoir généraliser les résultats obtenus.

1.1. Modèles et théories existants

La disposition spatiale des sources visuelles d'information est un facteur essentiel pour prédire le partage attentionnel. Les expériences du chapitre 4 ont montré que, dans un jeu qui requiert une anticipation visuelle d'un mouvement de bas en haut, l'anticipation des joueurs et la performance augmentaient significativement quand le score était affiché en dessous de la fenêtre de jeu au lieu d'être affiché en haut de cette fenêtre. Bien que suivre le score fasse partie du jeu, une meilleure performance était obtenue pour cette position du score, qui n'implique aucun chevauchement entre la position du score et la zone d'anticipation visuelle, que quand le score était placé en bas de la fenêtre de jeu directement dans la zone d'anticipation.

Ce résultat montre que le principe de proximité-compatibilité (Wickens & Carswell, 1995) s'applique aux jeux vidéo dynamiques, mais seulement dans des contextes d'absence de chevauchement visuel entre la zone d'anticipation et les informations qui doivent être suivies simultanément. La proximité spatiale entre ces sources d'informations faciliterait le partage de l'attention seulement en l'absence d'interférence provoquée par le chevauchement. Ce résultat semble également être en accord avec le modèle des ressources multiples de Wickens (2002, 2008). La modalité visuelle du modèle inclut deux canaux attentionnels différents, à savoir les canaux focal et ambiant (voir la description du modèle p. 32). Dans un jeu vidéo tel que celui utilisé dans le chapitre 4, des informations disposées à des endroits différents de l'interface de jeu pourraient être traitées par des canaux attentionnels différents. Le canal focal traiterait l'information principale (e.g., la scène d'action) et le canal ambiant traiterait les informations contextuelles (e.g., le score). Aussi, lorsque les sources d'information ne sont pas superposées, le modèle des ressources multiples serait suffisant pour prédire le partage attentionnel entre ces deux sources. Des ressources différentes étant allouées à ces deux canaux, aucune baisse de performance n'est attendue.

La prédiction devient plus difficile lorsque les sources d'information sont superposées dans l'espace et ne peuvent pas être séparées. Les expériences des chapitres 5 et 6 ont étudié ce problème. L'unique prédiction possible à partir du modèle des ressources multiples est que les ressources attentionnelles sont partagées entre les sources. Cependant, savoir comment ce partage affecte le suivi de chacune des sources nécessite de prendre en compte d'autres théories. Dans les interfaces des jeux vidéo, les objets avec lesquels le joueur doit interagir et les arrière-plans ne peuvent pas être séparés spatialement. Ils sont donc obligatoirement traités visuellement par le même canal attentionnel, que ce soit par le canal visuel focal ou par le canal visuel ambiant. Dans les expériences de la thèse, la mesure de la dégradation de la tâche principale de jeu a permis de comprendre comment l'attention est partagée entre les arrière-plans et les objets superposés en fonction de la complexité de la tâche et de deux caractéristiques visuelles importantes de l'arrière-plan, son mouvement et sa complexité. Pour comprendre la dégradation de la performance dans ces

situations, deux cadres théoriques supplémentaires ont été nécessaires : les connaissances sur l'OKN et les conditions de son annulation (e.g., Ilg, 1997 ; Rubinstein & Abel, 2011 ; Wyatt & Pola, 1984) et la théorie de la recherche visuelle guidée adaptée aux scènes complexes (Wolfe, 1994 ; Wolfe et al., 2002).

1.2. Modèle de partage attentionnel entre deux sources visuelles d'information superposées

Les résultats des chapitres 5 et 6 ont permis d'élaborer un modèle simple de prédiction d'allocation de ressources lorsque deux sources d'information (un arrière-plan et des objets fixes superposés) sont traitées par le même canal visuel, a priori le canal visuel focal (tableau 4). Ce modèle prend en compte les caractéristiques du mouvement et de la structure associée de l'arrière-plan, ainsi que la complexité de la tâche principale à réaliser avec les objets à l'écran. Quatre éléments fondamentaux, issus des expériences de la thèse et des recherches antérieures, guident ce modèle.

Tableau 4. *Prédiction de la diminution des ressources attentionnelles allouées aux objets (fixes) impliqués dans la tâche principale au bénéfice des éléments de l'arrière-plan, en fonction des caractéristiques du mouvement et de la structure associée de l'arrière-plan sur lequel ils sont superposés, ainsi que de la complexité de la tâche.*

	Arrière-plan statique		**Arrière-plan en mouvement**	
	Structure latérale	**Structure radiale**	**Structure latérale**	**Structure radiale**
Tâche simple	0	-1	-1	*-1*
Tâche complexe	0	-1	*0*	*-1*

Note. Le score de prédiction 0 signifie qu'aucune baisse de ressources allouées aux objets n'est prédite par rapport au niveau de référence (situation de tâche simple et arrière-plan sans motif ni structure). Le score de prédiction -1 signifie qu'une baisse de ressources est prédite par rapport au niveau de référence. Les scores de prédiction en italique signifient qu'une baisse ou une baisse plus forte est empêchée du fait de l'annulation de l'OKN par une attention soutenue sur des points fixes présents dans la scène.

Premièrement, ce modèle part du principe que dans la condition la plus simple, telle que celle utilisée dans l'expérience 4 (tâche simple et arrière-plan statique sans motif ni structure), la majorité des ressources attentionnelles est allouée à la tâche principale. La quantité de ressources allouée dans ce cas constitue le niveau de référence dans le modèle.

Deuxièmement, un mouvement de l'arrière-plan nécessite d'allouer des ressources au traitement de ce mouvement au détriment de la tâche principale. Les expériences 4 et 6 ont montré que dans une tâche simple de visée de cible typique des jeux vidéo, la performance était dégradée par un mouvement latéral de l'arrière-plan. Ces résultats confirment de précédentes recherches sur l'OKN (Harrison et al., 2010 ; Kaminiarz et al., 2007 ; Tozzi et al., 2007) qui ont montré que le mouvement d'ensemble de l'arrière-plan diminuait la performance dans des tâches de détection exigeant peu de ressources attentionnelles, quand la structure de l'arrière-plan était latérale. Cette baisse de performance peut aussi être expliquée par le fait que le mouvement de l'arrière-plan serait un élément distracteur qui perturberait l'attention portée à la tâche principale. La conclusion et l'utilisation de ce résultat n'en sont cependant pas modifiées.

Troisièmement, une structuration complexe de l'arrière-plan nécessite également d'allouer des ressources au traitement des éléments constituant cet arrière-plan, au détriment de la tâche principale. Les expériences du chapitre 5 ont montré que quelle que soit la tâche principale du joueur, la complexité structurale avait une influence sur la performance du joueur. La performance du joueur était dégradée par une forte complexité structurale (structure radiale) par rapport à une complexité nulle (arrière-plan uni). Par contre, la performance n'était pas dégradée significativement lorsque la complexité était faible (structure latérale). Ces résultats sont en accord avec les théories de guidage de l'attention (Treisman & Gelade, 1980 ; Wolfe, 1994 ; Wolfe et al., 2002), qui montrent que le traitement des informations est ralenti quand la complexité de la scène visuelle, et notamment de son arrière-plan, augmente. Comme pour le mouvement de l'arrière-plan, la complexité structurale peut être considérée comme un élément distracteur qui perturberait l'attention portée à la tâche principale et impliquerait cette baisse de performance.

Quatrièmement, lorsque plusieurs points fixes sont présents sur un arrière-plan en mouvement, les ressources disponibles restent allouées au traitement de ces points fixes. Ceux-ci peuvent être liés à la complexité de la tâche (plusieurs objets à traiter séquentiellement) ou à la complexité structurale

du mouvement de l'arrière-plan (point et lignes de fuite). Dans ce cas, aucune ressource n'est allouée au traitement du mouvement de l'arrière-plan (ou l'attention portée à la tâche principale ne serait pas perturbée par ce mouvement). Les expériences du chapitre 5 ont montré que, dans une tâche simple de visée de cible typique des jeux vidéo, la performance était dégradée par un mouvement latéral de l'arrière-plan, mais pas par un mouvement radial. Par contre, ni le mouvement latéral, ni le mouvement radial ne dégradaient significativement la performance de recherche visuelle, qui est une tâche plus exigeante en termes de ressources attentionnelles. Pour cette tâche, seule la plus grande complexité structurale de l'arrière-plan (structure radiale) avait une influence sur la performance. De même, l'expérience 6 a montré que dans une tâche complexe de visée de cible, la performance n'était pas dégradée par le mouvement latéral de l'arrière-plan. Aussi, plus la quantité de ressources allouées aux éléments fixes superposés à l'arrière-plan latéral ou aux éléments fixes constitutifs de la structure radiale est importante, plus les effets négatifs de l'OKN et de la sensation de vection induits par le mouvement de l'arrière-plan sur la performance pourraient être atténués, voire annulés (Ilg, 1997 ; Menozzi & Koga, 2004 ; Rubinstein & Abel, 2011 ; Seno et al., 2011).

1.3. Eléments non pris en compte dans le modèle

Deux résultats des expériences de la thèse n'ont pas été pris en compte dans l'élaboration de ce modèle, en raison des limites détaillées dans les paragraphes suivants. Ces résultats sont d'une part l'influence de la complexité intrinsèque de l'arrière-plan (expérience 5) et d'autre part l'influence de la présence d'un indice sonore sur la performance du joueur (expérience 4).

L'expérience 5 a montré que dans un environnement de jeu vidéo, un mouvement de type latéral de l'arrière-plan dégrade la performance du joueur quelle que soit la complexité visuelle de cet arrière-plan, mais seulement si la tâche est simple, et seulement si cette complexité est intrinsèque à l'arrière-plan et n'augmente pas le nombre d'éléments fixes de la scène visuelle. Si les points saillants qui créent la complexité de l'arrière-plan sont en mouvement avec l'ensemble de l'arrière-plan, le joueur ne peut pas annuler l'OKN et les effets négatifs sur la performance demeurent. Ces résultats

sont en accord avec les théories de guidage de l'attention (Treisman & Gelade, 1980 ; Wolfe, 1994 ; Wolfe et al., 2002), qui montrent que le traitement des informations est ralenti quand la complexité de la scène visuelle augmente. Ce résultat n'a pas été pris en compte dans le modèle pour deux raisons.

Premièrement, l'influence de la complexité intrinsèque de l'arrière-plan sur la performance des joueurs n'a pas été montrée quel que soit le type de décor utilisé. Quand le décor était non-naturel (composé d'objets abstraits), l'effet simple de la complexité de l'arrière-plan était significatif. Quand le décor était naturel (composé d'arbres), l'effet simple n'était pas significatif, mais l'interaction avec le mouvement de l'arrière-plan l'était. Cependant, dans les deux cas, quand la complexité de l'arrière-plan n'implique pas de points fixes lorsqu'il est en mouvement, l'OKN et ses effets négatifs ne sont pas annulés. Ils confirment les conclusions du chapitre 5 en démontrant que seule une forte complexité liée à la structure de l'arrière-plan, qui présente des points fixes saillants, permet d'annuler l'OKN et ses effets négatifs. Des expériences supplémentaires sont nécessaires pour déterminer l'influence des caractéristiques visuelles des scènes complexes sur le traitement attentionnel de ces éléments et, par conséquent, sur la performance des individus. Ces expériences seraient basées sur des théories élaborées notamment en psychophysique, dans lesquelles la perception des formes des objets, des contrastes, de la luminosité ou encore des couleurs, est mieux prise en compte (e.g., Graham, 2011 ; Kingdom, 2011).

Deuxièmement, l'influence de la complexité intrinsèque de l'arrière-plan (non liée à la structure visuelle qui génère son mouvement apparent) n'a pas été testée dans le cas d'un mouvement de type radial. L'expérience 5 a montré que dans le cas d'un mouvement de type latéral, cette complexité a une influence sur la performance des joueurs. Une expérience supplémentaire est nécessaire pour vérifier que cette complexité, qui s'ajouterait à la complexité structurale, a également une influence dans le cas d'un mouvement radial. D'après les résultats des expériences de la thèse et de la littérature exposés plus haut, cette influence semble toutefois très probable.

Enfin, les résultats de l'expérience 4 n'ont pas montré d'influence particulière d'un indice sonore sur la performance quand l'arrière-plan est en mouvement et/ou complexe. Pourtant, un des objectifs de cette expérience était de proposer une aide de type sonore et verbale à l'opérateur dans une situation dynamique complexe. Même si la présence d'un indice sonore améliorait la performance des participants, cette amélioration n'était pas modulée par la complexité ou le mouvement de l'arrière-plan. Les données des mouvements oculaires laissent toutefois penser que cette modulation pourrait exister dans des conditions expérimentales différentes. Des expériences supplémentaires sont nécessaires pour étudier l'apport de la multimodalité dans des environnements dynamiques complexes.

Malgré ces limites, les résultats obtenus dans la thèse sont suffisamment généralisables pour permettre plusieurs recommandations aux concepteurs d'environnements virtuels, et plus précisément aux concepteurs de jeux vidéo.

2. Recommandations pour la conception des jeux vidéo

Dans le domaine scientifique de l'interaction personne-système, les études telles que celles présentées dans la thèse participent à l'enrichissement des connaissances théoriques, mais également à l'amélioration des méthodes de conception des systèmes existants. Après un rappel des principales méthodes utilisées dans la conception des jeux vidéo centrée sur le joueur, quelques recommandations sont proposées.

Actuellement, plusieurs méthodes peuvent être utilisées par les concepteurs de jeux vidéo dans le cadre d'une approche centrée-utilisateur. Même s'il n'existe pas de consensus en ce qui concerne les critères d'étude de l'interaction personne-système ou de l'interaction joueur-jeu vidéo (e.g., utilisabilité, expérience utilisateur, jouabilité), la plupart des méthodes, comme le test utilisateur et l'évaluation heuristique, restent malgré tout communes à tous les concepteurs (Baccino, Bellino, &

Colombi, 2005 ; Hornbæk, 2006). Le test utilisateur, très utilisé (Bach & Scapin, 2010 ; Bastien, 2010), met en situation des participants (e.g., joueurs-cibles) face au système testé (Barendregt & Bekker, 2006 ; Barendregt, Bekker, Bouwhuis, & Baauw, 2006 ; Höysniemi, Hämäläinen, & Turkki, 2003). L'évaluation heuristique est toutefois une méthode plus accessible, notamment parce qu'elle ne nécessite pas l'intervention d'utilisateurs (Schaffer, 2008). Elle permet d'éprouver un système selon des règles (heuristiques, « guidelines ») testées et validées par ailleurs. Cependant, les règles utilisées pour les jeux vidéo (e.g., Desurvire, Caplan, & Toth, 2004 ; Desurvire & Wiberg, 2009 ; Koeffel et al., 2010 ; Pinelle, Wong, Stach, & Gutwin, 2009) sont en majorité des adaptations des règles traditionnelles appliquées aux autres interactions personne-système (e.g., heuristiques de Nielsen, 1993), ou issues d'observations subjectives de joueurs en situation.

Plusieurs recommandations, décrites dans les paragraphes suivants, ont été élaborées à partir des résultats des expériences de la thèse. Elles sont suffisamment générales pour être utilisées dans la conception de la plupart des jeux existants. Pour la majorité d'entre elles, elles sont également adaptées pour la conception d'environnements virtuels ou d'environnements dynamiques complexes.

2.1. Anticipation visuelle du mouvement et positionnement des informations contextuelles

Pour les jeux vidéo utilisés dans le chapitre 4 (expériences 1 et 2), positionner l'affichage du score en-dessous de la fenêtre de jeu aidait les joueurs à anticiper le mouvement des obstacles et à obtenir une meilleure performance. Ce résultat peut être généralisé aux autres jeux vidéo qui requièrent une anticipation visuelle des mouvements. La recommandation suivante est proposée.

- **Recommandation 1**

  Pour les jeux qui requièrent une anticipation visuelle des mouvements, les principales informations contextuelles doivent être placées dans la direction attendue de l'anticipation, mais ne doivent pas chevaucher les zones d'action principale et d'anticipation, ou leur être superposées.

De manière plus générale, les résultats de l'étude complètent le principe de proximité-compatibilité de Wickens et Carswell (1995) pour la conception d'interfaces visuelles dynamiques simples. La recommandation proposée ci-dessus peut être généralisée à tous les environnements interactifs qui requièrent de l'anticipation des opérateurs humains, par exemple pour la conception des postes de conduite automobile ou de pilotage d'avions.

2.2. Mouvement de l'arrière-plan, structure visuelle, et complexité de la tâche

Les résultats obtenus dans le chapitre 5 (expériences 3a, 3b et 4) permettent eux aussi de proposer des recommandations pour la conception des jeux vidéo et des environnements virtuels en général. Les structures de mouvement radial des arrière-plans sont de plus en plus utilisées dans les environnements virtuels pour simuler un déplacement de l'opérateur en profondeur (vue en « première personne »). Les résultats des expériences ont montré qu'elles rendent plus difficile la détection ou l'utilisation d'objets fixes. Par conséquent, la mise en place d'une troisième dimension a un coût pour la performance de l'opérateur. Cependant, dans ce cas, le mouvement de l'arrière-plan ne génère pas de coût supplémentaire. A l'inverse, une structure d'arrière-plan moins complexe (latérale), en mouvement ou non, a un coût inférieur sur la performance quand la tâche requiert beaucoup de ressources attentionnelles visuelles. Quand la tâche requiert peu de ressources ou ne permet pas d'annuler un OKN en l'absence de points fixes suffisamment pertinents, le mouvement latéral de l'arrière-plan doit être évité.

Ces recommandations nécessitent une adaptation quand l'environnement virtuel à concevoir est un jeu vidéo. En effet, les recommandations ergonomiques sont généralement proposées pour améliorer les interactions personne-système dans des tâches productives (Nielsen, 1993). Or, les jeux vidéo sont avant tout des objets de divertissement et le but principal du joueur n'est pas de produire quelque chose, mais de vivre une expérience spécifique qui inclut entre autres plaisir, amusement et émotions (Pagulayan et al., 2008). Pour permettre au joueur d'atteindre ce but, une recommandation générale de conception est de créer un challenge optimal pour le joueur (Qin et al., 2010).

Les résultats obtenus dans le chapitre 5 ont montré que les variations des propriétés visuelles de bas niveau des arrière-plans peuvent être utilisées pour faire varier la difficulté d'un jeu, comme Jie et Clark (2008) l'ont déjà montré avec la complexité visuelle. La combinaison du mouvement et de la complexité de la structure associée peut aussi être utilisée. Par ailleurs, les résultats de l'expérience 6 ont confirmé ceux des expériences 3a, 3b et 4, et montré les interactions entre la difficulté d'une tâche de jeu vidéo et les conséquences d'un mouvement latéral de l'arrière-plan. Les résultats de l'expérience 5 ont quant à eux montré que la complexité de la structure qui génère le mouvement apparent de l'arrière-plan et la complexité propre à l'arrière-plan (qui se déplace « avec » lui) doivent être considérées différemment dans le cas d'une tâche de jeu vidéo simple. Les recommandations suivantes ont été élaborées à partir de l'ensemble de ces résultats.

- **Recommandation 2a**

  Pour garder un niveau de difficulté faible quand la tâche principale est simple (en termes de demandes attentionnelles), les arrière-plans en mouvement, complexes ou avec une structure complexe (e.g., radiale) doivent être évités.

- **Recommandation 2b**

  Pour garder un niveau de difficulté faible quand la tâche principale est complexe, une structure complexe doit également être évitée, mais afficher un arrière-plan en mouvement ne devrait pas modifier la difficulté du jeu.

- **Recommandation 3a**

  Pour garder un niveau de difficulté plus élevé quand la tâche principale est simple, un arrière-plan en mouvement, complexe ou de structure visuelle complexe doit être affiché. Afficher un arrière-plan en mouvement ET complexe peut multiplier la difficulté. Par contre, afficher un arrière-plan en mouvement ET de structure complexe (e.g., radiale) n'augmente pas plus la difficulté.

- **Recommandation 3b**

  Pour garder un niveau de difficulté plus élevé quand la tâche principale est complexe, un arrière-plan à la structure complexe doit être affiché. Afficher un arrière-plan en mouvement n'augmente pas plus la difficulté.

3. Perspectives

Ce travail de thèse ouvre finalement plusieurs axes de recherche futurs qui peuvent être partagés en trois catégories : (a) valider les résultats obtenus dans des environnements interactifs simples avec des environnements plus complexes, similaires à des environnements réels, (b) étudier les effets des facteurs manipulés dans la thèse avec des environnements restitués sur des écrans de petite taille de

terminaux mobiles, et (c) explorer l'influence des caractéristiques des interfaces visuelles de jeux vidéo sur l'expérience subjective des joueurs.

Premièrement, les environnements utilisés dans les différentes études de la thèse ont été créés dans le but de reproduire de manière simplifiée des environnements virtuels ou de jeux vidéo réels. Ils restent malgré tout simples, et assez éloignés des jeux commerciaux. Par exemple, la complexité et le mouvement des arrière-plans utilisés dans la thèse ne sont pas représentatifs de l'ensemble des jeux existants. La complexité peut être plus hétérogène, et le mouvement n'est pas forcément constant ou uniquement de type latéral ou radial. Dans des jeux où le monde représenté est en 3 dimensions, le mouvement général à l'écran peut être une composition complexe entre des mouvements de type latéral et de type radial. De futures recherches pourraient étudier les caractéristiques visuelles des interfaces dans des jeux ou des environnements virtuels proches de ceux existants, notamment ceux proposant des mondes visuels complexes en 3 dimensions. Une de ces recherches pourrait avoir pour objectif de valider les différentes recommandations de conception proposées dans la thèse dans un jeu vidéo écologique plus complexe visuellement que le jeu simple utilisé dans la thèse, comme un jeu de tir en vue subjective.

Deuxièmement, les joueurs peuvent jouer depuis plusieurs années avec des systèmes mobiles spécifiques aux jeux vidéo, comme les consoles de jeux portables, ou avec d'autres systèmes comme les téléphones mobiles (Browne & Anand, 2012). La particularité de ces terminaux est qu'ils intègrent l'artefact, et l'écran qui restitue l'interface visuelle du jeu. Ces systèmes mobiles sont de plus en plus utilisés dans le cadre du jeu vidéo (Feijoo, Gómez-Barroso, Aguado, & Ramos, 2012 ; Liang & Yeh, 2011 ; Y. Liu & Li, 2011). Une des raisons de ce succès est que les possibilités techniques de conception sont de plus en plus développées, notamment en ce qui concerne les aspects graphiques des interfaces (Chehimi, Coulton, & Edwards, 2008). Toutefois, la réduction des écrans de jeux de ces systèmes par rapport aux systèmes fixes (e.g., ordinateurs, téléviseurs) pose des problématiques supplémentaires concernant la conception des interfaces visuelles. La thèse a montré que les

caractéristiques de la scène principale d'action, notamment l'arrière-plan, et de l'affichage tête-haute, ont un impact sur le partage attentionnel entre les sources d'information et, par conséquent, sur la performance du joueur. De nouvelles études devraient explorer l'impact de ces caractéristiques sur des interfaces visuelles de petite taille. Concernant le positionnement de l'affichage tête-haute, le chevauchement d'une information contextuelle sur la scène d'action principale peut, par exemple, s'avérer inévitable. Plusieurs expériences devraient être menées pour manipuler différentes techniques d'affichage de cette information (e.g., affichage par intermittence, intégration de l'information dans la scène principale) tout en gardant la performance et l'expérience du joueur optimales. Concernant les caractéristiques des arrière-plans, le mouvement ou la complexité peuvent ne pas perturber ou perturber différemment l'attention du joueur. Par exemple, toujours du fait de la petite taille de l'écran, un mouvement d'ensemble de l'arrière-plan pourrait ne pas déclencher de perception de déplacement de soi (vection) ou de mouvements réflexes du regard (OKN).

Enfin, troisièmement, les différentes études de la thèse se sont focalisées sur l'influence des caractéristiques des interfaces visuelles sur la performance des joueurs. Ceci a permis notamment de proposer des solutions de conception qui sont utilisables par les concepteurs d'autres environnements virtuels que les jeux vidéo. Toutefois, un élément important des jeux vidéo n'a pas été étudié dans les expériences de la thèse. Le joueur de jeux vidéo ne cherche pas systématiquement à être performant, mais cherche également à vivre une expérience particulière. Du point de vue de l'expérience utilisateur, les ressentis subjectifs ont une part importante dans l'interaction joueur-jeu vidéo. De futures études pourraient étudier l'influence des caractéristiques visuelles des interfaces des jeux vidéo sur les ressentis subjectifs des joueurs, comme le plaisir de jeu ou les émotions. Le plaisir de jeu correspond aux réactions positives des joueurs générées par une session de jeu (Fang, Chan, Brzezinski, & Nair, 2010 ; Klimmt, Hartmann, & Frey, 2007 ; Trepte & Reinecke, 2010). Comme dans toute activité de loisir et de divertissement, il est généralement lié aux émotions (positives et négatives) ressenties par le joueur pendant et après une session de jeu

(Mandryk, Inkpen, & Calvert, 2006). Plusieurs auteurs ont étudié ces émotions pour améliorer la conception des jeux vidéo. Ils ont principalement utilisé des mesures psychophysiologiques, comme la réponse électrodermale, l'électrocardiographie ou l'électromyographie, pour étudier et observer les émotions que peuvent ressentir les joueurs en situation de jeu (Chanel et al., 2011 ; Mandryk & Atkins, 2007 ; Mandryk et al., 2006 ; Nacke et al., 2010 ; Yannakakis et al., 2010). Une étude future pourrait avoir pour objectif principal de comprendre l'influence de la composition et de l'organisation spatiale des informations contextuelles sur l'expérience utilisateur dans les jeux vidéo, à l'aide de ce type de mesures. Les résultats de l'étude permettraient de mieux comprendre l'importance, pour une expérience utilisateur optimale, d'une conception centrée sur l'utilisateur des interfaces visuelles des jeux vidéo, mais également des environnements virtuels.

Conclusion

Les choix de conception des interfaces visuelles des jeux vidéo ont une influence sur le comportement du joueur. La manipulation des caractères dynamiques et de la complexité des scènes visuelles a un impact sur le comportement du regard du joueur et sur sa performance, qui dépend de la difficulté de la tâche principale à accomplir. Les jeux vidéo, comme l'ensemble des environnements virtuels et des environnements dynamiques complexes, sont mieux conçus en prenant en compte ce type de résultats. Ceci confirme que la connaissance de l'humain et de son fonctionnement est indispensable aujourd'hui pour la conception des systèmes interactifs. La psychologie ergonomique, et plus généralement l'ergonomie, y a toute sa place, et la prise en compte de notions comme l'utilisabilité est pertinente lorsque ces systèmes sont utilisés dans le cadre de tâches liées à la productivité.

La question reste posée lorsque les systèmes à concevoir sont liés au divertissement, comme les jeux vidéo. L'intégration de l'ergonomie dans la conception des jeux vidéo peut être un problème, puisque le but d'un joueur est différent de celui d'un utilisateur d'un système lié à la productivité. Ce qui est vrai pour un système lié à la productivité (e.g., un système optimal est un système simple d'utilisation) ne l'est pas forcément pour un jeu vidéo. En réalité, L'objectif principal du concepteur de jeux vidéo est de divertir le joueur en lui proposant un challenge suffisamment motivant pour

garder son engagement dans le jeu, et non de le rendre plus simple d'utilisation. L'objectif est atteint si les difficultés rencontrées par le joueur sont liées à ce challenge, et non à des éléments non prévus par le concepteur et issus d'une mauvaise conception du jeu.

Une des problématiques actuelles des concepteurs est la manière dont le challenge doit s'adapter au joueur (Qin et al., 2010). Par exemple, les joueurs n'ont pas tous le même niveau d'expertise de jeu ou les mêmes attentes. Une question des chercheurs dans ce domaine est de savoir comment les « modèles utilisateurs » doivent être implémentés lors de la conception d'un jeu pour permettre une adaptation optimale du challenge au joueur, notamment au cours du jeu. Cette question a été abordée récemment avec l'étude du processus de « difficulté ajustée dynamiquement », déjà intégré dans certains jeux (C. Liu et al., 2009). Dans ce cas, la difficulté varie au cours du temps en fonction des succès et des échecs du joueur.

Aussi, la psychologie ergonomique et l'ergonomie ont toute leur place dans la conception des jeux vidéo, comme pour les autres systèmes interactifs. Même si les problématiques diffèrent, l'objectif du concepteur est toujours de proposer une interaction optimale entre un humain et un système.

Le défi actuel des ergonomes du jeu vidéo est différent, et peut-être plus complexe que la pertinence de l'ergonomie dans la conception d'un jeu vidéo. Même si l'ergonomie a sa place dans la conception d'un jeu vidéo, la question de savoir à quel moment elle doit intervenir est posée. Deux manières de répondre à cette question existent. L'ergonomie peut être présente dès l'origine de la création du jeu, ou seulement dans le cadre de la production du jeu. Dans le premier cas, l'ergonomie est présente dès que les idées préliminaires du jeu prennent forme dans l'esprit du créateur. Les capacités du futur joueur sont prises en compte dès les bases de la mécanique du jeu posées. Concrètement, il s'agirait d'intégrer les rôles du « game designer », qui est le responsable de ces aspects du jeu dans un processus habituel de conception, et de l'ergonome. Un argument fort dans ce sens est que, dans une équipe de conception, l'ergonome est celui qui connait le mieux l'humain et qui peut anticiper le comportement du joueur. Aussi, l'ergonome est le plus à même de prédire la

viabilité des idées du créateur. Dans le second cas, l'ergonomie est au service de la création, au même titre que les autres spécialités impliquées dans la conception d'un jeu (e.g., conception sonore, conception graphique, programmation informatique). Concrètement, à la manière d'un programmeur qui transcrit les idées des créateurs en exploitant le plus possible les capacités techniques des machines existantes, l'ergonome rend réalisables ces idées en optimisant les capacités mentales et physiques des joueurs et en compensant leurs limites.

Choisir entre ces deux possibilités n'est pas évident. Une analogie avec l'industrie cinématographique peut aider à se positionner. Schématiquement, deux types de conception d'un film existent aujourd'hui, le cinéma dit « d'auteur » et le cinéma dit « commercial ». Dans le premier cas le film est créé uniquement selon les idées du créateur, sans tenir compte des attentes du public et des conséquences commerciales. Dans le second cas, le but est de concevoir un film qui rassemble le plus de spectateurs possibles, en répondant à leurs attentes, et qui garantit le plus gros succès commercial. La conception d'un jeu vidéo peut être abordée sous chacun de ces deux angles. Dans un cas, si la vision du créateur est primordiale, l'ergonome, comme les autres membres de l'équipe de conception, travaille au service du créateur. Dans l'autre, si le jeu est créé pour répondre au désir de divertissement du plus grand nombre de joueurs, la conception est purement centrée-utilisateur et l'ergonome est présent dès la phase de création. En somme, ces deux visions pourraient coexister dans l'industrie du jeu vidéo, à la manière des deux types de créations cinématographiques exposés.

La réponse à la question du moment où l'ergonomie doit intervenir dans la conception d'un jeu vidéo a également une forte influence sur le panorama de la conception des jeux vidéo. Les deux points de vue discutés plus haut impliquent deux modes de répartition des connaissances en ergonomie entre les différents métiers des jeux vidéo. La vision actuelle est de garder ces connaissances propres à l'ergonome. Dans ce cas, l'ergonome maitrise parfaitement les théories et méthodes actuelles de la conception et de l'évaluation des systèmes interactifs en général (e.g., l'utilisabilité), et des systèmes liés au divertissement en particulier (e.g., l'expérience utilisateur). Ces connaissances permettent à

l'ergonome d'être au service de la création du jeu. Une vision alternative serait que l'ergonomie doit faire partie des connaissances communes à tous les membres d'une équipe de conception, y compris les créateurs. Les créateurs (i.e., ceux qui ont les idées initiales) devraient alors particulièrement maîtriser ces connaissances qu'ils prendraient en compte lors de leur processus créatif.

Au final, la psychologie ergonomique doit prendre sa place dans la conception d'un jeu vidéo, quelle qu'en soit la forme. Le jeu vidéo s'adresse à un joueur humain, et la psychologie ergonomique est la discipline qui connait le mieux l'humain et ses comportements lors d'interactions finalisées avec des systèmes. A partir du moment où des personnes se donnent pour objectif de concevoir un système, un objet ou un produit avec lequel un humain doit interagir, celles-ci devraient avoir à l'esprit les contraintes que cela implique. La conception des jeux vidéo centrée sur le joueur est d'autant plus importante qu'elle permet d'optimiser l'expérience du joueur, autrement dit le divertissement recherché.

Bibliographie

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly, 24*, 665-694.

Alexander, A. L., Wickens, C. D., & Hardy, T. J. (2005). Synthetic vision systems: The effects of guidance symbology, display size, and field of view. *Human Factors, 47*, 693-707.

Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., & Moret-Bonillo, V. (2010). A context-of-use taxonomy for usability studies. *International Journal of Human-Computer Interaction, 26*, 941-970.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., et al. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries: A meta-analytic review. *Psychological Bulletin, 136*, 151-173.

Ang, C. S., Zaphiris, P., & Mahmood, S. (2007). A model of cognitive loads in massively multiplayer online role playing games. *Interacting with Computers, 19*, 167-179.

Aponte, M.-V., Levieux, G., & Natkin, S. (2011). Measuring the level of difficulty in single player video games. *Entertainment Computing, 2*, 205-213.

Baccino, T., Bellino, C., & Colombi, T. (2005). *Mesure de l'utilisabilité des interfaces*. Paris: Hermes Science Publications-Lavoisier.

Bach, C., & Scapin, D. L. (2010). Comparing inspections and user testing for the evaluation of virtual environments. *International Journal of Human-Computer Interaction, 26*, 786-824.

Barcenilla, J., & Bastien, J. M. C. (2009). L'acceptabilité des nouvelles technologies: Quelles relations avec l'ergonomie, l'utilisabilité et l'expérience utilisateur? *Le Travail Humain, 72*, 311-331.

Barendregt, W., & Bekker, M. (2006). Developing a coding scheme for detecting usability and fun problems in computer games for young children. *Behavior Research Methods, 38*, 382-389.

Barendregt, W., Bekker, M. M., Bouwhuis, D. G., & Baauw, E. (2006). Identifying usability and fun problems in a computer game during first use and after some practice. *International Journal of Human-Computer Studies, 64*, 830-846.

Barr, P., Noble, J., & Biddle, R. (2007). Video game values: Human-computer interaction and games. *Interacting with Computers, 19*, 180-195.

Bastien, J. M. C. (2010). Usability testing: A review of some methodological and technical aspects of the method. *International Journal of Medical Informatics, 79*, e18-e23.

Bavelier, D., Green, C. S., Han, D. H., Renshaw, P. F., Merzenich, M. M., & Gentile, D. A. (2011). Brains on video games. *Nature Reviews Neuroscience, 12*, 763-768.

Beck, M. R., Lohrenz, M. C., & Trafton, J. G. (2010). Measuring search efficiency in complex visual search tasks: Global and local clutter. *Journal of Experimental Psychology: Applied, 16*, 238-250.

Bernhard, M., Stavrakis, E., & Wimmer, M. (2010). An empirical pipeline to derive gaze prediction heuristics for 3D action games. *ACM Transactions on Applied Perception, 8*, article 4, 1-30.

Bernhaupt, R. (Ed.). (2010). *Evaluating user experience in games: Concepts and methods*. London: Springer-Verlag.

Bonnardel, N., Piolat, A., & Le Bigot, L. (2011). The impact of colour on Website appeal and users' cognitive processes. *Displays, 32*, 69-80.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*, 387-398.

Bostan, B. (2009). Player motivations: A psychological perspective. *ACM Computers in Entertainment, 7*, article 22, 1-26.

Boyle, E., Kennedy, A.-M., Traynor, O., & Hill, A. D. K. (2011). Training surgical skills using nonsurgical tasks - Can Nintendo Wii improve surgical performance? *Journal of Surgical Education, 68*, 148-154.

Boyle, E. A., Connolly, T. M., Hainey, T., & Boyle, J. M. (2012). Engagement in digital entertainment games: A systematic review. *Computers in Human Behavior, 28*, 771-780.

Brandouy, O. (2001). Laboratory incentive structure and control-test design in an experimental asset market. *Journal of Economic Psychology, 22*, 1-26.

Broadbent, D. (1958). *Perception and communications*. New York: Pergamon.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45*, 624-634.

Brooksby, A. (2008). Exploring the representation of health in videogames: A content analysis. *Cyberpsychology & Behavior, 11*, 771-773.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. In E. Dykstra-Erickson & M. Tscheligi (Eds.), *CHI '04 extended abstracts on Human factors in computing systems* (pp. 1297-1300). New York: ACM.

Browne, K., & Anand, C. (2012). An empirical evaluation of user interfaces for a mobile video game. *Entertainment Computing, 3*, 1-10.

Bubka, A., & Bonato, F. (2010). Natural visual-field features enhance vection. *Perception, 39*, 627-635.

Calvillo Gámez, E. H., Cairns, P., & Cox, A. L. (2010). Assessing the core elements of the gaming experience. In R. Bernhaupt (Ed.), *Evaluating user experience in games* (pp. 47-71). London: Springer-Verlag.

Caroux, L., Le Bigot, L., & Vibert, N. (2011). Variations des stratégies d'anticipation visuelle en fonction de la règle dans un jeu vidéo. In C. Bastien, J. Cegarra, A. Chevalier, J. Dinet, N. Gregori & V. Grosjean (Eds.), *Actes du sixième colloque de psychologie ergonomique EPIQUE'2011* (pp. 405-418). Nancy, France: Presses Universitaires de Nancy.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*, 1484-1525.

Castel, A. D., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica, 119*, 217-230.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*, 349-354.

Chanel, G., Rebetez, C., Betrancourt, M., & Pun, T. (2011). Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Transactions on Systems Man and Cybernetics Part A: Systems and Humans, 41*, 1052-1063.

Chang, C.-H., Pan, W.-W., Tseng, L.-Y., & Stoffregen, T. (2012). Postural activity and motion sickness during video game play in children and adults. *Experimental Brain Research, 217*, 299-309.

Chehimi, F., Coulton, P., & Edwards, R. (2008). Evolution of 3D mobile games development. *Personal and Ubiquitous Computing, 12*, 19-25.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*, 661-686.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper and Row.

Dahan, D., & Tanenhaus, M. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review, 12*, 453-459.

Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In E. Dykstra-Erickson & M. Tscheligi (Eds.), *CHI '04 extended abstracts on Human factors in computing systems* (pp. 1509-1512). New York: ACM.

Desurvire, H., & Wiberg, C. (2009). Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration. In A. A. Ozok & P. Zaphiris (Eds.), *Proceedings of the 3rd International Conference on Online Communities and Social Computing: Held as Part of HCI International 2009* (pp. 557-559). Heidelberg, Germany: Springer-Verlag.

Diels, C., Ukai, K., & Howarth, P. A. (2007). Visually induced motion sickness with radial displays: Effects of gaze angle and fixation. *Aviation Space and Environmental Medicine, 78*, 659-665.

Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human Factors, 40*, 354-375.

Dumas, J. S., & Salzman, M. C. (2006). Usability assessment methods. *Reviews of Human Factors and Ergonomics, 2*, 109-140.

Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia, 47*, 1780-1789.

Dyson, B. J. (2010). "She's a waterfall": Motion aftereffect and perceptual design in video games involving virtual musicianship. *Perception, 39*, 131-132.

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision, 11*(5), article 14, 1-36.

Egenfeldt-Nielsen, S. (2006). Overview of research on the educational use of video games. *Digital Kompetanse, 1*, 184-213.

Fabricatore, C., Nussbaum, M., & Rosas, R. (2002). Playability in action videogames: A qualitative design model. *Human-Computer Interaction, 17*, 311-368.

Fadden, S., Ververs, P. M., & Wickens, C. D. (2001). Pathway HUDs: Are they viable? *Human Factors, 43*, 173-193.

Fang, X., Chan, S., Brzezinski, J., & Nair, C. (2010). Development of an instrument to measure enjoyment of computer game play. *International Journal of Human-Computer Interaction, 26*, 868-886.

Fehd, H. M., & Seiffert, A. E. (2010). Looking at the center of the targets helps multiple object tracking. *Journal of Vision, 10*(4), article 19, 1-13.

Feijoo, C., Gómez-Barroso, J.-L., Aguado, J.-M., & Ramos, S. (2012). Mobile gaming: Industry challenges and policy implications. *Telecommunications Policy, 36*, 212-221.

Fleetwood, M. D., & Byrne, M. D. (2006). Modeling the visual search of displays: A revised ACT-R model of icon search based on eye-tracking data. *Human-Computer Interaction, 21*, 153-197.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1030-1044.

Frey, A., Hartig, J., Ketzel, A., Zinkernagel, A., & Moosbrugger, H. (2007). The use of virtual environments based on a modification of the computer game Quake III Arena (R) in psychological experimenting. *Computers in Human Behavior, 23*, 2026-2039.

Gee, J. P. (2006). Why game studies now? Video games: A new art form. *Games and Culture, 1*, 58-61.

Giannopulu, I., & Lepecq, J.-C. (1998). Linear-vection chronometry along spinal and sagittal axes in erect man. *Perception, 27*, 363-372.

Gibson, B., Eberhard, K., & Bryant, T. (2005). Linguistically mediated visual search: The critical role of speech rate. *Psychonomic Bulletin & Review, 12*, 276-281.

Gilden, D. L., Thornton, T. L., & Marusich, L. R. (2010). The serial process in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 533-542.

Golding, J. F. (1998). Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain Research Bulletin, 47*, 507-516.

Golding, J. F. (2006). Predicting individual differences in motion sickness susceptibility by questionnaire. *Personality and Individual Differences, 41*, 237-248.

González Sánchez, J. L., Zea, N. P., & Gutiérrez, F. L. (2009). From usability to playability: Introduction to player-centered video game development process. In M. Kurosu (Ed.), *Proceedings of the 1st International Conference on Human Centered Design: Held as Part of HCI International 2009* (pp. 65-74). Heidelberg, Germany: Springer-Verlag.

Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors, 36*, 387-405.

Graham, N. V. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. *Vision Research, 51*, 1397-1430.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534-537.

Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1465-1478.

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18*, 88-94.

Harrison, W. J., Thompson, M. B., & Sanderson, P. M. (2010). Multisensory integration with a head-mounted display: Background visual motion and sound motion. *Human Factors, 52*, 78-91.

Hassenzahl, M., Diefenbach, S., & Göritz, A. (2010). Needs, affect, and interactive products - Facets of user experience. *Interacting with Computers, 22*, 353-362.

Hassenzahl, M., & Monk, A. (2010). The Inference of perceived usability from beauty. *Human-Computer Interaction, 25*, 235-260.

Hassenzahl, M., & Tractinsky, N. (2006). User experience - A research agenda. *Behaviour & Information Technology, 25*, 91-97.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*, 188-194.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*, 498-504.

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science, 16*, 219-222.

Henderson, J. M., Chanceaux, M., & Smith, T. J. (2009). The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of Vision, 9*(1), article 32, 1-8.

Hertzum, M. (2010). Images of usability. *International Journal of Human-Computer Interaction, 26*, 567-600.

Hoc, J.-M., & Amalberti, R. (1994). Diagnostic et prise de décision dans les situations dynamiques. *Psychologie Française, 39*, 177-192.

Hoc, J.-M., & Ceillier, J.-M. (2001). La gestion d'environnements dynamiques. *Psychologie Française, 46*, 103-106.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies, 64*, 79-102.

Horrey, W. J., & Wickens, C. D. (2004). Driving and side task performance: The effects of display clutter, separation, and modality. *Human Factors, 46*, 611-624.

Horrey, W. J., Wickens, C. D., & Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied, 12*, 67-78.

Hou, J., Nam, Y., Peng, W., & Lee, K. M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Computers in Human Behavior, 28*, 617-623.

Höysniemi, J., Hämäläinen, P., & Turkki, L. (2003). Using peer tutoring in evaluating the usability of a physically interactive computer game with children. *Interacting with Computers, 15*, 203-225.

Hubert-Wallander, B., Green, C., Sugarman, M., & Bavelier, D. (2011). Changes in search rate but not in the dynamics of exogenous attention in action videogame players. *Attention, Perception, & Psychophysics, 73*, 2399-2412.

Ilg, U. J. (1997). Slow eye movements. *Progress in Neurobiology, 53*, 293-329.

Institut de sondage GfK ISL. (2012, 5 juillet). Les Français et les jeux vidéo. Page web consultée le 17 juillet 2012, http://www.afjv.com/news.php?id=1288&title=francais_jeux_video

Iordanescu, L., Grabowecky, M., Franconeri, S., Theeuwes, J., & Suzuki, S. (2010). Characteristic sounds make you look at target objects more quickly. *Attention, Perception, & Psychophysics, 72*, 1736-1741.

Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review, 15*, 548-554.

Isbister, K., & Schaffer, N. (Eds.). (2008). *Game usability: Advice from the experts for advancing the player experience*. San Francisco: Morgan Kaufmann Publishers.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology, 17*, 259-273.

Itti, L. (2005). Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition, 12*, 1093-1123.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489-1506.

Jans, B., Peters, J. C., & De Weerd, P. (2010). Visual spatial attention to multiple locations at once: The jury is still out. *Psychological Review, 117*, 637-682.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., et al. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*, 641-661.

Jie, L., & Clark, J. J. (2008). Video game design using an eye-movement-dependent model of visual attention. *ACM Transactions on Multimedia Computing Communications and Applications, 4*, article 22, 1-16.

Jin, S.-A. A. (2009). Avatars mirroring the actual self versus projecting the ideal self: The effects of self-priming on interactivity and immersion in an exergame, Wii Fit. *Cyberpsychology & Behavior, 12*, 761-765.

Johnson, D., & Wiles, J. (2003). Effective affective user interface design in games. *Ergonomics, 46*, 1332-1345.

Jun, H. K., Daniel, V. G., Eric, S., Bruce, P., Randy, J. P., & Dennis, W. (2008). Tracking real-time user experience (TRUE): A comprehensive instrumentation solution for complex systems. In M. Czerwinski & A. Lund (Eds.), *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems CHI'08* (pp. 443-452). New York: ACM.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Kaminiarz, A., Krekelberg, B., & Bremmer, F. (2007). Localization of visual targets during optokinetic eye movements. *Vision Research, 47*, 869-878.

Kennedy, R. S., Drexler, J., & Kennedy, R. C. (2010). Research in visually induced motion sickness. *Applied Ergonomics, 41*, 494-503.

Kim, J., & Palmisano, S. (2010). Eccentric gaze dynamics enhance vection in depth. *Journal of Vision, 10*(12), article 7, 1-11.

Kingdom, F. A. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research, 51*, 652-673.

Klimmt, C., Hartmann, T., & Frey, A. (2007). Effectance and control as determinants of video game enjoyment. *Cyberpsychology & Behavior, 10*, 845-848.

Koeffel, C., Hochleitner, W., Leitner, J., Haller, M., Geven, A., & Tscheligi, M. (2010). Using heuristics to evaluate the overall user experience of video games and advanced interaction games. In R. Bernhaupt (Ed.), *Evaluating user experience in games* (pp. 233-256). London: Springer-Verlag.

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2009). Auditory and visual capture during focused visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1303-1315.

Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research, 51*, 1457-1483.

Kuze, J., & Ukai, K. (2008). Subjective evaluation of visual fatigue caused by motion images. *Displays, 29*, 159-166.

Lavie, T., Oron-Gilad, T., & Meyer, J. (2011). Aesthetics and usability of in-vehicle navigation displays. *International Journal of Human-Computer Studies, 69*, 80-99.

Law, E. L. C., & van Schaik, P. (2010). Modelling user experience - An agenda for research and practice. *Interacting with Computers, 22*, 313-322.

Le Meur, O., Le Callet, P., & Barba, D. (2007). Predicting visual fixations on video based on low-level visual features. *Vision Research, 47*, 2483-2498.

Lee, H.-W., Legge, G. E., & Ortiz, A. (2003). Is word recognition different in central and peripheral vision? *Vision Research, 43*, 2837-2846.

Lee, K. M. (2004). Presence, explicated. *Communication Theory, 14*, 27-50.

Léger, L., Fouquereau, N., & Tijus, C. (2009). GAME_MOC: Méthode pour évaluer les jeux vidéo avec les mouvements oculaires. In B. Cahour, F. Anceaux & A. Giboin (Eds.), *Actes du 5ème colloque de psychologie ergonomique EPIQUE'2009* (pp. 16-23). Paris: Télécom ParisTech.

Léger, L., Tijus, C., & Baccino, T. (2005). La discrimination visuelle et sémantique : Pour la conception ergonomique du contenu de sites web. *Revue d'Interaction Homme-Machine, 6*, 81-106.

Lepecq, J.-C., De Waele, C., Mertz-Josse, S., Teyssèdre, C., Huy, P. T. B., Baudonnière, P.-M., et al. (2006). Galvanic vestibular stimulation modifies vection paths in healthy subjects. *Journal of Neurophysiology, 95*, 3199-3207.

Liang, T.-P., & Yeh, Y.-H. (2011). Effect of use contexts on the continuous use of mobile services: the case of mobile games. *Personal and Ubiquitous Computing, 15*, 187-196.

Liu, C., Agrawal, P., Sarkar, N., & Chen, S. (2009). Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. *International Journal of Human-Computer Interaction, 25*, 506-529.

Liu, Y., & Li, H. (2011). Exploring the impact of use context on mobile hedonic services adoption: An empirical study on mobile gaming in China. *Computers in Human Behavior, 27*, 890-898.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*, 6-14.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning, and instruction: III. Conative and affective process analysis* (pp. 223-253). Hillsdale, NJ: Lawrence Erlbaum.

Mandryk, R. L., & Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies, 65*, 329-347.

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology, 25*, 141-158.

Martin-Emerson, R., & Wickens, C. D. (1997). Superimposition, symbology, visual attention, and the head-up display. *Human Factors, 39*, 581-601.

McIntire, J. P., Havig, P. R., Watamaniuk, S. N. J., & Gilkey, R. H. (2010). Visual search performance with 3-D auditory cues: Effects of motion, target location, and practice. *Human Factors, 52*, 41-53.

Menozzi, M., & Koga, K. (2004). Visual information processing in augmented reality: Some aspects of background motion. *Swiss Journal of Psychology, 63*, 183-190.

Merhi, O., Faugloire, E., Flanagan, M., & Stoffregen, T. A. (2007). Motion sickness, console video games, and head-mounted displays. *Human Factors, 49*, 920-934.

Moshagen, M., Musch, J., & Göritz, A. S. (2009). A blessing, not a curse: Experimental evidence for beneficial effects of visual aesthetics on performance. *Ergonomics, 52*, 1311-1320.

Mosimann, U. P., Felblinger, J., Colloby, S. J., & Müri, R. M. (2004). Verbal instructions and top-down saccade control. *Experimental Brain Research, 159*, 263-267.

Murphy, S. (2010). Can video games be art? *New Scientist, 207*(2778), 46-47.

Nacke, L. E., Grimshaw, M. N., & Lindley, C. A. (2010). More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers, 22*, 336-343.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature, 434*, 387-391.

Neider, M. B., & Zelinsky, G. J. (2011). Cutting through the clutter: Searching for targets in evolving complex scenes. *Journal of Vision, 11*(14), article 7, 1-16.

Nielsen, J. (1993). *Usability engineering*. San Francisco: Morgan Kaufmann.

Nørgaard, M., & Sørensen, J. R. (2008). Organizational challenges for user research in the videogame industry: Overview and advice. In K. Isbister & N. Schaffer (Eds.), *Game usability: Advice from the experts for advancing the player experience* (pp. 7-27). San Francisco: Morgan Kaufmann Publishers.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. New York: Basic Books.

Norme ISO 9241-11. (1998). *Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV) -- Partie 11: Lignes directrices relatives à l'utilisabilité*.

Norme ISO 9241-210. (2010). *Ergonomie de l'interaction homme-système -- Partie 210: Conception centrée sur l'opérateur humain pour les systèmes interactifs*.

Ojanpää, H., Näsänen, R., & Kojo, I. (2002). Eye movements in the visual search of word lists. *Vision Research, 42*, 1499-1512.

Olds, E. S., & Fockler, K. A. (2004). Does previewing one stimulus feature help conjunction search? *Perception, 33*, 195-216.

Orvis, K. A., Horn, D. B., & Belanich, J. (2008). The roles of task difficulty and prior videogame experience on performance and motivation in instructional videogames. *Computers in Human Behavior, 24*, 2415-2433.

Osman, M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin, 136*, 65-86.

Owsley, C., & McGwin Jr, G. (2010). Vision and driving. *Vision Research, 50*, 2348-2361.

Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., & Fuller, T. (2008). User-centered design in games. In A. Sears & J. A. Jacko (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies, and emerging applications* (pp. 741-759). New York: Lawrence Erlbaum Associates.

Park, N., Lee, K. M., Jin, S.-A. A., & Kang, S. (2010). Effects of pre-game stories on feelings of presence and evaluation of computer games. *International Journal of Human-Computer Studies, 68*, 822-833.

Pearson, R., & van Schaik, P. (2003). The effect of spatial layout of and link colour in web pages on performance in a visual search task and an interactive search task. *International Journal of Human-Computer Studies, 59*, 327-353.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision, 7*(2), article 20, 1-36.

Peters, R. J., & Itti, L. (2008). Applying computational tools to predict gaze direction in interactive visual environments. *ACM Transactions on Applied Perception, 5*, article 9, 1-19.

Pierno, A. C., Caria, A., Glover, S., & Castiello, U. (2005). Effects of increasing visual load on aurally and visually guided target acquisition in a virtual environment. *Applied Ergonomics, 36*, 335-343.

Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009). Usability heuristics for networked multiplayer games. In S. Teasley & E. Havn (Eds.), *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 169-178). New York: ACM.

Pola, J., Wyatt, H., & Lustgarten, M. (1992). Suppression of optokinesis by a stabilized target: Effects of instruction and stimulus frequency. *Perception & Psychophysics, 52*, 186-200.

Pola, J., Wyatt, H. J., & Lustgarten, M. (1995). Visual fixation of a target and suppression of optokinetic nystagmus: Effects of varying target feedback. *Vision Research, 35*, 1079-1087.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25.

Qin, H., Rau, P. L. P., & Salvendy, G. (2009). Measuring player immersion in the computer game narrative. *International Journal of Human-Computer Interaction, 25*, 107-133.

Qin, H., Rau, P. L. P., & Salvendy, G. (2010). Effects of different scenarios of game difficulty on player immersion. *Interacting with Computers, 22*, 230-239.

Raj, R., Geisler, W. S., Frazor, R. A., & Bovik, A. C. (2005). Contrast statistics for foveated visual systems: Fixation selection by minimizing contrast entropy. *Journal of the Optical Society of America A: Optics Image Science and Vision, 22*, 2039-2049.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human-performance models. *IEEE Transactions on Systems Man and Cybernetics, 13*, 257-266.

Ravaja, N., Saari, T., Turpeinen, M., Laarni, J., Salminen, M., & Kivikangas, M. (2006). Spatial presence and emotions during video game playing: Does it matter with whom you play? *Presence: Teleoperators and Virtual Environments, 15*, 381-392.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*, 1457-1506.

Reali, F., Spivey, M., Tyler, M., & Terranova, J. (2006). Inefficient conjunction search made efficient by concurrent spoken delivery of target identity. *Perception & Psychophysics, 68*, 959-974.

Renshaw, J. A., Finlay, J. E., Tyfa, D., & Ward, R. D. (2004). Understanding visual influence in graph design through temporal and spatial eye movement characteristics. *Interacting with Computers, 16*, 557-578.

Riecke, B. E., Schulte-Pelkum, J., Avraamides, M. N., Von Der Heyde, M., & Bülthoff, H. H. (2006). Cognitive factors can influence self-motion perception (vection) in virtual reality. *ACM Transactions on Applied Perception, 3*, 194-216.

Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2007). Task and context determine where you look. *Journal of Vision, 7*(14), article 16, 1-20.

Royden, C. S., Wolfe, J. M., & Klempen, N. (2001). Visual search asymmetries in motion and optic flow fields. *Perception & Psychophysics, 63*, 436-444.

Rubinstein, N. J., & Abel, L. A. (2011). Optokinetic nystagmus suppression as an index of the allocation of visual attention. *Investigative Ophthalmology & Visual Science, 52*, 462-467.

Sabri, A. J., Ball, R. G., Fabian, A., Bhatia, S., & North, C. (2007). High-resolution gaming: Interfaces, notifications, and the user experience. *Interacting with Computers, 19*, 151-166.

Schaffer, N. (2008). Heuristic evaluation of games. In K. Isbister & N. Schaffer (Eds.), *Game usability: Advice from the experts for advancing the player experience* (pp. 79-90). San Francisco: Morgan Kaufmann Publishers.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision, 11*(5), article 9, 1-30.

Seno, T., Ito, H., & Sunaga, S. (2011). Attentional load inhibits vection. *Attention, Perception, & Psychophysics, 73*, 1467-1476.

Shackel, B. (2009). Usability - Context, framework, definition, design and evaluation. *Interacting with Computers, 21*, 339-346.

Sonderegger, A., & Sauer, J. (2010). The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied Ergonomics, 41*, 403-410.

Sonderegger, A., Zbinden, G., Uebelbacher, A., & Sauer, J. (2012). The influence of product aesthetics and usability over the course of time: A longitudinal field experiment. *Ergonomics, 55*, 713-730.

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science, 12*, 282-286.

Stanney, K. M., Mollaghasemi, M., Reeves, L., Breaux, R., & Graeber, D. A. (2003). Usability engineering of virtual environments (VEs): Identifying multiple criteria that drive effective VE system design. *International Journal of Human-Computer Studies, 58*, 447-481.

Stanney, K. M., Mourant, R. R., & Kennedy, R. S. (1998). Human factors issues in virtual environments: A review of the literature. *Presence: Teleoperators and Virtual Environments, 7*, 327-351.

Stanney, K. M., & Salvendy, G. (1998). Aftereffects and sense of presence in virtual environments: Formulation of a research and development agenda. *International Journal of Human-Computer Interaction, 10*, 135-187.

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics, 48*, 1294-1313.

Stoffregen, T. A., Faugloire, E., Yoshida, K., Flanagan, M. B., & Merhi, O. (2008). Motion sickness and postural sway in console video games. *Human Factors, 50*, 322-331.

Syndicat National du Jeu Vidéo. (2011, 19 octobre). Le jeu vidéo en France en 2011 : éléments clés. Page web consultée le 17 juillet 2012, http://www.snjv.org/data/document/jeu-video-france-2011.pdf

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision, 11*(5), article 5, 1-23.

Thornton, T. L., & Gilden, D. L. (2007). Parallel and serial processes in visual search. *Psychological Review, 114*, 71-103.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*, 766-786.

Tozzi, A., Morrone, M. C., & Burr, D. C. (2007). The effect of optokinetic nystagmus on the perceived position of briefly flashed targets. *Vision Research, 47*, 861-868.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Trepte, S., & Reinecke, L. (2010). Avatar creation and video game enjoyment. *Journal of Media Psychology: Theories, Methods, and Applications, 22*, 171-184.

Trutoiu, L. C., Mohler, B. J., Schulte-Pelkum, J., & Bülthoff, H. H. (2009). Circular, linear, and curvilinear vection in a large-screen virtual environment with floor projection. *Computers & Graphics, 33*, 47-58.

Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics, 50*, 1235-1249.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics, 46*, 629-646.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology: General, 123*, 394-409.

Van Loon, E. M., Hooge, I. T. C., & Van den Berg, A. V. (2003). Different visual search strategies in stationary and moving radial patterns. *Vision Research, 43*, 1201-1209.

Vicente, K. J., Roth, E. M., & Mumaw, R. J. (2001). How do operators monitor a complex, dynamic work domain? The impact of control room technology. *International Journal of Human-Computer Studies, 54*, 831-856.

Virvou, M., & Katsionis, G. (2008). In the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education, 50*, 154-178.

Waespe, W., & Schwarz, U. (1987). Slow eye movements induced by apparent target motion in monkey. *Experimental Brain Research, 67*, 433-435.

Washburn, D. (2003). The games psychologists play (and the data they provide). *Behavior Research Methods, Instruments, & Computers, 35*, 185-193.

Webb, N. A., & Griffin, M. J. (2003). Eye movement, vection, and motion sickness with foveal and peripheral vision. *Aviation Space and Environmental Medicine, 74*, 622-625.

Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., & Groner, R. (2008). Playing online games against computer- vs. human-controlled opponents: Effects on presence, flow, and enjoyment. *Computers in Human Behavior, 24*, 2274-2291.

Weinstein, L. F., & Wickens, C. D. (1992). Use of nontraditional flight displays for the reduction of central visual overload in the cockpit. *International Journal of Aviation Psychology, 2*, 121-142.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 62-102). New York: Academic Press.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*, 159-177.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*, 449-455.

Wickens, C. D., Alexander, A. L., Ambinder, M. S., & Martens, M. (2004). The role of highlighting in visual search through maps. *Spatial Vision, 17*, 373-388.

Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors, 37*, 473-494.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. *Human Factors, 45*, 360-380.

Wickens, C. D., & Liu, Y. L. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors, 30*, 599-616.

Wickens, C. D., & Long, J. (1995). Object versus space-based models of visual-attention: Implications for the design of head-up displays. *Journal of Experimental Psychology: Applied, 1*, 179-193.

Wickens, C. D., & McCarley, J. S. (2008). *Applied attention theory*. Boca-Raton, FL: Taylor & Francis.

Williams, I. M., Mulhall, L., Mattingley, J., Lueck, C., & Abel, L. (2006). Optokinetic nystagmus as an assessment of visual attention to divided stimuli. *Journal of Clinical Neuroscience, 13*, 828-833.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*, 225-240.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1*, 202-238.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9*, 33-39.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7*, 70-76.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual-search. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 419-433.

Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J., & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research, 42*, 2985-3004.

Wolfson, S., & Case, G. (2000). The effects of sound and colour on responses to a computer game. *Interacting with Computers, 13*, 183-192.

Wood, R. T. A., Griffiths, M. D., Chappell, D., & Davies, M. N. O. (2004). The structural characteristics of video games: A psycho-structural analysis. *Cyberpsychology & Behavior, 7*, 1-10.

Wyatt, H. J., & Pola, J. (1984). A mechanism for suppression of optokinesis. *Vision Research, 24*, 1931-1945.

Wyatt, H. J., Pola, J., & Lustgarten, M. (1988). "Passive suppression" of optokinesis by stabilized targets. *Vision Research, 28*, 1023-1029.

Yannakakis, G., Martínez, H., & Jhala, A. (2010). Towards affective camera control in games. *User Modeling and User-Adapted Interaction, 20*, 313-340.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

Yeh, M., & Wickens, C. D. (2001). Attentional filtering in the design of electronic map displays: A comparison of color coding, intensity coding, and decluttering techniques. *Human Factors, 43*, 543-562.

Zaphiris, P., & Ang, C. S. (2007). HCI issues in computer games. *Interacting with Computers, 19*, 135-139.

Zelinsky, G. J. (2008). A Theory of eye movements during target acquisition. *Psychological Review, 115*, 787-835.

Annexes

Table des annexes

Annexe 1 : Prépublication de l'article présenté au colloque EPIQUE 2011 182
Annexe 2 : Questionnaire pré-expérimental des expériences 1 et 2 191
Annexe 3 : Questionnaire pré-expérimental des expériences 3 à 6 192
Annexe 4 : Formes géométriques utilisées dans les expériences 3a et 3b 195
Annexe 5 : Cible et curseur utilisés dans l'expérience 4 196
Annexe 6 : Créatures utilisées comme cible ou distracteur et curseur des expériences 5 et 6 ... 197
Annexe 7 : Table des figures 198
Annexe 8 : Table des tableaux 200
Annexe 9 : Table des matières détaillée 201

Annexe 1 : Prépublication de l'article présenté au colloque EPIQUE 2011

Caroux, L., Le Bigot, L., & Vibert, N. (2011). Variations des stratégies d'anticipation visuelle en fonction de la règle dans un jeu vidéo. In C. Bastien, J. Cegarra, A. Chevalier, J. Dinet, N. Gregori & V. Grosjean (Eds.), *Actes du sixième colloque de psychologie ergonomique EPIQUE'2011* (pp. 405-418). Nancy, France: Presses Universitaires de Nancy.

Variations des stratégies d'anticipation visuelle en fonction de la règle dans un jeu vidéo

Loïc CAROUX, Ludovic LE BIGOT, Nicolas VIBERT
Centre de Recherches sur la Cognition et l'Apprentissage (CeRCA)
UMR 6234 CNRS – Université de Poitiers – Université de Tours
5 rue Théodore Lefebvre
86000 Poitiers
loic.caroux@univ-poitiers.fr, ludovic.le.bigot@univ-poitiers.fr, nicolas.vibert@univ-poitiers.fr

Symposium

RÉSUMÉ

Dans certains jeux vidéo, le joueur a besoin d'anticiper le mouvement d'objets, d'obstacles ou de cibles pour atteindre son but. Cependant, des recherches ont montré que les stratégies visuelles utilisées dans un environnement visuel dynamique variaient selon le but des individus. Le but de l'expérience exploratoire présentée dans cet article était de comparer les stratégies oculaires de participants jouant dans un même environnement de jeu nécessitant de l'anticipation visuelle selon deux règles du jeu différentes. Les deux règles du jeu vidéo utilisé consistaient soit à « éviter », soit à « toucher » un maximum d'obstacles avec une balle. L'enregistrement des mouvements oculaires montre que l'anticipation est plus importante et le balayage de l'écran de jeu plus large quand les joueurs doivent éviter les obstacles que quand ils doivent les toucher. Les implications de ce résultat sur les stratégies oculaires optimales et la conception de jeux vidéo sont discutées.

MOTS-CLÉS
Scènes visuelles dynamiques, évitement, poursuite oculaire, obstacles, mouvements oculaires.

INTRODUCTION

Ergonomie des jeux vidéo

La complexité grandissante des jeux vidéo (JV) impose aujourd'hui une conception centrée sur le joueur pour optimiser les chances d'un succès commercial (Pagulayan, Keeker, Wixon, Romero, & Fuller, 2008). En général, ce sont des recommandations ergonomiques ou des heuristiques non spécifiques (par ex. Nielsen, 1993) aux JV qui sont actuellement utilisées pour améliorer l'utilisabilité de ces systèmes ou l'expérience utilisateur des joueurs. Cependant, du point de vue de l'activité, les JV nécessitent des règles de conception différentes de celles utilisées pour les outils informatiques classiques comme les pages web ou les logiciels liés à la productivité. Les études spécifiques à l'utilisabilité ou l'expérience utilisateur des JV se sont récemment multipliées (Bernhaupt, 2010 ; Caroux, Le Bigot, & Vibert, 2010 ; Isbister & Schaffer, 2008 ; Zaphiris & Ang, 2007). Ces études s'intéressent généralement aux caractéristiques des JV partagées avec l'ensemble des environnements virtuels (Stanney, Mollaghasemi, Reeves, Breaux, & Graeber, 2003) comme l'immersion (Jennett et al., 2008), ou à celles partagées avec l'ensemble des divertissements comme le plaisir de jeu (Fang, Chan, Brzezinski, & Nair, 2010) ou les interactions entre joueurs (Ho & Huang, 2009). Cependant, les études concernant la conception des interfaces visuelles des JV, une de leurs spécificités importantes, sont peu présentes dans la littérature.

Conception des interfaces visuelles des jeux vidéo

Les interfaces visuelles des JV contiennent généralement une combinaison d'éléments graphiques complexes comme des objets (par ex. avatars, ennemis ou cibles) et des fonds structurés (par ex.

décors, paysages) en mouvement. Des informations contextuelles peuvent également être superposées à la scène d'action. Ces informations sont des indicateurs permanents (par ex. score, points de vie, carte du monde virtuel) ou ponctuels (par ex. messages d'alarmes) sur la situation en cours de jeu. Celles-ci sont composées de mots, liste de mots, données chiffrées et/ou symboles. Cet assemblage complexe d'éléments visuels dynamiques et/ou en mouvement implique des règles de perception visuelle différentes de celles observées dans le cadre de la perception de scènes visuelles statiques (Caroux, Vibert, & Le Bigot, 2009 ; Dyson, 2010).

Certains auteurs ont étudié l'influence d'un élément isolé des interfaces visuelles des JV, comme la couleur du fond d'écran (Wolfson & Case, 2000) ou la position de la caméra dans des JV présentant un monde en 3 dimensions (Yannakakis, Martínez, & Jhala, 2010), sur la performance et l'expérience du joueur. D'autres auteurs ont étudié plus globalement la disposition des différents éléments constitutifs des interfaces visuelles des JV. Par exemple, Sabri, Ball, Fabian, Bhatia et North (2007) ont montré que sur des interfaces de JV en haute résolution (par ex. plusieurs moniteurs), les informations contextuelles importantes (les plus regardées ou utilisées) devaient se trouver proches du curseur contrôlé par le joueur. Ces auteurs ont observé que sur une configuration à plusieurs moniteurs, les joueurs étaient plus performants quand les informations contextuelles étaient affichées sur le moniteur effectivement utilisé. Caroux, Le Bigot et Vibert (2011) ont étudié en enregistrant les mouvements oculaires l'influence de la position à l'écran d'une information contextuelle importante (le score) sur la performance du joueur dans un jeu nécessitant une anticipation visuelle des mouvements d'obstacles. Selon le principe de proximité compatibilité de Wickens et Carswell (1995), le score devrait se trouver le plus proche possible de la zone d'anticipation des obstacles pour faciliter l'intégration de toutes les informations utiles au joueur. Caroux et al. (2011) ont montré que positionner le score à proximité de la zone d'anticipation du mouvement des obstacles facilitait effectivement l'anticipation de ce mouvement, mais que la performance des joueurs était meilleure quand le score était positionné juste à l'extérieur plutôt qu'à l'intérieur de cette zone d'anticipation.

D'autres études récentes ont utilisé le suivi des mouvements oculaires pour comprendre précisément les mécanismes de la perception des scènes visuelles dynamiques présentes dans les JV. Par exemple, certains auteurs ont utilisé les mouvements oculaires comme outil d'évaluation de la sensation d'immersion du joueur dans un JV (Jennett et al., 2008) ou de la qualité générale du JV (Léger, Fouquereau, & Tijus, 2009). D'autres auteurs ont observé et modélisé les mouvements du regard dans les scènes visuelles des JV en fonction des propriétés de la scène visuelle et des intentions du joueur (Bernhard, Stavrakis, & Wimmer, 2010 ; Jie & Clark, 2008 ; Peters & Itti, 2008). Les résultats de ces études sont utiles pour concevoir des éléments d'aide ou de challenge dans le but d'améliorer l'expérience du joueur. Par exemple, Peters et Itti (2008) ont montré que la prise en compte des caractéristiques visuelles de la scène visuelle dans un JV présentant un monde en 3 dimensions (saillances de couleurs, différences de contraste, entropie, mouvements des objets) aidait à la prédiction de l'orientation de l'attention, et donc du regard. Cependant, ces observations étaient conditionnées par le type de jeu utilisé. Dans un JV qui nécessite de l'exploration visuelle (par ex. jeu d'aventure), où la vision du joueur est libre, les propriétés de la scène visuelle sont prédominantes. Par contre, dans un JV où la vision est contrainte par le temps (par ex. jeu de course), la prédiction à l'aide des caractéristiques visuelles de la scène est plus difficile. De plus, Bernhard et al. (2010) ont montré que dans un jeu de tir à la première personne (FPS), la prédiction de l'orientation du regard était biaisée à cause de la présence d'un réticule de visée souvent placé au centre de l'écran.

Ces différentes études ont montré qu'il était difficile de généraliser un comportement oculaire quel que soit le type de JV. Comme l'activité de JV est dirigée par un but différent selon le type du jeu, l'attention est orientée selon des contraintes différentes. Les propriétés de la scène visuelle du JV ne sont généralement pas suffisantes pour prédire les mouvements du regard du joueur. C'est pourquoi il est important de prendre en compte les règles du jeu et les instructions données au joueur pour concevoir les différents éléments d'aide ou de challenge des interfaces visuelles d'un JV.

Anticipation et suivi d'objets visuels en mouvement

Généralement, les JV qui présentent des scènes visuelles en mouvement nécessitent une anticipation visuelle du mouvement d'éléments à éviter (par ex. les JV de course) ou à toucher (par ex. les JV de tir) par un avatar pour réussir une tâche (Caroux et al., 2011). Par exemple, quand des éléments (cibles ou ennemis) d'une scène visuelle en 2 dimensions sont en mouvement du bas de

l'écran de jeu vers le haut, le joueur doit porter son attention régulièrement sur l'apparition de ces éléments en bas de l'écran afin de les détecter le plus rapidement possible. De plus, en présence de plusieurs éléments en mouvement à l'écran, il doit être capable de porter son attention sur l'ensemble de ces éléments en parallèle. Le joueur doit donc être capable d'adopter une stratégie visuelle optimale en fonction du but du jeu. Par exemple, Yarbus (1967) a démontré que durant l'exploration d'une scène visuelle statique, les mouvements oculaires dépendaient des instructions qui étaient données et donc des buts des observateurs.

Plusieurs études sur l'anticipation visuelle dans des situations dynamiques existent, notamment dans les domaines de la conduite automobile (Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003) et du pilotage d'avion (Wickens, Goh, Helleberg, Horrey, & Talleur, 2003). Par exemple, Underwood et al. (2003) ont montré qu'en situation de conduite naturelle, les conducteurs experts et novices avaient des stratégies d'anticipation visuelle différentes. Les conducteurs experts portaient surtout leur attention aux endroits les plus éloignés de la route pour anticiper les sources éventuelles d'accident, alors que les novices regardaient plus souvent les zones de la route proches de leur véhicule, signe que leur conduite n'était pas encore automatisée.

Le suivi multiple d'objets en mouvements à quant à lui été peu étudié dans des conditions naturelles (Cavanagh & Alvarez, 2005 ; Jans, Peters, & De Weerd, 2010). Des études en laboratoire ont toutefois permis d'observer des stratégies différentes selon la tâche demandée. Par exemple, le comportement de suivi de multiples objets en mouvement se caractérise par des séquences alternées de saccades et de poursuite oculaire lente des objets (Fedh et Seiffert, 2010). Rothkopf, Ballard et Hayhoe (2007) ont montré que dans un même environnement dynamique les individus adoptaient des stratégies de suivi visuel différentes selon la tâche. Par exemple, quand la tâche était de toucher des objets en mouvement, les fixations oculaires étaient localisées majoritairement sur les centres des objets, tandis que quand la tâche était d'éviter ces mêmes objets, les fixations oculaires visaient majoritairement leurs extrémités. La tâche « éviter les objets » induisait un comportement du regard d'évitement des objets, caractérisé plutôt par des séquences de saccades et fixations oculaires de balayage de tout l'environnement, tandis que la tâche « toucher les objets » induisait un comportement de suivi successif des différents objets.

Présentation de l'étude

Le but de l'étude était d'observer les modifications du comportement oculaire de joueurs en fonction des instructions de jeu (ou règles du jeu) pour un même environnement dynamique de JV. Une expérience exploratoire utilisant un JV simple contenant une scène visuelle dynamique a été réalisée. Le principe du jeu, basé sur celui utilisé par Caroux et al. (2011), était de contrôler une balle dans un environnement visuel dynamique toujours identique, mais selon deux règles différentes. La première règle était d'éviter, avec la balle, un maximum d'obstacles. La deuxième était de toucher, avec la balle, un maximum de ces mêmes obstacles. Comme ces deux instructions impliquent des buts différents pour le joueur, l'hypothèse était que les comportements oculaires des joueurs devaient être différents. Même si les deux instructions nécessitent un comportement d'anticipation (Caroux et al., 2011), l'instruction « éviter » devait amener un comportement de simple évitement des obstacles, et l'instruction « toucher » un comportement de suivi successif des obstacles (Fehd & Seiffert, 2010 ; Rothkopf et al., 2007).

METHODE

Participants

Trente-six participants droitiers (16 femmes et 20 hommes) ont été recrutés en tant que volontaires non rémunérés. Leur âge moyen était de 23.75 ans (*min*. = 17, *max*. = 47; écart-type *ET* = 5.75). Neuf participants étaient non-joueurs (ne jouaient jamais à des jeux vidéo), huit étaient des joueurs occasionnels (qui jouaient moins de 2 heures par semaine à des jeux vidéo), onze étaient des joueurs réguliers (entre 2 et 6 heures par semaine), et huit étaient des joueurs très réguliers (plus de 6 heures par semaine).

Matériel

Un JV similaire à celui utilisé par Caroux et al. (2011), a été créé à l'aide du logiciel *« Game Maker 7.0 »*. Sur un écran de jeu de 480x672 pixels (≈ 161x226 millimètres sur un écran 17 pouces de résolution 1024x768 pixels), des obstacles (carrés noirs de 32x32 pixels) défilaient de bas en haut sur un fond gris (couleurs achromatiques). Environ 48 obstacles étaient présents en permanence sur l'écran. Ils apparaissaient par lignes horizontales de 2 ou 3 en bas de l'écran, à des positions définies aléatoirement. Une balle de couleur verte était contrôlable uniquement sur l'axe horizontal au milieu de l'écran de jeu. Le score était affiché dans un rectangle bleu dans le coin haut-droit de l'écran (figure 1).

Deux règles du jeu distinctes ont été proposées. La première règle était de faire éviter les obstacles à la balle en la déplaçant vers la gauche ou vers la droite de l'écran. Le score était initialisé à 99 et diminuait de 1 point par obstacle touché. La seconde règle consistait au contraire à toucher le maximum d'obstacles possibles avec la balle. Le score était initialisé à 0 et augmentait d'un point par obstacle touché. Dans chaque cas, l'objectif était d'atteindre le meilleur score possible pour chaque partie. A part la règle présentée au début de l'expérimentation, l'environnement de jeu était exactement similaire dans les deux conditions. La durée d'une partie n'était pas contrôlée par le joueur et était indépendante de la réussite au jeu. Elle était comprise aléatoirement entre 37 et 52 secondes, mais seules 30 secondes de chaque partie ont été utilisées pour l'analyse (de la 6ème à la 35ème seconde). Aucune condition de fin de partie ne distinguait les deux conditions expérimentales puisque le jeu n'était jamais interrompu, même en cas de faible performance.

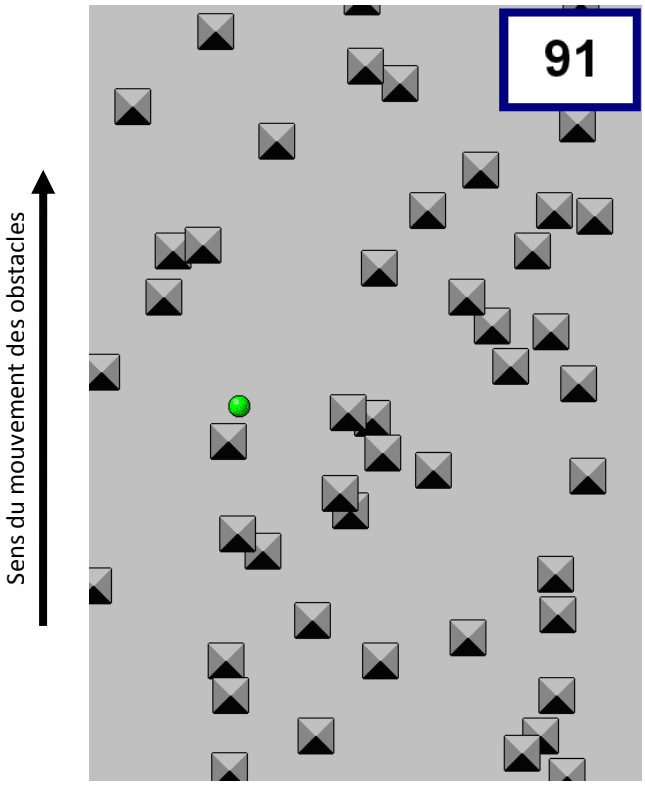


Figure 1. Capture d'écran du jeu utilisé. Le participant devait contrôler la balle verte sur un axe horizontal. Deux règles différentes étaient proposées aux participants. La première était d'éviter les obstacles qui se déplaçaient du bas de l'écran vers le haut avec la balle. La seconde était de toucher le maximum possible d'obstacles avec la balle.

Appareillage

Les données de mouvements oculaires ont été collectées par un oculomètre Tobii 1750. Il s'agit d'un oculomètre portable qui utilise la technique de vidéo-oculographie binoculaire. Sa précision est de

0,5° et sa vitesse d'échantillonnage de 50 Hz. Un écran TFT de 17 pouces, de résolution 1280x1024 pixels est intégré au système. Pour l'expérimentation, la résolution du moniteur a été limitée à 1024x768 pixels. Etant non-invasif, l'oculomètre a permis d'expérimenter en situation naturelle. Le système était relié à un ordinateur qui enregistrait les données collectées par l'oculomètre et faisait fonctionner le programme informatique du jeu. Le contrôle du jeu et des mouvements de la balle s'effectuait à l'aide d'un pavé numérique.

Design expérimental et procédure

Les participants passaient l'expérimentation individuellement. La règle du jeu, présentée avant son démarrage, était manipulée entre les participants. Les participants étaient assignés aléatoirement dans l'un des deux groupes, en veillant à l'équilibre en termes d'expérience personnelle des jeux vidéo. Chaque participant jouait 12 parties (soit entre 444 et 624 secondes de jeu effectif pour 360 secondes analysées).

Variables dépendantes

Le nombre total moyen et la durée moyenne des fixations oculaires ont été calculés pour chaque participant sur l'ensemble des parties. De plus, deux indicateurs proposés par Caroux et al. (2011) ont été calculés à partir de la coordonnée verticale (axe Y) de la position sur l'écran de chaque fixation oculaire : la position verticale moyenne et son écart-type, utilisé comme indicateur d'amplitude de la dispersion. Les fixations oculaires ont été définies selon un seuil temporel et un seuil spatial. Le seuil temporel retenu était de 60 millisecondes minimum et le seuil spatial était de 1° d'angle visuel maximum (diamètre : 30 pixels). En d'autres termes, une fixation était définie dès que le regard restait à moins de 1° d'angle d'un point initial pendant au moins 60 ms. Ces valeurs ont été choisies dans le but de les distinguer au maximum des saccades, ainsi que pour pallier au moins partiellement une importante variabilité interindividuelle généralement constatée dans la littérature (Rayner, 1998).

Tableau 1. Position verticale moyenne du regard et dispersion des points de fixation des participants sur l'ensemble des parties pour les deux conditions « Eviter » et « Toucher ».

	coordonnée Y	
Variable	Eviter	Toucher
Position moyenne		
Moyenne	470.8	447.0
Ecart-type	26.3	22.4
Dispersion		
Moyenne	82.8	56.8
Ecart-type	22.1	38.2

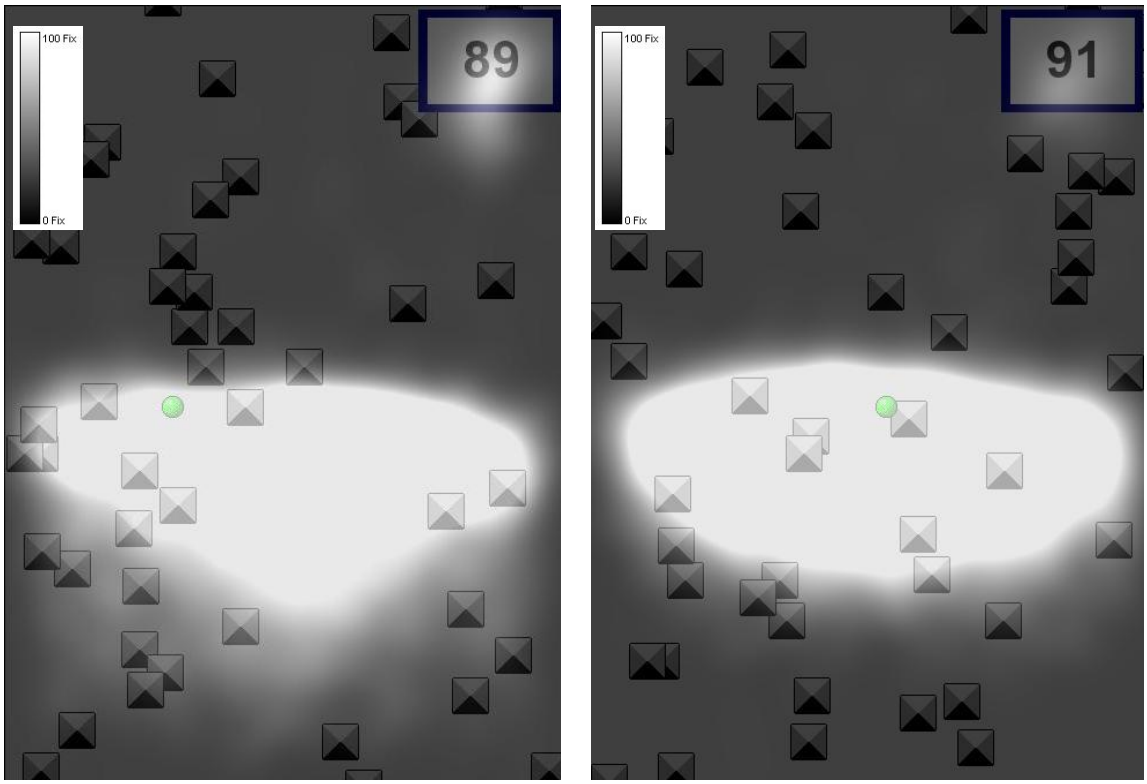
Note. Les données sont exprimées en pixels. L'écran de jeu avait pour dimensions 480x672 pixels. La coordonnée Y=0 correspond à l'extrémité haute de l'écran de jeu et Y= 672 à l'extrémité basse. Sur l'écran 17 pouces utilisé, de résolution 1024x768 pixels, 1 millimètre équivaut à 2.98 pixels.

RESULTATS

Des tests t de Student ont été utilisés pour comparer les comportements oculaires des deux groupes de participants (éviter/toucher). Les données sont présentées dans le tableau 1 (position verticale moyenne et dispersion des fixations), pour l'ensemble des 12 parties (6 minutes). La figure 2 représente sous forme de cartes de chaleurs la répartition des fixations sur l'écran de jeu.

Le nombre de fixations oculaires, $t(34) = -0.93$, ns, et leur durée moyenne, $t(34) = -0.20$, ns, n'étaient pas significativement différents entre les deux groupes. Par contre, la position verticale moyenne des fixations oculaires était significativement plus basse dans la condition *« éviter »* que dans la condition *« toucher »*, $t(34) = 2.93$, $p < .01$, $d = 1.00$. L'amplitude de la dispersion de la position

verticale moyenne des fixations oculaires était significativement plus grande dans la condition *« éviter »* que dans la condition *« toucher »*, $t(34) = 2.49$, $p < .05$, $d = 0.86$. En d'autres termes, la position des fixations dans la condition *« éviter »* était plus basse et leur dispersion plus grande que dans la condition *« toucher »*.



Condition « Eviter » Condition « Toucher »

Figure 2. Cartes de chaleur représentant le nombre total de fixations oculaires (sur 6 minutes d'enregistrement) pour l'ensemble des participants de chaque groupe (soit 18 par groupe). L'échelle de couleurs située dans le coin haut-gauche de chaque carte de chaleur correspond au nombre de fixations à un endroit précis de l'écran. La couleur blanche correspond à un nombre de total de 100 fixations ou plus, la couleur noire à une absence de fixation.

DISCUSSION

L'hypothèse a été confirmée. Même si le nombre de fixations moyen et leur durée moyenne n'étaient pas significativement différents selon la tâche, les deux types d'instructions ont induit chez les participants des comportements oculaires différents. Les participants anticipaient plus le mouvement des obstacles et balayaient plus l'écran de jeu lorsqu'ils devaient éviter les obstacles que lorsqu'ils devaient les toucher. Pour la tâche « éviter », les joueurs étaient capables grâce à un balayage visuel rapide et ample de traiter l'information sur une large portion de l'écran, y compris vers le bas. Pour la tâche « toucher » en revanche, la nécessité d'estimer précisément la vitesse de la balle et des obstacles les plus proches contraignait probablement les joueurs à suivre les obstacles plus longuement. Cette stratégie devait réduire leurs capacités d'anticipation à cause du temps passé à suivre des yeux les obstacles arrivant au niveau de la balle. L'ensemble de ces résultats sont bien cohérents avec ceux de la littérature (Fehd & Seiffert, 2010 ; Rothkopf et al., 2007).

La principale limite de cette étude est liée à l'outil utilisé pour analyser les mouvements oculaires des participants (ClearView 2.7.0 de Tobii). Il était en effet impossible d'obtenir des données précises concernant les fixations oculaires sur les obstacles en mouvement. La possibilité d'analyser les mouvements oculaires à l'aide d'AOIs (aires d'intérêt) pour chaque objet en mouvement aurait permis

d'obtenir des données supplémentaires sur la manière dont chacun d'entre eux était fixé en fonction de sa position à l'écran. Par exemple, connaitre les moments et les durées des fixations d'un obstacle en fonction de son éloignement de la balle pourrait apporter des informations nouvelles sur la stratégie visuelle du joueur selon la règle du jeu. Deux hypothèses pourraient alors être formulées en fonction de la littérature et des résultats observés. Puisque avec l'instruction « toucher », les participants étaient amenés à adopter un comportement de suivi successif des obstacles (Fehd & Seiffert, 2010 ; Rothkopf et al., 2007), la première hypothèse serait que les obstacles sont plus et plus longtemps fixés dans la condition « toucher » que dans la condition « éviter ». De plus, puisque les participants étaient capables de traiter l'information dans une zone plus vaste avec la tâche « éviter » qu'avec la tâche « toucher », la seconde hypothèse serait que la zone de l'écran où les obstacles sont les plus longuement fixés est située plus bas dans la condition « éviter » que dans la condition « toucher ». De nouvelles études, accompagnées d'analyses plus pointues, doivent être menées pour observer plus précisément ces stratégies visuelles.

De plus, ces nouvelles études permettraient de déduire des différentes stratégies visuelles observées des préconisations pour la conception centrée-utilisateur des JV. Par exemple, il est important de savoir que la distance optimale d'anticipation peut être différente selon la règle du jeu si l'on souhaite mettre en place des aides à l'anticipation optimale pour améliorer la performance des joueurs, par exemple en modifiant la position d'une information contextuelle importante (comme le score). De la même manière, des éléments de challenge pourraient être intégrés à l'interface visuelle pour rendre volontairement plus difficile la tâche du joueur, par exemple en cachant temporairement la zone optimale d'anticipation. Dans les deux cas, la connaissance précise des mécanismes d'attention visuelle en fonction de la règle du jeu créerait des outils utiles pour une meilleure prise en compte des capacités du joueur et de son expérience du jeu lors de la conception des JV (Jie & Clark, 2008).

REMERCIEMENTS

Loïc Caroux est financé par une allocation de recherche doctorale de la Direction Générale de l'Armement (DGA), et suivi dans ce cadre par Didier Bazalgette.

BIBLIOGRAPHIE

Bernhard, M., Stavrakis, E., & Wimmer, M. (2010). An empirical pipeline to derive gaze prediction heuristics for 3D action games. *ACM Transactions on Applied Perception, 8*(1), article 4, 1-30.

Bernhaupt, R. (Ed.). (2010). *Evaluating user experience in games: Concepts and methods*. London: Springer.

Caroux, L., Le Bigot, L., & Vibert, N. (2010). Vers une utilisabilité spécifique des jeux vidéo ? In *Annexes des actes d'IHM'10, 22e Conférence Francophone sur l'Interaction Homme-Machine* (pp. 5-8). Luxembourg: IHM'10.

Caroux, L., Le Bigot, L., & Vibert, N. (2011). Maximizing players' anticipation by applying the proximity-compatibility principle to the design of video games. *Human Factors, 53*, 103-117.

Caroux, L., Vibert, N., & Le Bigot, L. (2009). Détecter l'apparition d'objets sur un fond visuel en mouvement. In B. Cahour, F. Anceaux & A. Giboin (Eds.), *Actes du 5ème colloque de psychologie ergonomique EPIQUE'2009* (pp. 293-298). Paris: Télécom ParisTech.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*, 349-354.

Dyson, B. J. (2010). "She's a waterfall": Motion aftereffect and perceptual design in video games involving virtual musicianship. *Perception, 39*, 131-132.

Fang, X., Chan, S., Brzezinski, J., & Nair, C. (2010). Development of an Instrument to Measure Enjoyment of Computer Game Play. *International Journal of Human-Computer Interaction, 26*, 868 - 886.

Fehd, H. M., & Seiffert, A. E. (2010). Looking at the center of the targets helps multiple object tracking. *Journal of Vision, 10*(4), article 19, 1-13.

Ho, S. H., & Huang, C. H. (2009). Exploring success factors of video game communities in hierarchical linear modeling: The perspectives of members and leaders. *Computers in Human Behavior, 25*, 761-769.

Isbister, K., & Schaffer, N. (Eds.). (2008). *Game usability: Advice from the experts for advancing the player experience*. San Francisco: Morgan Kaufmann Publishers.

Jans, B., Peters, J. C., & De Weerd, P. (2010). Visual spatial attention to multiple locations at once: The jury is still out. *Psychological Review, 117*, 637-682.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., et al. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*, 641-661.

Jie, L., & Clark, J. J. (2008). Video game design using an eye-movement-dependent model of visual attention. *ACM Transactions on Multimedia Computing Communications and Applications, 4*(3), article 22, 1-16.

Léger, L., Fouquereau, N., & Tijus, C. (2009). GAME_MOC: Méthode pour évaluer les jeux vidéo avec les mouvements oculaires. In B. Cahour, F. Anceaux & A. Giboin (Eds.), *Actes du 5ème colloque de psychologie ergonomique EPIQUE'2009* (pp. 16-23). Paris: Télécom ParisTech.

Nielsen, J. (1993). *Usability engineering*. Boston: Academic press.

Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., & Fuller, T. (2008). User-centered design in games. In A. Sears & J. A. Jacko (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications* (pp. 741-759). New York: Lawrence Erlbaum Associates.

Peters, R. J., & Itti, L. (2008). Applying Computational Tools to Predict Gaze Direction in Interactive Visual Environments. *ACM Transactions on Applied Perception, 5*(2), article 9, 1-19.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2007). Task and context determine where you look. *Journal of Vision, 7*(14), article 16, 1-20.

Sabri, A. J., Ball, R. G., Fabian, A., Bhatia, S., & North, C. (2007). High-resolution gaming: Interfaces, notifications, and the user experience. *Interacting with Computers, 19*, 151-166.

Stanney, K. M., Mollaghasemi, M., Reeves, L., Breaux, R., & Graeber, D. A. (2003). Usability engineering of virtual environments (VEs): identifying multiple criteria that drive effective VE system design. *International Journal of Human-Computer Studies, 58*, 447-481.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: sequences of eye fixations made by experienced and novice drivers. *Ergonomics, 46*, 629-646.

Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors, 37*, 473-494.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. *Human Factors, 45*, 360-380.

Wolfson, S., & Case, G. (2000). The effects of sound and colour on responses to a computer game. *Interacting with Computers, 13*, 183-192.

Yannakakis, G., Martínez, H., & Jhala, A. (2010). Towards affective camera control in games. *User Modeling and User-Adapted Interaction, 20*, 313-340.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

Zaphiris, P., & Ang, C. S. (2007). HCI issues in computer games. *Interacting with Computers, 19*, 135-139.

Annexe 2 : Questionnaire pré-expérimental des expériences 1 et 2

No

QUESTIONNAIRE

Initiales : ……

Age : ……

Sexe : Masculin Féminin *(entourez la bonne réponse)*

Main dominante : Gauche Droite Les deux *(entourez la bonne réponse)*

Profession :……

Formation :……

Niveau d'études :……

Jouez-vous à des jeux vidéo ? Oui Non *(entourez la réponse de votre choix)*

Si oui, **depuis combien d'années ?** ……

Si oui, **à quelle fréquence ?** *(Entourez la réponse de votre choix)*

1 fois/mois 2-4 fois/mois 2-6 h/semaine 6-12h/semaine + de 12h/semaine

Si oui, **à quel type de jeux ?** *(Entourez la(les) réponse(s) de votre choix)*

Action Aventure Course Plate-forme Sport Jeu de rôle

Jeu de tir Combat Puzzle Réflexion Gestion/stratégie

Annexe 3 : Questionnaire pré-expérimental des expériences 3 à 6

Questionnaire préliminaire

Age : ..... ans

Sexe : Masculin Féminin (*entourez la bonne réponse*)

Main dominante : Gauche Droite Les deux (*entourez la bonne réponse*)

Profession :.....(*Ex. : étudiant, informaticien…*)

Formation initiale:.....(*Ex. : psychologie, mécanique…*)

Niveau d'études :.....(*Ex. : CAP, BAC+2…*)

Etes-vous sujet au strabisme divergent ? oui non *(entourez la bonne réponse)*

Etes-vous sujet au strabisme convergent ? oui non *(entourez la bonne réponse)*

Questionnaire sur votre expérience des jeux vidéo

Ce questionnaire a été conçu pour déterminer votre degré d'expérience des jeux vidéo.

Le terme « jeux vidéo » regroupe globalement tous les jeux qui utilisent un dispositif informatique et un écran. Ainsi, pour répondre aux questions de ce questionnaire, vous devrez prendre en compte tous les jeux vidéo auxquels vous avez pu jouer à l'aide d'un ordinateur, un système connecté à une télévision (par ex. Wii, Playstation, NES), un système portable dédié (par ex. DS, PSP, Game Boy), un téléphone portable…

Ce questionnaire est constitué de 3 parties et d'une question finale :

Partie A : votre expérience actuelle des jeux vidéo (ces 12 derniers mois)
Partie B : votre expérience passée des jeux vidéo (avant ces 12 derniers mois)
Partie C : Votre plus importante période d'expérience de jeux vidéo

Il est important de répondre à chaque question (sauf mention contraire).

Pour répondre, vous devrez soit entourer une ou plusieurs bonnes réponses, soit écrire directement la réponse.

Partie A : Votre expérience actuelle des jeux vidéo (ces 12 derniers mois)

Ces 12 derniers mois, à quelle fréquence avez-vous joué à des jeux vidéo ? (*Entourez la réponse de votre choix*)

Tous les jours Toutes les semaines Tous les mois Moins d'1 fois par mois Jamais

> *Si jamais, passez à la partie B*

Ces 12 derniers mois, quelle était la durée d'une session typique de jeu vidéo ? (*Indiquez la durée en temps*)

.....

Ces 12 derniers mois, combien d'heures par semaine en moyenne avez-vous joué à des jeux vidéo ?
(*Entourez la réponse de votre choix*)

0-1 h 1-2h 3-5h 5-10h + de 10h

Ces 12 derniers mois, à quel(s) type(s) de jeux vidéo avez-vous joué? *(Entourez la(les) réponse(s) de votre choix)*

Action Aventure Course Plate-forme Sport Jeu de rôle – RPG Party game

Tir – FPS Combat Puzzle Réflexion Gestion Stratégie

Autres :

Partie B : Votre expérience passée des jeux vidéo (avant ces 12 derniers mois)

Avant ces 12 derniers mois, avez-vous joué à des jeux vidéo ? (*entourez la réponse de votre choix*)

Oui Non

> *Si Non, le questionnaire est terminé*

Depuis quel âge jouez-vous à des jeux vidéo ? (*Indiquez l'âge en années*)

..... ans

Avant ces 12 derniers mois, à quelle fréquence avez-vous joué à des jeux vidéo ? (*Entourez la réponse de votre choix*)

Tous les jours Toutes les semaines Tous les mois Tous les ans Moins d'1 fois par an

(Suite de la partie B)

Avant ces 12 derniers mois, à quel(s) type(s) de jeux vidéo avez-vous joué? *(Entourez la(les) réponse(s) de votre choix)*

Action Aventure Course Plate-forme Sport Jeu de rôle – RPG Party game

Tir – FPS Combat Puzzle Réflexion Gestion Stratégie

Autres :

Partie C : Votre plus importante période d'expérience des jeux vidéo

A quelle période de votre vie avez-vous le plus joué à des jeux vidéo ?
(Indiquez les âges en années correspondant au début et à la fin de la période)

A partir de l'âge de ..... ans, jusque l'âge de ..... ans

Sur cette période, à quelle fréquence avez-vous joué à des jeux vidéo ? (*Entourez la réponse de votre choix*)

Tous les jours Toutes les semaines Tous les mois Moins d'1 fois par mois

Sur cette période, quelle était la durée d'une session typique de jeu vidéo ? (*Indiquez la durée en temps*)

.....

Sur cette période, combien d'heures par semaine en moyenne avez-vous joué à des jeux vidéo ?
(*Entourez la réponse de votre choix*)

0-1 h 1-2h 3-5h 5-10h + de 10h

Sur cette période, à quel(s) type(s) de jeux vidéo avez-vous joué? *(Entourez la(les) réponse(s) de votre choix)*

Action Aventure Course Plate-forme Sport Jeu de rôle – RPG Party game

Tir – FPS Combat Puzzle Réflexion Gestion Stratégie

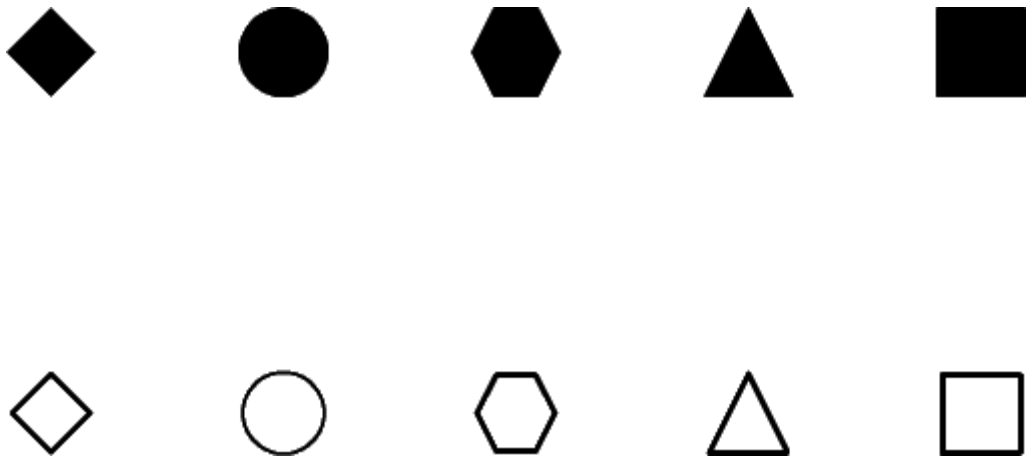
Autres :

Question finale

Sur toute votre vie, combien d'années au total (toutes fréquences confondues) avez-vous joué à des jeux vidéo ?

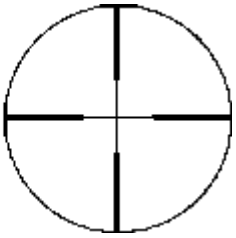
..... années

Annexe 4 : Formes géométriques utilisées dans les expériences 3a et 3b



Les formes sont à l'échelle 1 : 1.

Annexe 5 : Cible et curseur utilisés dans l'expérience 4



La cible et le curseur sont à l'échelle 1 : 1.

Annexe 6 : Créatures utilisées comme cible ou distracteur et curseur des expériences 5 et 6

Les créatures et le curseur sont à l'échelle 1 : 1.

Le curseur de l'expérience 5 ne comportait pas de liseré noir.

Annexe 7 : Table des figures

Figure 1. Captures d'écran des interfaces visuelles de deux jeux vidéo commerciaux récents. 22

Figure 2. Représentation graphique (adaptée de l'anglais) des quatre dimensions de la théorie des ressources multiples selon Wickens (2002, 2008). 33

Figure 3. Captures d'écran des interfaces du jeu dans les quatre conditions de positions du score des expériences 1 et 2. 72

Figure 4. Captures d'écran du jeu vidéo utilisé dans l'expérience 1 dans les conditions Haut et Bas.. 74

Figure 5. Côtés droits des captures d'écran du jeu vidéo utilisé dans l'expérience 2 dans les quatre conditions de position du score. 83

Figure 6. Mesures moyennes obtenues pour les trois variables dépendantes dans chaque condition de l'expérience 2. 84

Figure 7. Les trois arrière-plans utilisés dans l'expérience 3a avec des exemples de champs de recherche..... 95

Figure 8. Coordonnées moyennes de la fixation initiale (en pixels) selon le type d'arrière-plan utilisé dans l'expérience 3a..... 99

Figure 9. Coordonnées moyennes de la fixation initiale (en pixels) selon le type d'arrière-plan utilisé dans l'expérience 3b..... 104

Figure 10. Les cinq arrière-plans utilisés dans l'expérience 4 avec des exemples de position initiale de cible et de curseur. 108

Figure 11. Nombre d'ennemis touchés par partie dans l'expérience 4 dans chaque condition (structure visuelle x nature de l'arrière-plan). 112

Figure 12. Temps moyen (en ms) pour toucher chaque cible dans l'expérience 4 dans chaque condition (structure visuelle x nature de l'arrière-plan). 112

Figure 13. Durée des moyennes des fixations oculaires en fonction du type de structure de l'arrière-plan et de la présence ou de l'absence de l'indice sonore dans l'expérience 4. 113

Figure 14. Coordonnées moyennes de la fixation initiale (en pixels) selon la structure et la nature de l'arrière-plan dans l'expérience 4..... 115

Figure 15. Les deux types de décor utilisés pour les arrière-plans dans l'expérience 5 avec différents exemples de dispositions d'une cible, d'un distracteur et du curseur. 126

Figure 16. Temps de réponse moyen (en ms) pour toucher chaque cible dans l'expérience 5 dans chaque condition (taux de recouvrement X mouvement)..... 129

Figure 17. Les deux types de décor utilisés pour les arrière-plans dans l'expérience 6 et des exemples de dispositions de la cible, des distracteurs et du curseur en fonction du niveau de difficulté de la tâche..... 137

Figure 18. Temps de réponse moyen (en ms) pour toucher chaque cible dans l'expérience 6 dans chaque condition (difficulté de la tâche X mouvement)..... 139

Annexe 8 : Table des tableaux

Tableau 1. *Résumé des caractéristiques principales des quatre conditions de position du score des expériences 1 et 2.* 70

Tableau 2. *Nombre et proportion moyens de fixations oculaires faites par les participants dans chaque AOI et dans chaque condition de l'expérience 1.* 76

Tableau 3. *Temps de réponse moyens (en ms) dans les expériences 3a et 3b.* 98

Tableau 4. *Prédiction de la diminution des ressources attentionnelles allouées aux objets (fixes) impliqués dans la tâche principale au bénéfice des éléments de l'arrière-plan, en fonction des caractéristiques du mouvement et de la structure associée de l'arrière-plan sur lequel ils sont superposés, ainsi que de la complexité de la tâche.* 148

Annexe 9 : Table des matières détaillée

TABLE DES MATIERES **5**

INTRODUCTION **9**

1. Les jeux vidéo aujourd'hui **9**

2. Vers une utilisabilité spécifique des jeux vidéo ? **12**
2.1. De l'utilisabilité à l'expérience utilisateur 12
2.2. Modèles d'utilisabilité adaptés aux jeux vidéo 14

3. Présentation de la thèse **16**

CHAPITRE 1 : LES INTERACTIONS JOUEUR-JEU VIDEO **19**

1. Les jeux vidéo comme environnements virtuels **20**
1.1. Caractéristiques des interfaces des jeux vidéo 20
1.1.1. Les informations visuelles 20
1.1.2. Les informations auditives et haptiques 23
1.2. Le point de vue de l'utilisateur 24
1.2.1. Engagement, immersion, présence 24
1.2.2. Les sensations ou effets ressentis 27

2. Les jeux vidéo comme environnements dynamiques complexes **28**
2.1. Interaction avec des environnements dynamiques complexes 28
2.2. Partage attentionnel lors de l'interaction avec des environnements dynamiques complexes 31
2.2.1. L'attention dans des situations multitâches 31
2.2.2. La théorie des ressources multiples 32
2.2.3. Intégration de plusieurs sources d'information 35

3. Synthèse **37**

CHAPITRE 2 : PERCEPTION DES ENVIRONNEMENTS VISUELS DYNAMIQUES COMPLEXES **39**

1. La perception visuelle de scènes complexes **40**
1.1. Le guidage de l'attention visuelle 40
1.1.1. Guidage ascendant 41
1.1.2. Guidage descendant 42
1.2. Les mouvements oculaires en tant qu'indicateurs de l'attention visuelle 43
1.3. La recherche visuelle 44
1.3.1. Principes 45
1.3.2. Recherche visuelle et multimodalité 46

2. Perception et interaction personne-système dans des scènes visuelles complexes et/ou dynamiques **48**
2.1. Scènes visuelles dynamiques 48
2.1.1. Perception d'objets en mouvement 48
2.1.2. Les effets d'arrière-plans en mouvement 49
2.2. Scènes visuelles complexes 54
2.2.1. Principes 54
2.2.2. Influence de la complexité sur la performance d'un individu 55

3. Synthèse **58**

CHAPITRE 3 : SYNTHESE ET HYPOTHESES GENERALES 61

1. Synthèse des principaux éléments théoriques 61

2. Raisonnement et hypothèses générales 63

CHAPITRE 4 : MAXIMISER L'ANTICIPATION VISUELLE DES JOUEURS EN APPLIQUANT LE PRINCIPE DE PROXIMITE-COMPATIBILITE A LA CONCEPTION DES JEUX VIDEO 67

1. Introduction 67

2. Expérience 1 : Score en haut ou en bas de l'écran de jeu 70
2.1. Méthode 70
2.2. Résultats 75
2.3. Discussion 77
2.3.1. Hypothèse préliminaire: balayage visuel de l'interface visuelle du jeu 77
2.3.2. Amplitude réduite du balayage visuel dans la condition Bas 77
2.3.3. Influence d'une amplitude réduite du balayage visuel sur la performance 78
2.3.4. Raisonnement de l'expérience 2 79

3. Expérience 2 : Score en haut, en bas, au milieu ou en dessous de l'écran de jeu 80
3.1. Méthode 80
3.2. Résultats 82
3.3. Discussion 84

4. Discussion générale 86

CHAPITRE 5 : LA PERFORMANCE A UN JEU VIDEO EST PLUS DEGRADEE PAR LA COMPLEXITE STRUCTURALE DE L'ARRIERE-PLAN QUE PAR SON MOUVEMENT 89

1. Introduction 89

2. Expérience 3a : Tâche de recherche visuelle 92
2.1. Méthode 94
2.2. Résultats 97
2.3. Discussion 99

3. Expérience 3b : Tâche de recherche visuelle et structure radiale 101
3.1. Méthode 101
3.2. Résultats 102
3.3. Discussion 104
3.4. Discussion générale des expériences 3a et 3b 105
3.4.1. La tâche de recherche visuelle 105
3.4.2. Influence de la complexité structurale de l'arrière-plan sur la recherche visuelle 105

4. Expérience 4 : Tâche de tir 106
4.1. Méthode 107
4.2. Résultats 110
4.3. Discussion 115

5. Discussion générale 118

CHAPITRE 6 : COMPLEXITE VISUELLE DE L'ARRIERE-PLAN ET DIFFICULTE DE LA TACHE INTERAGISSENT AVEC LE MOUVEMENT DE L'ARRIERE-PLAN DANS UN JEU VIDEO 121

1. Introduction 121

2. Expérience 5 : Complexité et mouvement de l'arrière-plan 123
2.1. Méthode 124
2.2. Résultats 128
2.2.1. Temps de réponse moyen 128
2.2.2. Durée moyenne des fixations oculaires 131
2.2.3. Fixation initiale 131
2.3. Discussion 131
2.4. Conclusion 134

3. Expérience 6 : Difficulté de la tâche et mouvement de l'arrière-plan 135
3.1. Méthode 136
3.2. Résultats 138
3.2.1. Temps de réponse moyen 139
3.2.2. Durée moyenne des fixations oculaires 140
3.2.3. Fixation initiale 140
3.3. Discussion 141

4. Discussion générale 142

CHAPITRE 7 : DISCUSSION GENERALE 145

1. Partage attentionnel dans les interfaces visuelles de jeux vidéo 146
1.1. Modèles et théories existants 146
1.2. Modèle de partage attentionnel entre deux sources visuelles d'information superposées 148
1.3. Eléments non pris en compte dans le modèle 150

2. Recommandations pour la conception des jeux vidéo 152
2.1. Anticipation visuelle du mouvement et positionnement des informations contextuelles 153
2.2. Mouvement de l'arrière-plan, structure visuelle, et complexité de la tâche 154

3. Perspectives 156

CONCLUSION 161

BIBLIOGRAPHIE 165

ANNEXES 181
Annexe 1 : Prépublication de l'article présenté au colloque EPIQUE 2011 182
Annexe 2 : Questionnaire pré-expérimental des expériences 1 et 2 191
Annexe 3 : Questionnaire pré-expérimental des expériences 3 à 6 192
Annexe 4 : Formes géométriques utilisées dans les expériences 3a et 3b 195
Annexe 5 : Cible et curseur utilisés dans l'expérience 4 196
Annexe 6 : Créatures utilisées comme cible ou distracteur et curseur des expériences 5 et 6 ... 197
Annexe 7 : Table des figures 198
Annexe 8 : Table des tableaux 200
Annexe 9 : Table des matières détaillée 201

Résumé

Les jeux vidéo occupent une place importante dans notre société. Cependant, leur conception ne prend aujourd'hui que peu en compte les spécificités de l'interaction joueur-jeu vidéo, qui sont déterminantes pour une expérience de jeu optimale. L'objectif de cette thèse était de comprendre l'influence de différents choix de conception des interfaces visuelles des jeux vidéo sur la performance et le comportement du regard des joueurs. Ces interfaces, généralement dynamiques et complexes, sont composées de trois sources d'information : les objets avec lesquels le joueur interagit et l'arrière-plan, qui composent la scène d'action principale, ainsi que les informations contextuelles superposées à la scène principale. Sept expériences ont été réalisées pour comprendre le partage attentionnel entre ces sources dans le cadre d'une activité de jeu vidéo. Les caractéristiques de l'arrière-plan, des informations contextuelles et de la tâche ont été manipulées. La performance et les mouvements du regard des joueurs ont été enregistrés. Les résultats ont montré que le partage attentionnel entre deux sources d'information (e.g., scène d'action et informations contextuelles) est facilité lorsqu'elles ne se chevauchent pas. Lorsque qu'elles se chevauchent nécessairement (e.g., objets et arrière-plan), les caractéristiques de mouvement et de complexité de l'arrière-plan et la difficulté de la tâche modulent très largement la dégradation de la performance. Un modèle théorique de partage attentionnel entre deux sources visuelles d'information superposées est proposé. Des recommandations sont établies pour la conception des jeux vidéo, mais aussi des environnements virtuels en général.

Mots-clés : Interaction joueur-jeu vidéo, Attention visuelle, Ressources attentionnelles, Mouvements oculaires, Expérience joueur.

Abstract

Video games take an increasingly important place in our society. Nowadays, however, their design doesn't take enough into account the specificity of player-game interaction, which is essential for an optimal play experience. The thesis' goal was to understand the impact of different design choices for video games visual interfaces on players' performance and gaze behavior. These dynamic and complex interfaces include three information sources: the objects that the player interacts with and the background, which make up together the main action scene, and the contextual information superimposed on the main scene. Seven experiments were made to understand how attention is shared between these sources during video game activity. The characteristics of the background, of the contextual information and of the game task were manipulated. The performance was measured while participants' eye movements were recorded. The results showed that attention sharing between two sources (e.g., the action scene and contextual information) is optimal when the two sources do not overlap at all. When they necessarily overlap (e.g., objects superimposed on a background), the nature of background movement, the background visual complexity and the difficulty of the task can strongly modulate the performance decrease. A theoretical model of attention sharing between two superimposed sources of visual information is proposed. Practical recommendations are drawn up for video game design, but also for virtual environments design.

Keywords: Player-video game interaction, Visual attention, Attentional resources,
Eye movements, Player experience.