

Université Paris 8 - Vincennes/Saint-Denis
École doctorale Cognition, Langage, Interaction
Spécialité : Psychologie Cognitive

EA 4004 - CHArt - Cognitions Humaine et Artificielle
FED 4246 - LUTIN - Laboratoire des Usages en Technologies d'Information Numériques

Bertin Ergonomie et Facteur Humain

Thèse

pour obtenir le grade de Docteur en Psychologie

ALLOCATION DE L'ATTENTION VISUELLE EN CONDUITE AUTOMOBILE : L'IMPACT DE LA REALITE AUGMENTEE

Présentée et soutenue publiquement par

Robin EYRAUD

le 23 décembre 2016

Sous la co-direction de

Thierry BACCINO et Elisabetta ZIBETTI

Membres du jury :

Thierry BACCINO	<i>Directeur</i>	<i>Professeur</i>	<i>Univ. Paris 8- CHArt</i>
Jean-Michel BOUCHEIX	<i>Rapporteur</i>	<i>Professeur</i>	<i>Univ. de Bourgogne – LEAD</i>
Patrice TERRIER	<i>Rapporteur</i>	<i>Professeur</i>	<i>Univ. de Toulouse - CLLE-LTC</i>
Charles TIJUS	<i>Président</i>	<i>Professeur</i>	<i>Univ. Paris 8- CHArt</i>
Marie VIAN	<i>Responsable Entreprise</i>	<i>Ergonome</i>	<i>Bertin Ergonomie</i>
Elisabetta ZIBETTI	<i>Co-Directeur</i>	<i>Maître de Conférences</i>	<i>Univ. Paris 8- CHArt</i>

Recherche réalisée en contrat CIFRE entre l'Université Paris 8 et Bertin Ergonomie
CONVENTION CIFRE N° 2011/0363



Une partie de ce manuscrit a fait l'objet d'une publication parue :

Eyraud, R., Zibetti, E., Baccino, T. (2015). Allocation of Visual Attention while Driving with Simulated Augmented Reality. *Transportation Research Part F: Traffic Psychology and Behaviour*, 32. 46-55.

Titre : Allocation de l'attention visuelle en conduite automobile : l'impact de la réalité augmentée.

Résumé : L'attention est le processus qui consiste à filtrer les informations utiles à l'activité, de celles qui lui sont inutiles. La réalité augmentée (RA) de surlignage guide ce processus de sélection de l'information, en mettant en valeur certains éléments par rapport à d'autres. Telle qu'envisagée actuellement pour la conduite automobile, la RA de surlignage met en évidence des éléments liés à l'activité générale de conduite (e.g. panneaux en cas de mauvaise visibilité, direction à emprunter), mais indépendamment des manœuvres. Or, la littérature sur l'attention visuelle en activité nous montre que les parcours oculaires sont très spécifiques aux buts et sous-buts immédiats. Une RA qui ne respecte pas cette priorisation « naturelle » du traitement de l'information risque donc de perturber la prise d'informations.

Le premier objectif de cette recherche est de déterminer dans quelle mesure l'allocation de l'attention visuelle en conduite automobile se concentre sur les informations liées à la manœuvre. Le second objectif est d'étudier l'impact de la RA sur cette allocation de l'attention.

Nous avons mis en place trois expérimentations dans lesquelles les participants visualisaient des scènes statiques et dynamiques de conduite automobile, et devaient décider s'ils pouvaient réaliser une manœuvre. Nous avons analysé les variations d'allocation de l'attention visuelle selon les manœuvres notifiées et les conditions de RA à l'aide d'enregistrements oculométriques. Nos résultats montrent que l'attention visuelle est fortement allouée aux indices permettant la prise de décision, mais qu'elle ne néglige pas les indices permettant la compréhension générale de la scène. La RA optimise l'attention visuelle lorsqu'elle met en évidence des indices liés à la manœuvre, mais elle perturbe l'attention visuelle dans les autres conditions. Ces résultats permettent d'identifier et de caractériser différents risques inhérents à la RA de surlignage, et de discuter des pistes de conception pour les prendre en compte.

Mots-clés : conduite automobile, allocation de l'attention visuelle, réalité augmentée.

Title: Allocation of visual attention while driving: the impact of augmented reality.

Abstract: Attention involves distinguishing information which is useful for an activity from that which is not. Augmented reality (AR) highlighting guides this process of information selection by increasing the salience of high-value elements. In such systems, “value” is typically seen as linked to the overall activity of driving (e.g. traffic signs in case of poor visibility, or indications of direction). However, several studies have shown that during an activity eye movements are specific to the immediate goal. AR which does not respect this “natural” prioritizing of information thus risks interfering with information-processing.

The first objective of this study is to determine to what extent the allocation of visual attention in automobile driving is focused on information related to a maneuver. The second objective is to study the impact of AR on this allocation of attention.

We set up three experiments in which participants viewed static and dynamic driving situations and had to decide whether they could perform a maneuver. We analyzed the variations in allocation of visual attention in relation to the maneuvers in question and the AR conditions using eye movements recording. Our results show that visual attention is strongly linked to cues which permit decision-making, but that it does not overlook cues allowing a general comprehension of the situation. AR optimizes visual attention when it highlights cues related to the maneuver, but it disturbs visual attention in other cases. These findings make it possible to identify and categorize various risks inherent in AR highlighting, and to discuss possible approaches to dealing with them through more effective design.

Key-words: driving, allocation of visual attention, augmented reality.

Remerciements

Je tiens à remercier mes directeurs de thèse, Thierry Baccino et Elisabetta Zibetti, qui m'ont permis de réaliser cette recherche. Merci à Thierry pour son suivi exigeant et bienveillant, et pour l'autonomie qu'il m'a accordé tout au long de ce travail. Merci à Elisabetta pour ses encouragements et ses relectures aiguisées (avec toujours le souci d' « élargir ma pensée compactée » !). Plus largement, merci pour son accompagnement tout au long de mon parcours universitaire.

Je remercie également les membres du jury, Jean-Michel Boucheix, Patrice Terrier et Charles Tijus, d'avoir accepté d'examiner ce travail.

Je remercie Dominique Soler de m'avoir accueilli au sein de Bertin Ergonomie dans le cadre du contrat CIFRE ; et de m'avoir permis - à chaque étape de ce travail - d'échanger avec des acteurs de l'industrie automobile concernés par notre problématique de recherche.

Je remercie Marie Vian pour tout ce qu'elle m'a appris durant mon passage à Bertin Ergonomie, professionnellement et personnellement. Merci pour sa confiance, et merci de m'avoir donné confiance ...

Merci aussi à toute l'équipe de Bertin Ergonomie pour son accueil. Ce fut trois années remplies de bons moments et de belles rigolades ! Merci à l'équipe du Lutin pour son aide lors de la mise en place des passations expérimentales.

Merci à Jocelyn et Adrien pour leur aide lors des prises de vue en conduite automobile, à Mike pour m'avoir formé à Adobe After Effects et à Sebastian pour son aide en anglais.

Pour finir, un grand merci à ma famille et à ma belle famille pour leur soutien. Et merci à Julie, sans qui rien n'est jamais possible.

TABLES DES MATIERES

Préambule	12
1 INTRODUCTION.....	15
1.1 LA REALITE AUGMENTEE.....	15
1.1.1 Définition.....	15
1.1.2 Applications.....	16
1.1.3 Bénéfices potentiels de la RA en conduite automobile	20
1.1.4 Risques potentiels de la RA en conduite automobile	20
1.2 L'ATTENTION EN CONDUITE AUTOMOBILE	21
1.2.1 L'activité de conduite automobile	21
1.2.2 L'attention du conducteur.....	22
1.2.3 Les défauts d'attention.....	23
1.3 L'ORIENTATION DE L'ATTENTION VISUELLE.....	25
1.3.1 L'orientation exogène de l'attention.....	26
1.3.2 L'orientation endogène de l'attention.....	27
1.3.3 Intégration des facteurs endogènes et exogènes : le modèle de Wickens et al. (2001) 28	
1.3.4 L'attention visuelle optimale	30
1.3.5 La Valeur d'une information visuelle.....	31
1.3.6 Spécificité de l'attention visuelle au but de la tâche.....	33
2 EXPERIMENTATIONS	38
2.1 HYPOTHESES ET EXPERIMENTATIONS	38
2.1.1 Hypothèses.....	38
2.1.2 Choix Méthodologiques	39

2.1.3	Expérimentations	40
2.2	EXPERIMENTATION I : ALLOCATION DE L'ATTENTION VISUELLE SUR DES SCENES STATIQUES	41
2.2.1	Hypothèses et prédictions	41
2.2.2	Méthode	42
2.2.2.1	Participants.....	42
2.2.2.2	Matériel	42
2.2.2.3	Dispositif expérimental.....	47
2.2.2.4	Procédure	48
2.2.2.5	Plan expérimental.....	48
2.2.3	Résultats.....	50
2.2.3.1	Fréquences de fixations sur les AOIs.....	50
2.2.3.2	Pourcentage de temps de fixation sur les AOIs	51
2.2.3.3	Pourcentage d'AOIs fixées	52
2.2.4	Discussion.....	53
2.3	EXPERIMENTATION II : ALLOCATION DE L'ATTENTION VISUELLE SUR DES SCENES DYNAMIQUES	57
2.3.1	Méthode	57
2.3.1.1	Participants.....	57
2.3.1.2	Matériel	58
2.3.1.3	Procédure	58
2.3.2	Résultats.....	59
2.3.2.1	Fréquences de fixations sur les AOIs.....	60
2.3.2.2	Pourcentage de temps de fixation sur les AOIs	61
2.3.2.3	Proportion d'AOIs fixés.....	61
2.3.3	Discussion.....	62
2.4	EXPERIMENTATION III : ALLOCATION DE L'ATTENTION VISUELLE SUR DES SCENES DYNAMIQUES AVEC REALITE AUGMENTEE.....	65

2.4.1	Introduction	65
2.4.2	Méthode	68
2.4.2.1	Participants.....	68
2.4.2.2	Matériel	69
2.4.2.3	Dispositif expérimental.....	72
2.4.2.4	Procédure	72
2.4.2.5	Plan expérimental.....	73
2.4.2.6	Variables dépendantes.....	73
2.4.3	Résultats.....	74
2.4.3.1	Fréquences de fixations.....	75
2.4.3.2	Pourcentage de temps de fixation	77
2.4.4	Discussion.....	79
3	DISCUSSION GENERALE	84
3.1	RESUME DU CONTEXTE ET DES ENJEUX.....	84
3.2	L'ALLOCATION DE L'ATTENTION SUR DES SCENES VISUELLES DE CONDUITE AUTOMOBILE	84
3.3	RISQUES ET BENEFICES POTENTIELS DE LA REALITE AUGMENTEE DE SURLIGNAGE.....	88
3.3.1	RA et prise d'information.....	88
3.3.2	RA et risques de distraction.....	89
3.3.3	RA et risques de distraction selon les phases de l'activité de conduite.....	90
3.3.4	RA et optimisation de l'allocation attentionnelle	91
3.4	CONCLUSION ET PERSPECTIVES.....	92
4	BIBLIOGRAPHIE	94
5	ANNEXES	104
5.1	Exemples de cartes de saillance statiques (Expérimentation I)	105
5.2	Exemples de cartes de saillance dynamiques (Expérimentations II et III)	106
5.3	Exemples de scènes dynamiques avec RA (Expérimentation III)	107

Préambule

On observe aujourd'hui une multiplication des technologies nomades utilisées en conduite automobile, comme le GPS, les avertisseurs de radars, le téléphone, les applications pour smartphone, et plus récemment les systèmes embarqués qui synchronisent les différents appareils et/ou intègrent toutes ces fonctions sur la même interface. Si certaines de ces technologies rajoutent des tâches qui ne sont pas liées à la conduite, d'autres ont pour objectif de supplanter et de rendre plus sûrs des usages déjà existants. Par exemple, l'utilisation du GPS évite de lire une carte tout en conduisant. Mais la frontière entre les bénéfices et les effets néfastes d'un système est parfois mince.

Parmi les systèmes d'aide à la conduite, la réalité augmentée (RA) est une technologie dont les effets (positifs ou négatifs) sur l'activité de conduite sont encore mal connus. A ce jour, les systèmes de RA dans le domaine automobile sont au début de leur développement, et ne sont pas encore implantés dans des véhicules commercialisés. Nous manquons donc de recul sur leurs effets potentiels sur la conduite. Les recherches applicatives sur le sujet proviennent principalement du domaine de l'aviation, où des technologies de RA sont déjà mises en œuvre (e.g. Caird, Horrey, & Edwards, 2001; Wickens & Long, 1995). En revanche dans le domaine automobile, les travaux sur le sujet sont encore prospectifs (Narzt & Pomberger, 2004). Pourtant l'enjeu est crucial, car en modifiant l'environnement visuel du conducteur, la RA peut impacter très directement son attention visuelle et sa prise d'informations.

L'impact d'un dispositif sur l'attention en conduite automobile est généralement analysé de deux façons : les mesures de performances et les mesures d'allocation attentionnelle.

Les mesures de performances portent par exemple sur les temps de réaction ou les écarts de trajectoire. Certaines sont prévues par des normes, comme l'ISO 26022:2010 (« *Simulated lane change test to assess in-vehicle secondary task demand* »). La tâche décrite par cette norme consiste en une simulation de conduite, durant laquelle le conducteur doit changer de voie à chaque injonction (à intervalle régulière), tout en réalisant une tâche secondaire. La moyenne des écarts par rapport à une trajectoire de référence donne une indication sur la compatibilité de la tâche secondaire avec la conduite.

L'allocation attentionnelle correspond à la manière dont le conducteur répartit son attention visuelle entre les différentes informations. Un indicateur fréquemment utilisé est le TGT (*Total Glance Time*) : il s'agit du temps cumulé de prise d'information visuelle pour la réalisation d'une tâche secondaire. Ainsi des seuils d'acceptabilité ont été établis par certains constructeurs ou consortiums de constructeurs. Ces seuils d'acceptabilité varient fortement selon les organismes : par exemple le TGT maximum acceptable pour la réalisation d'une tâche secondaire à la conduite est de 8 secondes pour la JAMA¹ [*Japan Automobile Manufacturers Association*] et de 20 secondes pour l'AAM² [*Alliance Automobile Manufacturers –USA*]).

Cependant, la RA présente une particularité : elle ne « s'ajoute » pas au cockpit du véhicule comme c'est le cas pour un écran, mais se fond dans l'environnement de conduite. Les modifications d'allocation d'attention visuelle s'en trouvent ainsi difficilement identifiables. C'est particulièrement le cas de la RA dite « **de surlignage** » (Narzt & Pomberger, 2004) qui consiste à mettre en saillance des éléments de l'environnement (par exemple un panneau) sans réellement y ajouter d'informations.

L'objectif de notre recherche est d'étudier les modifications d'allocation visuelle potentiellement induites par la RA « de surlignage ». Il s'agira dans un premier temps de trouver un angle de vue permettant d'analyser les variations d'allocation. Puis, nous pourrons étudier plus précisément l'impact de ce type de RA sur l'allocation attentionnelle en conduite automobile. Dans cette perspective, nous utiliserons l'analyse des mouvements oculaire pour rendre compte de la répartition de l'attention visuelle.

Ce travail de thèse est organisé de la manière suivante. L'**Introduction**, présente la technologie de réalité augmentée, ses applications en conduite automobile puis les bénéfices et risques inhérents à cette technologie. Ensuite seront abordées les notions d'attention et de défaut d'attention en activité conduite. Nous présenterons enfin le modèle théorique d'allocation de l'attention visuelle (SEEV, Wickens et al., 2001) sur lequel nous nous appuierons tout au long de cette recherche. Le chapitre **Expérimentation** expose nos problématiques et hypothèses de recherche. Il décrit les trois expérimentations que nous avons mis place pour y répondre. Dans la **Discussion générale** nous considérerons nos résultats au vu de l'état de l'art et du modèle théorique (SEEV) qui nous a servi de référence. Nous mènerons enfin une réflexion sur les implications pratiques de notre étude.

¹<http://www.jama-english.jp/>

²<http://www.autoalliance.org/>

1 INTRODUCTION

1.1 LA REALITE AUGMENTEE

1.1.1 Définition

Un système de réalité augmentée (RA) est un système qui complète le monde réel avec des éléments virtuels qui semblent coexister dans le même espace que le monde réel (Azuma et al., 2001). C'est une technologie émergente, que beaucoup considèrent comme capable de révolutionner l'interaction homme-machine (Krevelen & Poelman, 2010). En 2010, Krevelen et Poelman relèvent trois types de procédés :

(1) La RA à travers la vidéo est le dispositif le plus élémentaire. Il consiste à implémenter en temps réel des objets virtuels à l'intérieur d'une vidéo de l'environnement. Ainsi, l'opérateur regarde son environnement à travers une caméra (e.g. à travers un casque vidéo ou un smartphone) et des éléments virtuels qui s'y greffent.

(2) Le second type de procédé est optique : il consiste à projeter à l'aide d'un faisceau laser (ou LED) les objets virtuels sur un miroir transparent (e.g. des lunettes, un objectif, un pare-brise). Ainsi, l'opérateur voit les objets se superposer directement dans l'environnement réel. Une alternative consiste à projeter directement l'image sur la rétine à l'aide d'un laser de faible puissance.

(3) Le dernier type de procédé consiste à utiliser un vidéoprojecteur afin de projeter les objets virtuels dans l'environnement (e.g. sur un mur, sur une maquette).

Quels que soient les procédés, la problématique principale pour le développement de ces technologies consiste à localiser les éléments virtuels par rapport au monde réel. Il s'agit de construire une représentation visuelle où les deux mondes sont vus suivant la même perspective. Pour cela, ces technologies utilisent par exemple des données GPS, des données issues de radars ou des algorithmes de reconnaissance d'images.

1.1.2 Applications

C'est dans le domaine militaire et aérien que les premières applications de RA sont apparues, notamment avec la généralisation des dispositifs tête-haute (*Head-Up Display*, HUD) dans les avions de combat. Des informations relatives au paramètre de vol (e.g. cap, inclinaison) ou à la mission (e.g. cibles) sont projetées sur un miroir transparent. L'utilisation de cette technologie a été ensuite étendue aux avions de ligne, par exemple pour faciliter le repérage de la piste d'atterrissage la nuit ou en cas de brouillard. Les HUD les plus récents affichent des données issues d'une caméra infrarouge ou des images radar en les superposant à l'image extérieure, afin de donner une vision « augmentée » au pilote lorsque la visibilité est nulle (Figure 1).

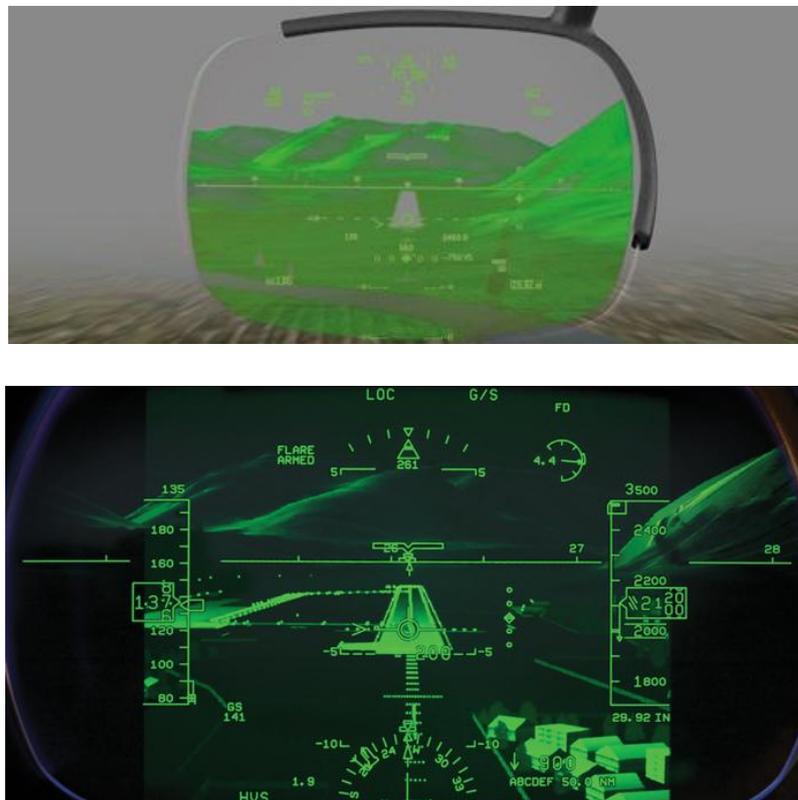


Figure 1. HUD EVS-3000 de Rockwell Collins destiné aux avions de ligne. Ici, les données sont projetées sur un miroir transparent placé devant les yeux du pilote. Ils sont particulièrement utiles par temps de brouillard (image du haut) ou de nuit (image du bas).³

Dans le domaine automobile, plusieurs constructeurs commencent à proposer des HUD. Ces appareils projettent sur un miroir transparent ou sur le pare-brise des informations relatives aux paramètres de conduite du véhicule (e.g. vitesse) ou à la navigation (e.g. instructions du GPS, limitation de vitesse) (Figure 2 ; Figure 3). Cependant, ces HUD

³Source : <http://www.rockwellcollins.com>

n'intègrent pas encore de technologie de réalité augmentée : les informations sont dans le champ de vision du conducteur, mais elles sont indépendantes de l'environnement.



Figure 2. HUD de la Lexus RX. Affichage des instructions de navigation et de la vitesse du véhicule



Figure 3. HUD de la BMW série 7. Affichage de la distance inter-véhicule, de la vitesse, et des instructions de navigation.

Il reste encore des freins technologiques à la généralisation de véritables systèmes de réalité augmentée dans les automobiles. Notamment en ce qui concerne les aspects d'affichage (projection sur l'ensemble du pare-brise, adaptabilité aux conditions de luminosité, gestion des perspectives) et les aspects d'analyse de la situation (détection et analyse des panneaux, des véhicules, etc.).

Malgré le manque de maturité de ces technologies, de nombreuses applications et modalités de RA ont été envisagées, et cela dès le début des années 2000. George et al. (2011) en proposent un inventaire ; les principales modalités sont les suivantes :

- « La route augmentée » consiste à surligner la voie à emprunter (Figure 4 ; Figure 5). C'est une aide à la navigation à court terme (car sur une distance limitée). Cela permettrait par exemple au conducteur de remarquer une sortie cachée qu'il aurait pu manquer (Narzt, Pomberger et al., 2003).
- « Le surlignage » consiste à mettre en valeur certains éléments du champ visuel du conducteur, dans le but d'attirer son attention. Il est par exemple possible de surligner un véhicule, un piéton ou les contours des voies (Narzt & Pomberger, 2004).
- « L'annotation » consiste à indiquer des points d'intérêt dans l'environnement (e.g. station essence, prix de l'essence) (Narzt, Pomberger et al., 2006).

On peut également citer « la voiture à suivre » qui consiste à insérer devant le véhicule une voiture virtuelle se comportant comme un vrai véhicule, accélérant, freinant et activant ses clignotants (Narzt, Pomberger et al., 2003). Ici aussi, c'est une aide à la navigation, à plus court terme encore que la route augmentée. On peut également évoquer le prototype de la société Virtual Cable TM qui consiste à créer l'illusion 3D d'un câble de tramway se trouvant au-dessus du conducteur, indiquant le chemin à emprunter (Figure 6). En 2014, Land Rover a dévoilé un concept de « capot transparent » qui consiste à afficher sur le pare-brise une vidéo du terrain caché par le capot, ce qui a pour effet de rendre l'avant du véhicule "virtuellement invisible" depuis l'habitacle (Figure 7). Cette technologie permettrait de mieux visualiser le terrain, la position et le braquage des roues, et ainsi de faciliter les manœuvres sur les sentiers difficiles.



Figure 4. Concept de la route augmentée (Narzt, Pomberger et al., 2003)



Figure 5. Projet de réalité augmentée de PSA : le marquage au sol est augmenté afin de mettre en évidence la trajectoire à suivre.⁴



Figure 6. Prototype du Virtual Cable™. Un faisceau laser simule un câble indiquant la voie à prendre.⁵



Figure 7. Concept du « capot transparent » de Land rover.⁶

⁴Source : <http://www.psa-peugeot-citroen.com/fr/>

⁵Source : <http://www.mvs.net>

⁶Source : <http://www.landrover.com>

1.1.3 Bénéfices potentiels de la RA en conduite automobile

Les bénéfices potentiels de la RA se situent à deux niveaux. D'abord, elle peut permettre de réduire les déplacements du regard sur la route car les informations sont affichées directement dans l'environnement du conducteur. Ensuite, elle peut permettre d'orienter l'attention du conducteur, en mettant en valeur les éléments utiles à la conduite, par exemple en cas de mauvaise visibilité ou de danger. Une étude sur simulateur de conduite a montré que la RA permet d'améliorer le temps de détection d'éléments critiques sur le bord de la route, comme des piétons ou des panneaux signalant un danger (Rusch et al., 2013). Le même bénéfice a été observé pour des conducteurs âgés (Schall et al., 2013; Rush et al., 2014).

D'ailleurs, même dans le cas des véhicules autonomes, la RA constitue un moyen sérieusement envisagé pour supporter les transitions entre les modes automatiques et manuels. La RA serait utilisée pour aider le conducteur à se construire rapidement une représentation de la situation adéquate lors des reprises en main du véhicule (e.g. Projet Localisation et Réalité Augmentée conduit par l'Institut de Recherche Technologique).

1.1.4 Risques potentiels de la RA en conduite automobile

Toutefois, la RA n'est pas une technologie absoute de risques pour la conduite. A ces bénéfices potentiels, se greffent des risques potentiels. La RA implique la modalité visuelle du système cognitif; mais plus une tâche demande un contrôle visuel, plus elle entrave le contrôle visuel de la scène routière (Noy, Lemoine, Klachan, & Burns, 2003). Même si les informations en RA sont intégrées à l'environnement, elles peuvent constituer des distracteurs ou masquer certaines informations routières. Dans le domaine de l'aviation, les études sur le HUD ont montré qu'un surplus d'informations ajoutées dans l'environnement peut gêner la prise d'informations (Ververs & Wickens, 1998) et notamment gêner l'appréciation des distances (Wickens, 1997). Le risque lié à l'excès d'informations est particulièrement présent pour la conduite automobile. On peut aisément imaginer la multiplication d'informations sur le pare-brise : des informations relatives à la navigation (e.g. guidage GPS), à l'environnement (e.g. points d'intérêts), au véhicule (e.g. vitesse) ou aux systèmes embarqués (e.g. appels, radio).

Il existe aussi le risque d'attirer l'attention du conducteur vers un élément qu'il a déjà perçu et traité (Schall, Rusch, Gavin, Lee, Vecera & Rizzo, 2013). Egalement, des éléments

ajoutés peuvent masquer des informations importantes, ou encore détourner l'attention des conducteurs de ces informations importantes (Schall, Rusch, Lee, Vecera & Rizzo, 2010).

Enfin, la modalité de RA de surlignage, qui consiste à rendre saillant des éléments de l'environnement extérieur (e.g. panneaux, piétons, voie à emprunter) oriente l'attention visuelle, et de ce fait peut rentrer en conflit avec le processus « naturel » de prise d'information nécessaire à un moment de l'activité. Pour identifier et caractériser ce type de risque, il paraît nécessaire de prendre en considération les mécanismes d'orientation de l'attention visuelle en activité de conduite.

1.2 L'ATTENTION EN CONDUITE AUTOMOBILE

1.2.1 L'activité de conduite automobile

Selon Neboit (1980), «conduire un véhicule, c'est effectuer un déplacement dans un environnement en constante évolution. Ce déplacement est orienté vers des buts, soumis à différentes règles (explicites et implicites) et s'effectue au moyen d'un outil particulier : le véhicule ». Pour l'auteur, l'activité de conduite est une activité globale aux sous-tâches multiples qui demande une parfaite coordination entre des « sous-tâches cognitives » (e.g. compréhension de l'environnement, gestion des buts) et des « sous-tâches sensori-motrices » (e.g. tenue du volant, le changement de vitesse).

Dans la même veine, Michon (1985) décrit l'activité de conduite comme un ensemble de tâches simultanées qui ont des exigences temporelles et cognitives différentes. Elle se compose de 3 niveaux hiérarchiques (Figure 8). Le premier niveau de tâche est le niveau stratégique. Il correspond à la planification générale du trajet : l'itinéraire, la gestion du temps, des coûts et risques liés au trajet. Cette planification prend place dans une temporalité large (comptabilisée en minutes ou en heures). Ensuite le niveau tactique, qui correspond aux manœuvres. Une manœuvre est une action permettant de s'adapter à la situation immédiate. Les manœuvres peuvent découler du niveau stratégique (e.g. tourner ou changer de voie pour aller en direction de Paris) ou elles peuvent être imposées par le contexte (e.g. éviter un obstacle, dépasser un véhicule trop lent). Elles prennent place dans une temporalité courte, de l'ordre de la seconde. Le dernier niveau, dit opérationnel, correspond au contrôle immédiat du

véhicule, par exemple la gestion de la boîte de vitesse ou de la pédale de frein. Ici la temporalité est de l'ordre de la milliseconde.

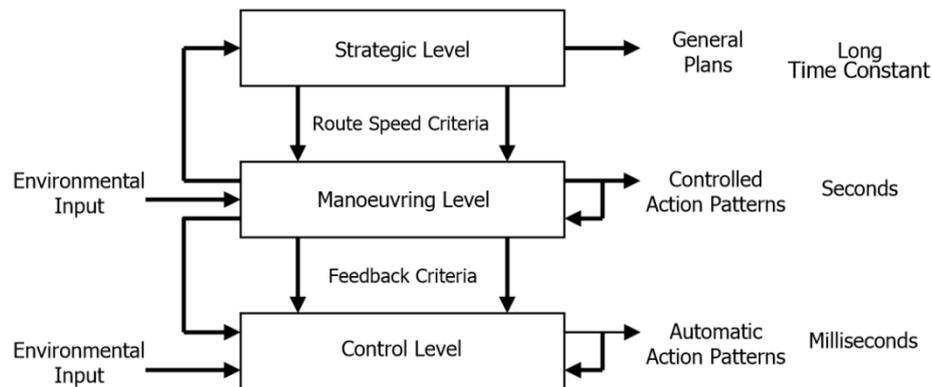


Figure 8. Les trois niveaux hiérarchiques de l'activité de conduite (Michon, 1985).

Le degré de complexité de l'activité de conduite dépend fortement de la situation. Lorsque la route est droite et à sens unique, la tâche du conducteur consistera principalement à maintenir l'allure, la position du véhicule sur la chaussée, et à maintenir une attention suffisamment soutenue pour être capable de réagir en cas d'évènement imprévisible. En revanche, la conduite en ville implique la gestion de nombreuses configurations routières et demande des prises de décision fréquentes, qui se réalisent sous une forte contrainte temporelle. La gestion d'une grande quantité d'informations est complexe et nécessite de fortes ressources cognitives et attentionnelles pour s'adapter correctement aux situations (Saad et al., 1992).

1.2.2 L'attention du conducteur

Classiquement, on définit l'attention comme un processus de traitement de l'information permettant à l'homme de s'adapter à son environnement (Broadbent, 1958). Si de nombreux modèles théoriques de l'attention ont été développés (e.g. Broadbent, 1958 ; Deutsh & Deutsh, 1963 ; Treisman & Gelade, 1980 ; Wickens, 1984 ; Posner, 1988), les auteurs s'accordent sur la fonction sélective de l'attention. Dès 1890, William James écrivait : « *Tout le monde sait ce qu'est l'attention. C'est quand l'esprit prend possession, sous une forme claire et active, d'un objet ou d'une pensée parmi plusieurs qui sont présent(e)s au même moment. La focalisation, la concentration et la conscience en sont l'essence. Elle implique le*

retrait de certaines choses pour mieux traiter d'autres choses. ». Broadbent (1958) propose une définition de l'attention qui reprend ce principe de sélection. Il la définit comme le processus permettant de faire le tri entre les nombreuses stimulations sensorielles auxquelles l'être humain est confronté, afin d'en limiter la quantité à traiter. Ainsi, l'attention filtre les informations utiles à son activité, de celles qui lui sont inutiles.

On conçoit aisément que l'attention soit un processus central dans une activité dynamique et complexe comme la conduite automobile. En effet, les informations de l'environnement peuvent être nombreuses (e.g. véhicules, piétons, lignes, signalétiques, panneaux publicitaires) et les situations se renouvellent à tous instants (e.g. feux tricolores, déplacement de son véhicule, déplacement des autres acteurs de l'environnement). Le conducteur devra alors constamment répartir ses ressources attentionnelles entre les différentes stimulations. Or, les capacités attentionnelles sont limitées, et pour ne pas s'épuiser elles doivent être distribuées de la manière la plus optimale possible (Amalberti, 2001).

1.2.3 Les défauts d'attention

Lorsqu'un déséquilibre advient dans la répartition des ressources entre les différentes sources d'informations de l'environnement, le conducteur subit une défaillance attentionnelle. C'est par exemple le cas s'il se focalise sur une information peu utile à sa tâche de conduite (e.g. écouter la radio) au détriment d'une information indispensable. Les défauts d'attention sont l'une des principales causes des accidents de la route (Hendricks, Fell, & Freedman, 2001 ; Lee & Strayer, 2004). 37,8% des accidents routiers seraient ainsi directement imputables à un défaut d'attention du conducteur ou à des erreurs de perception (Hendricks et al., 2001).

Lemercier et Cellier (2008) distinguent trois types de défaut d'attention en conduite automobile. Le premier est l'*inattention*, qui correspond à une réorientation interne de l'attention. Le focus attentionnel passe de l'environnement externe vers les pensées internes, par exemple le conducteur pense à son dîner et ne fait pas attention au feu rouge. Une étude de Lagarde & al. (2004) révèle par exemple une augmentation significative des risques d'accident chez des sujets ayant subi un stress profond comme une séparation ou un divorce. Les auteurs en déduisent que les pensées du conducteur sont happées par leurs préoccupations, et de ce fait l'attention portée à la conduite diminue. À l'inverse, la *distraction* est une réorientation externe de l'attention. Le focus attentionnel passe de la scène routière vers un élément intrus, externe lui aussi. Par exemple, la sonnerie téléphonique est une source de

distraction pour le conducteur, car elle peut réorienter l'attention du conducteur de la scène routière vers la source de l'émission du son (Posner, 1980). Enfin, l'*interférence* est un défaut de l'attention dû à une tâche secondaire. Elle peut consister en un défaut du contrôle visuel, par exemple lorsque le conducteur détourne le regard de la route alors qu'il rédige un SMS. Plus une tâche secondaire demande un contrôle visuel, plus elle entrave le contrôle visuel de la scène routière (Noy, Lemoine, Klachan, & Burns, 2003).

Si les défauts de l'attention peuvent être provoqués par un élément intrus (interne ou externe), ils peuvent également découler des conflits d'attention entre les différentes sous-tâches qui constituent l'activité de conduite elle-même. Alexander et Lunenfeld (1986) ont modélisé cette compétition d'attention entre les tâches de conduite (Figure 9). La compétition peut être verticale entre les tâches de différents niveaux hiérarchiques, par exemple un conducteur qui est perdu et qui concentre son attention sur les panneaux directionnels, pourra ne pas voir un véhicule qui s'insère devant lui (conflit entre des tâches de navigation et d'interaction avec le trafic). La compétition peut aussi être horizontale entre deux tâches de même niveau hiérarchique, par exemple un conducteur focalisé sur son rétroviseur peut ne pas voir le véhicule qui freine devant lui (conflit entre deux tâches d'interaction avec le trafic). Les ressources attentionnelles se distribuent donc entre ces différentes sous-tâches, parfois en défaveur de l'une d'elles.

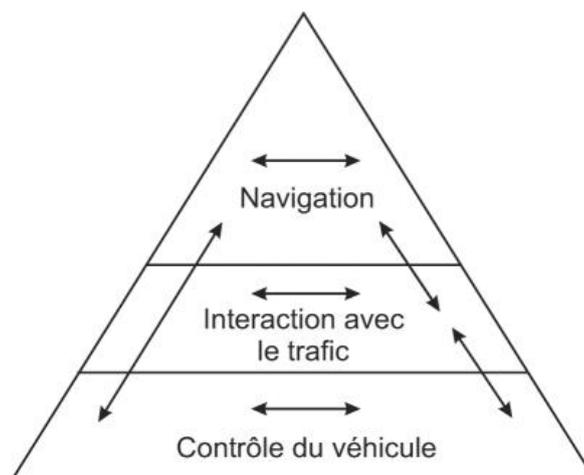


Figure 9. Compétitions d'attention horizontales et verticales entre les tâches de conduite (d'après Alexander et Lunenfeld, 1986).

Des défauts d'attention peuvent également être provoqués par l'utilisation d'outils numériques pendant la conduite. Par exemple, la consultation d'un écran oblige le conducteur

à détourner son regard de l'environnement routier, et l'utilisation des commandes tactiles implique encore plus de mobilisation visuelle, car en l'absence de retour haptique le conducteur doit « viser juste » sur l'écran.

La mise en évidence de ces risques par de nombreuses études (pour une revue de littérature exhaustive sur les défauts d'attention du conducteur, voir Young et Regan, 2007) et le durcissement de la législation sur l'utilisation d'outils numériques au volant, ont contribué à rendre central la problématique de la compatibilité de ces technologies à la conduite. Dans cette optique, les concepteurs mettent en œuvre des avancées, à la fois en termes d'ergonomie et en termes de technologie. Par exemple la synthèse vocale permet d'écouter un SMS, et la commande vocale permet d'en dicter un sans quitter la route des yeux. Mais si ces technologies permettent de soulager la mobilisation du regard, elles ne mettent pas à l'abri d'une réorientation interne de l'attention (vers les pensées). De plus, des technologies sensées réduire les défauts d'attention peuvent en créer d'autres. Par exemple le GPS permet d'éviter les détournements du regard provoqués par la lecture d'une carte, mais peut provoquer des conflits d'attention entre une tâche de navigation et une tâche d'interaction avec le trafic : un conducteur qui cherche sur son GPS la direction à emprunter peut ignorer un piéton qui traverse (Benedetto, Pedrotti, Minin, Baccino, Re et Montanari, 2011). De la même manière, la RA de surlignage, en mettant en valeur certains éléments de l'environnement, oriente l'attention vers l'une ou l'autre de ces sous-tâches, et risque ainsi de créer des conflits d'attention entre elles (navigation / interaction avec le trafic / contrôle du véhicule). Pour anticiper ces risques et être en mesure de concevoir des systèmes de RA qui ne rentrent pas en conflit avec l'orientation de l'attention du conducteur, il devient nécessaire de s'intéresser plus précisément à la manière dont l'attention visuelle est allouée en activité.

1.3 L'ORIENTATION DE L'ATTENTION VISUELLE

Classiquement, on explique l'orientation de l'attention visuelle à la fois par des facteurs exogènes- c'est-à-dire dirigée par le stimulus - et des facteurs endogènes - c'est-à-dire dirigée de façon volontaire.

1.3.1 L'orientation exogène de l'attention

L'attention visuelle exogène (ou *bottom-up*) est involontaire, et est dirigée par le stimulus. Les éléments saillants vont ainsi attirer irrésistiblement le regard. Le guidage exogène de l'attention a été largement étudié, et plusieurs facteurs de saillance visuelle ont été identifiés. Ainsi, le facteur principal est le degré de ressemblance / dissemblance entre un élément et les autres. Un élément qui se différencie des autres (un *singleton*) attire plus fortement l'attention. La différenciation peut concerner la dimension de la couleur (Theeuwes, 1994), de la forme (Theeuwes, 2005), du mouvement (McLeod, Driver, & Crisp, 1988), ou encore de l'orientation (Treisman, 1998). Également, il a été montré que l'apparition soudaine d'un élément dans la scène visuelle est aussi un facteur important de saillance visuelle (Jonides & Yantis, 1988 ; Peterson & Kramer, 2001 ; Theeuwes, 1994).

La plupart de ces études consiste en des tâches de recherche visuelle, dans lesquelles le participant doit détecter une cible, généralement un singleton (e.g. un carré parmi des cercles). Or, l'élément saillant est aussi un singleton (e.g. un élément rouge parmi des éléments verts) et de fait partage cette caractéristique avec la cible. Pour Yantis et Egeth (1999), il est donc possible que l'effet des éléments saillants sur l'attention serait dû à une stratégie des participants qui peut se résumer ainsi : « puisque la cible est un singleton, je cherche un singleton ». Dans ce cas, l'orientation de l'attention sur les éléments saillants serait guidée par les connaissances du participant, et ne serait pas exogène, mais endogène (interne).

L'étude de Most et al. (2005) montre assez clairement que des facteurs endogènes peuvent moduler l'influence des facteurs exogènes de l'attention. Les auteurs ont mis en place une expérimentation dans laquelle les participants devaient suivre du regard des points noirs se déplaçant aléatoirement sur un écran. Pendant la tâche, un élément singleton traversait l'écran de part en part. Il en ressort que les participants ont perçu cet élément uniquement lorsqu'il partageait une caractéristique de la cible (e.g. la couleur ou la forme). Lorsqu'il ne partageait aucune caractéristique commune, l'élément n'était pas détecté alors même qu'il se distinguait par sa couleur, son mouvement ou son apparition brusque. Pour les auteurs, la tâche a engendré une « préparation attentionnelle » aux propriétés de la cible, et la présence d'un distracteur, même extrêmement saillant du point de vue perceptif, est ignorée des traitements attentionnels dès lors qu'il se différencie de la cible. Ici, l'allocation de l'attention sur des items saillants résulte donc de l'utilisation de connaissances sur les caractéristiques de la cible, et semble donc dirigée par les demandes de la tâche. Ainsi, bien que l'attention puisse

être guidée de manière purement exogène par certains stimuli saillants, le déploiement de l'attention au sein d'une scène visuelle serait également dépendant des processus endogènes, eux-mêmes dépendants de la tâche à réaliser (Boloix, 2007).

1.3.2 L'orientation endogène de l'attention

L'attention visuelle endogène (ou *top-down*) est guidée de manière volontaire vers certaines informations. D'abord, le regard s'oriente préférentiellement vers certaines zones de la scène. Les scènes naturelles ont des structures relativement stables et régulières (Chun, 2000 ; Schyns & Oliva, 1994; Torralba & Oliva, 2003), et le système cognitif peut tirer parti de ces régularités pour déterminer quelles zones de la scène sont susceptibles de contenir des informations pertinentes qu'il serait utile de traiter, et, au contraire quelles zones de la scène ont peu de chance d'être informatives (Chun, 2000). Par exemple dans une scène routière, il y a plus de chance de trouver des informations pertinentes sur la route que dans la zone du ciel.

Egalement, le regard s'oriente rapidement et précisément vers certains objets de la scène. Dans le cas d'exploration libre de scènes visuelles (i.e. sans consigne), il a été montré très tôt que le regard se pose de façon systématique et récurrente sur certains éléments de la scène et en ignore d'autres (Yarbus, 1967). Ce sont les objets les plus sémantiquement informatifs – c'est-à-dire les plus importants pour la compréhension de la scène - qui recueillent le plus grand nombre de fixations oculaires. (Yarbus, 1967 ; Mackworth & Morandi, 1967; O'Regan et al., 2000).

Toutefois l'observation « passive » d'une scène est finalement une situation assez rarement rencontrée, et l'attention visuelle s'oriente encore différemment en activité. Car en plus d'être guidée par la sémantique de la scène, l'attention est aussi guidée de manière endogène par les buts de la tâche. Edmonds (1993) souligne que pour une tâche donnée, certaines propriétés ne sont saillantes que si elles sont relatives au but. Par exemple dans une tâche de description d'un itinéraire routier, la propriété de taille est saillante lorsque la tâche est la désignation d'un immeuble, mais ne l'est pas lorsqu'il s'agit de désigner une intersection de rues. Pour Landragin (2012), la saillance liée à l'intention recouvre deux aspects : (i) les fonctionnalités des objets et (ii) leur incongruence compte tenu de l'intention. Pour illustrer le facteur fonctionnel, l'auteur donne l'exemple d'un ordinateur allumé dans une pièce contenant d'autres ordinateurs éteints, qui sera particulièrement saillant si l'observateur est un utilisateur potentiel. Dans cet exemple, l'ordinateur allumé est plus saillant que les ordinateurs éteints car il est congruent avec l'intention de l'observateur (i.e. utiliser un

ordinateur). Pour illustrer le facteur de l'incongruence, l'auteur explique qu'une chaise renversée sera plus saillante qu'une chaise à l'endroit, si l'on a besoin de s'asseoir. Ici, l'objet est rendu saillant par son incongruence avec le but. Ainsi, un objet attire plus l'attention s'il est congruent avec l'intention (i.e. utile à la tâche) ou s'il est incongruent avec l'intention (i.e. entrave la tâche). En revanche, l'auteur ne précise pas comment s'organisent les deux aspects de la saillance liée à l'intention. L'organisation pourrait être hiérarchique : si le but principal *utiliser l'ordinateur* est prioritaire au sous-but *s'asseoir* alors l'ordinateur serait plus saillant que la chaise. Elle pourrait aussi être temporelle : s'il faut s'asseoir avant d'utiliser l'ordinateur alors la chaise serait plus saillante que l'ordinateur.

Deux processus d'orientation de l'attention – endogène et exogène – sont donc à l'œuvre en situation d'activité. Pour cette raison certains modèles d'attention visuelle intègrent ces facteurs conjointement pour expliquer l'orientation du regard.

1.3.3 Intégration des facteurs endogènes et exogènes : le modèle de Wickens et al. (2001)

La plupart des travaux de modélisation des processus d'orientation du regard s'inscrivent généralement dans une approche *bottom-up* : ils considèrent l'orientation du regard comme un processus résultant d'effets de saillance perceptive dans la scène visuelle, pour lequel le regard est guidé par le signal visuel (Koch & Ullman, 1985 ; Itti & Koch, 2001). Les effets de saillance résultent par exemple des couleurs ou des contrastes (Parkhurst & Niebur, 2003). Une grande partie de ces recherches porte sur des tâches de visualisation libre (les participants visualisent des images sans consignes particulières), les buts des participants ne sont donc pas définis et les effets potentiels de la tâche sur l'orientation du regard ne sont pas pris en considération. Selon Rothkopf et al. (2007) l'orientation du regard a été presque exclusivement étudiée indépendamment de tâches spécifiques, alors que dans la vie de tous les jours la vision humaine est dévouée à la réalisation d'une succession de tâches.

Certains modèles ont tout de même intégré des facteurs *top-down*. C'est par exemple le cas du modèle de Navalpakkam & Itti (2005) qui intègre la connaissance des caractéristiques de la cible ou celui de Michael et al. (2006) pour lequel les effets de saillance visuelle sont augmentés ou inhibés selon la pertinence de la zone. Dans ces modèles, des facteurs liés à la sémantique de l'image et à la tâche modulent la carte de saillance, mais restent des facteurs de second plan (Rothkopf et al., 2007).

Wickens et al. (2001) ont élaboré le modèle théorique SEEV (*Saliency, Effort, Expectancy, Value*) qui se distingue des autres dans la mesure où il explique que l'allocation de l'attention visuelle en activité doit prendre en compte de manière équivalente les facteurs *top-down* et *bottom-up*.

Dans le modèle SEEV, le regard est guidé par 4 facteurs : la Saillance, l'Effort, l'Attente et la Valeur (Figure 10).

La *Saillance* constitue le principal facteur *bottom-up* de l'attention visuelle : le regard est guidé par des informations purement visuelles. C'est d'ailleurs le facteur central dans la plupart des modèles de recherche visuelle (Tatler et al., 2011). Plusieurs caractéristiques de saillance ont été identifiées, comme la couleur, la forme ou le mouvement (Itti & Koch, 2000).

L'*Effort* constitue le second facteur qui est traité de manière *bottom-up* par le modèle. Il correspond à l'angle visuel entre les informations, c'est-à-dire à la distance que parcourt l'œil pour atteindre une zone de la scène visuelle. Cette distance, si elle est trop importante, peut inhiber la prise d'information (Kvalseth, 1997 ; Sheridan, 1970). Par exemple, un conducteur négligera de regarder son angle mort car il nécessite un large déplacement du regard et de la tête (Wickens 2007).

L'*Attente* est un facteur *top-down* et correspond à la probabilité de voir un événement dans une zone. Plus précisément, elle est déterminée par la fréquence d'occurrence d'événements dans la zone. Par exemple la probabilité d'événement est plus forte dans une zone d'intersection : un véhicule peut tourner, il peut y avoir un ralentissement ou un feu tricolore (Werneke & Volrath, 2012).

Enfin la *Valeur*, également un facteur *top-down*, correspond à l'importance d'un élément visuel pour la réalisation d'une tâche. Ce facteur prend également en compte l'importance de la tâche à laquelle est lié l'élément visuel. Ce facteur était déjà présent dans le modèle de Carbonell (1966), où la valeur d'une zone dépend du bénéfice si elle est consultée, et du coût si elle est ignorée. Pour reprendre l'exemple de Horrey et al. (2006), consulter les lignes au sol apporte un fort bénéfice lors d'une tâche de changement de voie, et ignorer les informations du tableau de bord ne représente pas un coût important pour l'exécution du changement de voie.

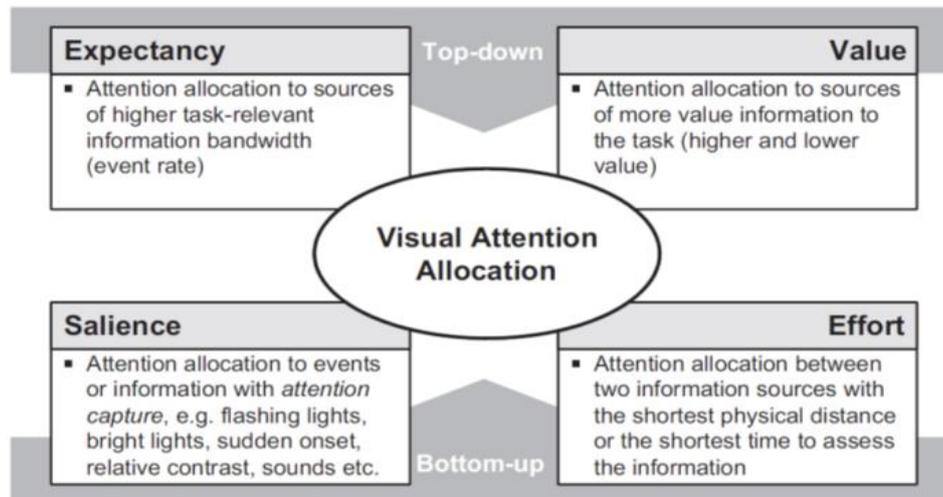


Figure 10. Le modèle SEEV (Wickens et al., 2003, schéma d'après Werneke & Vollrath, 2012).

Wickens et al. (2003) ont adapté le modèle SEEV en un outil prédictif, qui a été expérimenté avec succès sur l'activité de pilotage d'avion (Wickens et al., 2003), puis sur la conduite automobile (Horrey et al., 2006). La principale application de ce modèle est l'optimisation du guidage attentionnel : chacun des 4 facteurs peut être mis à profit pour aider le conducteur (ou l'opérateur en général) à atteindre les informations pertinentes avec un moindre effort.

1.3.4 L'attention visuelle optimale

Il n'est pas possible de traiter toutes les informations visuelles d'une scène (Ullman, 1984). Sélectionner les informations pertinentes parmi toutes les informations présentes est indispensable pour être en mesure de traiter une scène (Brooks, 1986; Bajcsy, 1988; Ballard, 1991). Ainsi lors d'une activité orientée vers la réalisation d'un but, la stratégie de prise d'information doit être spécifique au but pour être optimale (Shinoda et al., 2001). Ce type de stratégie est non seulement nécessaire pour la mise en œuvre de la tâche et l'atteinte du but, mais elle permet également de simplifier la représentation de la situation et d'en faciliter le traitement. Notons que l'attention visuelle doit également rester disponible pour capter des événements imprévisibles qui sont susceptibles de changer le cours de la tâche (Ullman, 1984 ; Shinoda et al., 2001). Par exemple, le conducteur qui a l'intention de doubler un véhicule concentre son attention visuelle sur le rétroviseur et l'angle mort (informations nécessaires à la réalisation du son but) mais doit rester disponible pour détecter un éventuel piéton qui traverse (événement imprévisible).

D'après le modèle SEEV l'attention visuelle est guidée par 4 facteurs, mais pour être optimale - c'est-à-dire spécifique au but - elle doit principalement être guidée par les facteurs *top-down* (Attente et Valeur), plutôt que par les facteurs *bottom-up* (Saillance et Effort) (Wickens, 2003). En effet, un élément saillant risque de constituer une distraction visuelle s'il n'est pas utile pour atteindre un but. Egalement, un élément perceptif utile mais nécessitant un effort trop important pour être consulté, aura plus de chance d'être négligé (par exemple l'angle mort).

L'optimisation de l'allocation de l'attention visuelle est par conséquent fortement dépendante de la manière avec laquelle le dispositif est conçu. Ainsi, les facteurs *bottom-up* peuvent être utilisés pour optimiser l'attention visuelle. Il est possible de concevoir des interfaces dans lesquelles les informations pertinentes pour la tâche de conduite sont mises en valeur, par exemple avec un indicateur de panne qui clignote (augmentation de la Saillance). Il est aussi possible de situer les informations importantes dans le champ de vision du conducteur, par exemple avec un détecteur d'angle mort au niveau du rétroviseur (diminution de l'Effort).

Toutefois, contrairement à l'environnement visuel d'une interface (i.e. cockpit de la voiture), l'environnement routier est non contrôlable, dynamique et relativement aléatoire. Ainsi, pour favoriser chez le conducteur une allocation de l'attention optimale dans l'environnement routier, seul le facteur *Saillance* peut être utilisé. Dans cette optique, la RA est envisagée pour diriger l'attention visuelle du conducteur vers les informations de l'environnement ayant une forte valeur informative.

1.3.5 La Valeur d'une information visuelle

Pour Wickens, la Valeur d'une information correspond à son utilité, que Carbonnel et al. (1966) expriment par le bénéfice à consulter, et le coût à ignorer. Dans la formule de prédiction du parcours oculaire du modèle SEEV (Wickens et al., 2003), la Valeur d'un indice est exprimée par le produit de la pertinence de l'indice pour une tâche, et de la priorité de la tâche associée :

$$\text{Valeur d'un indice} = \text{pertinence de l'indice pour une tâche} * \text{priorité de la tâche associée.}$$

Horrey et al. (2006) ont utilisé cette formule afin de modéliser l'attention visuelle du conducteur utilisant une technologie embarquée. Dans leur étude, les conducteurs devaient conduire sur une ligne droite tout en lisant et en numérotant vocalement des numéros de

téléphone qui s'affichent sur un écran dans le cockpit du simulateur. Les participants avaient pour consigne soit de privilégier la tâche de conduite, soit de privilégier la tâche de numérotation, soit de privilégier les deux tâches. Il y avait 2 AOI⁷s : l'environnement extérieur et l'écran affichant les numéros de téléphone. Les auteurs ont ensuite entrepris de modéliser l'attention visuelle des conducteurs afin de confronter leurs prédictions aux résultats de l'expérimentation. Pour calculer le coefficient de Valeur de chacun des AOIs, les auteurs ont considéré que :

- La tâche de conduite est toujours plus importante que la tâche de numérotation. Ils ont ainsi attribué 2 points à l'environnement extérieur, et 1 point à l'écran.
- L'environnement extérieur est plus pertinent que l'écran pour une tâche de conduite. A l'inverse, l'écran est plus pertinent que l'environnement pour la tâche de numérotation. Ils ont ainsi ajouté 1 point lorsque l'AOI était pertinente pour la tâche et retiré un point à l'autre AOI.

Les coefficients des AOIs sont reportés dans le Tableau 1

AOIs	Tâche prioritaire		
	Conduite	Conduite + numérotation	Numérotation
Environnement extérieur	3	2	1
Ecran	0	1	2

Tableau 1. Coefficients des valeurs attribués aux AOIs dans l'étude de Horrey et al. (2006). Notons que pour leur calcul, les auteurs ont rajouté un point à chaque note afin d'éviter les valeurs nulles.

A l'issue de ce calcul, l'environnement extérieur a une valeur supérieure à l'écran dans les conditions de priorité à la conduite (3 vs. 0) et de priorité équivalente (2 vs. 1). Et l'écran a une valeur supérieure à l'environnement, uniquement dans la condition de priorité à la tâche de numérotation (2 vs. 1). Les auteurs ont montré que les prédictions élaborées sur la base de ce calcul sont fortement corrélées à l'allocation de l'attention visuelle observée lors des expérimentations.

Cette manière d'appréhender la Valeur est basée sur une distinction entre plusieurs tâches (la conduite et la numérotation) auxquelles correspondent un environnement visuel (respectivement la route et l'écran). Dans la perspective où l'on s'intéresse uniquement à

⁷ Area Of Interest

l'activité de conduite (et donc qu'à l'environnement extérieur), comment différencier les Valeurs des différents éléments visuels ? Par exemple, comment juger la Valeur d'un panneau par rapport à un autre ? Les études s'intéressant au parcours oculaire comme indicateur de l'attention visuelle pendant la réalisation d'une tâche peuvent fournir des pistes de réponse.

1.3.6 Spécificité de l'attention visuelle au but de la tâche

Les travaux de Yarbus (1967) ont montré très tôt la spécificité du parcours oculaire selon le but de la tâche. Par exemple dans une tâche d'observation d'une peinture, les participants orientaient leur regard différemment selon le but qui leur était assigné. Le recueil des parcours oculaires montre que les participants focalisent davantage leur attention sur les visages lorsqu'ils doivent estimer l'âge des protagonistes, et à l'inverse ils focalisent davantage leur attention vers les objets lorsqu'ils doivent estimer le confort de la maison. D'autres études sur les parcours oculaires lors de la réalisation de tâches quotidiennes ont montré que les fixations se concentrent sur des informations très liées au but en cours. Lors d'une tâche de préparation de thé par exemple, Land, Mennie, et Rusted (1998) remarquent que les objets de la scène visuelle non nécessaires à la réalisation du but sont très peu fixés. Reprenant ce constat, Hayhoe et al. (2003) a étudié le parcours oculaire de participants pour une tâche de préparation de sandwich. Il en ressort que les fixations sont très spécifiques à chaque étape de la procédure pour la réalisation du but, et que seules 2% des fixations concernent des objets de la scène non pertinents (ingrédients ou outils non utilisés).

Dans le domaine automobile, Wallis et Bühlhoff (2000) ont montré des différences d'attention visuelle entre des participants selon les tâches qui leur étaient assignées. Les participants étaient placés dans un simulateur de conduite, soit en tant que conducteur, soit en tant que passager. Dans les scènes visuelles projetées, des blocs étaient disposés sur la route, près de la route ou loin de la route. Les participants devaient détecter les changements concernant ces blocs (orientation, position, absence/présence, couleur). Il en ressort que les participants conducteurs détectent mieux les changements de blocs situés sur la route que les changements des autres blocs. Pour les participants passagers, la position des blocs n'a pas d'influence sur la détection des changements. Les auteurs en déduisent que les participants conducteurs ont focalisé leur attention sur la route, et ont donc traité prioritairement les éléments pertinents pour la finalité de la tâche (i.e. la conduite). Dans cette expérimentation l'environnement visuel était le même pour tous, et on voit donc l'influence du but dans l'orientation de l'attention lors de la réalisation de tâches.

Shinoda et al. (2001) ont utilisé un procédé expérimental similaire pour étudier l'influence de la tâche sur l'attention visuelle. Dans leur étude, des participants placés dans un simulateur de conduite avaient pour consigne de suivre le véhicule précédent. La moitié des participants avaient une tâche supplémentaire : respecter les règles de conduite. Dans l'environnement, des panneaux *stationnement interdit* placés dans des intersections se changeaient en panneaux *stop* pendant de courtes périodes. Il en ressort que les participants ayant la double tâche ont eu de meilleurs taux de détection des changements. Les analyses du regard ont montré que ces participants faisaient beaucoup plus de fixations sur les zones d'intersections, où se trouvaient les panneaux. Ainsi, pour une même scène visuelle, les exigences de la tâche conduisent à des différences dans l'allocation de l'attention.

D'autres auteurs ont utilisé le paradigme de la détection de changement. Il consiste à introduire un changement dans une scène visuelle entrecoupée par un masque, et d'observer si les participants détectent ce changement. Rensink (1997) a utilisé ce paradigme sur des images de scènes de tous les jours. L'auteur avait préalablement défini dans les images des « zones d'intérêt ». Ces zones ont été déterminées lors d'une expérimentation indépendante, dans laquelle des participants devaient décrire oralement les scènes visuelles. Les « zones d'intérêt central » étaient celles mentionnées par la majorité des participants, et les « zones d'intérêt marginal » étaient celles qui n'étaient mentionnées par aucun participant. Il résulte de cette expérimentation que les participants détectent plus facilement des changements dans les « zones d'intérêt central ».

Il est intéressant de noter que pour Triesch et al. (2003), la notion « d'intérêt » est très floue et ne suffit pas à décrire la distribution des ressources attentionnelles. Ces derniers font l'hypothèse qu'une variable essentielle de la détection du changement est l'importance d'un élément de la scène visuelle pour la tâche, et à un moment donné de la tâche. Dans leur expérimentation, il était demandé aux participants de déplacer des briques dans un environnement virtuel, et l'élément qui changeait était la taille de certaines briques. La tâche nécessitait de prendre deux décisions : (i) quelle brique sélectionner et (ii) où la déplacer. Dans la première condition les participants doivent déplacer les briques en commençant par les petites et en finissant par les grandes (la taille des briques est ici une information utile pour la première décision : sélectionner la brique à déplacer). Dans la seconde condition les participants doivent également commencer par les petites briques, et ils doivent en plus les déplacer dans des zones différentes selon leurs tailles (ici la taille des briques est une information utile pour les deux décisions : sélectionner et déplacer). Dans la

troisième condition ils n'ont pas de contrainte particulière (la taille des briques n'est donc pas une information utile). Il en ressort que les participants ont détecté plus souvent le changement lorsqu'il était utile pour la décision en cours. Leur explication est que seule une quantité très limitée d'informations visuelles est intégrée, « juste à temps », quand elles sont nécessaires pour atteindre le but en cours. A l'inverse, les autres informations connexes à la tâche mais secondaires pour atteindre le but immédiat seraient temporairement négligées. Les enregistrements de l'activité oculaire des participants montrent que les fixations étaient plus longues sur les éléments pour lesquels un changement a été détecté. Leurs résultats ne permettent toutefois pas de déterminer clairement si cette augmentation du temps de fixation est consécutive au changement des éléments ou à la pertinence des éléments pour la tâche (pour lesquels les changements ont été plus souvent détectés).

Dans le prolongement de Triesch et al. (2003), Boloix (2006) a couplé une tâche de détection de changements avec une tâche de résolution de problème. L'objectif était d'observer les différences de détection de changements, selon la pertinence de l'élément par rapport à la tâche. Dans son expérience, les participants devaient déplacer un véhicule dans le plan d'une ville (sur un écran). Pour réaliser le trajet, les participants doivent passer par plusieurs étapes et sous-étapes. Au cours de la réalisation du trajet, des étapes sont modifiées (déplacement ou disparition). Soit l'étape modifiée ne fait pas partie du trajet (donc de la tâche), soit elle en fait partie. Les étapes faisant partie du trajet peuvent être de quatre types : l'étape immédiatement à venir (sous-but immédiat), une étape précédente (sous-but antérieurement atteint), une étape prochaine (sous-but à atteindre ultérieurement) ou la destination (but final). Les résultats montrent que les participants détectent plus fréquemment les changements sur les étapes qui font partie de leur trajet que sur les étapes qui ne sont pas sur leur trajet. De plus, ils détectent plus fréquemment les changements concernant les étapes immédiatement à venir (sous-but immédiat) que les changements concernant les autres étapes (antérieures, ultérieures, destination). L'auteur en déduit que les informations visuelles sont d'autant mieux représentées en mémoire qu'elles sont immédiatement impliquées dans le sous-but à atteindre.

Les travaux de Boloix (2006) dépassent la dichotomie *pertinence /non pertinence* en distinguant deux niveaux de représentations : *sémantique /fonctionnel*. D'après l'auteur, les différents traitements visuels et les connaissances stockées en mémoire à long terme contribuent à un niveau de représentation visuelle « *sémantique* » de la scène, c'est-à-dire qui consiste à comprendre la scène. En revanche, lorsqu'un individu doit réaliser une tâche sur

cette scène (e.g. une tâche de recherche visuelle), les buts de la tâche, stockés en mémoire de travail, guident l'attention visuelle sur les éléments de la scène nécessaires pour l'exécution de la tâche en cours et la réalisation du but. C'est ce qui constitue le niveau de représentation visuelle « fonctionnel ». Ce niveau fonctionnel domine le niveau sémantique le temps de la tâche, sans pour autant l'inhiber. C'est-à-dire que l'attention se concentre en priorité sur les éléments nécessaires à la réalisation du but - au moment de sa réalisation - mais n'ignore pas les autres éléments nécessaires à la compréhension générale de la scène.

Ces derniers résultats de recherche soulignent que l'attention s'oriente prioritairement sur les éléments liés au but (ou sous-but) en cours de réalisation. Nous proposons dans la Figure 11 une représentation schématique de la priorisation du traitement des informations visuelles selon les travaux de Triesch et al. (2003) et Boloix (2006).

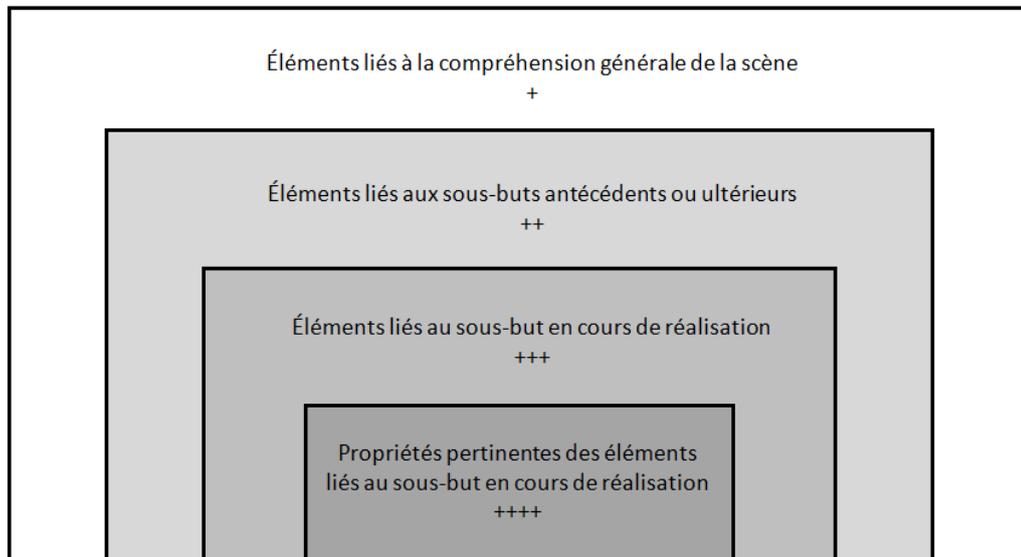


Figure 11. Priorisation du traitement des informations visuelles (selon les travaux de Triesch et al., 2003 et Boloix, 2006).

Si l'attention visuelle se focalise sur les éléments pertinents pour la réalisation d'un but, les recherches sur la cécité d'inattention (*inattentional blindness*) montrent également que les éléments de la scène non liés à la tâche tendent à être ignorés, et ce malgré leur saillance visuelle (Mack & Rock, 1998; Simons & Chabris, 1999). Ainsi, dans une expérience où les participants devaient déterminer lequel de deux segments constitutifs d'une croix est le plus grand, ceux-ci n'ont pas détecté l'apparition d'un stimulus non pertinent pour le but dans leur champ visuel (un carré), même si celui-ci apparaît en position fovéale (Mack & Rock, 1998). L'étude de Simons & Chabris (1999) a donné à ce phénomène une forte résonance

médiatique. Dans leur expérimentation, des participants visualisaient une vidéo dans laquelle deux équipes se livraient à des passes de basketball ; une équipe portait des maillots noirs et l'autre des maillots blancs. Les participants devaient réaliser une tâche de comptage : ils avaient pour but de compter le nombre de passes de l'équipe en blanc. Durant les échanges, un personnage déguisé en gorille (donc en noir) traversait la scène. Lorsqu'on demande aux participants s'ils avaient remarqué quelque-chose, la moitié d'entre eux n'ont pas rapporté l'évènement. Les auteurs expliquent cet échec par la focalisation de l'attention sur les personnages habillés en blanc. Ici l'attention visuelle des participants a été entièrement mobilisée par l'exécution de la tâche de comptage et a ignoré les éléments de la scène non liés à la réalisation du but.

2 EXPERIMENTATIONS

2.1 HYPOTHESES ET EXPERIMENTATIONS

2.1.1 Hypothèses

Nous avons vu que l'attention est le processus qui consiste à filtrer les informations utiles à l'activité, de celles qui sont inutiles (Broadbent, 1958). Un système de *RA de surlignage* guide ce processus de sélection de l'information, en mettant en valeur certains éléments par rapport à d'autres. Les cas d'usages envisagés dans le mode de RA de surlignage sont au nombre de trois : (i) la mise en saillance des lignes et panneaux en cas de mauvaise visibilité (e.g. nuit, brouillard, pluie), (ii) la mise en saillance de véhicules ou piétons présentant un risque de collision et (iii) la mise en saillance de la direction à emprunter. Dans les deux premiers cas d'usage, c'est le contexte qui définit la pertinence d'un élément, indépendamment de l'activité de conduite elle-même. Dans le troisième cas, c'est une mise en saillance d'éléments liés au niveau stratégique de l'activité de conduite (i.e. guidage). Dans ces trois cas d'usage, la RA met en évidence des éléments liés à l'environnement et au niveau stratégique de l'activité de conduite, mais est décorrélée de la manœuvre.

Or, nous avons vu que l'attention visuelle semble à la fois se concentrer sur les éléments les plus pertinents par rapport au but de la tâche (e.g. Rensink, 1997 ; Triesch et al., 2003 ; Boloix, 2006) et ignorer les informations les moins pertinentes (e.g. Mack & Rock, 1998; Simons & Chabris, 1999) ; ce qui se traduit par des parcours oculaires très spécifiques aux buts (et sous-but) immédiats (e.g. Land, Mennie et Rusted, 1998 ; Hayohe, 2003). Dans une activité de conduite automobile, le « but immédiat » s'apparente à la « manœuvre ». Rappelons que la manœuvre est une action permettant de s'adapter à la situation immédiate, par exemple dépasser un véhicule, changer de voie, tourner à un croisement. Ainsi, dans le cadre de notre recherche, la première hypothèse (**H1**) que nous posons est que l'allocation de l'attention visuelle est très spécifique aux manœuvres, c'est-à-dire qu'elle se dirige prioritairement sur les éléments les plus pertinents par rapport au but immédiat de la tâche.

De plus, l'activité de conduite se caractérise par une succession de prises de décision. Le conducteur doit identifier dans son environnement visuel les informations lui permettant de valider ou d'invalider ses buts et sous-but. Nous avons vu que l'incongruence d'un élément avec un but rend cet élément plus saillant (e.g. la chaise renversée si le but est de s'asseoir) (Conklin & McDonald, 1982 ; Landragin, 2012). Nous posons donc l'hypothèse (**H2**) que les éléments incongruents avec les manœuvres -c'est-à-dire les éléments qui invalident le but- focalisent encore davantage l'attention visuelle.

Dans un second temps, nous nous interrogerons sur l'impact que pourrait avoir une RA de surlignage sur l'allocation de l'attention visuelle. Nous posons l'hypothèse (**H3**) qu'une RA qui ne respecte pas la priorisation « naturelle » du traitement de l'information perturbera la prise d'informations. Plus précisément, l'hypothèse postule que l'allocation de l'attention visuelle sera dégradée par une RA de surlignage non spécifique à la manœuvre. Une information importante qui n'est pas mise en valeur par le système risquerait d'être ignorée ; et une information non pertinente mais qui est mise en valeur par le système pourrait focaliser l'attention au détriment des informations pertinentes.

2.1.2 Choix Méthodologiques

Comme nous l'avons vu, les technologies de RA pour l'automobile n'étant pas encore mûres, elles ne sont pas encore implantées dans des véhicules. Notre recherche se place dans une démarche prospective. Afin d'étudier en détail l'allocation de l'attention visuelle dans un environnement routier, nous avons choisi de travailler sur des images statiques et dynamiques. En effet, les images et les vidéos donnent un bon aperçu de la richesse et la complexité visuelle de la réalité. C'est dans cet environnement riche et complexe que la RA sera amenée à s'intégrer. Cependant, l'absence d'activité de conduite risque de mettre les participants en situation de visualisation passive. Pour pallier cette limite, nous avons décidé de mettre les participants en condition de prise de décision, afin de les engager dans la tâche.

D'autre part, dans le cadre de nos expérimentations nous avons utilisé l'oculométrie, qui est une technique d'enregistrement des mouvements des yeux. Elle consiste à repérer la position du regard au moyen d'une caméra vidéo. Les appareils les plus courants utilisent la technique du reflet cornéen : ils captent le reflet émis par un rayon infrarouge envoyé sur la cornée de l'œil pour repérer le centre de la pupille et connaître la position du regard. Cette technique permet de détecter les fixations (pendant lesquelles l'œil reste relativement immobile) et les saccades (sauts d'une fixation à l'autre). A partir des fixations et des

saccades peuvent être calculées des mesures spatiales (e.g. les localisations des fixations, les distances saccadiques) et des mesures temporelles (e.g. les durées des fixations). Lors de la visualisation d'une scène, ces mesures peuvent refléter des processus de traitement ou de recherche de l'information (Goldberg & Kotval, 1998). Par exemple les durées de fixations peuvent rendre compte de l'intérêt pour un élément ou de la difficulté à le traiter ; les tailles des saccades peuvent renseigner sur l'efficacité d'une recherche d'information. Dans le cadre de cette recherche, nous nous concentrons sur **l'allocation de l'attention visuelle**. Nous avons vu que les ressources attentionnelles sont limitées : le regard s'oriente préférentiellement vers certaines informations au dépens d'autres informations. Pour rendre compte de cette allocation, nous devons donc être en mesure de déterminer (i) quelles informations focalisent l'attention, et (ii) au dépens de quelles autres informations. L'oculométrie, qui permet de rendre compte de l'orientation de l'attention visuelle, constitue donc un outil privilégié pour tenter de répondre à cette problématique.

2.1.3 Expérimentations

Pour tester nos hypothèses, nous avons mis en place 3 expérimentations dans lesquelles des participants visualisaient des scènes routières et devaient prendre des décisions.

La première expérimentation vise à déterminer dans quelle mesure l'attention visuelle est spécifique aux buts immédiats lors d'une prise de décision dans une situation de conduite. Les participants devaient prendre connaissance d'une manœuvre, visualiser pendant quelques secondes une image de situation de conduite automobile, puis décider si la situation permettait la réalisation de la manœuvre. Pour chaque situation, la manœuvre et les indices visuels liés à cette manœuvre variaient selon les groupes expérimentaux. Nous avons également fait varier la congruence des indices avec les buts, c'est-à-dire le fait que les indices permettent la réalisation du but ou non.

La seconde expérimentation réplique la précédente, cette fois-ci en prenant en compte l'aspect dynamique de l'activité de conduite. Les participants visualisaient des scènes vidéo de conduite automobile, Au cours des vidéos, une voix indiquait une manœuvre et le participant devait décider si la situation lui permettait de la réaliser ou non.

La troisième expérimentation vise à déterminer l'impact de la Réalité Augmentée sur l'allocation de l'attention visuelle en conduite. Les participants visualisaient des scènes vidéos dans lesquelles certains indices de la scène visuelle étaient augmentés graphiquement : soit

des indices généraux liés à la conduite, soit des indices liés à une manœuvre en particulier, soit tous les indices, soit aucun indice. Egalement, nous nous sommes intéressés à deux phases de conduite : la phase « de navigation » où le conducteur roule tout droit et n'a pas de but particulier; et la phase « de manœuvre » où le conducteur a l'intention de faire une manœuvre, et regarde si la situation le lui permet.

2.2 EXPERIMENTATION I : ALLOCATION DE L'ATTENTION VISUELLE SUR DES SCENES STATIQUES

2.2.1 Hypothèses et prédictions

La première expérimentation vise à déterminer dans quelle mesure l'attention visuelle est spécifique aux buts immédiats dans une situation de conduite automobile. Nous avons mis en place une procédure expérimentale dans laquelle des participants prennent connaissance d'une manœuvre, visualisent une image de situation de conduite automobile, puis décident si la situation permet la réalisation de la manœuvre.

Nous avons vu que l'attention visuelle semble à la fois se concentrer sur les éléments les plus pertinents par rapport à la tâche (e.g. Rensink, 1997 ; Triesch et al., 2003 ; Boloix, 2006) et ignorer les informations les moins pertinentes (e.g. Mack & Rock, 1998; Simons & Chabris, 1999) ; ce qui se traduit par des parcours oculaires très spécifiques au but immédiat (e.g. Land, Mennie et Rusted, 1998 ; Hayohe, 2003). Dans une activité de conduite automobile, le « but immédiat » s'apparente à la « manœuvre ». Nous avons également vu qu'un élément incongruent par rapport à un but gagne en saillance visuelle (Conklin & McDonald, 1982 ; Landragin, 2012). Ainsi, nous posons les deux hypothèses suivantes :

[H1] L'allocation de l'attention visuelle sur des scènes de conduite automobile est spécifique aux manœuvres envisagées. Deux prédictions [P1 et P2] découlent de cette hypothèse. En premier lieu [P1], on s'attend à ce que l'ensemble des indices (i.e. les informations sémantiques) focalisent davantage l'attention visuelle que les zones de saillance visuelle. En second lieu [P2], on s'attend à ce que l'attention visuelle soit davantage allouée aux indices liés aux manœuvres (permettant la prise de décision) qu'aux indices généraux (permettant la compréhension de la scène).

[H2] Un indice qui interdit une manœuvre (selon le code de la route) ou empêche une manœuvre (selon la situation), focalise davantage l'attention visuelle qu'un indice qui autorise ou permet une manœuvre.

2.2.2 Méthode

2.2.2.1 Participants

Quarante participants volontaires ont été recrutés (21 femmes, âge moyen = 32, min=23, max = 51, SD = 7,5). Tous avaient le permis de conduire et une expérience de conduite depuis plus de 3 ans (moy = 12, max =26, SD = 6). Aucun des participants n'avait de problème de vue et ne portait de lunettes ou lentilles. Les participants ont donné leur autorisation pour l'enregistrement de leurs mouvements oculaires et de leurs réponses, et ils ont été dédommagés avec une carte cadeau. Les 40 participants ont été repartis aléatoirement dans 2 groupes.

2.2.2.2 Matériel

Soixante images de situations de conduite automobile ont été utilisées. Les prises de vue ont été réalisées en condition de circulation urbaine, à l'aide d'une caméra embarquée à grand angle fixée sur le pare-brise d'une voiture.

Avant chaque image, un écran notifie une manœuvre au participant. Les manœuvres peuvent être de six types (Figure 12) : (i) tourner à gauche/droite (ii) se déporter sur la voie de gauche/droite, (iii) dépasser, (iv) se garer, (v) s'arrêter et (vi) continuer tout droit.

Chacune des manœuvres peut être congruente ou non avec la situation de la scène visuelle projetée. La congruence d'une manœuvre découle des règles de circulation, par exemple la manœuvre *tourner à gauche* est non congruente s'il y a un panneau *sens interdit* à gauche ; en revanche elle est congruente si aucun indice ne l'empêche. Sur les 60 situations présentées à chaque participant, la moitié (30 images) est précédée par l'affichage d'une notification de manœuvre congruente et l'autre moitié est précédée d'une notification de manœuvre non congruente.

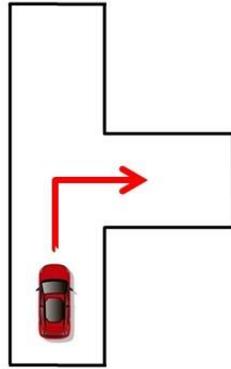
Lors de l'élaboration du matériel, deux conducteurs experts ont noté la difficulté qu'ils percevaient pour chaque situation (« la prise de décision est : simple – moyenne – difficile »). Toutes les situations retenues pour l'expérimentation (congruentes ou non congruentes)

étaient de difficulté équivalente: elles correspondaient à des situations de conduite normale (sans pièges ni ambiguïté) dans lesquelles la prise de décision était simple.

Sur chaque image, trois types de zones d'intérêts (AOIs) ont été définies:

- Les *indices liés à la manœuvre* correspondent aux indices permettant de décider la réalisation de la manœuvre donnée (e.g. la ligne continue qui permet de décider de ne pas changer de voie) (Figure 15). Les indices peuvent être des éléments de signalisation (panneaux, lignes au sol, feux tricolores) ou des événements (véhicules, piétons) (Figure 13). Deux participants experts (10 et 13 ans d'expérience de conduite) ont indiqué pour chaque image et pour chaque manœuvre, quels indices leur permettent de prendre la décision.
- Les *indices généraux* permettent la compréhension de la scène, mais ne sont pas directement liés à la manœuvre. Dans notre plan expérimental, ils correspondent à des indices liés à une autre manœuvre que celle notifiée. Ainsi, les *indices liés à la manœuvre* du groupe A sont les *indices généraux* du groupe B, et inversement (Figure 15).
- Les *zones de saillance visuelle* ont été définies sur la base de cartes de saillances statiques (basée sur les contrastes et les couleurs, pour plus d'informations voir Belardinelli, Pirri et Carbone, 2009) (Figure 14 ; voir Annexe 5.1 pour plus d'exemples).

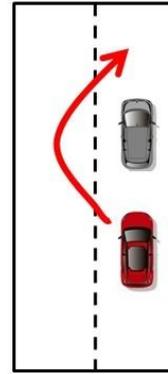
Puisque les prises de vue des situations ont été réalisées en condition de circulation urbaine le nombre de zones d'intérêts (AOI) de chaque catégorie peut varier entre 2 et 3 par scène, selon les situations.



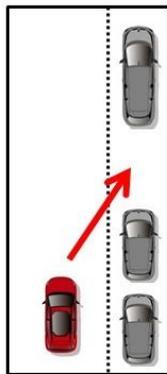
Tourner
à gauche/droite



Se déporter
à gauche/droite



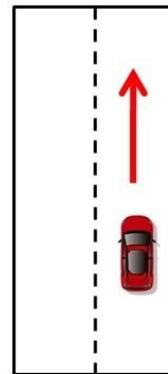
Dépasser



Se garer



S'arrêter

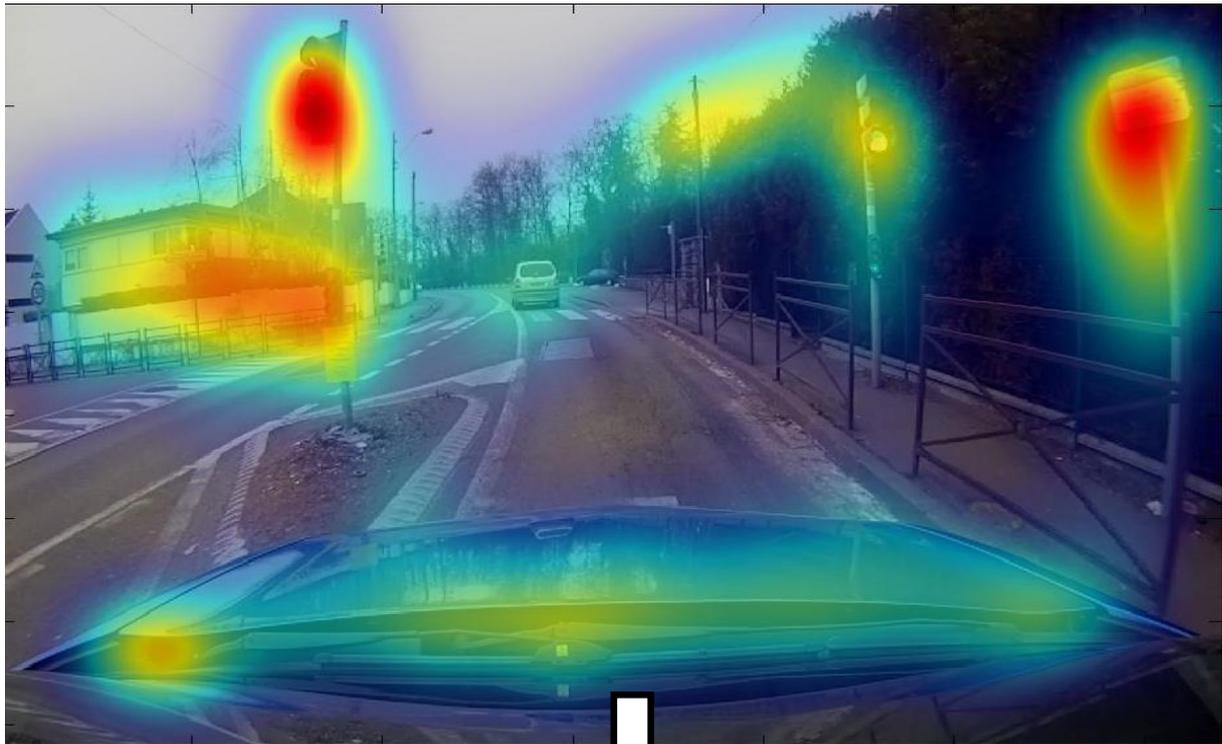


Continuer
tout droit

Figure 12. Les types de manœuvres notifiées aux participants.

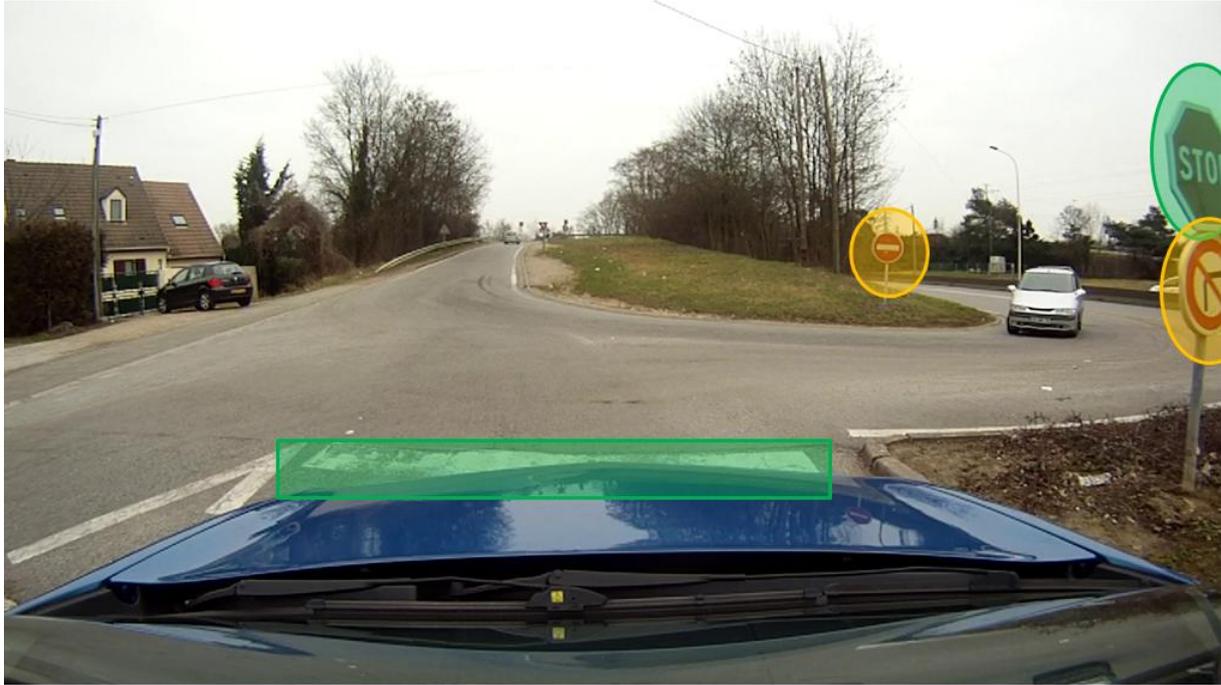
Eléments de signalisation		
<p>Feux tricolore</p> 	<p>Marquage au sol</p> 	<p>Panneau</p> 
Evénements		
<p>Piéton</p> 	<p>Véhicule</p> 	

Figure 13. Exemples des différents types d'indices présents dans les scènes visuelles.



Zones de forte saillance visuelle

Figure 14. Exemples de zones de saillance définies à partir de cartes de saillance calculées par les algorithmes de Belardinelli, Pirri et Carbone, 2009.



Légende :

Buts Indices	« S'arrêter »	« Tourner à droite »
	<p><u>Indices liés à la manœuvre</u></p> <p>Les indices sont congruents : la ligne STOP et le panneau STOP imposent au conducteur de s'arrêter.</p>	<p><u>Indices généraux</u></p> <p>Les indices ne sont pas liés à la manœuvre « tourner à droite ».</p>
	<p><u>Indices généraux</u></p> <p>Les indices ne sont pas liés à la manœuvre « s'arrêter ».</p>	<p><u>Indices liés à la manœuvre</u></p> <p>Les indices sont incongruents : les deux panneaux interdisent au conducteur de tourner à droite.</p>

Figure 15. Exemples d'indices dans une scène visuelle.

2.2.2.3 Dispositif expérimental

L'expérience a été construite avec le logiciel Experiment Center de SMI®. Nous avons utilisé un oculomètre SMI RED4™ réglé à 50 Hertz et les calibrages ont été faits en 9 points. Les participants étaient placés à 60cm de l'oculomètre et les stimuli étaient affichés sur un écran 17 pouces (résolution 1280x768).

2.2.2.4 Procédure

La procédure se déroule en 3 étapes (Figure 16). D'abord, l'écran notifie à l'écrit une manœuvre à réaliser. L'affichage dure 4s, et lorsqu'il contient une indication latéralisée (gauche/droite), il est accompagné de la flèche associée afin de pallier les éventuelles difficultés de latéralisation des participants. Ensuite, une image de situation de conduite automobile est affichée durant 4s. Après l'affichage de l'image, le participant doit indiquer s'il peut réaliser la manœuvre ou pas dans la situation présentée. Pour cela, il sélectionne sa réponse sur un formulaire affiché à l'écran (réponse binaire) ; le temps de réponse nécessaire à la prise de décision était libre. La passation comporte 60 séquences dont l'ordre de présentation a été randomisé.

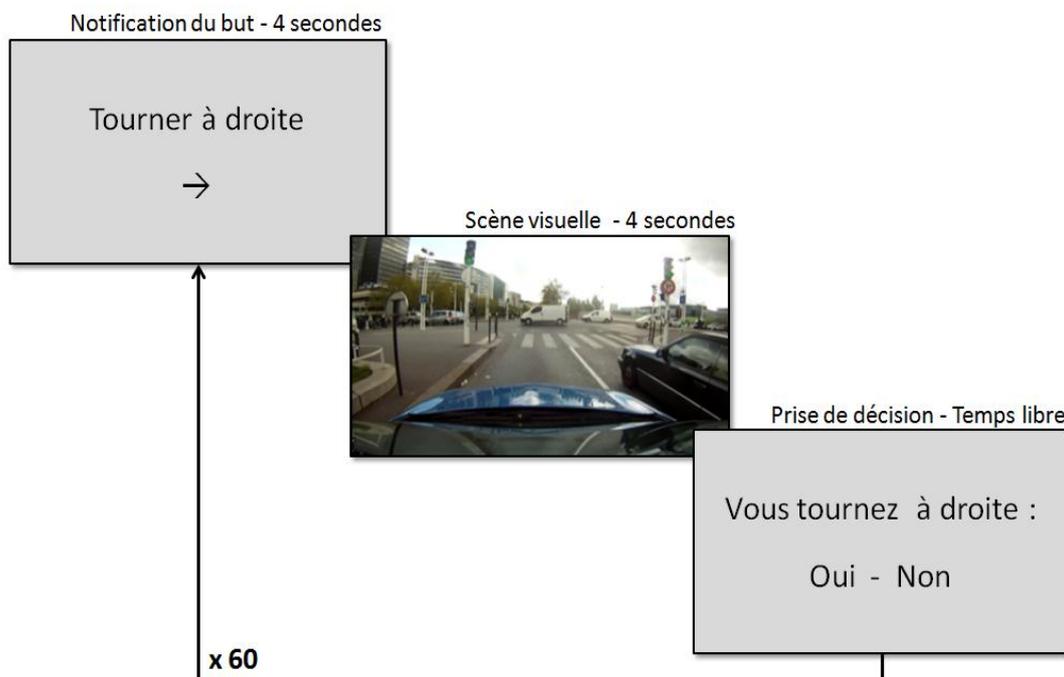


Figure 16. Déroulement de la procédure.

2.2.2.5 Plan expérimental

Nous avons mis en place un plan mixte 2x2. Les 40 participants ont tous visionné les 60 images de situation de conduite. Le facteur principal est *la manœuvre notifiée* aux participants. C'est un **facteur inter-sujet à 2 modalités** qui définit deux groupes dans lesquels les participants ont été affectés aléatoirement: 20 participants ont eu des manœuvres A (groupe A) et 20 participants ont eu des manœuvres B (groupe B). Ainsi pour une même image, une manœuvre différente était notifiée selon le groupe. Les manœuvres A et B étaient

composées de manière équivalente, il y avait par exemple autant de manœuvres « dépasser » dans les deux groupes. Les manœuvres A et B ont été définies de manière à ce que les indices permettant de prendre la décision soient différents d'un groupe à l'autre : les *indices liés à la manœuvre* du groupe A sont les *indices généraux* du groupe B, et inversement (Figure 15). Le second facteur est la *congruence entre la manœuvre et la situation*. C'est un **facteur intra-sujet à deux modalités**. Sur la totalité des 60 images de situation de conduite qu'ils avaient à visionner, les participants ont chacun reçu 30 manœuvres congruentes et 30 manœuvres non congruentes avec la situation.

Trois variables dépendantes ont été sélectionnées afin de rendre compte de l'allocation de l'attention visuelle : (i) la fréquence de fixation sur les AOIs, (ii) le pourcentage de temps de fixation (Percentage Dwell Time - PDT) sur les AOIs, et (iii) le pourcentage d'AOIs fixés.

La *fréquence de fixation sur les AOIs* correspond au nombre de fixations par seconde sur les AOIs ; c'est un indicateur classique qui témoigne de l'intérêt pour une information (Fitts, Jones & Milton, 1950). Pour une activité de conduite automobile, il est généralement admis que lorsque les fixations se focalisent sur une zone, alors l'attention est aussi focalisée sur cette zone (Mourant & Rockwell, 1970 ; voir Shinar, 2008, pour une discussion).

Le *pourcentage de temps de fixation (PDT) sur les AOIs* correspond à la proportion du temps passé par le conducteur à regarder une zone, par rapport aux autres zones. C'est l'indicateur communément utilisé pour rendre compte de l'allocation de l'attention visuelle (e.g. Wickens & al., 2001 ; Horrey & al., 2006 ; Lemonnier, Brémond & Baccino. 2015). Une allocation de l'attention visuelle optimale est principalement guidée par des facteurs *top-down*, et l'effet des facteurs *bottom-up* doit être minimisé (Wickens, 2003) ; ce qui se caractérise par un pourcentage de temps de fixation plus important sur les indices liés à la tâche que sur les zones de saillance visuelle.

Le *pourcentage d'AOIs fixés* correspond au nombre d'AOIs fixés dans une scène, par rapport au nombre total d'AOIs du même type dans la scène. C'est un indicateur du degré d'exhaustivité de la prise d'information, et par conséquent de l'optimalité du parcours visuel : plus le conducteur est exhaustif dans ses fixations sur les indices pertinents pour la réalisation de son but, moins il risque d'ignorer un indice important pour sa prise de décision.

Afin d'analyser le parcours visuel immédiat après la notification du but, seules les données issues des 2 premières secondes des stimuli ont été analysées (rappelons que les images restaient affichées pendant 4 secondes afin de ne pas mettre les participants dans une

situation de forte contrainte temporelle, ce qui pourrait constituer un biais). De plus, la première fixation a été retirée des données pour chaque scène visuelle afin de contourner le biais de fixation central.

Une ANOVA mixte a été opérée avec un facteur inter-sujets (manœuvre) et deux facteurs intra-sujets (congruence et type d'AOI). Le facteur Manœuvre comportait deux modalités (A et B), la congruence pouvait être de deux types (congruente ou incongruente) et il y avait trois types de zone d'intérêt (indices liés à la manœuvre, indices généraux et zones de saillance visuelle). L'ANOVA 2x2x3 a été réalisée sur les trois variables dépendantes : les fréquences de fixations, les pourcentages de temps de fixations, et les proportions d'AOIs fixés.

2.2.3 Résultats

Les moyennes des mesures ont été reportées dans le Tableau 2. Un filtrage a été opéré sur les durées de fixations pour traiter les données aberrantes : les durées supérieures à 2 écarts types ont été remplacées par la moyenne. L'analyse des fréquences de fixations a porté sur les données brutes.

Mesures	Manœuvre	Manœuvres congruentes			Manœuvres incongruentes		
		Indices liés à la manœuvre	Indices généraux	Zones saillantes	Indices liés à la manœuvre	Indices généraux	Zones saillantes
Fréquences des fixations (nombre/s)	A	0,62 (0,14)	0,23 (0,04)	0,19 (0,06)	0,67 (0,11)	0,25 (0,05)	0,19 (0,06)
	B	0,46 (0,12)	0,27 (0,11)	0,19 (0,06)	0,63 (0,19)	0,29 (0,12)	0,22 (0,06)
Pourcentages des temps de fixation (%)	A	0,60 (0,06)	0,23 (0,06)	0,17 (0,05)	0,63 (0,05)	0,21 (0,04)	0,16 (0,05)
	B	0,53 (0,08)	0,28 (0,07)	0,19 (0,07)	0,59 (0,06)	0,23 (0,06)	0,19 (0,05)
Pourcentages d'AOIs fixés (%)	A	0,59 (0,09)	0,33 (0,06)	0,37 (0,12)	0,46 (0,07)	0,32 (0,07)	0,32 (0,09)
	B	0,6 (0,11)	0,3 (0,09)	0,29 (0,09)	0,48 (0,1)	0,32 (0,09)	0,33 (0,07)

Tableau 2. Moyennes et SD des fréquences de fixations et des pourcentages de temps de fixations pour les deux premières secondes des stimuli.

2.2.3.1 Fréquences de fixations sur les AOIs

Nous observons un effet principal du type d'AOI ($F(2, 76) = 404.17, p < .001, \eta^2 = .91$). Les participants font plus de fixations sur les indices notés comme utiles à la réalisation de la manœuvre que sur les indices généraux ($F(1, 38) = 413.57, p < .001$). Les différences sont significatives pour le groupe A ($F(1, 38) = 279.34, p < .001$) et pour le groupe B ($F(1, 38) = 145.11, p < .001$). Egalement, ils font plus de fixations sur les indices généraux que sur les zones saillantes ($F(1, 38) = 26.54, p < .001$). Les différences sont significatives pour le groupe A ($F(1, 38) = 9.58, p < .004$) et pour le groupe B ($F(1, 38) = 17.56, p < .001$).

Un effet principal de la congruence est également observé : les participants font plus de fixations sur les zones d'intérêts lorsque leur manœuvre n'est pas congruente ($F(1, 38) = 25.47, p < .001, \eta^2 = .41$).

Enfin un effet d'interaction de la congruence et du type d'AOI est observé ($F(2, 76) = 15.86, p < .001, \eta^2 = .29$) (Figure 17). Cet effet concerne uniquement les indices liés à la manœuvre ($F(1, 38) = 30.29, p < .001$), tant pour le groupe A ($F(1, 38) = 5.03, p < .031$) que pour le groupe B ($F(1, 38) = 56.36, p < .001$). Les fréquences de fixation sur les autres AOIs ne sont pas impactés par la congruence ($F < 1$).

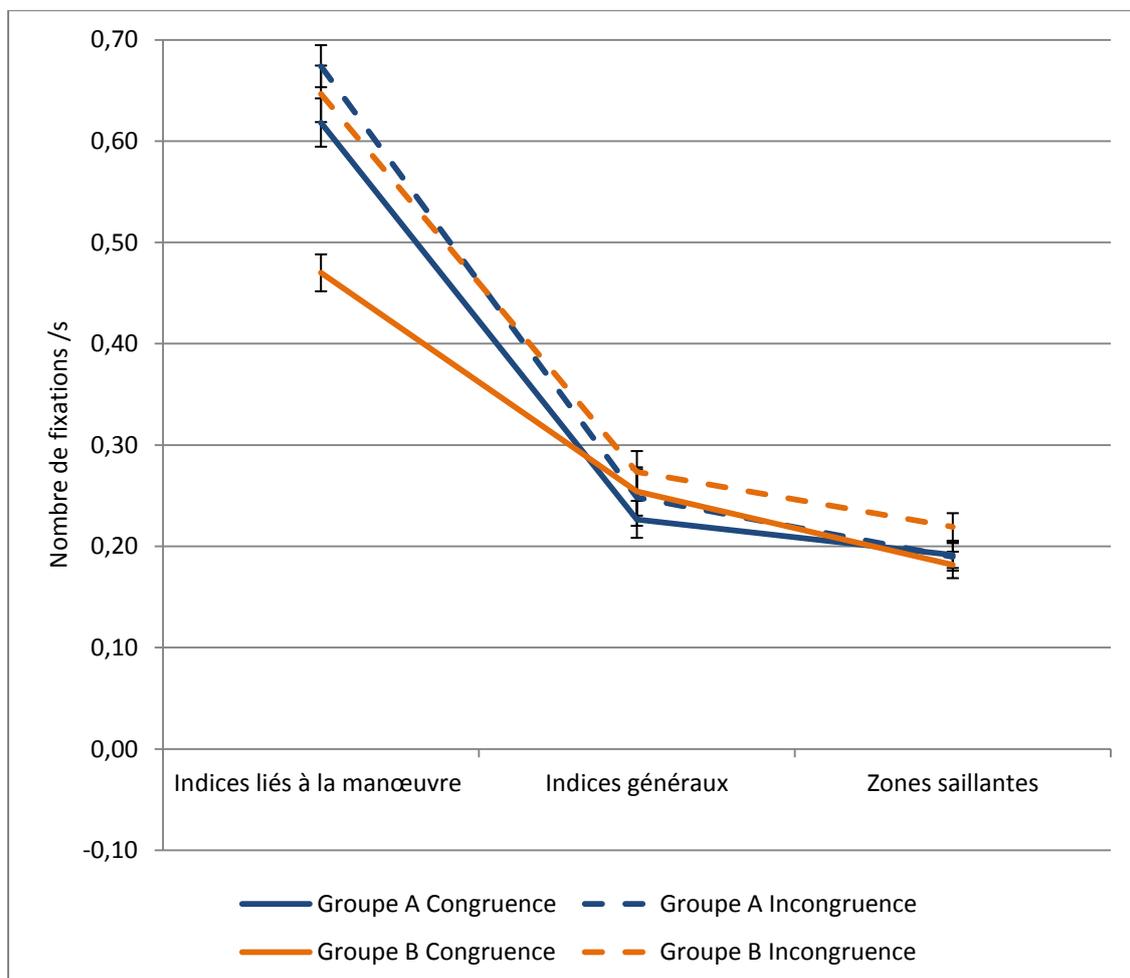


Figure 17. Fréquences de fixations sur les AOIs selon les groupes expérimentaux et la congruence des indices.

2.2.3.2 Pourcentage de temps de fixation sur les AOIs

L'ANOVA montre un effet principal du type d'AOI ($F(2, 76) = 555.76, p < .001, \eta^2 = .94$). Les indices spécifiques à la manœuvre recueillent des pourcentages de temps de fixations plus importants que les indices généraux ($F(1, 38) = 656.24, p < .001$) ; les différences sont

significatives dans le groupe A ($F(1, 38) = 420.51, p < .001$) et dans le groupe B ($F(1, 38) = 247.18, p < .001$). Egalement, les indices généraux recueillent des pourcentages de temps de fixations plus importants que les zones saillantes ($F(1, 38) = 24.05, p < .001$) ; les différences sont significatives dans le groupe A ($F(1, 38) = 9.72, p < .004$) et dans le groupe B ($F(1, 38) = 14.57, p < .001$).

Enfin, on observe un effet d'interaction de la congruence et du type d'AOI ($F(2, 76) = 10.67, p < .001, \eta^2 = .22$) (Figure 18). Les pourcentages de temps de fixations sur les indices liés à la manœuvre sont plus importants dans une situation d'incongruence ($F(1, 38) = 15.44, p < .05$) ; les différences sont significatives dans le groupe A ($F(1, 38) = 4.44, p < .05$) et dans le groupe B ($F(1, 38) = 13.17, p < .001$). Les pourcentages de temps de fixations sur les autres AOIs ne sont pas impactés par la congruence ($F < 1$).

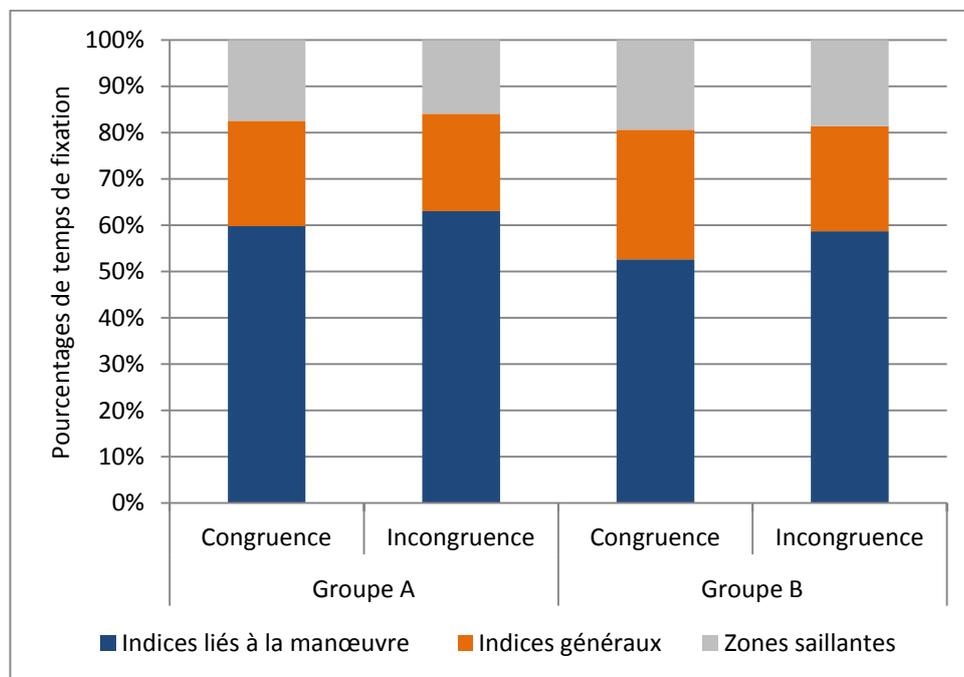


Figure 18. Pourcentages de temps de fixation sur les AOIs selon les groupes expérimentaux et la congruence des indices.

2.2.3.3 Pourcentage d'AOIs fixées

On observe un effet principal du type d'AOI ($F(2, 76) = 146.63, p < .001, \eta^2 = .79$). Les indices liés à la manœuvre sont fixés de manière plus exhaustive que les autres AOIs ($F(1, 38) = 337.29, p < .001$), la différence est significative pour le groupe A ($F(1, 38) = 143.03, p < .001$) et pour le groupe B ($F(1, 38) = 196.38, p < .001$). Il n'y a pas de différence significative de pourcentage d'indices fixés entre les indices généraux et les zones saillantes ($F < 1$).

Egalement, on relève un effet principal de la congruence ($F(1, 38) = 17.56, p < .001, \eta^2 = .32$) : le pourcentage d'indices fixés est plus important en situation d'incongruence. Il y a un effet d'interaction de la congruence et du type d'AOI ($F(2, 76) = 19.34, p < .001, \eta^2 = .34$) (Figure 19) ; les indices liés à la manœuvre sont fixés de manière plus exhaustive lorsqu'ils sont congruents que lorsqu'ils sont incongruents ($F(1, 38) = 58.48, p < .001$), les différences sont significatives pour le groupe A ($F(1, 38) = 29.58, p < .002$) et pour le groupe B ($F(1, 38) = 28.91, p < .001$). La congruence n'a d'effet ni sur les indices non liés à la manœuvre ($F < 1$), ni sur les zones saillantes ($F < 1$).

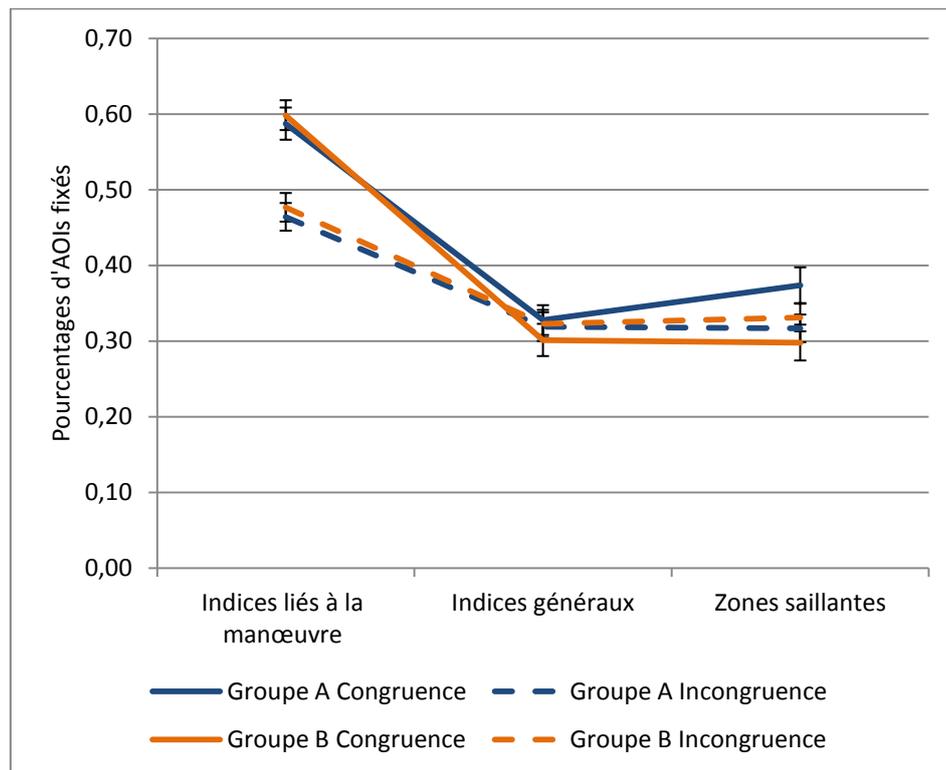


Figure 19. Pourcentages d'AOIs fixés selon les groupes expérimentaux et la congruence des indices.

2.2.4 Discussion

Dans cette première expérimentation, deux groupes de 20 participants ont visualisé 60 scènes visuelles statiques de situations de conduite automobile. Avant l'affichage de chaque image, il leur était notifié une manœuvre (e.g. « tourner à droite »). Les participants visualisaient ensuite l'image pendant 4s, puis devaient indiquer si la manœuvre était réalisable. Les manœuvres notifiées variaient d'un groupe expérimental à l'autre, mais les scènes visuelles étaient les mêmes. Nous avons contrôlé la congruence des manœuvres avec les situations, c'est-à-dire le fait que la manœuvre soit permise par la situation ou non. Nous

avons défini trois types de zones d'intérêts (AOIs) dans les scènes visuelles : (i) les indices liés à la manœuvre, (ii) les indices non liés à la manœuvre et (iii) les zones de saillance visuelle.

Nous avons fait l'hypothèse que les indices (i) et (ii), c'est-à-dire les informations sémantiques, canaliseront plus l'attention que les zones de saillance visuelle (iii). Parmi ces indices, on s'attend à ce que l'attention visuelle soit davantage allouée aux indices liés aux manœuvres (permettant la prise de décision) qu'aux indices généraux (permettant la compréhension de la scène). Enfin, parmi les indices liés aux manœuvres, on s'attend à ce que les indices incongruents (qui interdisent ou empêchent la manœuvre) focalisent encore davantage l'attention visuelle que les indices congruents.

Conformément à notre première hypothèse, nos résultats montrent que les indices liés aux informations sémantiques focalisent bien davantage l'attention visuelle que les zones de saillance visuelle, en termes de fréquence de fixation et de pourcentage de temps de fixation. Ce premier résultat va dans le sens d'une primauté des processus descendants (regard guidé par le but) sur les processus ascendants (regard guidé par les effets de saillance), et ce lors des toutes premières secondes de visualisation. Cette primauté et rapidité des processus ascendants lors des visualisations de scènes visuelles en situation d'activité a notamment été défendue par Rothkopf, Ballard et Hayhoe (2007) et Tatler, Hayhoe, Land et Ballard (2011).

Dans notre expérimentation, les indices liés à la manœuvre pour le groupe A correspondaient aux indices généraux pour le groupe B, et inversement. Ces indices étaient ainsi de même nature : par exemple des panneaux, des feux, des lignes au sol, des piétons ou d'autres véhicules. Ces indices avaient donc *a priori* la même valeur entre les deux groupes expérimentaux. Pourtant, comme nous l'avons prédit, nos résultats montrent que l'introduction d'un but (i.e. une manœuvre à réaliser) a conduit à une focalisation de l'attention sur les indices liés à ce but, qui se traduit par des fréquences de fixation plus importantes. Cette focalisation de l'attention est confirmée par les pourcentages de temps de fixation: on relève une répartition des temps de fixation d'environ 60% sur les indices liés à la manœuvre, contre seulement 20% sur les indices non liés à la manœuvre, et 20% sur les zones saillantes (Figure 18).

Egalement, les indices liés à la manœuvre sont fixés de manière plus exhaustive : environ 50% des indices liés au but recueillent des fixations, contre 30% des indices généraux et 30%

des zones saillantes. Tous nos indicateurs concourent à montrer qu'ici l'allocation de l'attention visuelle est orientée vers la manœuvre.

Cette focalisation de l'attention sur les indices liés à la manœuvre a été observée lors des deux premières secondes des stimuli, ce qui va dans le sens du principe d'immédiateté selon lequel le système visuel extrait les informations exactement au moment où elles sont nécessaires à la réalisation de la tâche afin de ne pas saturer les ressources cognitives (Ballard, Hayhoe, & Pelz, 1995; Ballard, Hayhoe, Pook, & Rao, 1997 ; Triesch, Ballard, Hayhoe, & Sullivan, 2003). Cependant, les scènes visuelles étaient affichées après la notification de la manœuvre, et nous ne pouvons pas conclure quant à l'immédiateté de la prise d'information sans analyser les visualisations précédant cette notification de la manœuvre. C'est une limite inhérente au protocole et à notre matériel. Il sera nécessaire lors des futures expérimentations d'observer la phase précédant l'introduction du but, afin d'avoir des mesures de référence et de pouvoir effectuer des comparaisons avant/après la notification.

Les indices généraux, qui sont utiles à la compréhension de la scène mais inutiles à la prise de décision, ne semblent pas plus attirer l'attention que les zones saillantes : les fréquences de fixation et les pourcentages de temps de fixation diffèrent peu entre ces deux types de zones. Ce résultat est cohérent avec ceux de Boloix (2006) qui a montré que les éléments non liés au but ne sont pas traités de façon prioritaire, alors même qu'ils peuvent aider à la compréhension générale de la scène. Cependant les indices généraux ne sont pas ignorés pour autant, puisque près d'un tiers d'entre eux ont recueilli des fixations.

Conformément à notre deuxième hypothèse, les indices incongruents ont recueilli plus de fixations et de pourcentage de temps de fixations que les indices congruents. L'incongruence accroît donc la Valeur de l'indice, ce que l'on pourrait expliquer par son rôle décisif dans la prise de décision : si le conducteur ne perçoit pas cet indice, il risque l'infraction ou l'accident. Mais paradoxalement, les indices incongruents sont fixés de manière moins exhaustive que les indices congruents. Une explication possible serait qu'un seul indice incongruent peut suffire à décider de ne pas réaliser une manœuvre (e.g. s'il y a un sens interdit à gauche alors le conducteur ne tourne pas à gauche). A l'inverse, le parcours visuel doit être exhaustif pour décider de réaliser la manœuvre (le conducteur vérifie qu'aucun indice ne l'empêche de tourner à gauche). On peut penser que cet effet ne sera pas nécessairement observé en situation réelle de conduite, car lorsqu'un indice est incongruent, le but est immédiatement réévalué, et les indices liés au but initial sont susceptibles de perdre leur pertinence.

L'utilisation de photographies a constitué une bonne approche pour mettre en place notre analyse de l'allocation de l'attention visuelle. Elles fournissent des situations de conduite réalistes et des scènes présentant un niveau de richesse visuelle proche de la réalité. En revanche, leur nature statique reste une limite importante. Effectivement, on ne peut pas omettre l'aspect dynamique de la conduite et l'importance du déplacement. La scène visuelle du conducteur se caractérise par le déplacement simultané (i) des indices et (ii) du conducteur par rapport aux indices. On peut donc raisonnablement penser que la prise d'information est plus complexe sur des scènes dynamiques, et que l'exigence d'optimalité du parcours oculaire en devient plus élevée. Il apparaît nécessaire de prolonger notre expérimentation afin de confirmer nos résultats sur des scènes visuelles dynamiques.

Expérimentation I - En résumé
Objectifs
<ul style="list-style-type: none"> • Déterminer dans quelle mesure l'allocation de l'attention visuelle est spécifique à la manœuvre. • Déterminer si l'incongruence d'un indice avec le but focalise davantage l'attention.
Principaux résultats
<ul style="list-style-type: none"> • L'ensemble des indices (i.e. les informations sémantiques) focalise davantage l'attention visuelle que les zones de saillance visuelle. • L'attention visuelle est davantage allouée aux indices liés aux manœuvres (permettant la prise de décision) qu'aux indices généraux (permettant la compréhension de la scène). • Les indices incongruents avec la manœuvre focalisent davantage l'attention visuelle que les indices congruents, mais sont consultés de manière moins exhaustive que les indices congruents.

2.3 EXPERIMENTATION II : ALLOCATION DE L'ATTENTION VISUELLE SUR DES SCENES DYNAMIQUES

La première expérimentation nous a montré que l'allocation de l'attention visuelle est très spécifique à la manœuvre. Nous avons alors utilisé des photographies de situation de conduite, afin d'obtenir un niveau de richesse visuelle proche de la réalité. La seconde expérimentation est une déclinaison de la première, dans laquelle nous utilisons cette fois-ci des vidéos de situations de conduite afin de prendre en compte **l'aspect dynamique** de l'activité. En effet, l'activité de conduite automobile a pour caractéristique d'être continuellement dynamique (par opposition à une succession de situations figées), en raison du déplacement simultané du conducteur et des autres usagers, ainsi que des événements de l'environnement (e.g. feux tricolores). Une situation de conduite est donc en permanente évolution et l'allocation de l'attention visuelle doit alors être optimale : le conducteur doit relever les bonnes informations au bon moment pour chaque prise de décision. De plus, le mouvement des éléments de l'environnement et le déplacement du conducteur complexifie la prise d'information.

Dans cette déclinaison de l'expérimentation I, les hypothèses et le plan expérimental sont conservés. La procédure expérimentale a été adaptée pour des scènes visuelles dynamiques ; dans un souci de clarté, seules les évolutions du matériel et de la procédure seront présentées ici.

2.3.1 Méthode

2.3.1.1 Participants

Vingt-quatre participants volontaires ont été recrutés (6 femmes, âge moyen = 33, min = 21, max = 59, SD = 10,6). Tous avaient le permis de conduire et une expérience de conduite depuis plus de 3 ans (moy = 10,5, max = 30, SD = 8,8). Aucun des participants n'avait une vue corrigée et ne portait ni lunettes, ni lentilles. Les participants ont donné leur autorisation pour l'enregistrement de leurs données oculaires et de leurs réponses, et ils ont été dédommagés d'une carte cadeau. Les 24 participants ont été repartis aléatoirement dans 2 groupes.

2.3.1.2 Matériel

Vingt séquences vidéo de situations de conduite automobile de 10 secondes ont été utilisées. Les prises de vue ont été réalisées en condition de circulation urbaine, à l'aide d'une caméra embarquée à grand angle fixée sur le pare-brise d'une voiture.

A la 7^{ème} seconde de chaque séquence, un enregistrement sonore notifie une manœuvre au participant (à la manière d'un GPS). La manœuvre est notifiée au cours de la vidéo, et n'interrompt pas son déroulement. Les manœuvres étaient analogues à celles notifiées dans l'expérimentation I (Voir partie 2.2.2.2 ; Figure 12). Chacune des manœuvres peut être congruente ou non avec la situation de la scène visuelle. Sur les 20 séquences présentées à chaque participant, la moitié (10 séquences) notifie des manœuvres congruentes et l'autre moitié notifie des manœuvres non congruentes.

Trois types de zones d'intérêt ont été définis. Les *indices liés à la manœuvre* et les *indices généraux* ont été définis selon les mêmes critères que pour l'expérimentation I (voir partie 2.2.2.2 ; Figure 15). Les *zones de saillance visuelle* ont été définies sur la base de cartes de saillances dynamiques calculées selon les critères de Belardinelli, Pirri et Carbone (2009) et de Belardinelli, Carbone et Schneider (2012) (voir Annexe 5.2 pour des exemples). Il s'agit d'une méthode prenant en compte à la fois les contrastes et les mouvements. Puisque les prises de vue des situations ont été réalisées en condition de circulation urbaine le nombre de zones d'intérêts (AOI) de chaque type peut varier entre 2 et 3 par scène, selon les situations.

2.3.1.3 Procédure

La procédure consiste à afficher sur un écran des séquences vidéo de situations de conduite automobile, d'une durée de 10 secondes chacune. Pendant les 7 premières secondes de chaque séquence, le participant doit regarder et se familiariser avec les scènes de conduite. L'annonce audio de la manœuvre, d'une durée d'environ 1 seconde, est notifiée à la 7^{ème} seconde des vidéos. A la fin de la séquence, le participant doit indiquer s'il peut réaliser la manœuvre ou pas. Pour cela, il sélectionne sa réponse sur un formulaire affiché à l'écran (réponse binaire) ; le temps de réponse nécessaire à la prise de décision était libre. La passation comporte 20 séquences dont l'ordre de présentation a été randomisé. Entre chaque séquence apparaît - durant 4s- un écran gris neutre (Figure 20).

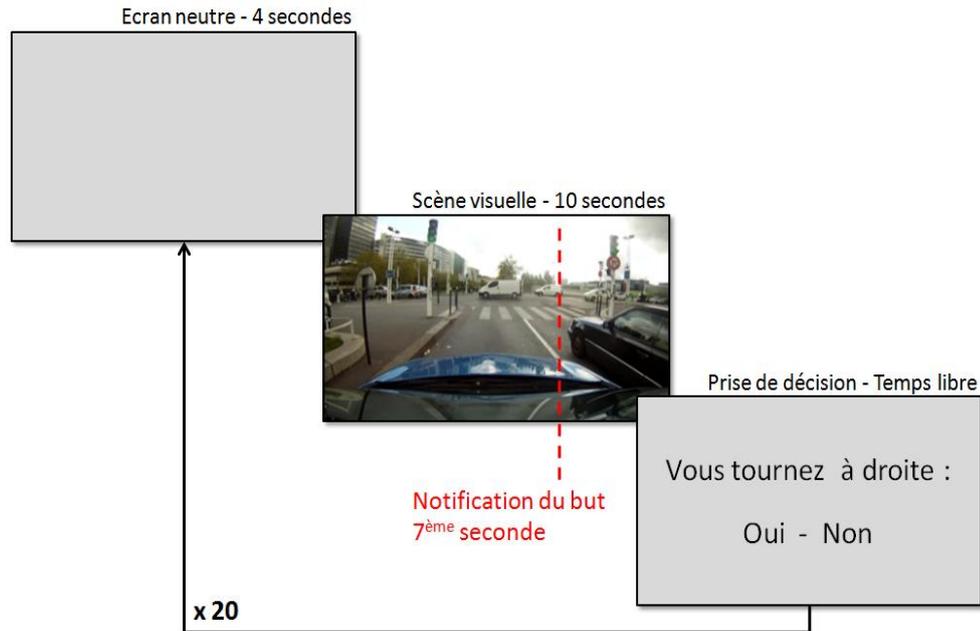


Figure 20. Déroulement de la procédure

2.3.2 Résultats

Les mesures ont été effectuées à partir de la notification du but, c'est-à-dire sur les 3 dernières secondes des vidéos (7-10s).

Une ANOVA mixte a été opérée avec un facteur inter-sujets (manœuvre) et deux facteurs intra-sujets (congruence et type d'AOI). Le facteur Manœuvre comportait deux modalités (A et B), la congruence pouvait être de deux types (congruente ou incongruente) et il y avait trois types de zone d'intérêt (indices liés à la manœuvre, indices non liés à la manœuvre et zones de saillance visuelle). L'ANOVA 2x2x3 a été réalisée sur les trois variables dépendantes : les fréquences de fixations, les pourcentages de temps de fixations, et les proportions d'AOIs fixés. Les moyennes des mesures ont été reportées dans le Tableau 3. Un filtrage a été opéré sur les durées de fixations pour traiter les données aberrantes : les durées supérieures à 2 écarts types ont été remplacées par la moyenne. L'analyse des fréquences de fixations a porté sur les données brutes.

Mesures	Manœuvre	Manœuvres congruentes			Manœuvres incongruentes		
		Indices liés à la manœuvre	Indices généraux	Zones saillantes	Indices liés à la manœuvre	Indices généraux	Zones saillantes
Fréquences des fixations (nombre/s)	A	0,21 (0,35)	0,07 (0,24)	0,07 (0,17)	0,17 (0,28)	0,05 (0,14)	0,14 (0,33)
	B	0,07 (0,22)	0,02 (0,09)	0,05 (0,17)	0,31 (0,40)	0,11 (0,26)	0,09 (0,22)
Pourcentages des temps de fixation (%)	A	0,58 (0,14)	0,25 (0,15)	0,17 (0,12)	0,44 (0,15)	0,16 (0,07)	0,40 (0,19)
	B	0,53 (0,31)	0,14 (0,16)	0,33 (0,34)	0,64 (0,13)	0,18 (0,12)	0,18 (0,11)
Pourcentages d'AOIs fixés (%)	A	0,52 (0,17)	0,3 (0,18)	0,25 (0,17)	0,55 (0,23)	0,27 (0,13)	0,35 (0,15)
	B	0,31 (0,16)	0,14 (0,13)	0,16 (0,14)	0,6 (0,18)	0,32 (0,19)	0,33 (0,17)

Tableau 3. Moyennes et SD des fréquences de fixations, des pourcentages de temps de fixations, et des pourcentages d'AOIs fixés.

2.3.2.1 Fréquences de fixations sur les AOIs

Nous observons un effet principal du type d'AOI ($F(2, 44) = 62.69, p < .004, \eta^2 = .74$). Les participants font plus de fixations sur les indices notés comme utiles à la réalisation de leurs buts, que sur les indices généraux ($F(1, 22) = 389.71, p < .001$); les différences sont significatives pour le groupe A ($F(1, 22) = 178.91, p < .001$) et pour le groupe B ($F(1, 22) = 213.71, p < .001$). En revanche il n'y a pas de différence de fréquence de fixations entre les indices généraux et les zones saillantes ($F < 1$) (Figure 21).

On observe également un effet principal de la congruence ($F(1, 22) = 38.22, p < .001, \eta^2 = .63$): les participants font plus de fixations sur les zones d'intérêts lorsque leur but n'est pas congruent. Cependant, cet effet de la congruence concerne le groupe B ($F(1, 22) = 78.82, p < .001$) mais n'est pas observé pour le groupe A ($F < 1$).

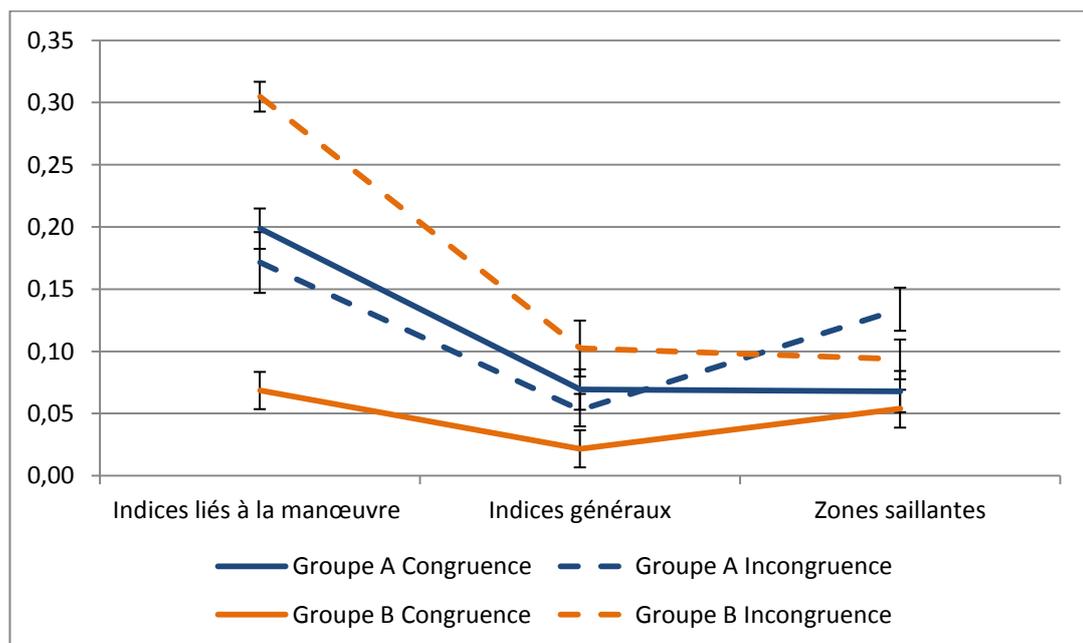


Figure 21. Fréquences de fixations sur les AOIs selon les groupes expérimentaux et la congruence des indices.

2.3.2.2 Pourcentage de temps de fixation sur les AOIs

L'ANOVA montre un effet principal du type d'AOI ($F(2, 44) = 32.62, p < .004, \eta^2 = .59$). Les indices liés à la manœuvre recueillent des pourcentages de temps de fixations plus importants que les autres indices ($F(1, 22) = 60.28, p < .001$); les différences sont significatives dans le groupe A ($F(1, 22) = 18.64, p < .001$) et dans le groupe B ($F(1, 22) = 45.85, p < .001$). En revanche on ne relève pas de différence de pourcentages de temps de fixations entre les indices généraux et les zones saillantes ($f < 1$) (Figure 22).

Enfin, il n'y a pas d'effet de la congruence sur le pourcentage de temps de fixation ($f < 1$).

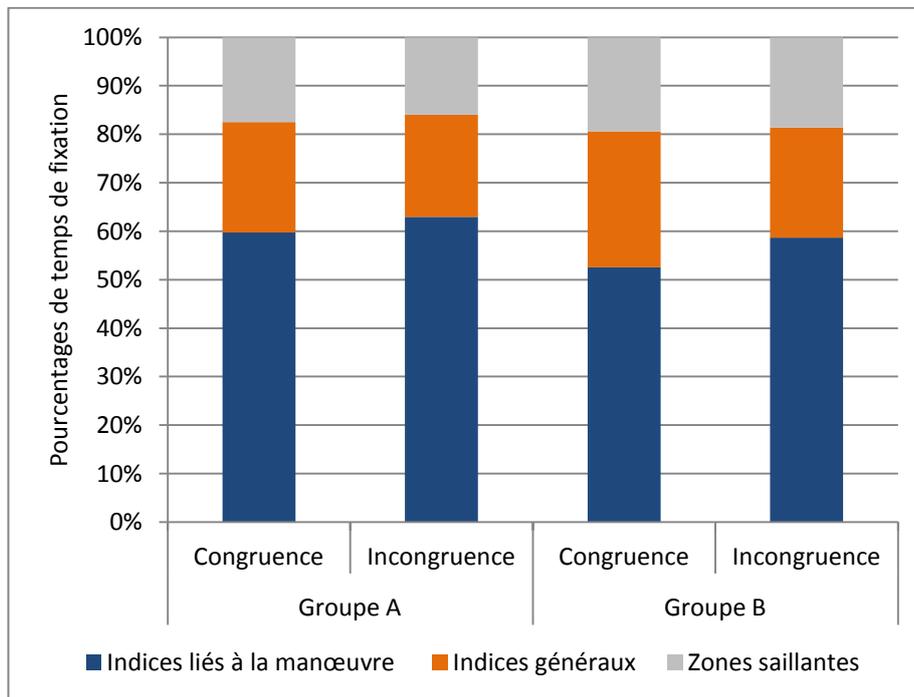


Figure 22. Pourcentages de temps de fixation sur les AOIs selon les groupes expérimentaux et la congruence des indices.

2.3.2.3 Proportion d'AOIs fixés

Nous observons un effet principal du type d'AOI ($F(2, 44) = 31.89, p < .001, \eta^2 = .59$). Les indices liés au but sont fixés de manière plus exhaustive que les autres AOIs ($F(1, 22) = 91.29, p < .001$); la différence est significative pour le groupe A ($F(1, 22) = 46.95, p < .001$) et pour le groupe B ($F(1, 22) = 44.42, p < .001$) (Figure 23).

On relève également un effet principal de la congruence ($F(1, 22) = 26.89, p < .001, \eta^2 = .55$): les indices sont fixés de manière plus exhaustives lorsqu'ils sont incongruents que lorsqu'ils sont congruents.

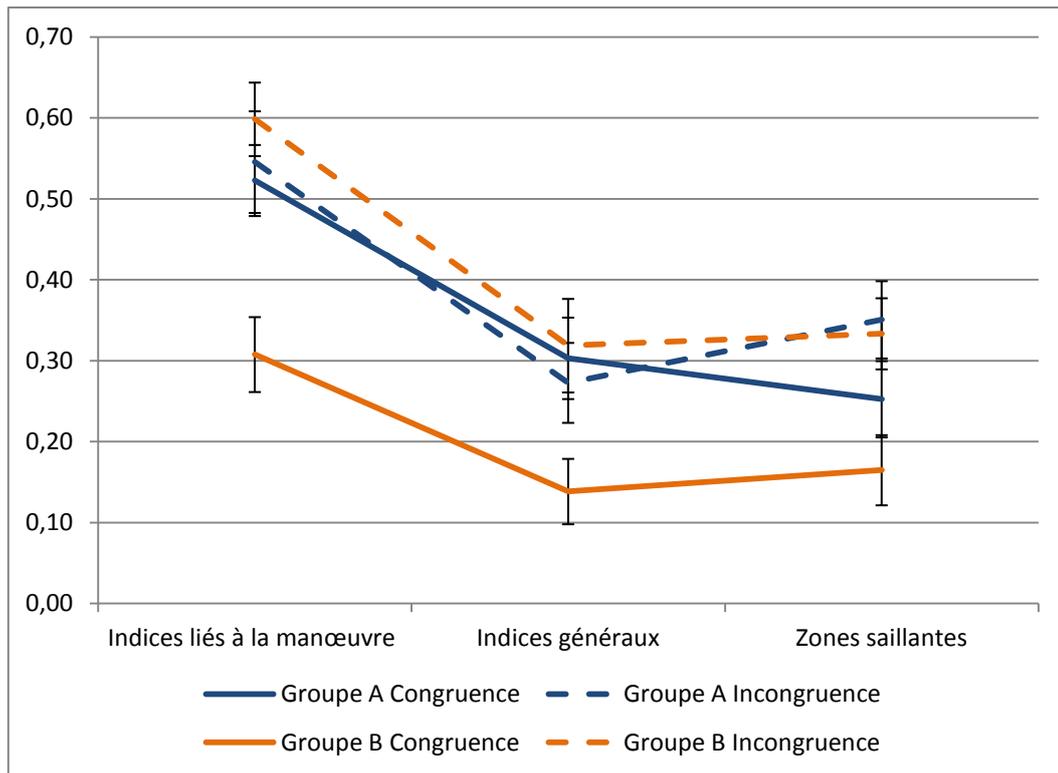


Figure 23. Pourcentages d'AOIs fixés selon les groupes expérimentaux et la congruence des indices.

2.3.3 Discussion

Afin de se rapprocher de l'environnement visuel naturel de l'activité de conduite, nous avons répliqué le protocole de l'expérimentation I sur des scènes visuelles dynamiques. Nos hypothèses sont identiques (cf., partie 2.2.1) : les indices (i.e. les informations sémantiques) devraient focaliser l'attention par rapport aux zones de saillance visuelle. Parmi ces indices, on s'attend à ce que l'attention visuelle soit davantage allouée aux indices liés aux manœuvres (permettant la prise de décision) qu'aux indices généraux (permettant la compréhension de la scène) [H1]. Enfin, parmi les indices liés aux manœuvres, on s'attend à ce que les indices incongruents (qui interdisent ou empêchent la manœuvre) focalisent encore davantage l'attention visuelle que les indices congruents [H2].

Globalement, nos résultats vont dans le sens de ceux obtenus lors de l'expérimentation I. Comme dans les scènes statiques, les indices (c'est-à-dire les informations sémantiques) présents dans les scènes dynamiques focalisent davantage l'attention visuelle que les zones saillantes, en termes de fréquence de fixation et de pourcentage de temps de fixation.

Egalement, on observe une focalisation de l'attention sur les indices liés à la manœuvre, qui se traduit par des fréquences de fixations plus importantes, ainsi qu'un pourcentage de

temps de fixations plus important. Les indices liés à la manœuvre sont aussi fixés de manière plus exhaustive : environ 50% des indices liés au but recueillent des fixations, contre 25% des indices généraux et 25% des zones saillantes. Notons que dans les deux expérimentations, les proportions de fixations des indices liés à la manœuvre sont comparables (environ 50%). Que la scène soit statique ou dynamique, les participants sont tout autant exhaustifs dans les fixations sur ces indices.

Notons qu'au contraire, l'effet de la congruence diffère entre les deux expérimentations. Sur les scènes statiques, le pourcentage d'AOIs fixés diminue lorsque la situation était incongruente avec le but. Un seul indice incongruent suffit à invalider le but, et les participants n'ont alors pas besoin d'en rechercher d'autres (e.g. on doit s'arrêter s'il y a un feu rouge). A l'inverse, sur les scènes dynamiques, le pourcentage d'AOIs fixés augmente lorsque la situation est incongruente avec le but: les participants sont alors plus exhaustifs dans leur prise d'information. L'explication peut résider dans le caractère évolutif des scènes dynamiques. En effet, puisque dans les scènes dynamiques la situation évolue dans le temps, les participants n'abandonnent pas nécessairement leur but au 1^{er} indice incongruent. Une situation incongruente à un temps t est susceptible d'être congruente à $t+1$ (et inversement). Par exemple un feu rouge peut passer au vert, ou une voiture empêchant de changer de voie peut se déplacer et permettre la manœuvre. Le caractère évolutif de la scène peut donc amener les participants à continuer de relever les indices qui sont incongruents (mais susceptibles de devenir congruents).

Les expérimentations I et II ont permis d'établir un « profil » d'allocation d'attention visuelle sur notre matériel. C'est à dire la manière dont elle se répartie parmi les différents types d'indices et les zones saillantes, en fonction d'un but. Dans l'expérimentation suivante, nous nous interrogerons sur la manière dont la RA de surlignage, en mettant en valeur certains indices, peut modifier cette allocation attentionnelle.

Expérimentation II - En résumé

Objectifs

Déclinaison de l'expérimentation I, cette fois-ci sur des scènes dynamiques.

Les objectifs sont les mêmes que ceux de l'expérimentation I, à savoir :

- Déterminer dans quelle mesure l'allocation de l'attention visuelle est spécifique à la manœuvre.
- Déterminer si l'incongruence d'un indice avec le but focalise davantage l'attention.

Principaux résultats

Résultats analogues à ceux de l'expérimentation I :

- L'ensemble des indices (i.e. les informations sémantiques) focalisent davantage l'attention visuelle que les zones de saillance visuelle.
- L'attention visuelle est davantage allouée aux indices liés aux manœuvres (permettant la prise de décision) qu'aux indices généraux (permettant la compréhension de la scène).

Résultats différents de ceux de l'expérimentation I :

- Les indices incongruents sont consultés de manière plus exhaustive que les indices congruents.

2.4 EXPERIMENTATION III : ALLOCATION DE L'ATTENTION VISUELLE SUR DES SCENES DYNAMIQUES AVEC REALITE AUGMENTEE

La troisième expérimentation a fait l'objet d'une publication dans la revue *Transportation Research part F : Traffic Psychology and Behaviour*. L'article est traduit en français et présenté ci-dessous.

2.4.1 Introduction

L'utilisation des systèmes embarqués pour l'automobile (radio, GPS, kit mains libres, etc.) multiplie les sources attentionnelles pour le conducteur, et de ce fait modifie profondément son attention visuelle (McGehee, 2001 ; Benedetto, Pedrotti, Minin, Baccino, Re & Montanari, 2011). Lors de l'utilisation de tels systèmes, l'attention visuelle est allouée à la fois à l'environnement routier et à l'intérieur de la voiture. Le récent développement des systèmes de réalité augmentée (RA) laisse entrevoir des changements majeurs. En effet, ces systèmes permettront d'intégrer des informations directement dans l'environnement extérieur, et de ce fait modifieront l'attention d'une manière encore différente. Pour mieux comprendre l'impact de tels systèmes, nous ferons référence à un modèle général de l'allocation de l'attention visuelle (SEEV) développé par Wickens, Goh, Helleberg et Horrey (2001).

Le modèle SEEV de Wickens et al. (2001) explique l'allocation de l'attention visuelle en activité en prenant compte à la fois des processus *top-down* et *bottom-up*. Il a initialement été appliqué à l'activité de pilotage d'avion (Wickens, Goh, Helleberg, Horrey & Talleur, 2003), avant d'être appliqué à la conduite automobile (Horrey, Wickens & Consalus, 2006). Dans ce modèle, le regard est guidé par 4 facteurs : la Saillance, l'Effort, l'Attente et la Valeur. La Saillance constitue le principal facteur *bottom-up* de l'attention visuelle (Tattler, Hayhoe, Land & Ballard, 2011), plusieurs facteurs de saillance visuelle ont été identifiés, comme par exemple la couleur, la forme ou le mouvement (Itti & Koch, 2000). L'Effort constitue le second facteur *bottom-up*, il correspond à l'angle visuel entre les informations, (i.e. la distance que parcourt l'œil pour atteindre une zone). Cette distance, si elle est trop importante, peut inhiber la prise d'information (Kvalseth, 1997 ; Sheridan, 1970 ; Wickens, 2007). L'Attente est un facteur *top-down* et correspond à la probabilité de voir un événement dans une zone. Enfin la Valeur, également *top-down*, correspond à l'importance d'un élément visuel pour une tâche, et à l'importance de la tâche elle-même. Ce facteur était déjà présent dans le modèle de

Carbonell et al. (1966), où la valeur d'une zone dépend du bénéfice à la consulter, et du coût si elle est ignorée. Pour reprendre l'exemple de Horrey (2006), consulter les lignes au sol apporte un fort bénéfice lors d'un changement de voie, et ignorer les informations du tableau de bord ne représente pas un coût important pour l'exécution de cette tâche.

Il n'est pas possible de traiter toutes les informations visuelles d'une scène (Ullman, 1984), et une stratégie de prise d'information doit être spécifique à la tâche pour être optimale (Shinoda, Hayhoe & Shrivastava, 2001). Selon le modèle SEEV l'attention visuelle est guidée par 4 facteurs, mais pour être optimale elle doit principalement être guidée par les facteurs *top-down* (Attente et Valeur). A l'inverse les facteurs *bottom-up* (Saillance et Effort) doivent être minimisés ou mis à profit (Wickens, 2003). En effet, un élément saillant guide le regard s'il est pertinent, mais il risque de constituer une distraction visuelle s'il n'est pas pertinent. Egalement, un élément pertinent nécessitant un effort trop important pour être consulté, aura plus de chance d'être négligé. Une allocation de l'attention visuelle optimale est par conséquent fortement dépendante de la manière dont le dispositif est conçu. Il est possible de concevoir des interfaces dans lesquelles les informations importantes sont mises en valeur, par exemple avec un indicateur de panne qui clignote (augmentation de la Saillance). Il est aussi possible de placer les informations importantes dans le champ de vision du conducteur, par exemple avec un détecteur d'angle mort au niveau du rétroviseur (diminution de l'Effort). Mais contrairement au cockpit de la voiture, l'environnement routier est non contrôlable, dynamique et relativement aléatoire. Ainsi, seul le facteur de la Saillance peut être utilisé pour favoriser chez le conducteur une allocation de l'attention optimale.

Mettre en évidence des indices permet d'attirer l'attention (Jonides, 1980) et donc de détecter une cible plus rapidement (Fisher and Tan, 1989). Notamment, dans le domaine de l'aviation, cela permet d'améliorer la détection d'indices mal visibles (Yeh & Wickens, 2001). Les dispositifs tête-haute (Head-up display – HUD) permettent de diriger l'attention vers des indices visuels, notamment lorsqu'ils sont basés sur le principe de « conformité », c'est-à-dire lorsqu'ils lient les éléments de l'interface à l'environnement (Wickens & Long, 1995 ; Caird, Horrey & Edwards, 2001). La Réalité Augmentée (RA) est basée sur ce principe de conformité. Un système de réalité augmentée (RA) est un système qui complète le monde réel avec des éléments virtuels qui semblent coexister dans le même espace (Azuma et al., 2001). Telle qu'envisagée aujourd'hui, la RA en conduite automobile propose d'incruster des éléments virtuels sur un HUD ou sur le pare-brise (Krevelen & Poelman, 2010 ; Narzt, et al., 2003). On distingue deux types de modalité de RA. La première consiste à ajouter des

informations (e.g. une carte, une direction ou un point d'intérêt) dans l'environnement (e.g. Kim & Dey, 2009). La seconde modalité consiste à augmenter des éléments présents dans l'environnement, par exemple un panneau, une ligne ou un piéton (e.g. Narzt & al., 2006). Pour cette seconde modalité, la RA consiste à accroître la visibilité des éléments à forte valeur pour le conducteur, afin d'optimiser son attention visuelle. Des études sur simulateur ont montrés que la RA permet d'améliorer la détection de dangers, comme des piétons ou des panneaux d'avertissement (Rusch et al., 2013). Ces mêmes bénéfices ont été observés pour des conducteurs âgés (Schall et al., 2013 ; Rush et al., 2014). Mais si la RA a le potentiel d'optimiser l'attention visuelle, il subsiste des risques de distraction (Ververs & Wickens, 1998). Par exemple, des éléments en RA peuvent masquer ou détourner l'attention d'autres informations importantes (Schall, Rusch, Lee, Vecera, & Rizzo, 2010).

Les cas d'usages envisagés aujourd'hui pour la RA consistent à mettre en évidence des éléments liés à l'environnement et à la tâche générale de conduite, mais pas à la manœuvre. Dans le modèle de conduite - classiquement utilisé - de Michon (1985), la manœuvre se situe à un niveau hiérarchique intermédiaire : entre le niveau stratégique (planification de l'activité de conduite, gestion de l'itinéraire) et le niveau opérationnel (e.g. contrôle de la pédale de frein). Or, certains auteurs ont montrés que le parcours visuel en activité est très spécifique à la tâche immédiate. Plus précisément, des études sur les parcours oculaires pour des tâches quotidiennes ont montré que les fixations extraient des informations très spécifiques au but de la tâche en cours (Rothkopf, et al., 2007). Lors d'une tâche de préparation de thé par exemple, Land, Mennie, et Rusted (1998) remarquent que les objets de la scène visuelle non pertinents à l'action sont très peu fixés. Reprenant ce constat, Hayhoe, Shrivastava, Myruczek et Pelz (2003) ont étudié le parcours oculaire de participants pour une tâche de préparation de sandwich. Il en ressort que les fixations sont très spécifiques à chaque étape de la procédure pour la réalisation du but, et que seules 2% des fixations concernent des objets de la scène non pertinent (ingrédients ou outils non utilisés). Si l'attention visuelle se concentre sur les éléments pertinents par rapport aux buts de la tâche, les recherches utilisant le paradigme de la cécité attentionnelle montrent également que les autres éléments de la scène tendent à être ignorés malgré leur saillance visuelle (Mack & Rock, 1998; Simons & Chabris, 1999). Dans la même veine, Triesch et al. (2003) ont utilisé le paradigme de changement dans une expérimentation où il était demandé aux participants de classer des briques dans un environnement virtuel. Les participants ont détectées plus souvent le changement lorsqu'il concernait une propriété pertinente pour la réalisation du but.

En définitive, les études que nous venons de citer montrent que les mouvements des yeux sont influencés par le but de la tâche. Le parcours oculaire semble très spécifique au but en cours, et tend à ignorer les objets non pertinents à un moment t. Il paraît alors important qu'un système de RA guide le regard d'une manière qui respecte ce processus de prise d'information, afin de limiter les problèmes de distraction pendant la conduite. Cette recherche vise justement à étudier les modifications d'allocation de l'attention visuelle induites par un système de RA simulé. Pour cela, nous avons mis en place une expérimentation dans laquelle les participants visualisaient des scènes vidéo de conduite automobiles. Au cours des vidéos, une voix indiquait une manœuvre et le participant devait décider si la situation lui permettait de la réaliser. Certains indices de la scène visuelle pouvaient être augmentés graphiquement : soit des indices généraux liés à la conduite, soit des indices liés à une manœuvre en particulier, soit tous les indices, soit aucun indice. De plus, nous avons contrôlé la congruence des indices avec les manœuvres, c'est-à-dire le fait que les indices permettent leurs réalisations ou non. Enfin, nous nous sommes intéressés à deux phases de conduite : une première phase dans laquelle le conducteur roule tout droit et n'a pas de but particulier ; et une deuxième phase dans laquelle le conducteur a l'intention de faire une manœuvre et doit décider si la situation le lui permet.

Dans la condition naturelle (sans RA), on s'attend à ce que l'allocation de l'attention visuelle soit très spécifique à la manœuvre. De manière générale, l'attention visuelle devrait être accrue sur les indices augmentés. La condition dans laquelle les indices spécifiques à la manœuvre sont augmentés devrait induire une allocation de l'attention optimale, et plus orientée vers la manœuvre. A l'inverse, l'augmentation des indices généraux devrait détériorer l'allocation de l'attention visuelle portée sur la scène et détourner l'attention des indices pertinents pour la manœuvre.

2.4.2 Méthode

2.4.2.1 Participants

Quarante-huit participants volontaires ont été recrutés (21 femmes, âge moyen = 34, min = 21, max = 53, SD = 11). Tous avaient le permis de conduire et une expérience de conduite depuis plus de 3 ans (moy = 12, max = 30, SD = 8). Aucun des participants n'avait de problème de vue et ne portait de lunettes ou de lentilles de contact. Les participants ont donné leur autorisation pour l'enregistrement de leurs mouvements oculaires et de leurs réponses, et

ils ont été dédommagés d'une carte cadeau. Les 48 participants ont été repartis aléatoirement dans 4 groupes.

2.4.2.2 Matériel

Vingt séquences vidéo de situations de conduite automobile de 10 secondes ont été utilisées. Les prises de vue ont été réalisées en condition de circulation urbaine, à l'aide d'une caméra embarquée à grand angle fixée sur le pare-brise d'une voiture. A la 7^{ème} seconde de chaque séquence, un enregistrement sonore (à la manière d'un GPS) notifie une manœuvre au participant. La manœuvre est notifiée au cours de la vidéo, et n'interrompt pas son déroulement, comme dans une situation de conduite naturelle guidée par la voix d'un GPS. Les manœuvres peuvent être de quatre types : (i) tourner à gauche/droite (ii) se déporter sur la voie de gauche/droite, (iii) dépasser et (iv) se garer (cf., Figure 12). Chacune des manœuvres peut être congruente ou non avec la situation de la scène visuelle projetée. La congruence d'une manœuvre découle des règles de circulation, par exemple la manœuvre *tourner à gauche* est non congruente s'il y a un panneau *sens interdit* à gauche ; en revanche elle est congruente si aucun indice ne l'empêche.

Sur les 20 séquences présentées à chaque participant, 10 d'entre elles notifient des manœuvres congruentes et les 10 autres notifient des manœuvres non congruentes. Toutes les séquences (congruentes ou incongruentes) sont de difficulté équivalente⁸ : ce sont des situations de conduites normales (sans pièges ni ambiguïtés) pour lesquelles la prise de décision est simple.

Nous avons définis trois types de zones d'intérêt (AOIs) :

- Les *indices généraux* [G] correspondent aux indices liés à la conduite (e.g. feux tricolores, panneaux, lignes, etc.). Ils sont utilisés dans les situations de conduite urbaine comme des indications de base.
- Les *indices liés à la manœuvre* [M] sont définis par les règles de circulation (par exemple la ligne continue interdit le dépassement) ou par la situation (par exemple la présence d'un véhicule sur la voie de droite empêche de se déporter sur cette même voie). Ils peuvent être utilisés pour décider si la manœuvre est réalisable.
- Les *zones de saillance visuelle* ont été définies sur la base de cartes de saillance dynamique, calculées selon la méthode de Belardinelli, Pirri et Carbone (2009) et de

⁸Un contrôle de la difficulté a été effectué. Deux conducteurs experts ont évalué la difficulté (facile – moyen – difficile) d'un ensemble de séquences. Les 20 séquences retenues ont été notées comme « faciles ».

Belardinelli, Carbone et Schneider (2012). C'est une extension du cadre computationnel proposé par Adelson et Bergen (1985) et par Heeger (1987) prenant en compte l'aspect spatio-temporel de la scène visuelle ; c'est-à-dire à la fois les mouvements et leur vitesse. Il y avait 2 à 4 zones d'intérêt de chaque catégorie dans chaque scène, selon les situations présentées.

Sur les vidéos, certains indices ont été *augmentés* graphiquement pour simuler la Réalité Augmentée. Les indices augmentés sont mis en valeur avec des formes semi-opaques superposées sur l'image (Figure 24 ; voir Annexe 5.3 pour plus d'exemples) ; les effets ont été réalisés sur le logiciel Adobe After Effects ®. Les indices sont augmentés pendant la durée de leur présence dans la scène visuelle, c'est-à-dire environ 3s. Dans la première phase des vidéos (Phase 1 - avant la notification de la manœuvre : 0-7s) 1 à 2 indices sont augmentés ; dans la seconde phase (Phase 2 - après la notification de la manœuvre : 7-10s) 2 à 4 indices sont augmentés. Durant la Phase 2, le type et le nombre d'indices augmentés varient selon les conditions expérimentales (Figure 24).



[M⁺G] Seul l'indice lié à la manœuvre est augmenté : la ligne centrale. Cet indice est incongruent avec la manœuvre "se déporter sur la voie de droite".

[MG⁺] Seuls les indices généraux (non liés à la manœuvre) sont augmentés : les panneaux.



[M⁺G⁺] Tous les indices sont augmentés : la ligne centrale (indice lié à la manœuvre) et les panneaux (indices généraux).

[MG] Aucun indice n'est augmenté – condition contrôle.

Figure 24. Exemples de scènes visuelles. Les quatre conditions d'une même scène sont représentées. La manœuvre notifiée aux participants était « se déporter sur la voie de droite ». [M] = Indices liés à la manœuvre ; [G] = Indices généraux. Les indices augmentés graphiquement sont désignés par le symbole ⁽⁺⁾.

2.4.2.3 Dispositif expérimental

La manipulation a été construite sous SMI Experiment Center ®. Nous avons utilisé un oculomètre SMI RED4™ réglé à 50 Hertz avec 9 points de calibration. Les participants étaient placés à 60cm de l'oculomètre et les stimuli étaient affichés sur un écran 17 pouces (résolution 1280x768). Les fixations ont été calculées avec un algorithme de dispersion (logiciel SMI Begaze, version 3.1.77) utilisant les paramètres suivants : 80 ms de durée de fixation minimum et 100 pixels de dispersion maximum. Les AOIs étaient suffisamment grandes pour être limitées aux bordures des éléments.

2.4.2.4 Procédure

La procédure consiste à afficher sur un écran des séquences vidéo de situations de conduite automobile, au cours desquelles était notifiée une manœuvre. A la fin de la séquence, le participant doit indiquer s'il peut réaliser la manœuvre ou pas, selon la situation présentée dans la vidéo. Pour cela, il sélectionne sa réponse sur un formulaire affiché à l'écran (oui/non) ; le temps de réponse nécessaire à la prise de décision est libre.

La passation comporte vingt séquences dont l'ordre de présentation a été randomisé. Entre chaque séquence apparaît un écran gris neutre durant 4s (Figure 25). Pendant les 7 premières secondes de chaque séquence, les participants devaient regarder et se familiariser avec la scène de conduite (Phase 1). La notification audio (annonce de la manœuvre), qui durait environ 1 seconde, était donnée au début de la Phase 2 (phase de prise de décision) car les mouvements des yeux sont susceptibles d'être instantanément spécifiques au but (e.g. pour la notification « tourner à droite », la parcours oculaire peut être spécifique au but dès que le mot « tourner » est entendu). La phase 2 dure trois secondes.

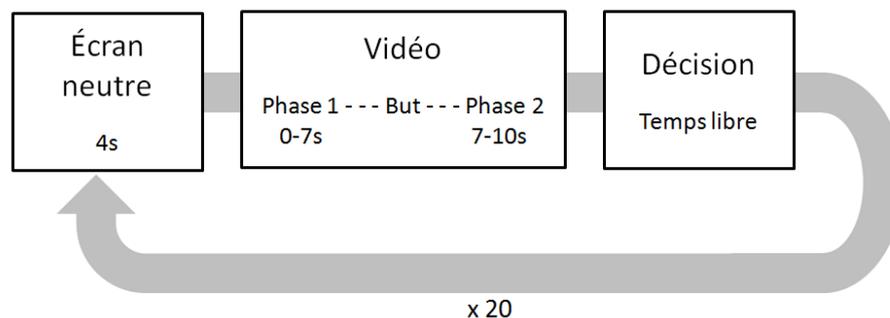


Figure 25. Etapes de la procédure.

2.4.2.5 Plan expérimental

Nous avons mis en place un plan mixte qui combine des facteurs inter-sujets et intra-sujets.

Le facteur principal est le type de *Réalité Augmentée (RA)* dans les vidéos. C'est un facteur inter-sujet à 4 modalités : (1) Seuls les indices liés à la manœuvre sont augmentés [M+G] ; (2) Seuls les indices généraux sont augmentés [MG+] ; (3) Tous les indices sont augmentés [M+G+] ; (4) Aucun indice n'est augmenté [MG] – groupe contrôle. Les quarante-huit participants ont été répartis aléatoirement dans ces quatre conditions (12 participants pour chacune).

Le second facteur est la *congruence entre la manœuvre et la situation*. C'est un facteur intra-sujet à deux modalités : sur la totalité des 20 séquences qu'ils avaient à visionner, les participants ont chacun eu 10 manœuvres congruentes et 10 manœuvres incongruentes avec la situation.

Le troisième facteur est la *phase des séquences*. C'est un facteur intra-sujet à deux modalités : Phase 1 (i.e. avant la notification de manœuvre) et Phase 2 (après la notification de manœuvre).

2.4.2.6 Variables dépendantes

Deux variables dépendantes ont été sélectionnées conformément à la littérature, afin d'évaluer l'impact de la RA sur l'attention visuelle : le pourcentage de temps de fixation sur les AOIs et la fréquence de fixation.

Le *pourcentage de temps de fixation sur les AOIs* est communément utilisé pour rendre compte des variations de l'allocation de l'attention visuelle : il correspond au temps passé par le conducteur à regarder une AOI, par rapport à son temps de fixation total sur toutes les AOIs. Une allocation de l'attention optimale est principalement guidée par des processus *top-down* et le guidage *bottom-up* doit être minimisé (Wickens, 2003), ce qui se doit se traduire par un pourcentage de temps de fixation élevé sur les indices liés à la tâche et par un moindre pourcentage de temps de fixation sur les zones saillantes.

La *fréquence de fixation* correspond au nombre de fixation par seconde, permet de rendre compte de l'optimalité de la prise d'information, c'est-à-dire de l'effort déployé pour atteindre l'information. Une prise d'information optimale se caractérise ainsi par des fixations moins

nombreuses (Goldberg & Kotval, 1998; Rayner, 1998) et spécifiques à la tâche (Shinoda et al., 2001).

2.4.3 Résultats

Afin d'étudier l'impact de la RA sur l'allocation de l'attention visuelle, selon la congruence et le type d'AOI, nous avons effectué une ANOVA mixte 4x2x3, avec un facteur inter-sujets (type de RA) et deux facteurs intra-sujets (congruence et type d'AOI). Le facteur *type de RA* comprend 4 modalités [MG ; M+G ; MG+ ; M+G+]. Le facteur *congruence* a deux modalités (manœuvres congruentes et incongruentes) et trois types d'AOI ont été utilisés (indices liés à la manœuvre, indices généraux, zones de saillance visuelle). Les analyses ont porté sur les données de la phase 2 (après la notification de la manœuvre). L'ANOVA a été réalisée sur les deux variables dépendantes (fréquence de fixation et pourcentage de temps de fixation). Les moyennes sont reportées dans le Tableau 4. Le LSD de Fisher a été utilisé pour les tests post hoc.

Pour analyser l'impact de la RA en fonction des deux Phases des séquences, nous avons effectué une ANOVA mixte 3x2x2, avec un facteur intra-sujet (type de RA) et deux facteurs inter-sujets (Phase et type d'AOI). Concernant le facteur de *type de RA*, seuls trois modalités - dans lesquelles des indices étaient augmentés graphiquement- ont été considérées [M^+G ; MG^+ ; M^+G^+]. Le facteur *Phase* comprend deux modalités : la Phase 1 (avant la notification de la manœuvre) et la Phase 2 (après la notification de la manœuvre). Le troisième facteur de cette analyse est le type d'AOI ; nous avons distingué les indices augmentés des indices non augmentés indépendamment de leurs liens avec la manœuvre (puisqu'elle n'est pas encore notifiée en Phase 1). Seule la fréquence de fixation a été sélectionnée comme variable dépendante pour cette ANOVA. Les moyennes sont reportées dans le Tableau 5.

L'analyse des fréquences de fixation a été réalisée sur les données brutes. Un filtre a été appliqué aux durées de fixation pour éliminer les données extrêmes : les temps supérieurs à deux écarts types ont été remplacés par la moyenne (4,97% des données). Les fréquences de fixation et les pourcentages de temps de fixation ont été pondérés en fonction du taux de couverture des AOIs dans les scènes visuelles (tailles de durées des AOIs).

Mesures	Type de RA	Manœuvres congruentes			Manœuvres incongruentes		
		Indices liés à la manœuvre	Indices généraux	Zones saillantes	Indices liés à la manœuvre	Indices généraux	Zones saillantes
Fréquences de fixation (nombre/s)	RA sur les indices liés à la manœuvre [M ⁺ G]	0,35 (0,09)	0,08(0,04)	0,05(0,04)	0,51 (0,09)	0,02(0,03)	0,06(0,03)
	RA sur les indices généraux [MG ⁺]	0,21 (0,05)	0,20(0,09)	0,07(0,03)	0,34 (0,13)	0,13(0,07)	0,08(0,05)
	RA sur tous les indices [M ⁺ G ⁺]	0,29 (0,12)	0,18(0,08)	0,05(0,04)	0,38 (0,13)	0,09(0,06)	0,08(0,03)
	Pas de RA [MG]	0,17 (0,08)	0,08(0,04)	0,05(0,03)	0,29 (0,11)	0,06(0,06)	0,07(0,04)
Pourcentages de temps de fixation (%)	RA sur les indices liés à la manœuvre [M ⁺ G]	70,6 (13,3)	15,9 (7,7)	7,9 (5,8)	83,6 (7,7)	3,4 (3,8)	8,5 (4,2)
	RA sur les indices généraux [MG ⁺]	40,2 (10,4)	37,1 (12,3)	12,6 (5,4)	59,8 (11,7)	20,8 (9,5)	11,7 (6,4)
	RA sur tous les indices [M ⁺ G ⁺]	50,3(16,6)	33,3 (10,6)	8,6 (5,8)	66,4 (14,7)	13,6 (8,6)	12,6 (6,1)
	Pas de RA [MG]	53,3 (14,2)	25,3 (9,9)	12,6 (9,5)	65,5 (16,1)	13,3 (12,1)	13,4 (8,1)

Tableau 4. Moyennes et écarts types des fréquences de fixations et des pourcentages de temps de fixation (Phase 2).

Mesures	Type de RA	Phase 1		Phase 2	
		Indices augmentés	Indices non augmentés	Indices augmentés	Indices non augmentés
Fréquence de fixation (nombre/s)	RA sur les indices liés à la manœuvre [M ⁺ G]	0,30 (0,09)	0,22 (0,06)	0,78 (0,13)	0,21 (0,09)
	RA sur les indices généraux [MG ⁺]	0,26 (0,11)	0,23 (0,05)	0,32 (0,14)	0,67 (0,18)
	RA sur tous les indices [M ⁺ G ⁺]	0,29 (0,12)	0,20 (0,06)	0,88 (0,25)	0,18 (0,09)

Tableau 5. Moyennes et écarts types des fréquences de fixation sur les indices (Phases 1 et 2).

2.4.3.1 Fréquences de fixations

L'ANOVA 4x2x3 (moyennes dans le Tableau 4) montre un effet principal du type de RA ($F(3, 44) = 10.86, p < .001, \eta^2 = .42$). Les participants en condition de RA ont fait plus de fixations sur les AOIs que ceux de la condition contrôle [MG] ($F(1, 44) = 31.88, p < .001$). En particulier, les participants ont eu des fréquences de fixations plus élevées dans les conditions de RA sur les indices généraux [MG⁺] que dans les trois autres conditions ($F(1, 44) = 5.84, p < .02$).

Nous observons également un effet d'interaction entre le type de RA et le type d'AOI ($F(6, 88) = 11.64, p < .001, \eta^2 = .46$) (Figure 26). Un test post hoc (LSD) nous montre que les

participants en condition de RA font plus de fixations sur les indices liés à la manœuvre que ceux de la condition contrôle (tous les $p < .002$).

Il y a un effet d'interaction entre la congruence et le type d'AOI ($F(2, 88) = 55.22, p < .001, \eta^2 = .57$) (Figure 27). Un test post hoc (LSD) montre que la fréquence de fixation sur les indices liés à la manœuvre est plus importante dans les situations incongruentes ($p < .001$). L'effet est inverse pour les indices généraux : la fréquence de fixation est plus faible dans les situations incongruentes ($p < .001$). Il n'y a pas d'interaction entre la congruence et le type de RA ($F < 1$).

L'ANOVA 3x2x2 (moyennes dans le Tableau 5) montre un effet principal de la Phase ($F(1, 33) = 369.45, p < .001, \eta^2 = .92$) : les participants ont une fréquence de fixation sur les AOIs plus élevée durant la phase 2 que durant la Phase 1. Il y a un effet d'interaction entre la Phase et le type d'AOI ($F(1, 33) = 43.03, p < .001, \eta^2 = .56$) (Figure 28). Un test post hoc (LSD) montre que les participants font plus de fixations sur les indices augmentés dans la Phase 1 que dans la Phase 2 ($p < .001$). En revanche, il n'y a pas de différence de fréquence de fixation sur les indices non augmentés entre les Phases 1 et 2.

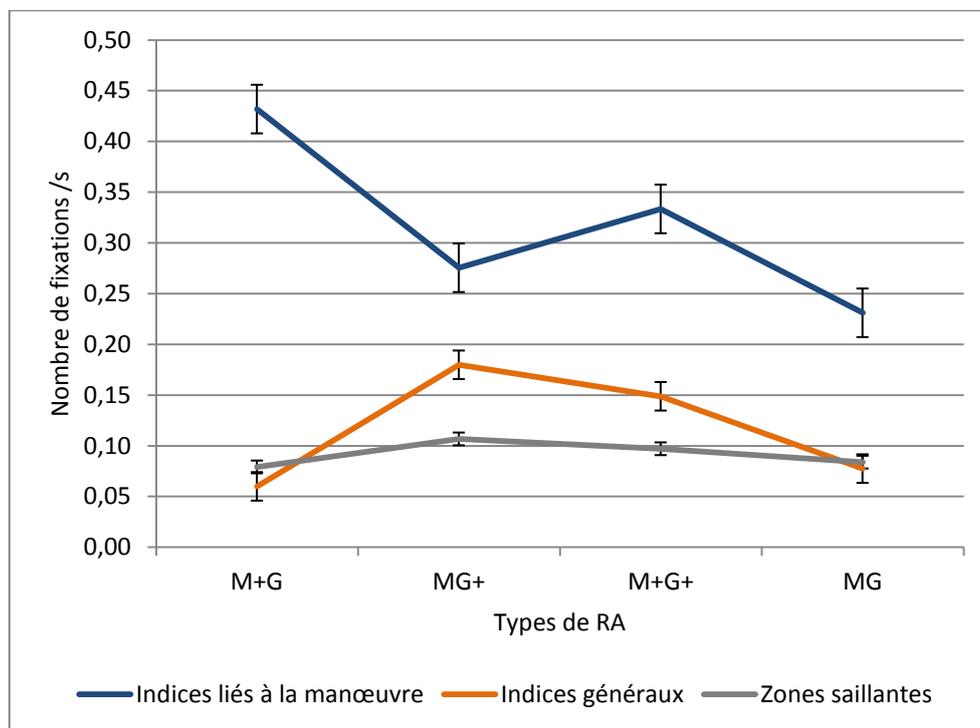


Figure 26. Fréquences de fixations sur les AOIs selon les types de RA.

Types de RA:

[M⁺G] RA sur les indices liés à la manœuvre

[MG⁺] RA sur les indices généraux

[M⁺G⁺] RA sur tous les indices

[MG] Pas de RA (condition contrôle)

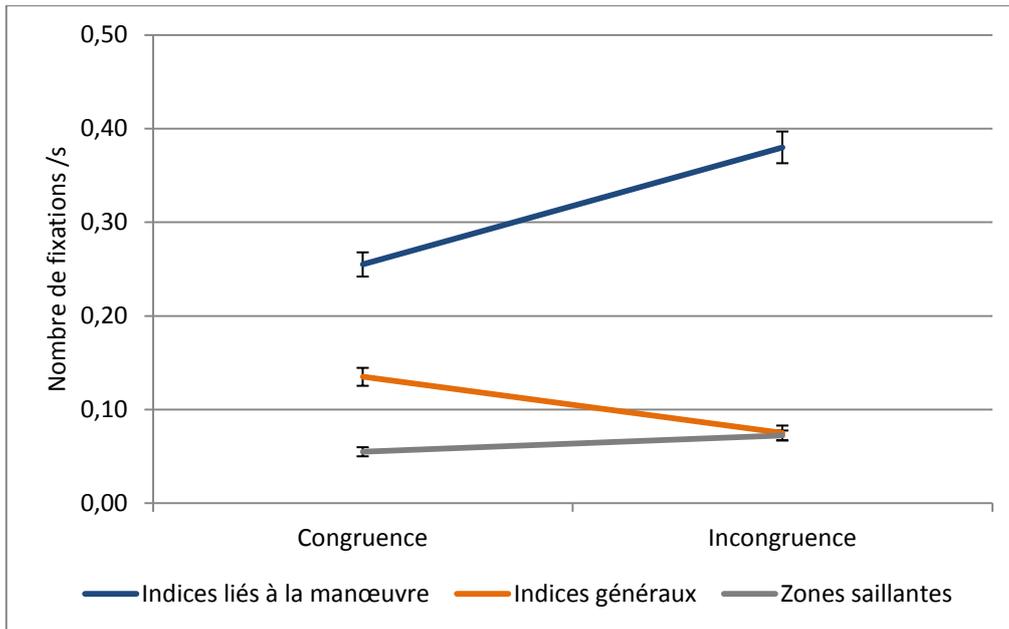


Figure 27. Fréquences de fixations sur les AOIs selon les conditions de congruence.

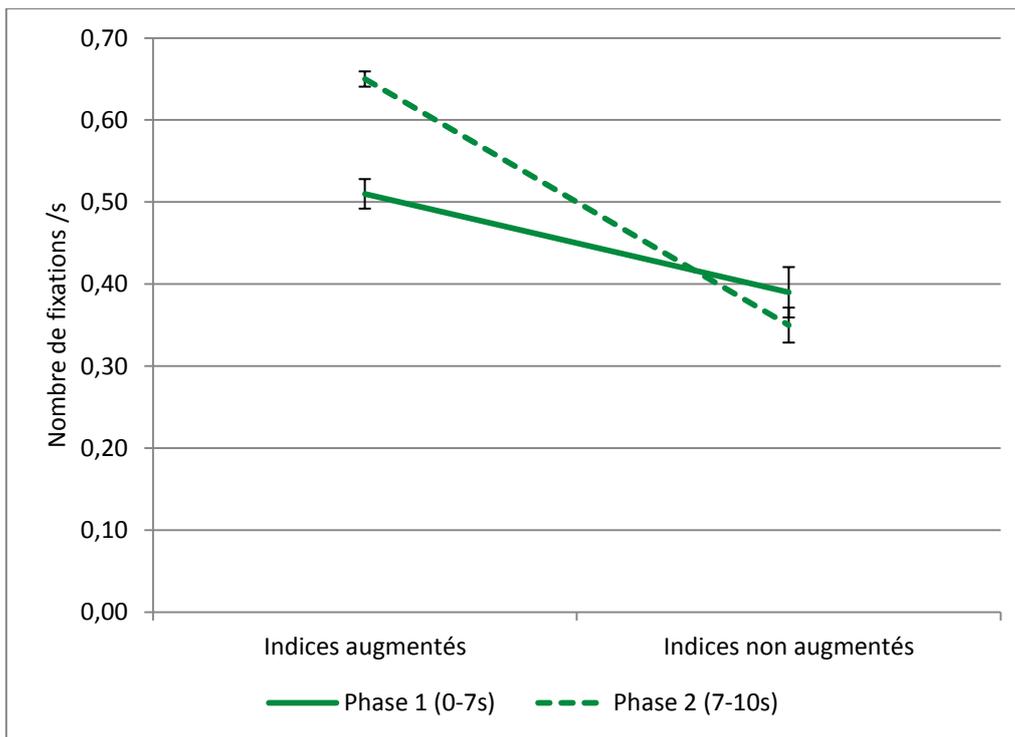


Figure 28. Fréquences de fixations sur les indices selon les 2 phases.

2.4.3.2 Pourcentage de temps de fixation

L'ANOVA 4x2x3 (moyennes dans le Tableau 4) montre un effet principal du type d'AOI ($F(2, 88) = 346.51, p < .001, \eta^2 = .89$) : les indices liés à la manœuvre ont des pourcentages de

temps de fixation plus élevés que les indices généraux ($F(1,44) = 282.89, p<.001$), et les indices généraux ont des pourcentages de temps de fixation plus élevés que les zones saillantes ($F(1, 44) = 48.35, p<.001$).

On relève un effet d'interaction entre le type de RA et le type d'AOI $F(6, 88) = 12.07, p<.001, \eta^2 = .45$ (Figure 29). Quand les indices liés à la manœuvre sont augmentés [M^+G], les pourcentages de temps de fixations sur ces indices sont plus importants qu'en condition contrôle ($F(1, 44) = 15.89, p<.001$). Inversement, lorsque les indices généraux sont augmentés [MG^+] les pourcentages de temps de fixation sur les indices liés à la manœuvre sont inférieurs à ceux de la condition contrôle ($F(1, 44) = 4.51, p<.05$). Lorsque tous les indices sont augmentés [M^+G^+], les pourcentages de temps de fixation sur les indices liés à la manœuvre ne diffèrent pas de la condition contrôle ($F<1$).

Lorsque les indices liés à la manœuvre sont augmentés [M^+G] il y a une diminution des pourcentages de temps de fixation sur les zones saillantes par rapport à la condition contrôle [MG] ($F(1, 44) = 4.87, p<.05$). Dans les deux autres conditions [MG^+ et M^+G^+] les pourcentages de temps de fixation sur les zones saillantes ne diffèrent pas de ceux de la condition contrôle ($F<1$).

Enfin, on relève un effet principal de la congruence ($F(1, 44) = 5.03, p<.05, \eta^2 = .10$) : le pourcentage de temps de fixation sur les indices est légèrement supérieur dans les situations incongruentes. Il n'y a pas d'effet d'interaction entre la congruence et le type de RA ($F<1$).

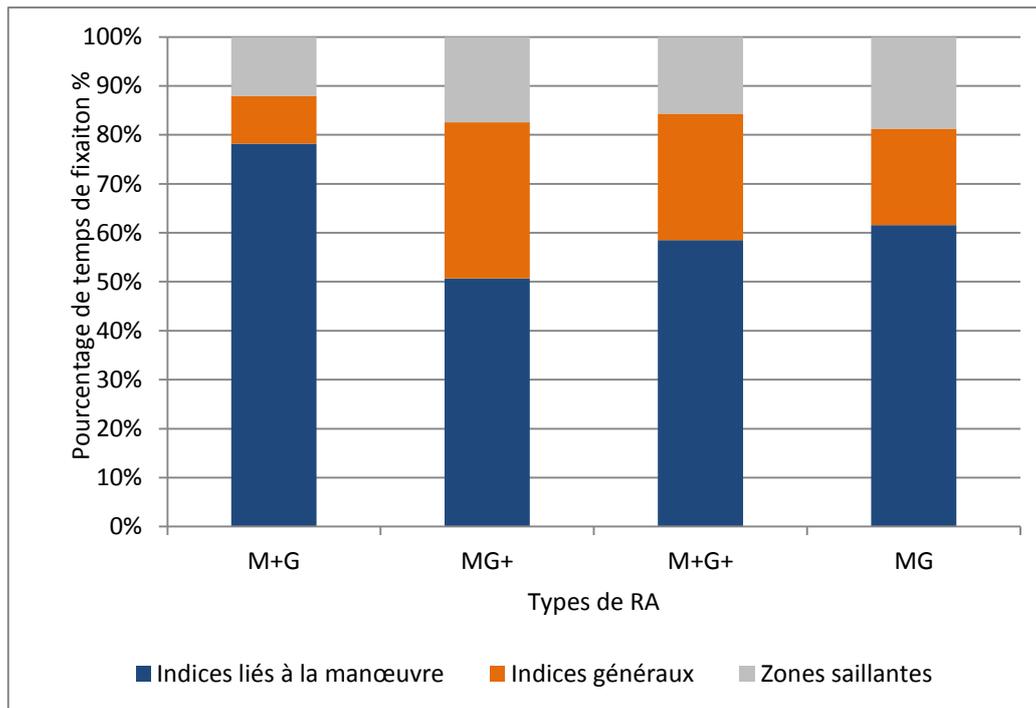


Figure 29. Pourcentages de temps de fixation sur les AOIs selon les types de RA.

Types de RA:

[M⁺G] RA sur les indices liés à la manœuvre

[MG⁺] RA sur les indices généraux

[M⁺G⁺] RA sur tous les indices

[MG] Pas de RA (condition contrôle)

2.4.4 Discussion

Nous avons mis en place une expérimentation dans laquelle des participants visualisaient des vidéos de situations de conduite automobile, au cours desquelles ils devaient prendre des décisions. Nous avons 4 conditions expérimentales : les indices liés à la manœuvre étaient augmentés [M+G], les indices généraux étaient augmentés [MG+], tous les indices étaient augmentés [M+G+], et aucun indice n'était augmentés (condition contrôle) [MG]. Les résultats montrent que la RA a un impact sur l'allocation de l'attention visuelle uniquement lors d'une phase de prise de décision. L'augmentation graphique d'indices généraux n'affecte pas la détection des indices nécessaires à la réalisation de la manœuvre, mais elle altère l'allocation de l'attention visuelle: les fixations sont plus nombreuses et moins spécifiques à la manœuvre. A l'inverse, l'augmentation graphique d'indices liés à la manœuvre induit une allocation de l'attention visuelle optimale: les fixations sont très spécifiques à la manœuvre et les proportions de temps de fixations sur les zones saillantes diminuent.

Dans la formule de prédiction du parcours oculaire du modèle SEEV (Wickens et al. 2003), la Valeur d'un indice est exprimée par le produit de la pertinence de l'indice et de la

priorité de la tâche associée. La priorité correspond ici à l'importance d'une tâche par rapport à une autre ; par exemple la conduite est prioritaire à l'utilisation de l'autoradio (Horrey et al., 2006). Notre étude ne comportait pas de multiples tâches, mais elle envisageait la tâche à deux niveaux hiérarchiques : celui de la navigation (le véhicule roule tout droit et le participant n'a pas de décision à prendre) et celui de la manœuvre. Nos résultats montrent une importante spécificité du parcours oculaire à la manœuvre : en condition contrôle on relève une répartition des temps de fixations d'environ 60% sur les indices liés à la manœuvre, contre seulement 20% sur les indices généraux, et 20% sur les zones saillantes (Figure 29). Ainsi, la valeur d'un élément d'information semble d'autant plus élevée qu'elle est liée à une manœuvre. Ce résultat va dans le sens de ceux de Triesch et al. (2003) et Rothkopf et al. (2007) qui montrent que les fixations extraient des informations très spécifiques au but immédiat de la tâche. Dans les études de Land et al. (1998) et Hayohe (2003), la répartition des fixations est encore plus polarisée vers le but de la tâche, puisque dans leur étude seuls 2% des fixations concernent un des objets de la scène non pertinents pour la réalisation du sous-but immédiat. Les écarts entre les fréquences de fixations observées dans les études précédentes et dans la notre peuvent s'expliquer par la nature statique et non évolutive des environnements étudiés (e.g. table de cuisine et ingrédients dans l'étude de Hayohe, 2003). Lors d'une tâche de conduite automobile, la nature dynamique et changeante de l'environnement routier rajoute à la tâche une contrainte de recherche d'information, ce qui implique pour le participant un tâtonnement et un ajustement du regard (Lemonnier, Bremond & Baccino, 2014).

Les séquences vidéo comportaient deux phases : la première phase avant la notification de la manœuvre (0 à 7s) puis la seconde phase après la notification de la manœuvre (7 à 10s). Nous n'avons pas observé de différence de fréquences de fixations sur les indices non augmentés entre les deux phases. En revanche, les participants ont fait plus de fixations sur les indices augmentés durant la seconde phase. C'est donc lors d'une prise de décision ou d'une manœuvre imminente (phase 2) que les fixations sur les indices se multiplient et que la RA attire l'attention visuelle du conducteur. Ainsi, le conducteur est plus vulnérable à la RA pendant une phase de prise de décision. Durant cette phase critique, un système de RA qui mettrait en valeur des éléments sans liens avec la manœuvre risquerait d'engendrer de la distraction visuelle.

Pour être optimale, l'allocation de l'attention visuelle doit principalement être guidée par les facteurs *top-down* (Attente et Valeur), et les facteurs *bottom-up* doivent être soit

minimisés, soit mis en adéquation avec les éléments à forte valeur (Wickens, 2003). Nous avons fait l'hypothèse qu'une RA spécifique aux manœuvres - donc qui met en saillance des éléments à forte valeur - induirait une allocation de l'attention visuelle optimale. Nos résultats montrent que pour cette condition [M+G] les pourcentages de temps de fixation sur les indices liés à la manœuvre sont plus élevés qu'en condition contrôle (différence de 15% en moyenne) ; en conséquence les pourcentages de temps de fixations sur les indices généraux diminuent. Plus important, on observe une diminution notable du temps de fixations sur les zones saillantes (diminution d'environ 8%). Une RA spécifique à la manœuvre semble donc induire une allocation de l'attention visuelle plus optimale que la condition contrôle, qui se traduit par des visualisations très spécifiques à la manœuvre et une moindre sensibilité aux effets de saillance. Notons que ce bénéfice doit être tempéré, car pour les trois conditions avec RA la prise d'information semble plus difficile qu'en condition contrôle : les fréquences de fixations sont plus importantes. C'est cependant dans la condition de RA spécifique aux manœuvres [M+G] que cette altération est la moins marquée.

Puisque la RA met en valeur les indices augmentés, nous avons prédit qu'une RA non spécifique à la manœuvre dégraderait l'allocation de l'attention visuelle. Nos résultats confirment que l'allocation de l'attention s'en trouve altérée : les pourcentages de temps de fixations sur les indices liés à la manœuvre sont plus bas que pour le groupe contrôle (différence de 10% en moyenne), et en conséquence les indices généraux recueillent des pourcentages de temps de fixations plus élevés. Notons qu'en condition contrôle les indices généraux sont très peu fixés (autant que les zones saillantes), et que seule la RA sur les indices généraux [MG+] pousse les participants à faire des fixations sur ces indices. Egalement, la prise d'information est plus difficile que dans les trois autres conditions : les participants font des fixations plus nombreuses et plus éparpillées entre les trois AOIs. Toutefois, nos résultats montrent également que les participants qui visionnaient des séquences avec RA sur les indices généraux [MG+] ont fait légèrement plus de fixations sur les indices liés à la manœuvre par rapport à ceux du groupe contrôle. Ainsi, même si la RA non spécifique à la manœuvre altère l'allocation de l'attention visuelle, elle ne semble pas entraver la détection des indices utiles.

Dans la condition de RA où tous les indices sont augmentés [M+G+], l'allocation de l'attention est similaire à celle de la condition contrôle (Figure 29). Quand tous les indices sont augmentés, l'effet de la RA est atténué car elle ne met pas en valeur des indices par rapport à d'autres. En revanche, les fréquences de fixations montrent que la prise

d'information est plus difficile que dans la condition contrôle et que dans la condition de RA spécifique à la manœuvre [M+G].

Nous avons contrôlé la congruence des manœuvres avec la situation, c'est-à-dire la possibilité ou pas de réaliser la manœuvre. Les résultats montrent que lorsqu'une manœuvre est incongruente, la fréquence de fixation sur les indices liés à cette manœuvre augmente fortement alors que les indices généraux recueillent très peu de fixations. Une explication à cette focalisation de l'attention serait qu'un seul indice incongruent peut suffire à décider de ne pas réaliser une manœuvre (e.g. il y a un sens interdit à gauche donc le conducteur ne tourne pas à gauche). A l'inverse, le parcours visuel doit être exhaustif pour décider de réaliser la manœuvre (le conducteur vérifie qu'aucun indice ne l'empêche de tourner à gauche). On peut penser que cet effet ne sera pas nécessairement observé en situation réelle de conduite, car lorsqu'un indice est incongruent, le but est immédiatement réévalué, et les indices liés au but initial sont susceptibles de perdre leur pertinence.

Pour conclure, les résultats montrent que la RA impacte l'allocation de l'attention visuelle plus fortement lors d'une phase de prise de décision (Phase 2). A ce moment de l'activité, la RA semble dégrader l'allocation de l'attention visuelle lorsqu'elle est trop généraliste, et l'optimiser lorsqu'elle est spécifique aux manœuvres. Au regard de nos résultats on pourrait donc définir la qualité de la RA comme l'adéquation entre la *valeur* et la *saillance* : un système de RA performant devrait rendre saillant uniquement ou prioritairement les informations immédiatement pertinentes pour la réalisation de la manœuvre. Bien entendu, un système qui met en évidence des éléments critiques (e.g. un danger immédiat) peut être bénéfique même s'il interrompt la tâche du conducteur (Schall et al., 2013; Rush et al., 2014). Mais un système de RA qui met en évidence des éléments critiques devrait également tenir compte des manœuvres. Par exemple si le système met en évidence un piéton qui traverse brusquement la route à 50m (ce qui est une information potentiellement critique), mais que le conducteur doit tourner dans 10m, alors cette mise en évidence n'est pas pertinente, et peut éventuellement perturber sa prise d'information relative à la manœuvre.

Précisons que nos observations se limitent au mode de RA étudié ici, qui consiste à augmenter graphiquement des éléments déjà présents dans l'environnement. Pour ce type de RA, il s'agit d'augmenter la saillance d'indices existants, sans pour autant ajouter de nouvelles informations. D'autres modalités de RA, qui consistent à ajouter des informations à l'environnement (e.g. distance inter-véhicule, points d'intérêts), soulèvent d'autres questionnements quant à la modification de l'allocation de l'attention visuelle, proches de

ceux évoqués dans les études sur l'utilisation de technologies embarquées (e.g. Horrey et al., 2006).

De récentes études explorent les liens entre l'allocation de l'attention visuelle et les comportements routiers (e.g. Werneke & Vollrath, 2012) ; mais les liens entre l'allocation de l'attention visuelle, la RA et les performances de conduite sont encore à explorer. D'autre part, puisque dans cette étude les stimuli étaient des situations de conduite simples et que leurs difficulté a été contrôlée pour ne pas biaiser la tâche de prise de décision, les mesures de performance n'étaient pas pertinentes. Dans une prochaine étude, il serait intéressant d'intégrer le facteur Difficulté afin de mener des analyses de performance. En outre, un système de RA induit une hiérarchisation des informations : certains éléments sont mis en valeur par rapport à d'autres. On peut alors penser que la RA impacte la représentation qu'a le conducteur de la situation, et de futurs travaux pourraient caractériser cet effet ; par exemple déterminer si des informations pertinentes mais non augmentés sont considérées ou ignorées par le conducteur.

Expérimentation III - En résumé	
Objectifs	
<ul style="list-style-type: none"> • Caractériser les variations d'allocation de l'attention visuelle induites par la RA, selon : <ul style="list-style-type: none"> - Le type d'indices augmentés par la RA - La phase de conduite - La congruence du but avec la situation • Déterminer dans quelle mesure la RA peut optimiser ou dégrader l'allocation de l'attention visuelle. 	
Principaux résultats	
<ul style="list-style-type: none"> • La RA a un impact sur l'allocation de l'attention visuelle particulièrement lors d'une phase de prise de décision. • La mise en valeur (par la RA) des indices liés à la manœuvre induit une allocation de l'attention visuelle optimale : les fixations sont plus spécifiques à la manœuvre et les effets de saillance diminuent. • La mise en valeur (par la RA) d'indices généraux n'affecte pas la détection des indices liés à la manœuvre, mais elle altère l'allocation de l'attention visuelle. • Dans toutes les conditions de RA, la prise d'information est plus difficile. 	

3 DISCUSSION GENERALE

3.1 RESUME DU CONTEXTE ET DES ENJEUX

Des nombreux travaux de recherche en psychologie cognitive ont mis en évidence que l'allocation de l'attention découle à la fois de processus ascendants (effets de saillance visuelle) et descendants (buts de la tâche). La RA de surlignage a comme finalité principale d'optimiser l'attention visuelle du conducteur en créant une synergie entre ces deux processus : rendre saillant les éléments pertinents pour la tâche. Mais qu'est-ce qu'un élément pertinent pour une tâche ? Telle qu'envisagée dans les prototypes et les concepts de RA (voir partie 1.1.2), la pertinence est liée à la tâche générale de conduite. Par exemple, un élément peut être surligné parce qu'il n'est pas bien visible (e.g. un panneau ou un marquage au sol la nuit) ou parce qu'il représente un danger potentiel (e.g. un piéton qui traverse). Or, plusieurs études portant sur le parcours oculaire en situation d'activité ont montré que le système visuel se focalise sur des informations très spécifique au but immédiat (Land et al., 1999 ; Hayhoe et al., 2003 ; Triesch et al., 2003 and Rothkopf et al., 2007).

Dans un souci de sécurité, il apparaît donc nécessaire qu'un système de RA guidant l'attention ne rentre pas en conflit avec le processus naturel de prise d'information du conducteur. Dans ce contexte, notre recherche a visé à étudier l'allocation de l'attention visuelle sur des scènes visuelles de conduite automobile, et les modifications d'attention induites par un système de RA.

3.2 L'ALLOCATION DE L'ATTENTION SUR DES SCENES VISUELLES DE CONDUITE AUTOMOBILE

En 1890, William James écrivait que l'attention «*implique le retrait de certaines choses pour mieux traiter d'autres choses.* », et depuis, de nombreux auteurs ont également envisagé l'attention comme un « filtre » d'information (e.g. Broadbent, 1958 ; Deutsch & Deutsch, 1963 ; Treisman, 1980). L'attention, c'est donc la sélection d'une information **aux dépens d'une**

autre information. Lors de la prise de décision en conduite automobile, quels éléments focalisent l'attention, et aux dépens de quels autres éléments ? Etudier l'allocation attentionnelle c'est mettre cette question de la **répartition** au centre du problème. Le premier objectif de cette recherche a donc été de mieux comprendre comment est allouée l'attention visuelle du conducteur parmi les différents indices présents sur une scène visuelle.

Afin d'étudier l'allocation attentionnelle sur des scènes visuelles de conduite automobile, nous avons choisi comme cadre théorique de référence le modèle SSEV (Wickens et al., 2001), qui nous a paru le plus complet parmi ceux proposés jusqu'à maintenant. Ce modèle fait appel à quatre facteurs pour rendre compte de la répartition de l'allocation de l'attention visuelle, dont deux reliés à des processus *top-down* (i.e., la Valeur, l'Attente), et deux à des processus *bottom-up* (i.e., la Saillance et l'Effort). Dans le cadre de notre recherche, nous nous sommes particulièrement intéressés au facteur Valeur.

La Valeur d'une information correspond à l'importance de cette information pour la réalisation d'une tâche. Dans la formule de prédiction du parcours oculaire du modèle SEEV (Wickens et al., 2003), la Valeur d'un indice est exprimée par le produit de la pertinence de l'indice pour une tâche, par la priorité de la tâche associée. Par exemple, lorsque Horrey et al. (2006) ont utilisé cette formule afin de modéliser l'attention visuelle du conducteur utilisant une technologie embarquée, la valeur d'un indice variait selon s'il était relié à la tâche principale (conduite) ou à une tâche secondaire (numérotation).

La notion de Valeur a donc jusqu'à maintenant été définie sur la base d'une distinction entre plusieurs tâches. Mais le modèle SEEV ne permet pas de différencier les valeurs des indices visuels d'un environnement pour la seule tâche de conduite. Par exemple, comment juger la Valeur d'un panneau par rapport à un autre ? Les études qui ont montré que le parcours oculaire est très spécifique au but immédiat lors de la réalisation de tâches quotidiennes (e.g. Land et al., 1999 ; Hayhoe et al., 2003) nous ont amené à prendre en considération le rôle de la manœuvre. Nous avons alors proposé un paradigme expérimental dans lequel sont distingués les indices liés au but immédiat (à la manœuvre), les indices généraux permettant la compréhension de la scène et les zones de saillance visuelle.

Les résultats de nos deux premières expérimentations confirment— cette fois-ci dans le contexte de la conduite automobile — que l'attention visuelle est principalement endogène, c'est à dire qu'elle se focalise sur les éléments liés à la réalisation des buts immédiats (e.g. Land et al., 1999 ; Hayhoe et al., 2003). Cette importante spécificité des parcours oculaires au

but immédiat (c'est-à-dire à la manœuvre) a été observé à la fois sur des scènes visuelles statiques et dynamiques, et s'est exprimée à travers trois indicateurs du parcours oculaire. D'abord les fréquences de fixations : les participants ont fait près de deux fois plus de fixations sur les indices liés au but que sur les autres AOIs. Puis les pourcentages de temps de fixations : près de 60% des fixations sur les AOIs portaient sur des indices permettant de décider si la manœuvre notifiée était réalisable. Enfin, la focalisation de l'attention sur les indices liés à la manœuvre s'est traduite par le degré d'exhaustivité de la prise d'information : les participants ont fait des fixations sur environ la moitié des indices liés à la manœuvre. En toute logique, cette orientation endogène de l'attention se caractérise aussi par des effets de saillance visuelle très marginaux.

Cependant, nos résultats montrent que l'attention visuelle des participants n'était pas entièrement focalisée sur les indices liés au but, contrairement à ce qui a été observé par Hayhoe (2003). Dans leur expérimentation, seules 2% des fixations ne concernaient pas le but immédiat, alors que nous en avons observé au moins 40% (à la fois sur les scènes statiques et dynamiques). Cette différence entre nos résultats et ceux de Hayhoe (2003) peut s'expliquer par la différence des scènes visuelles proposées comme matériel expérimental dans les deux études. En effet, les scènes proposées dans l'étude de Hayhoe (2003) consistaient en un ensemble d'outils de cuisine et d'ingrédients posés sur une table. En comparaison, les scènes de conduite automobile sont visuellement plus complexes et leur compréhension générale requière nécessairement une prise d'information plus étendue, puisque même les indices généraux (qui ne sont pas directement liés à la prise de décision) sont indispensables pour la compréhension de la scène.

Ainsi, tout en confirmant que l'attention visuelle se focalise sur les indices liés au but immédiat (Land et al., 1999 ; Hayhoe et al., 2003), nos résultats montrent que sur des scènes visuelles complexes (scènes dynamiques comportant de nombreux indices), une partie non négligeable des ressources attentionnelles est allouée aux indices permettant une compréhension générale de la scène. Nous verrons que ce résultat permet d'identifier un risque lié à la RA : un système qui focalise l'attention sur les indices liés au but pourrait induire une prise d'information moins étendue, et ainsi perturber la compréhension générale de la scène.

D'autre part, nous avons testé un facteur susceptible de modifier la Valeur d'une information : la congruence de l'indice avec le but. L'incongruence d'un indice risquant de le rendre plus saillant (Conklin & McDonald, 1982 ; Landragin, 2012), nous avons prédit que

l'attention serait davantage focalisée sur les indices incongruents tant dans des scènes statiques que dynamiques. Les résultats ont montré que les indices incongruents avec la manœuvre sont visualisés de manière moins exhaustive lors de l'expérimentation I (scènes statiques) et de manière plus exhaustive lors de l'expérimentation II (scènes dynamiques). La différence entre ces résultats nous suggère que ce ne serait pas tant l'incongruence d'un indice avec le but qui focaliserait l'attention, que son caractère « changeant ». Un indice susceptible de changer (e.g. un piéton qui traverse, un véhicule qui change de voie, un feu qui change de couleur) peut à tout moment devenir congruent ou incongruent avec le but. Ainsi, le conducteur doit maintenir une attention continue sur ces éléments.

En définitive, nos résultats confirment l'importance du facteur Valeur apporté par le modèle SEEV, et contribuent à le spécifier : en conduite automobile la Valeur d'un indice est liée à sa charge sémantique (compréhension de la scène) et elle est d'autant plus élevée que l'indice est lié au but immédiat. Nous avons mis en évidence ces résultats à la fois sur des scènes visuelles statiques et dynamiques, issues de conditions de circulation réelles. Par ailleurs, le conducteur semble détecter les indices « changeants » (et donc susceptibles d'interférer avec le but) et y maintenir une attention continue. Toutefois, cette interprétation nécessiterait d'être confirmée ultérieurement dans le cadre d'une expérimentation dédiée.

Si dans le modèle SEEV, la notion de Valeur est définie sur la base d'une distinction entre plusieurs tâches (indices liés à une tâche prioritaire vs. indices liés à une tâche non prioritaire), nos résultats apportent la possibilité de différencier les Valeurs des indices visuels d'un environnement pour la seule tâche de conduite. Ces résultats devraient permettre d'affiner les modèles prédictifs d'allocation de l'attention visuelle en conduite (e.g. Horrey et al., 2006). Par exemple, on peut envisager de pondérer la Valeur des indices selon (i) leur niveau informationnel (compréhension de la scène), (ii) leur lien avec le but et (iii) leur propension à interférer avec le but (caractère « changeant » de l'indice).

Enfin, d'un point de vue méthodologique, le paradigme expérimental que nous avons élaboré s'est montré efficace pour révéler les variations d'allocation attentionnelle. Si dans le cadre de cette recherche nous l'avons essentiellement utilisé pour caractériser l'impact de la RA sur l'attention visuelle, son usage pourrait être étendu pour étudier l'impact d'autres systèmes d'aide à la conduite.

3.3 RISQUES ET BENEFICES POTENTIELS DE LA REALITE AUGMENTEE DE SURLIGNAGE

Trois types de risques liés à l'intégration de la RA peuvent être mis en évidence, notamment en termes de prise d'information, de distraction, et de distraction en relation aux phases de l'activité de conduite. Cependant, si notre objectif a surtout été d'identifier les risques inhérents à la RA, notre recherche a également permis de montrer que la RA de surlignage a toutefois le potentiel pour optimiser l'allocation attentionnelle. A la suite nous exposons et discutons, à l'égard de nos résultats expérimentaux, les risques potentiels et les bénéfices attendus de l'utilisation de la RA en vue de son utilisation en tant que dispositif d'aide à la conduite.

3.3.1 RA et prise d'information

Nos résultats montrent que dans toutes les conditions de RA, la prise d'information est plus difficile : la fréquence de fixation augmente sur tous les indices (augmentés ou non). Malgré un objectif de facilitation, la RA complexifie la scène visuelle. Puisqu'il est bien connu qu'un surplus d'informations ajoutées dans l'environnement peut gêner la prise d'informations (Ververs & Wickens, 1998 ; Schall, Rusch, Lee, Vecera & Rizzo, 2010), on pourrait penser que la RA de surlignage est moins concernée par ces risques car n'y a pas d'ajout de nouvelles informations, mais uniquement une mise en valeur d'informations existantes dans l'environnement. Or, la RA de surlignage ajoute des méta-informations : des indications sur la pertinence de certaines informations (e.g. en surlignant un panneau, elle indique que c'est une information importante). Ces méta-informations peuvent faire à leur tour l'objet d'un traitement cognitif (e.g. « pourquoi ce panneau est-il mis en valeur ? »). Pour éviter cet alourdissement cognitif - qui va à l'encontre de l'objectif de la RA- il est alors nécessaire que le critère de pertinence soit clairement identifié par le conducteur.

Nous proposons ainsi deux alternatives possibles : (i) le critère de pertinence est unique (e.g. mise en valeur uniquement des éléments liés à un risque de collision) ; (ii) les critères de pertinence sont explicites. Par exemple, on peut imaginer un code couleur : les dangers sont signalés en rouge, les informations mal visibles de nuit sont augmentées en bleu et les informations relatives à l'itinéraire sont augmentées en vert. Mais une telle transparence du

système rendrait-elle la prise d'information plus optimale ou au contraire l'alourdirait davantage ? Le rapport coût/bénéfice pour le conducteur devra alors être évalué.

3.3.2 RA et risques de distraction

Nos résultats ont montré que lorsque les indices généraux sont augmentés, alors l'allocation de l'attention est altérée : les fixations sont moins spécifiques à la manœuvre que lorsqu'il n'y a pas de RA. Le risque d'ignorer une information importante pour la prise de décision est donc accru. Dans ce cas de figure, la RA peut être un facteur de distraction. En revanche, nous avons observé que lorsque la RA met en valeur les indices liés à la manœuvre, l'allocation attentionnelle est optimale : elle se focalise sur ces mêmes indices et les effets de saillance diminuent. Le risque de distraction - c'est-à-dire que le focus attentionnel se déplace vers un élément intrus- est en principe diminué.

Mais même lorsque la RA guide l'attention vers les indices utiles, le risque de distraction demeure. En effet, il a été démontré que des défauts d'attention peuvent être provoqués par des conflits d'attention entre des tâches de différents niveaux hiérarchiques, par exemple entre une tâche de navigation et une tâche d'interaction avec le trafic (Alexander et Lunenfeld, 1986). La RA, en augmentant des indices liés à l'une ou l'autre des tâches peut donc être à l'origine d'un conflit d'attention. Par exemple, si le système met en évidence un panneau directionnel en lien avec l'itinéraire (tâche de navigation), le conducteur peut ignorer un véhicule s'arrêtant devant lui (interaction avec le trafic). Naturellement, dans une situation de danger imminent, un système de RA qui met en saillance l'élément critique peut être bénéfique même s'il détourne l'attention du conducteur de sa tâche initiale (Schall et al., 2013; Rush et al., 2014). Mais même dans ce cas de figure le risque de distraction subsiste. Par exemple, si un système met en évidence un piéton qui traverse brusquement à 50m mais que l'intention du conducteur est de tourner à 10m, alors la mise en évidence n'est plus pertinente et risque de perturber le conducteur dans la prise d'information relative à sa manœuvre. Des conflits d'attention peuvent donc apparaître à tous les niveaux entre les indices augmentés (Figure 30).

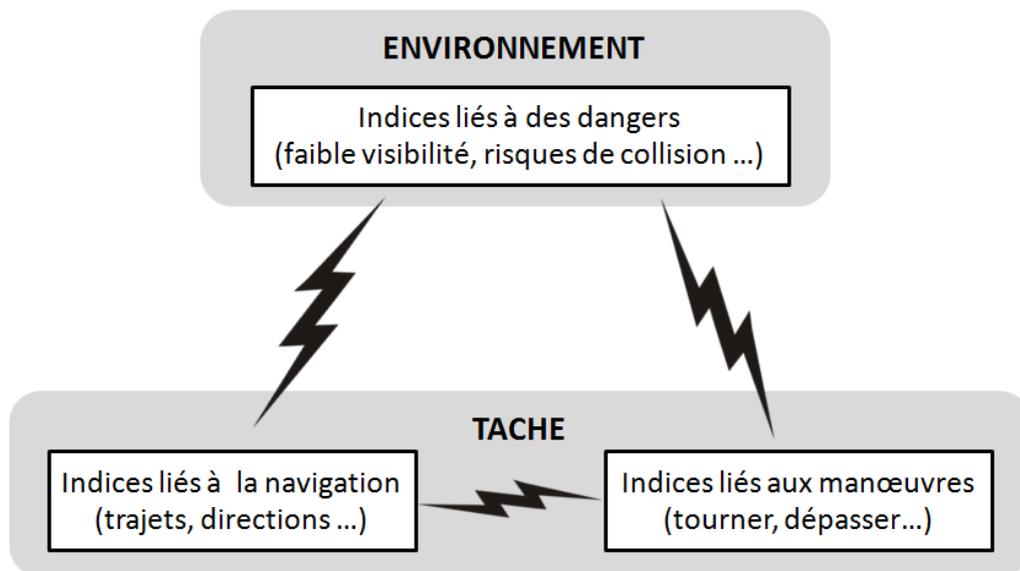


Figure 30. Conflits d'attentions potentiels entre différentes informations augmentées.

Pour limiter le risque de distraction, la RA de surlignage doit donc maintenir un bon niveau d'adéquation entre d'une part la valeur des informations, d'autre part leur mise en saillance. Ce niveau d'adéquation sera d'autant plus compliqué à obtenir que la valeur de l'information sera dépendante des manœuvres. Si le trajet est planifié, certaines manœuvres pourraient être anticipées par le système de guidage (e.g. sur tel trajet le conducteur doit tourner à droite à telle intersection) ; mais d'autres manœuvres ne pourront pas être anticipées (e.g. un dépassement).

Notons que le risque de distraction sera peut-être modulé par l'expertise du conducteur, c'est-à-dire son temps d'expérience de conduite avec de la RA. Dans nos expérimentations, les conducteurs étaient exposés à une nouvelle technologie (simulée). Mais comme pour de nombreuses activités, l'utilisateur expert pourrait s'adapter, c'est à dire faire abstraction d'une RA non pertinente et discriminer facilement les informations importantes à un moment t .

3.3.3 RA et risques de distraction selon les phases de l'activité de conduite

L'activité de conduite se compose d'une alternance de phases « de navigation » et de phases de prise de décision. Durant la phase de navigation, le conducteur ne gère que des tâches de niveau stratégique (e.g. gestion de l'itinéraire) et des tâches de niveau opérationnel (e.g. gestion de la vitesse). C'est principalement le cas lors de la conduite sur autoroute, par exemple. En revanche, lors des phases de prise de décision, le conducteur prépare une

manœuvre. Nos résultats montrent que lorsque le conducteur doit décider s'il peut réaliser un manœuvre, alors les fixations sur les indices se multiplient afin de trouver des informations validant ou invalidant la réalisation de son but. C'est lors de cette recherche visuelle que les indices augmentés attirent le plus fortement le regard. Par conséquent, le risque de distraction est d'autant plus important si un indice non pertinent est augmenté à ce moment de l'activité. Le conducteur est donc plus vulnérable à la distraction dans une phase de prise de décision.

Pour limiter ce risque, on peut imaginer différentes modalités d'adaptation de la RA à la phase de l'activité. Par exemple, la surbrillance qui n'est pas en lien avec une manœuvre peut s'estomper lors d'une phase critique afin de ne pas gêner la prise d'information. Ce type de procédé s'apparenterait aux « contres mesures cognitives » expérimentées dans le domaine aéronautique : des suppressions temporaires d'informations pour obliger le pilote à changer son focus attentionnel (e.g. Dehais, 2011).

3.3.4 RA et optimisation de l'allocation attentionnelle

Si notre objectif a surtout été d'identifier les risques inhérents à la RA, notre étude a également permis de montrer que la RA de surlignage a le potentiel pour optimiser l'allocation attentionnelle, car elle permet (i) d'orienter l'attention vers les éléments nécessaires à la prise de décision et (ii) de réduire la sensibilité aux effets de saillance visuelle. Il s'agit là de premiers résultats qui peuvent plaider en faveur des potentialités de la RA pour améliorer la sécurité en conduite automobile (e.g. détection de dangers, augmentation de la visibilité, guidage). Quelques recherches récentes ont d'ailleurs commencé à évaluer ces bénéfiques potentiels de la RA sur l'amélioration du temps de détection d'éléments critiques sur le bord de la route (Rusch et al., 2013), et notamment pour des conducteurs âgés (Schall et al., 2013; Rush et al., 2014). Dans ces études réalisées sur simulateur, les scènes étaient limitées dans leur complexité visuelle : les images de synthèse comportaient très peu d'indices. Nos résultats montrent que la RA permet de guider le regard dans une scène présentant une complexité visuelle proche de la réalité. De plus, nous avons montré que la RA permet de limiter les effets de saillance inhérents à ce type de scène (qui comporte de nombreux éléments en mouvements).

La prochaine étape consistera à vérifier que cette amélioration de l'allocation attentionnelle se traduise par une amélioration des performances de conduite. Dans cette optique, il sera indispensable d'effectuer des mesures de performances, par exemple sur simulateur. Par ailleurs, afin que la RA de surlignage puisse avoir un effet positif sur

l'attention visuelle, sa mise en œuvre devra probablement tenir compte des trois types de risques que nous venons d'identifier.

3.4 CONCLUSION ET PERSPECTIVES

Pour ce travail de recherche, nous nous sommes focalisés sur une dimension de l'activité de conduite automobile : l'allocation de l'attention visuelle. De cette manière nous avons pu identifier plus précisément les risques de distraction inhérents à la RA de surlignage. Les prolongements de nos travaux pourront s'étendre à la suite du processus : l'élaboration par le conducteur de son modèle de situation. En effet, la RA de surlignage opère une hiérarchisation de l'information qui impactera très certainement la représentation qu'a le conducteur de la situation. Dans cette optique, les travaux sur l'analyse et la mesure de la conscience de la situation (e.g. Endsley, 1988, 1995) offrent des perspectives tout à fait intéressantes.

Par ailleurs, l'activité cognitive du conducteur doit être envisagée conjointement avec sa mise en œuvre pratique. L'enjeu sous-jacent de cette recherche étant évidemment la sécurité, la suite de ces travaux devra prendre en compte la notion de performance de conduite, par exemple sur simulateur.

L'utilisation d'un simulateur permettrait en outre de prendre en compte l'aspect multitâches de l'activité de conduite pour l'analyse de l'allocation de l'attention visuelle, à l'instar de Horrey et al. (2006) qui distingue la tâche de conduite d'une tâche de numérotation sur téléphone ; ou de Lemonnier, Brémond et Baccino (2014, 2015) qui distinguent la tâche de contrôle du véhicule de la tâche d'interaction avec les usagers de la route. Cela permettrait d'étudier l'interaction de la RA avec la gestion simultanée de plusieurs tâches, qui est une caractéristique inhérente à l'activité de conduite.

D'autre part, la RA supprime une partie des processus *top-down* en présélectionnant les informations à forte valeur. Ainsi, comme l'ont démontré nos résultats, elle semble paradoxalement induire une prise d'information nettement plus guidée par un facteur *bottom-up*, puisque le conducteur oriente son attention vers les éléments mis en saillance par la RA. Ce changement risque-t-il de placer le conducteur dans une position de passivité, et ainsi de le désengager de son activité ? Ce risque, qui se rapproche de ceux identifiés dans les recherches

sur les effets pervers des technologies d'aide à la conduite (e.g. régulateur de vitesse, circulation autonome) mérite d'être pris en considération.

Enfin, nous avons montré que la RA de surlignage augmente le risque de conflits attentionnels, notamment lors des phases de recherche visuelle active, précédant une prise de décision. Pour limiter ce risque, un tel système devra prendre en compte finement l'activité du conducteur - ce qui constituera sans doute un obstacle technologique de taille. Les différentes modalités d'adaptation de la RA à l'activité du conducteur (aux tâches, aux phases de l'activité) sont autant de nouvelles pistes de recherche et d'expérimentation.

4 BIBLIOGRAPHIE

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2), 284–299.

Alexander, G. J. & Lunenfeld, H. (1986). Driver Expectancy in Highway Design and Traffic Operations. U.S. Department of Transportation, Federal Highway Administration, Washington, D.C.

Amalberti, R. (2001). La maîtrise des situations dynamiques. *Psychologie Française*, 46, 105-117.

Azuma, R., Baillet, Y., Behringer, R., Feiner, S., Julier, S. & MacIntyre, B. (2001). Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications*, 21(6), 34-47.

Bajcsy, R. (1988). Active perception. *Proceedings of the Institute of Electrical and Electronic Engineers*, 76, 996-1005.

Ballard, D., Hayhoe, M., & Pelz, J. (1995). Memory representations in natural tasks. *Cognitive Neuroscience*, 7, 66-80.

Ballard, D., Hayhoe, M., Pook, P., & Rao, R. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20, 723-767.

Ballard, D.H. (1991). Animate vision: An evolutionary step in computational vision. *J. of the Inst. of Electronics, Information, and Communication Engineers*, 74, 343-348.

Belardinelli, A., Carbone, A., & Schneider, W. X. (2012). Classification of multiscale spatiotemporal energy features for video segmentation and dynamic objects prioritization. *Pattern Recognition Letters*, 34(7), 713-722.

Belardinelli, A., Pirri, F., & Carbone, A. (2009). Motion Saliency Maps from Spatiotemporal Filtering. *Attention in Cognitive Systems* (pp. 112–123).

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14, 199-208.

Boloix, E. (2006). *Influence de la tâche sur le traitement des scènes visuelles complexes : une modélisation des niveaux de représentation*. Thèse. Université de Provence.

Boloix, E. (2007). Attention et perception de scènes visuelles. *L'année psychologique*, 107, p.113-151.

Broadbent, D. (1958). *Perception and communication*. London : Pergamon Press.

Brooks, R.A. (1986). A robust layered control system for a mobile robot. *IEEE J. Robotics and Automation*, 2, 14- 22.

Caird, J. K., Horrey, W. J., & Edwards, C. J. (2001). Effects of conformal and non-conformal vision enhancement systems on older driver performance. *Transportation Research Record*, 17(59), 38–45.

Carbonell, J. R. (1966). A queuing model of many-instrument visual sampling. *IEEE Transactions on Human Factors in Electronics*, 7, 157–164.

Chun, M.M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, 4, 170-178.

Conklin E. J., McDonald D. D. (1982). Saliency: The Key to the Selection Problem in Natural Language Generation, In *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics*, Toronto, pp. 129-135.

Dehais, F., Causse, M., Tremblay, S. (2011). Mitigation of Conflicts with Automation: Use of Cognitive Countermeasures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53, 448-460.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, 87, 272–300.

Edmonds P. G. (1993). *A Computational Model of Collaboration on Reference in Direction-Giving Dialogues*, Ms. Thesis, University of Toronto.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *In Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97–101). Santa Monica, CA: Human Factors Society.

Endsley. (1995). Measurement of situation awareness in dynamic system. *Human factors*, 37(1) 65-84.

Fisher, D. L., & Tan, K. C. (1989). Visual displays: The highlighting paradox. *Human Factors*, 31(1), 17–30.

Fitts, P. M., Jones, R.E. & Milton, J.L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review* , 9(2), 24-29.

George, P., Thouvenin, I., Fremont, V. & Cherfaoui, V. (2011). Réalité augmentée pour l'aide à la conduite intégrant l'observation du conducteur, *6èmes Journées de l'AFRV*, Biarritz.

Goldberg, J. H. & Kotval, X. P. (1998). Eye movement-based evaluation of the computer interface. In S.K.Kumar (Ed.), *Advances in Occupational Ergonomics and Safety* (pp. 529-532). Amsterdam: ISO Press.

Hayhoe, M., Shrivastava, A., Myruczek, R., & Pelz, J. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, 3, 49-63.

Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 4(8), 1455–1471.

Hendricks, D. L., Fell, J. C., & Freedman, M. (2001). *The Relative Frequency of Unsafe Driving Acts in Serious Traffic Crashes*. Washington (DC) : National Highway Traffic Safety Administration.

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(2), 67-78.

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

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. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, 2, 194-203.

James, W. (1890). *The principles of psychology*, New York: Holt, pp. 403-404.

Jonides, J. (1980). Voluntary versus automatic control over the mind's eye's movement. In: Long, JB.; Baddeley, AD., editors. *Attention and performance IX* (pp. 187-203). Erlbaum; Hillsdale.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, 43, 346-354.

Kim, S., & Dey, A. K. (2009). Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. *Proceedings of the 27th international conference on Human factors in computing system* (pp. 133-142). Boston, MA, USA.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4, 219 – 227.

Krevelen, D. W. F., & Poelman, R. (2010). A Survey of Augmented Reality Technologies, Applications and Limitations. *International Journal of Virtual Reality*, 9(2), 1-20.

Kvalseth, T. (1977). Human information processing in visual sampling. *Ergonomics*, 21, 439-454.

Lagarde, E., Chastang, J. F., Gueguen, A., Cœuret-Pellicer, M., Chiron, M., Lafont, S. (2004). Emotional stress and traffic accidents. The impact of separation and divorce. *Epidemiology*, 15, 762-766.

Land, M. F., Mennie, N., Rusted, J. (1998). Eye movements and the roles of vision in activities of daily living: making a cup of tea. *Investigative Ophthalmology & Visual Science* 39 (4) S457.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, 369, 742-744.

Land, M. F., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, 28, 1311-1328.

Landragin, F. (2012). Saillance physique et saillance cognitive. *Cognition, Représentation, Langage (CORELA)* 2(2).

Lee, D. L., & Strayer, D. L. (2004). Preface to special section on driver distraction. *Human Factors*, 46 (4), 583-586.

Leibowitz, H. W., & Post, R. M. (1982). The two modes of processing concept and some implications. In J. Beck (Ed.), *Organization and representation in perception* (pp. 343–363). Hillsdale, NJ: Erlbaum.

Lemercier, C. & Cellier, J-M. (2008). Les défauts de l'attention en conduite automobile: inattention, distraction, et interférence. *Le Travail Humain*, 71, 3, 271-296.

Lemonnier, S., Brémond, R., & Baccino, T. (2014). Discriminating cognitive processes with eye movements in a decision-making driving task. *Journal of Eye Movement Research*, 7(4), 1-15.

Lemonnier, S., Brémond, R., & Baccino, T. (2015). 35, 60-74.

Mack, A. & Rock, I. (1998). Inattention blindness. Cambridge, MA: MIT Press.

Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception and Psychophysics*, 2, 574-552.

McGehee, D. V. (2001). New design guidelines aim to reduce driver distraction. *Human Factors and Ergonomics Society Bulletin*, 44(10), 1–3.

McKee, S. P., & Nakayama, K. (1984). The detection of motion in the peripheral visual field. *Vision Research*, 24, 25–32.

McLeod, P., Driver, J., & Crisp, J. (1988). Visual Search for a conjunction of movement and form is parallel. *Nature*, 332,154–155.

Michael, G. A., Garcia, S., Fernandez, D. N., Sellal, F., & Boucart, M. (2006). The ventral premotor cortex (vPM) and resistance to interference. *Behavioral Neuroscience*, 120(2), 447-462.

Michon, J.A. (1985). A critical review of driver behavior models: what do we know, what should we do? In E. L. R.G., *Human Behaviour and Traffic Safety* (pp. 487-525). Plenum Press.

Most, S. B., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattention blindness and the capture of awareness. *Psychological Review*, *112*, 217-242.

Mourant, R. R. & Rockwell, T. H. (1970). Mapping eye-movement patterns to the visual scene in driving: An exploratory study. *Human Factors* *12*, 81-87.

Narzt, W., Pomberger, G., Ferscha, A., Kolb, D., Müller, R., Wieghardt, J., Hörtner, H. & Lindinger, C.(2004). A new visualization concept for navigation systems . Berlin, ALLEMAGNE, Springer.

Narzt, W., Pomberger, G., Ferscha, A., Kolb, D., Müller, R., Wieghardt, J., Hörtner, H., & Lindinger, C. (2006). Augmented Reality Navigation Systems; In: C. Stephanidis (ed.): *Universal Access in the Information Society* (pp. 177-187). Volume 4, Issue 3, Berlin Heidelberg: Springer.

Narzt, W., Pomberger, G., Ferscha, A., Kolb, D., Müller, R., Wieghardt, J., Hörtner, H., & Lindinger, C. (2003). Pervasive Information Acquisition for Mobile AR-Navigation Systems; In: *5th IEEE Workshop on Mobile Computing Systems & Applications* (pp. 13-20). Monterey, California, USA.

Navalpakkam V.; Itti., L.. (2005). Modeling the influence of task on attention. *Vision Research*, *45* (2), pp. 205–231.

Neboit, M. (1980). *L'exploration visuelle dans l'apprentissage de tâches complexes: l'exemple de la conduite automobile*. Ecole pratique des Hautes Etudes, Paris.

Noy, Y. I., Lemoine, T. L., Klachan, C., & Burns, P. C. (2003). Task interruptability and duration as measures of visual distraction. *Applied Ergonomics*, 207-215.

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Pictures change during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, *7*, 191-211.

Parkhurst, D. J., & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision*, *16* , 125–154.

Peterson, M. S., & Kramer, A. F. (2001). Contextual cueing reduces interference from task-irrelevant onset distractors. *Visual Cognition*, *8*, 843 859.

Posner, M. (1980). Orienting of Attention, *Exp Psychol*, *32*: 3-25.

Posner, M. I. (1988). Structures and functions of selective attention. *Clinical Neuropsychology and Brain Function: Research Measurement and Practice*, 173-202.

Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin*, 124, 123–164.

Rensink, R.A. (1997). How much of a scene is seen ? The Role of attention in scene perception. *Investigative Ophthalmology and Visual Science*, 38, S707.

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

Rusch, M. L., Schall, M. C., Gavin, P., Lee, J. D., Dawson, J. D., Vecera, S. & Rizzo. (2013). Directing driver attention with augmented reality cues. *Transportation Research Part F : Traffic Psychology and Behaviour*, 16(127).

Rusch, M. L., Schall, M. C., Lee, J. D., Dawson, J. D. & Rizzo, M. (2014). Augmented reality cues to assist older drivers with gap estimation for left-turns. *Accident Analysis and Prevention*, 71, 210-21.

Saad, F., Delhomme, P, Van Elslande, P., Gauje, T., Nachtergaele, C., & Picard-Dore, J. (1992). Analyse des comportements en situation réelle de conduite: le franchissement d'intersections. Paris: INRETS, 158.

Schall, M. C., Rusch, M. L., Lee, J. D., Dawson, J. D., Thomas, G., Aksan, N., & Rizzo, M. (2013). Augmented reality cues and elderly driver hazard perception. *Human Factors*. 55(3), 643-58.

Schall, M. C., Rusch, M., Lee, J., Vecera, S., & Rizzo, M. (2010). Attraction without distraction: Effects of augmented reality cues on driver hazard perception. *Journal of Vision*, 10(7), 236-236.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5, 195-200.

Sheridan, T. (1970). On how often the supervisor should sample. *IEEE Transactions on Systems Science and Cybernetics*, 6(2), 140-145.

Shinar, D. (2008). Looks are (almost) everything: where drivers look to get information. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50, 380-384.

Shinoda, H., Hayhoe, M. M., & Shrivastava, A. (2001). What controls attention in natural environments. *Vision Research*, 41, 3535-3546.

Shinoda, H., Hayhoe, M., & Shrivastava, A. (2001). What controls attention in natural environments. *Vision Research*, 41, 3535-3546.

Simons, D. J. & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentive blindness for dynamic events. *Perception*, 28, 1059-1074.

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

Theeuwes, J. (1994). Stimulus -driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception & Performance*, 20, 799- 806.

Theeuwes, J. (2005). Irrelevant Singletons Capture Attention. In Itti, L. & Rees, G. & Tsotsos, J. (Eds.) *Neurobiology of Attention*, pp. 418- 424. San Diego, CA: Elsevier

Torralba, A., & Oliva, A. (2003). Statistics of natural images categories. *Computation in Neural Systems*, 14, 391-412.

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society, London B*, 353(1373),1295–1306.

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

Triesch, J., Ballard, D., Hayhoe, M., & Sullivan, M. (2003). What you see is what you need. *Journal of Vision*, 3, 86-94.

Ullman, S. (1984). Visual routines. *Cognition*, 18, 97-159.

Van Krevelen, & Poelman, R. (2010). *International Journal of Virtual Reality*, 9 (2), 1.

Ververs, P. M. & Wickens, C. D. (1998). Head-up Displays: Effects of clutter, display intensity, and display location on pilot performance. *International Journal of Aviation Psychology*, 8 (4), 377-403.

Wallis, G., & Bühlhoff, H. H. (2000). What's scene and not seen: influences of movement and task upon what we see. *Visual Cognition*, 7, 175-190.

Werneke, J. & Vollrath, M. (2012). What does the driver look at? The influence of intersection characteristics on attention allocation and driving behavior. *Accident Analysis and Prevention*, 45, 610 - 619.

Wickens C. D. (2003). Pilot Actions and Tasks: Selections, Execution, and Control. In Tsang P. S., Vidulich M. A. (Eds.), *Principles and practice of aviation psychology* (pp. 239–263). Mahwah, N.J: Lawrence Erlbaum.

Wickens, C. D. (1997). Attentional issues in head-up displays. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics: Transportation systems*, 1, 3–21.

Wickens, C. D. (2007). Attention to attention and its applications: A concluding view. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice*. New York: Oxford University Press.

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(3), 179–193.

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(3), 360-380.

Wickens, C. D., Helleberg, J., Xu, X., & Horrey, W. J. (2001). Pilot task management: testing an attentional expected value model of visual scanning. In Tech. Rep. ARL-01-14/NASA-01-7. Aviation Research Lab.

Wickens, C.D. (1984). Processing resources in attention. Dans Parasuraman, R. et Davies, D.R. (Eds), *Varieties of attention*. New York : Academic Press.

Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 661-676.

Yarbus, A. L. (1967). *Eye movements and vision*. Springe.

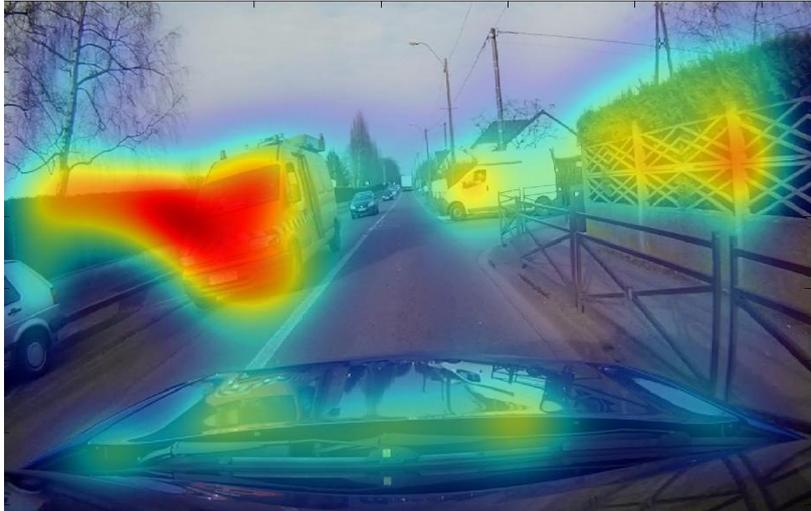
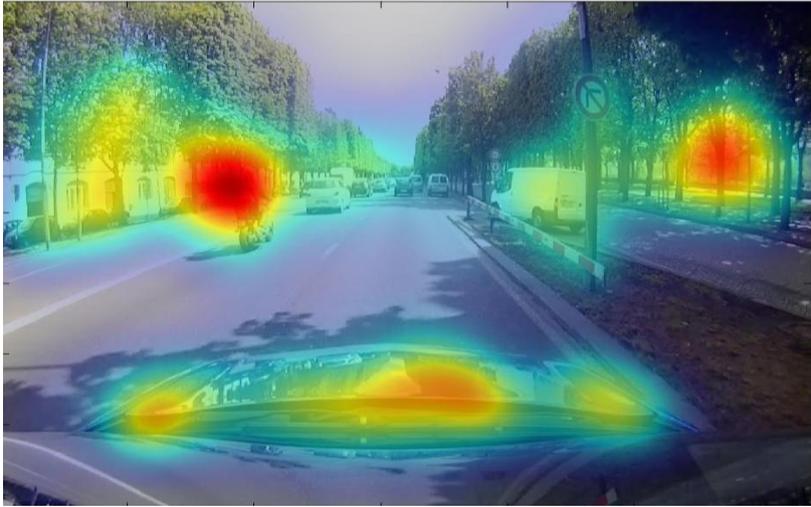
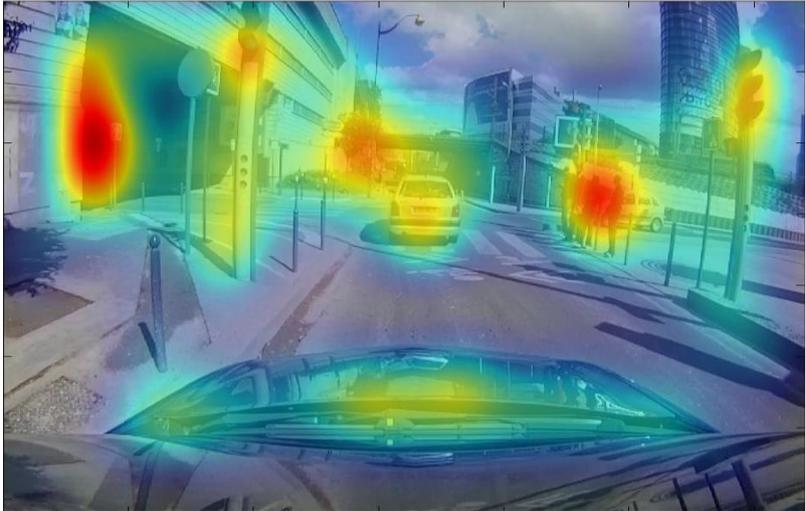
Yeh, M., & Wickens, C. D. (2001). Display signaling in augmented reality: Effects of cue reliability and image realism on attention allocation and trust calibration. *Human Factors*, 43(3), 355–365.

Young, K. & Regan, M. (2007). Driver distraction: A review of the literature. In: I.J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter & J.D. Irwin (Eds.). *Distracted driving*. Sydney, NSW: Australian College of Road Safety. 379-405.

5 ANNEXES

5.1 Exemples de cartes de saillance statiques (Expérimentation I)

Les cartes de saillance statiques ont été calculées selon la méthode de Belardinelli, Pirri et Carbone (2009).



5.2 Exemples de cartes de saillance dynamiques (Expérimentations II et III)

Les cartes de saillance dynamiques ont été calculées selon la méthode de Belardinelli, Pirri et Carbone (2009) et Belardinelli, Carbone et Schneider (2012).



5.3 Exemples de scènes dynamiques avec RA (Expérimentation III)

