PolyPublie
Polytechnique Montréal

POLYTECHNIQUE MONTRÉAL
UNIVERSITÉ D'INGÉNIERIE

Titre: Title:	Méthodologie de manipulation de l'immersion en environnement virtuel et ses effets sur des indicateurs de ressources attentionnelles
Auteur: Author:	Maxence Hébert-Lavoie
Date:	2023
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Hébert-Lavoie, M. (2023). Méthodologie de manipulation de l'immersion en environnement virtuel et ses effets sur des indicateurs de ressources attentionnelles [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/53388/

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/53388/
Directeurs de recherche: Advisors:	Philippe Doyon-Poulin, & Benoît Ozell
Programme: Program:	Doctorat en génie industriel

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal

https://publications.polymtl.ca

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Méthodologie de manipulation de l'immersion en environnement virtuel et ses effets sur des indicateurs de ressources attentionnelles

MAXENCE HÉBERT-LAVOIE
Département de mathématiques et de génie industriel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*
Génie industriel

Mai 2023

© Maxence Hébert-Lavoie, 2023.

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Cette thèse intitulée :

Méthodologie de manipulation de l'immersion en environnement virtuel et ses effets sur des indicateurs de ressources attentionnelles

présentée par **Maxence HÉBERT-LAVOIE**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Bruno AGARD, président
Philippe DOYON-POULIN, membre et directeur de recherche
Benoît OZELL, membre et codirecteur de recherche
Thomas HURTUT, membre
Stéphane BOUCHARD, membre externe

DÉDICACE

J'ai un système nerveux et un système confiant.
Apprendre ou à laisser...

REMERCIEMENTS

À ces personnes qui ont été et qui sont **présentes**. À ces personnes qui m'ont aidé, encouragé, enduré. À ces personnes qui se reconnaissent et aux autres qui le devraient, je vous dis merci. L'impression qu'on croyait en moi m'a donné la conviction que j'étais en mesure d'accomplir cette épreuve. Oui ! À vous toutes et tous et toutes et tous et toutes je dis des milliards de fois merci, preuve à l'appui : $Merci^{1000}$.

Un merci tout spécial à Benoit Ouellette et Jean-Marc Robert qui m'ont introduit à ce projet. À Bombardier, Marinvent et BinariesLid qui l'ont porté à ses débuts. À CRIAQ, CARIC et MITACS pour l'aide financière.

Et bien sûr beaucoup de gratitude, de la reconnaissance et un très grand merci à Philippe Doyon-Poulin et Benoît Ozell pour votre soutien et vos précieux conseils tout au long de mon cheminement.

RÉSUMÉ

Cette thèse étudie les méthodologies de mesure du concept de sentiment de présence en environnement virtuel. Elle vise à approfondir les connaissances liées à l'évaluation de ce concept.

Le concept de sentiment de présence peut être défini comme le réalisme perçu d'une expérience faisant l'objet de la médiation d'éléments externes à l'individu. Ce concept est majoritairement évalué à l'aide de questionnaires subjectifs puisqu'ils sont faciles à appliquer et à interpréter. Cependant, cette méthode de mesure a fait l'objet de plusieurs critiques liées à sa diagnosticité, à sa sensibilité et même à sa validité comme mesure du concept de présence. Depuis plusieurs années, les chercheurs témoignent de la nécessité de trouver une alternative objective aux questionnaires pour mesurer le sentiment de présence en environnement virtuel. Toutefois, la tâche est ardue puisque le concept de sentiment de présence est un quale, c'est-à-dire une expérience subjective. Le défi consiste donc à trouver une méthodologie permettant d'exprimer objectivement ce concept subjectif.

Le concept de sentiment de présence est caractérisé par deux facteurs : la cohérence et l'immersion. La cohérence désigne la capacité d'un environnement virtuel à agir comme l'utilisateur.rice s'y attend. L'immersion désigne l'étendue et la précision des modalités sensorielles simulées dans l'environnement virtuel.

L'immersion présente un intérêt particulier dans le cadre de cette recherche, car elle constitue une dimension objective du concept de sentiment de présence. Généralement, une variation de l'immersion a pour effet de faire varier le sentiment de présence ressentie. Ainsi, en faisant varier ce facteur, il est possible de réaliser une manipulation objective de la présence. Toutefois, on ne sait pas dans quelle gamme de valeur il est préférable de faire varier l'immersion. Théoriquement, l'immersion est maximale si chaque qualité sensorielle simulée peut être reproduite à son seuil perceptuel, c'est-à-dire le plus petit stimulus qu'un organe puisse sentir. On ne sait toutefois pas à quel point il est possible de diminuer la qualité de l'immersion sans affecter le sentiment de présence. Ainsi, on cherche à connaître le seuil fonctionnel des différentes qualités sensorielles. En d'autres mots, on cherche à savoir à quel point il est possible de dégrader l'immersion sans qu'un.e utilisateur.rice d'un environnement virtuel ne puisse le percevoir.

Ainsi, la première étape de cette thèse fut de réaliser un test psychophysique afin d'évaluer le seuil fonctionnel de trois qualités visuelles : la sensibilité au contraste, la taille du champ de vision et l'acuité visuelle. Pour ce faire, 30 participant.e.s ont exécuté un test de sélection

de cibles dans un environnement virtuel immersif alors que l'une des trois qualités visuelles se dégradait. Les participant.e.s ont rapporté le moment auquel ils étaient en mesure de percevoir une différence dans l'environnement virtuel. Cette évaluation a permis d'estimer les seuils fonctionnels des trois qualités visuelles étudiées. Des seuils au-dessus desquels un.e utilisateur.rice peut effectuer une tâche en environnement virtuel sans être affecté par la dégradation de ses qualités visuelles.

Ensuite, en nous basant sur ces seuils fonctionnels, nous avons évalué l'effet d'une dégradation de l'immersion sur des indicateurs de ressources attentionnelles : la charge cognitive, mesurée par un questionnaire subjectif et des indicateurs psychophysiologiques, et la conscience de la situation mesurée par la méthode SAGAT. Bien que la documentation scientifique ait déjà montré une relation entre ces indicateurs et le sentiment de présence, la nature de la relation entre ces indicateurs et les facteurs du sentiment de présence demeure inconnue. Ainsi, en effectuant cette manipulation, cette thèse visait à préciser la relation existant entre des indicateurs de ressources attentionnelles, la cohérence et l'immersion.

Soixante participant.e.s ont réalisé une tâche de conduite automobile dans un environnement virtuel immersif. Lors de celle-ci, ils furent soumis à trois des sept manipulations de conditions visuelles potentielles. Les sept conditions visuelles correspondaient à une condition sans dégradation, trois conditions où l'une des trois qualités visuelles ciblées par cette thèse était fixée au niveau du seuil fonctionnel et trois conditions où elles étaient fixées sous le seuil fonctionnel. Nos résultats ont montré que la charge cognitive et la conscience de la situation étaient uniquement affectées par l'immersion lorsque les conditions visuelles changeaient les circonstances dans lesquelles se déroulait l'expérience virtuelle. Ainsi, nous avons conclu que ces indicateurs de ressources attentionnelles devraient plutôt être utilisés comme mesure de la cohérence d'un environnement virtuel. Cette conclusion a mis en évidence l'impact direct de l'immersion sur la cohérence d'un environnement virtuel.

En somme, dans cette thèse, nous suggérons que la mesure objective du sentiment de présence doit se faire en évaluant l'immersion et la cohérence. Pour mesurer l'immersion, il est nécessaire d'identifier les modalités sensorielles simulées et la précision à laquelle elles le sont. La mesure de la cohérence n'est pas aussi bien définie. Néanmoins, pour ce faire, nos résultats montrent qu'il est pertinent de mesurer les ressources attentionnelles des utilisateur.rice.s d'un environnement virtuel grâce à des concepts pouvant être mesurés objectivement comme la charge cognitive et la conscience de la situation.

Cette thèse a permis de faire avancer les connaissances en 1) définissant une méthode objective et non arbitraire de manipulation du sentiment de présence et en 2) proposant d'évaluer des indicateurs de ressources attentionnelles comme mesures prometteuses pour évaluer ob-

jectivement le concept de sentiment de présence.

ABSTRACT

This thesis studies the measurement methodologies of the feeling of presence in a virtual environment. It aims at deepening the knowledge related to the evaluation of this concept.

Presence is defined as the perceived realism of an experience mediated by elements external to the individual. This concept is mainly evaluated using subjective questionnaires since they are easy to apply and interpret. However, this measurement method has been subject to several criticisms related to its diagnosticity, sensitivity, and even validity in measuring the concept of presence. Researchers have expressed the need to find an objective alternative to questionnaires to measure presence in a virtual environment for several years. However, this task is difficult since presence is a quale, i.e., a subjective experience. The challenge is to find a methodology to express this subjective concept objectively.

Presence is characterized by two factors: coherence and immersion. Coherence refers to the ability of a virtual environment to act as the user expects it to act. Immersion refers to the extent and accuracy of the sensory modalities simulated in the virtual environment.

Immersion is particularly interesting in this research, as it constitutes an objective factor presence. Generally, a variation in immersion impacts the presence felt by virtual environment users. Thus, it is possible to objectively manipulate presence by varying this factor. However, the recommended range of values that immersion can take without significantly affecting presence is undetermined. Theoretically, immersion is maximal if each simulated sensory quality can be reproduced at its perceptual threshold, i.e., the slightest stimulus an organ can sense. Since we have yet to determine to what extent it is possible to decrease the quality of immersion without affecting the feeling of presence, we seek to find the functional threshold of different sensory qualities. In other words, we want to know to what extent it is possible to degrade immersion without a virtual environment user being able to perceive it.

Thus, the first step of this thesis was to perform a psychophysical test to assess the functional threshold of three visual qualities: contrast sensitivity, field of view size, and visual acuity. To do this, 30 participants performed a target selection test in an immersive virtual environment while one of the three visual qualities deteriorated. Participants reported the moment at which they were able to perceive a difference in the virtual environment. This assessment allowed us to estimate the functional thresholds of the three visual qualities studied. Thresholds above which a user can perform a task in a virtual environment without being affected by the degradation of his visual qualities.

Then, based on these functional thresholds, we assessed the effect of varying the immersion level on indicators of attentional resources such as cognitive load, measured by a subjective questionnaire and psychophysiological indicators, and situational awareness, measured by the SAGAT method. Although the scientific literature has already shown a relationship between these indicators and presence, the relationship between these indicators and the factors of presence remains unknown. Thus, in performing this manipulation, this thesis aimed to clarify the relationship between attentional resources indicators, coherence, and immersion.

Sixty participants performed a driving task in an immersive virtual environment. During this task, they experienced three of seven potential visual condition manipulations. The seven visual conditions corresponded to a no-degradation condition, three conditions where one of the three visual qualities targeted by this thesis was set at the functional threshold, and three conditions where they were set below the functional threshold. Our results showed that cognitive load and situational awareness were only affected by immersion when the visual conditions changed the circumstances in which the virtual experience took place. Thus, we concluded that these indicators of attentional resources should instead be used to measure a virtual environment's coherence. Also, this conclusion highlighted the direct impact of immersion on the coherence of a virtual environment.

In sum, in this thesis, we suggest that the objective measurement of the feeling of presence should be done by assessing immersion and coherence. To measure immersion, it is necessary to identify the simulated sensory qualities and the accuracy to which they are simulated. The way to objectively measure coherence still needs improvements. Nevertheless, to do so, our results show that it is relevant to measure the attentional resources of the users of a virtual environment with concepts that can be objectively measured, such as cognitive load and situational awareness.

This thesis's main contributions are 1) defining an objective, non-arbitrary method of manipulating presence and 2) proposing to evaluate indicators of attentional resources to assess presence objectively.

TABLE DES MATIÈRES

DÉDICACE iii

REMERCIEMENTS iv

RÉSUMÉ v

ABSTRACT viii

TABLE DES MATIÈRES x

LISTE DES TABLEAUX xiv

LISTE DES FIGURES xv

LISTE DES SIGLES ET ABRÉVIATIONS xvii

CHAPITRE 1 INTRODUCTION 1
1.1 Mise en contexte 1
1.2 Éléments de la problématique 3
1.2.1 Facteurs du sentiment de présence 3
1.2.2 Modèle de traitement de l'information pour la présence en environnement virtuel 6
1.2.3 Méthodes d'évaluation du sentiment de présence 7
1.2.4 Variables influençant le sentiment de présence 11
1.3 Plan de la thèse 17

CHAPITRE 2 REVUE DE LITTÉRATURE 19
2.1 Qualités visuelles 19
2.1.1 Champ de vision 20
2.1.2 Acuité visuelle 22
2.1.3 Intensité lumineuse et contraste 23
2.2 Charge cognitive 25
2.2.1 Mesure de la charge cognitive 26
2.2.2 Évaluation de la charge cognitive en environnement virtuel 27
2.3 Mesures psychophysiologiques 27
2.3.1 Activité oculaire 28

2.3.2 Activité cardiovasculaire 29
2.4 Conscience de la situation 31
2.4.1 Mesure de la conscience de la situation 33
2.4.2 Conscience de la situation et sentiment de présence 34

CHAPITRE 3 DÉMARCHE DU TRAVAIL DE RECHERCHE 37
3.1 Synthèse de la problématique de recherche 37
3.2 Objectifs de recherche 38
3.3 Démarche de recherche 39
3.3.1 Expérience 1 : Identification de seuils fonctionnels 39
3.3.2 Expérience 2 : Évaluation de l'impact d'une manipulation de l'immersion sur des indicateurs physiologiques et cognitifs 40
3.3.3 Instruments de mesure 41

CHAPITRE 4 ARTICLE 1 : IDENTIFICATION OF VISUAL FUNCTIONAL THRESHOLD FOR IMMERSION ASSESSMENT IN VIRTUAL REALITY 46
4.1 Abstract 46
4.2 Introduction 46
4.2.1 Thresholds of Perception 47
4.3 Related Work 48
4.3.1 Field of View 48
4.3.2 Visual Acuity 50
4.3.3 Light and Contrast Sensitivity 51
4.3.4 Research Objective 52
4.4 Method 52
4.4.1 Participants 53
4.4.2 Materials 53
4.4.3 Procedure 54
4.4.4 Experimental Procedure 60
4.4.5 Experimental Design 61
4.4.6 Data Analysis 61
4.5 Results 62
4.5.1 Field of View 62
4.5.2 Visual Acuity 63
4.5.3 Contrast Sensitivity 64
4.5.4 Multi-Directional Selection Test Throughput Analysis 65
4.6 Discussion 66

4.6.1 Functional Thresholds . . . 67
4.6.2 Multi-Directional Selection Test . . . 69
4.6.3 Implication of Functional Thresholds . . . 71
4.7 Conclusion . . . 73
4.8 Acknowledgments . . . 73

CHAPITRE 5 ARTICLE 2 : THE RELATIONSHIP BETWEEN IMMERSION AND PSYCHOPHYSIOLOGICAL INDICATORS . . . 74
5.1 Abstract . . . 74
5.2 Introduction . . . 75
5.3 Related Work . . . 76
5.3.1 Presence and Related Concepts . . . 76
5.3.2 Psychophysiological Indicators . . . 79
5.3.3 Research Objectives and Hypothesis . . . 83
5.4 Method . . . 83
5.4.1 Participants . . . 83
5.4.2 Apparatus . . . 84
5.4.3 Self-Report Measures . . . 86
5.4.4 Procedure . . . 87
5.4.5 Data Analysis . . . 90
5.5 Results . . . 92
5.5.1 Presence Questionnaire . . . 92
5.5.2 NASA-TLX . . . 93
5.5.3 Cardiovascular Response . . . 93
5.5.4 Eye Response . . . 98
5.6 Discussion . . . 99
5.6.1 Effect of an Objective Degradation of Immersion on the PQ . . . 99
5.6.2 Effect of an Objective Degradation of Immersion on NASA-TLX . . . 100
5.6.3 Effect of an Objective Degradation of Immersion on Psychophysiological Indicators . . . 101
5.6.4 Theoretical Implications . . . 103
5.6.5 Limitations . . . 104
5.7 Conclusion . . . 106
5.8 Acknowledgments . . . 107

CHAPITRE 6 ARTICLE 3 : THE INFLUENCE OF IMMERSION ON SITUATIONAL AWARENESS IN A VIRTUAL ENVIRONMENT . . . 109

6.1 Abstract . . . 109
6.2 Introduction . . . 109
6.3 Related work . . . 110
6.3.1 Presence . . . 110
6.3.2 Situational awareness . . . 113
6.3.3 Research Objectives . . . 117
6.4 Method . . . 117
6.4.1 Participants . . . 117
6.4.2 Apparatus . . . 117
6.4.3 Procedure . . . 121
6.4.4 Data analysis . . . 123
6.5 Results . . . 124
6.5.1 Situational Awareness . . . 124
6.5.2 Presence Questionnaire . . . 127
6.5.3 Correlation between SA and PQ . . . 129
6.6 Discussion . . . 129
6.6.1 The effect of immersion's degradation on SAGAT scores . . . 129
6.6.2 The effect of immersion degradation on the PQ score . . . 132
6.6.3 Situational Awareness and Presence . . . 133
6.6.4 Functional thresholds . . . 134
6.6.5 Limitations . . . 136
6.7 Conclusion . . . 138
6.8 Acknowledgments . . . 138

CHAPITRE 7 DISCUSSION GÉNÉRALE . . . 139
7.1 Seuils fonctionnels . . . 140
7.2 Impacts d'une manipulation de l'immersion sur des indicateurs de ressources attentionnelles . . . 142
7.3 Implications théoriques . . . 147
7.4 Limites de la thèse . . . 149

CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS . . . 152
8.1 Perspectives futures . . . 153

RÉFÉRENCES . . . 155

LISTE DES TABLEAUX

Tableau 3.1	Distribution des scores du questionnaire NASA-TLX	42
Tableau 3.2	Distribution des scores du questionnaire NASA-TLX pour une tâche de conduite de véhicule	42
Table 4.1	Summary of post hoc pairwise comparison of task execution order	65
Table 4.2	Summary of post hoc pairwise comparison of throughput between visual degradation conditions	66
Table 4.3	Perceptual and functional thresholds of FOV, visual acuity and contrast sensitivity	72
Table 5.1	Descriptive statistics of Presence Questionnaire score for each degradation condition	93
Table 5.2	Descriptive Statistics of NASA-TLX Score for Each Degradation Condition	94
Table 5.3	Results of the Linear Mixed Model for Mean HR	94
Table 5.4	Results of the Linear Mixed Models for RMSSD, SDNN, and pNN50	96
Table 5.5	Results of the Linear Mixed Model for LF, HF, and LF/HF	97
Table 5.6	Results of the Linear Mixed Model for LHIPA	98
Table 6.1	SAGAT probes and answers for a driving task	120
Table 6.2	Descriptive statistics of SAGAT score for each degradation condition	125
Table 6.3	Results of the linear mixed model for SAGAT Score	125
Table 6.4	Results of the linear mixed model for SAGAT for each SA level	126
Table 6.5	Descriptive statistics of Presence Questionnaire score for each degradation condition	128

LISTE DES FIGURES

Figure 1.1	Relation entre les concepts du sentiment de présence	4
Figure 2.1	Représentation schématique de différentes portions du champ de vision	20
Figure 2.2	Modèle de la conscience de la situation	33
Figure 3.1	Environnement virtuel de l'expérience 1 selon trois perspectives	39
Figure 3.2	Environnement virtuel de l'expérience 2	40
Figure 4.1	Schematic representation of different sections of the FOV.	49
Figure 4.2	Virtual environment	54
Figure 4.3	Representation of different multi-directional selection test IDs	55
Figure 4.4	Representation of FOV degradation	57
Figure 4.5	Representation of visual acuity degradation	58
Figure 4.6	Representation of contrast degradation	59
Figure 4.7	Results of FOV degradation	63
Figure 4.8	Results of visual acuity degradation	64
Figure 4.9	Results of contrast sensitivity degradation	65
Figure 5.1	Relationship between presence concepts	76
Figure 5.2	Virtual environment	85
Figure 5.3	Virtual signs	86
Figure 5.4	Contrast sensitivity degradation conditions	88
Figure 5.5	FOV size degradation conditions	89
Figure 5.6	Visual acuity degradation conditions	89
Figure 5.7	Overview of the experimental procedure	89
Figure 5.8	Presence questionnaire score for each degradation condition	92
Figure 5.9	NASA-TLX Score for Each Degradation Condition	95
Figure 5.10	Revised Relationship Between Presence Concepts	105
Figure 6.1	Situational awareness model	114
Figure 6.2	Virtual environment	118
Figure 6.3	Overview of the experimental procedure	121
Figure 6.4	Virtual environment degradation conditions	122
Figure 6.5	SAGAT score for each degradation condition	124
Figure 6.6	Presence questionnaire score for each degradation condition	128
Figure 6.7	Correlation between the presence questionnaire score and the SAGAT score	129

Figure 7.1 Représentation schématique des trois qualités visuelles à des valeurs autour du seuil fonctionnel 141
Figure 7.2 Relation révisée entre les concepts du sentiment de présence 150

LISTE DES SIGLES ET ABRÉVIATIONS

ANOVA	ANalysis Of VAriance
CI	Confidence Intervals
FOV	Field Of View
HF	High Frequency
HMD	Head-Mounted Display
HRV	Heart Rate Variability
ID	Index of Difficulty
ICA	Index of Cognitive Activity
IPA	Index of Pupillary Activity
JND	Just Noticeable Difference
LHIPA	Low/High Index of Pupillary Activity
LF	Low Frequency
LF/HF	ratio of Low and High Frequencies
NASA-TLX	Nasa Task Load indeX
NN	Valeur normalisée du temps entre deux détections de battements de coeur
pNN50	proportion of the number of pairs of successive NNs that differ by more than 50 ms
PQ	Presence Questionnaire
PSE	Point of Subjective Equality
RMS	Root Mean Square
RMSSD	Root Mean Square of Successive Differences
SAGAT	Situation Awareness Global Assessment Technique
SDNN	Standard Deviation of NN intervals
UFOV	Useful Field Of View

CHAPITRE 1 INTRODUCTION

1.1 Mise en contexte

Les simulateurs sont appelés à prendre une place de plus en plus importante dans la formation de personnel, les processus de conception et la recherche. Ce sont des appareils permettant de représenter artificiellement le fonctionnement de leur équivalent réel. Par exemple, un simulateur de vol, imitant un avion, reproduit l'ensemble d'un environnement de vol incluant les différents systèmes de l'appareil et la manière dont ils interagissent avec divers facteurs externes comme la densité de l'air, le vent ou les précipitations. Les simulateurs offrent la possibilité d'exposer leurs utilisateur.rice.s à un environnement dépourvu des risques rencontrés dans l'environnement réel [1]. Ils présentent aussi l'avantage de permettre la sélection de scénarios précis, comme la simulation d'événements rares et critiques, plutôt que de se fier à leur disponibilité stochastique [1]. En somme, ils assurent un contrôle sur l'environnement simulé tout en permettant la représentation d'une situation réelle.

Les simulateurs incorporent généralement des instruments physiques authentiques disposés de manière à reproduire l'artefact réel [2]. Or, l'acquisition du matériel est coûteuse en ressources financières [3]. La taille des instruments, généralement considérable, impose de sévères contraintes quant à l'emplacement où peuvent être installés les simulateurs [4]. De plus, la complexité de leur configuration, de leur utilisation et de leur désassemblage engendre des coûts importants en ressources humaines [4]. Pour pallier ces contraintes, les simulateurs sont appelés à substituer leurs instruments physiques par des éléments virtuels en utilisant la réalité étendue.

La réalité étendue désigne une technologie permettant de créer et de visualiser un environnement dans lequel on perçoit à la fois des éléments réels et des éléments virtuels qui interagissent à plusieurs degrés [5]. Il s'agit d'un terme générique qui désigne à la fois la réalité virtuelle, la réalité augmentée et la réalité mixte [5]. Ce type de technologie est caractérisé par la substitution partielle ou complète des entrées sensorielles des utilisateur.rice.s par des données générées artificiellement permettant la perception d'éléments virtuels et offrant la possibilité d'interagir avec ces derniers [6].

Cette thèse s'inscrit dans ce contexte. Elle fait partie d'un projet de conception d'outils de réalité étendue visant à améliorer le développement et l'évaluation de postes de pilotage d'avion. L'emploi de la réalité étendue dans les simulateurs implique une tendance à utiliser un nombre croissant d'éléments virtuels pour simuler un environnement. Ainsi, la diminution

du nombre d'éléments physiques au profit d'éléments virtuels met en exergue la question du réalisme d'un environnement virtuel. Plus précisément, on cherche à savoir comment il est possible de mesurer le réalisme d'une expérience virtuelle. Dis autrement, comment est-il possible de savoir à quel degré des personnes impliquées dans une expérience virtuelle agissent de la même manière que dans un environnement réel ?

Le réalisme d'une expérience virtuelle est un concept subjectif lié à la façon dont un individu perçoit et comprend une expérience et ses impacts [7]. Évaluer le réalisme d'un environnement virtuel permet de déterminer si les éléments simulés sont fidèles à leur équivalent réel et de caractériser les similarités entre les deux environnements [8]. Ainsi, juger de la fidélité d'un environnement virtuel permet d'en déterminer la validité [8], d'être apte à savoir si les conclusions tirées des tests réalisés dans cet environnement peuvent être appliquées à son équivalent réel et d'identifier les éléments de la simulation qu'il serait possible d'améliorer.

Pour évaluer le réalisme d'une expérience virtuelle, il est indiqué de recourir à l'utilisation du concept de sentiment de présence [7]. Généralement, un réalisme plus élevé entraîne un sentiment de présence plus forte [7]. Le sentiment de présence est un indicateur de la validité écologique d'un environnement virtuel [2,9]. Elle permet de mesurer les variations pertinentes au niveau physiologique, comportemental, cognitif et émotif afin d'établir la fidélité d'un environnement virtuel [10]. Un fort sentiment de présence entraîne des comportements plus réalistes dans l'environnement virtuel et aide à réduire les biais comportementaux induits par la simulation [11]. Le sentiment de présence consiste en une condition nécessaire au transfert d'apprentissage acquis dans un environnement virtuel vers un environnement réel [2]. Ainsi, la présence peut contribuer à améliorer les performances, l'engagement et l'apprentissage des compétences [12] et est un des principaux facteurs associés à l'efficacité d'un environnement virtuel [13].

Cependant, à cause de sa nature subjective, le sentiment de présence est un phénomène dont la mesure est complexe. Pour mieux exposer les causes des difficultés liées à l'évaluation de ce concept, les prochaines sections offrent une définition du sentiment de présence et de ses principaux facteurs et présentent un modèle de traitement de l'information expliquant son émergence. Par la suite, elles exposent les méthodes de mesures actuelles du concept et les critiques qui leur sont adressées. Puis, elles présentent des variables d'intérêt pouvant être impliquées dans la mesure du sentiment de présence.

1.2 Éléments de la problématique

Le concept de présence a fait l'objet de plusieurs définitions et interprétations [14, 15]. Initialement, le sentiment de présence était considérée comme une extension du concept de téléprésence, un phénomène par lequel un opérateur sent qu'il est physiquement dans une autre localisation grâce aux rétroactions de machines contrôlées à distance [14]. Cette association est à l'origine d'une conceptualisation prépondérante du concept qui est décrit comme l'impression « d'être là », de se trouver physiquement dans un environnement virtuel [15, 16, 16–21]. Or, cette définition exclue un autre aspect important du sentiment de présence [22], c'est-à-dire l'impression que l'environnement virtuel devient réalité pour les participant.e.s [2, 9, 10, 15, 18, 23, 24]. Ainsi, le sentiment de présence combine deux aspects permettant de décrire et définir l'expérience résultante adéquatement et rendant possible une évaluation plus complète du concept [22] : le fait de se sentir dans un endroit différent de sa localisation physique et le réalisme perçu de l'expérience virtuelle.

Par conséquent, le sentiment de présence est définie comme le **réalisme perçu d'une expérience faisant l'objet de la médiation d'éléments externes à l'individu** [15]. Il s'agit d'une illusion perceptuelle [25]. Le fait de savoir qu'il s'agit d'une illusion ne change pas la perception ou la réponse à celle-ci [26]. Elle consiste en une forme de surveillance inconsciente de l'expérience basée sur une intégration multisensorielle efficace et congruente [25] qui permet l'expression de l'agentivité, le fait qu'une personne puisse se sentir responsable de ses propres actions [27].

1.2.1 Facteurs du sentiment de présence

Deux facteurs permettent le sentiment de présence : le fait de se sentir ailleurs que sa localisation physique ; et le réalisme perçu de l'expérience virtuelle. Ces facteurs sont importants, car ils englobent le concept, délimitant ce qui en fait et n'en fait pas partie.

Cette caractérisation se fonde sur l'affirmation de Slater [19], résultant de la méta-analyse d'études expérimentales sur le sentiment de présence, qui indique qu'il est constitué de deux composantes qui contribuent à une réponse réaliste des utilisateur.rice.s à un environnement virtuel, soit l'illusion de lieu et l'illusion de plausibilité. L'illusion de lieu se produit dans les contraintes imposées par l'immersion [19], alors que l'illusion de plausibilité dans celles imposées par la cohérence [15]. Une modélisation des relations entre immersion, cohérence et sentiment de présence est présentée à la figure 1.1.

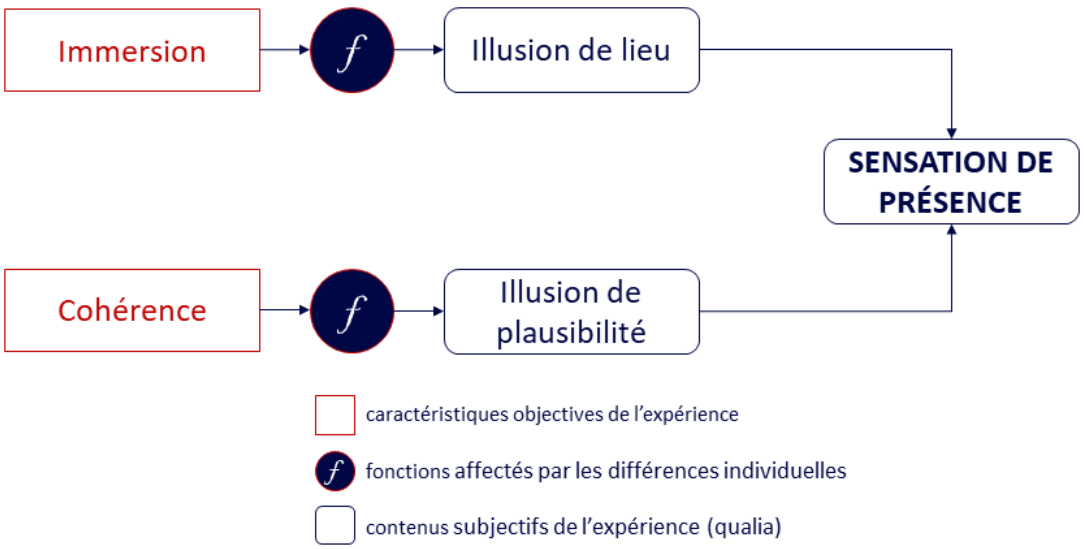


FIGURE 1.1 Relation entre les concepts du sentiment de présence

Note. Cette figure se base sur les relations entre les concepts du sentiment de présence établies par Skarbez et al. [15].

Immersion et illusion de lieu

L'illusion de lieu est responsable de l'impression « d'être là ». Elle fait référence au quale de se trouver physiquement dans un environnement virtuel malgré le fait que l'on sait pertinemment que l'on n'y est pas [19].

L'illusion de lieu est contrainte par l'immersion que permet un système. L'immersion est définie comme une caractéristique objective d'un environnement virtuel désignant l'étendue et la précision des modalités et qualités sensorielles simulées [20]. Notons que dans le domaine de la recherche sur la présence, le mot "immersion" fait l'objet d'une certaine ambiguïté. Une école de pensée considère l'immersion comme un état psychologique de l'utilisateur.rice, alors que l'autre considère l'immersion comme une caractéristique objective d'un système de réalité étendue [28]. Lors de l'emploi de ce mot, cette thèse fait référence à la seconde définition. L'immersion est objective et quantifiable puisqu'elle décrit les capacités d'un système particulier [20]. Elle fournit les limites à l'intérieur desquelles l'illusion de lieu peut se produire [19]. L'illusion de lieu est donc dépendante du système utilisé pour l'évoquer [19].

L'illusion de lieu se forme par une analyse introspective qu'une personne utilisant un environnement virtuel fait du cycle qui lie sa perception et ses actions [27]. La perception permet à une personne utilisant un environnement virtuel de savoir quelles actions elle est en mesure

de produire [27]. Puis, les actions sont effectuées de manière à pouvoir percevoir [19]. Plus la perception est reproduite de manière exacte, plus les actions qui en découlent peuvent être effectuées de manière précise. Ce lien entre perception et action est qualifié de contingence sensorimotrice [19].

Ainsi, l'illusion de lieu dépend des contingences sensorimotrices d'un système [19]. Les contingences sensorimotrices caractérisent les systèmes immersifs, car elles permettent de définir un ensemble d'actions valides pouvant être réalisées dans un environnement virtuel. Les actions valides sont définies comme les actes que les utilsateurs.rices d'un environnement virtuel peuvent faire qui résultent en un changement de leur perception ou qui ont un effet sur l'environnement. Un système aura un plus haut niveau d'immersion s'il permet d'effectuer plus d'actions valides. De manière à permettre un plus grand nombre d'actions valides, un système immersif doit donc être en mesure de reproduire la sensation et la perception de ses utilisateur.rice.s.

Cohérence et illusion de plausibilité

L'illusion de plausibilité fait référence à l'impression que les événements qui surviennent dans un environnement virtuel sont réels, qu'ils se produisent véritablement [19].

L'illusion de plausibilité est contrainte par la cohérence d'un environnement virtuel. La cohérence est analogue au concept d'authenticité [15] qui décrit le degré auquel un environnement virtuel répond aux attentes de ses utilisateur.ices de manière à ce que leurs intentions et leurs comportements soient compatibles avec l'expérience virtuelle [29]. Elle englobe des éléments subjectifs comme la qualité visuelle de l'environnement, la cohérence narrative et la pertinence d'un mode d'interaction [30]. La cohérence est définie comme l'ensemble de circonstances raisonnables qui peuvent être démontrées par un scénario sans y introduire de circonstances déraisonnables [15]. Une circonstance raisonnable représente un état d'un scénario virtuel qui est évident par rapport aux connaissances préalables de ses utilisateur.rice.s [15]. Les circonstances raisonnables sont différentes pour chaque individu [15].

Se produisant au sein des contraintes déterminées par la cohérence, l'illusion de plausibilité dépend du degré auquel un système produit des événements corrélés avec les comportements des utilisateur.rice.s d'un environnement virtuel [19]. L'illusion de plausibilité dépend des connaissances préalables qu'un.e utilisateur.rice a du monde réel et du monde potentiellement fictif représenté dans l'environnement virtuel [31]. Pour maintenir l'illusion de plausibilité, il est nécessaire que les événements qui ont lieu soient crédibles et qu'ils se conforment aux attentes de l'utilisateur.rice en fonction des circonstances présentées [19].

1.2.2 Modèle de traitement de l'information pour la présence en environnement virtuel

La formation et le maintien du sentiment de présence chez un.e utilisateur.rice d'un environnement virtuel peuvent être décrits par un modèle cyclique de traitement de l'information [25]. Selon ce modèle, le sentiment de présence résulte d'un traitement séquentiel de l'information comprenant deux étapes principales. À la suite de leur réception par les organes sensoriels, les stimuli provenant de l'environnement virtuel subissent un traitement implicite, puis un traitement explicite pouvant résulter en la formation des illusions de plausibilité et de lieu [25] et donc du sentiment de présence. Grâce aux propriétés interactives d'un environnement virtuel, le comportement implicite et explicite de ses utilisateur.rice.s influencera la façon dont l'environnement virtuel réagit et quels stimuli virtuels seront reçus et intégrés par le cerveau relançant le cycle de traitement de l'information [32].

Ainsi, la première étape correspond à un traitement implicite de l'information qui consiste en une intégration multisensorielle des stimuli. Il s'agit d'un processus par lequel plusieurs sens sont combinés pour créer une représentation cohérente du monde [25]. Trois phénomènes peuvent augmenter les chances que des stimuli soient intégrés. Les stimuli doivent notamment être présentés dans un court intervalle de temps et dans un espace rapproché [33], doivent correspondre aux actions anticipées par les utilisateur.rice.s et aux prédictions faites par le cerveau sur la base d'expériences antérieures [25].

De l'intégration congruente de stimuli découlera l'évaluation de la crédibilité de l'expérience virtuelle [25]. Cette deuxième étape est plus explicite que la précédente, car elle traite des informations plus complexes [25]. Elle implique par exemple des jugements sur les propriétés et comportements des éléments de l'environnement [25]. Cependant, le réalisme d'une expérience n'est pas une condition nécessaire à sa crédibilité [34]. Par exemple, même un environnement qui n'est pas photoréaliste peut évoquer un sentiment de présence [14]. En fait, une simulation serait jugée crédible lorsqu'elle permet à ses utilisateur.rice.s d'atteindre leurs objectifs [35]. Comme l'étape précédente, le résultat de l'évaluation de la crédibilité d'une expérience virtuelle est modulé par les attentes et les expériences antérieures des utilisateur.rice.s [25]. Un environnement virtuel, perçu comme un espace plausible, permet à ses utilisateur.rice.s de s'en former un modèle mental [36]. Lorsque les stimuli sont jugés crédibles et que les actions que les utilisateur.rice.s s'attendent à pouvoir effectuer peuvent être réalisées, les utilisateur.rice.s expriment leur agentivité dans l'environnement virtuel [27]. À ce moment, l'environnement virtuel, et non l'environnement réel, devient l'environnement de référence pour les expériences vécues par les utilisateur.rice.s [36]. Ainsi, si l'expérience virtuelle est jugée logiquement congruente, il pourra en résulter la formation des illusions de

plausibilité et de lieu [25].

Du produit de ces illusions découlera le sentiment de présence. L'illusion de plausibilité conduira à la sensation subjective que l'expérience se produit réellement et l'illusion de lieu mènera à l'impression de se trouver réellement dans le lieu représenté par l'environnement virtuel [25]. Puis, la signification subjective donnée aux lieux et aux objets expérimentés dans l'environnement virtuel aura un impact sur le comportement de l'utilisateur.rice [25]. À son tour, le comportement adopté influencera les stimuli parvenant à la boucle d'entrée sensorielle qui seront ensuite traités et intégrés par le cerveau [25].

Étant donné la subjectivité du sentiment de présence, deux utilisateur.rice.s faisant l'expérience du même environnement virtuel pourraient associer une signification différente à leur expérience et donc s'y comporter différemment [25].

1.2.3 Méthodes d'évaluation du sentiment de présence

La mesure du sentiment de présence permet d'avoir une indication de la qualité d'un environnement virtuel [13]. Dans la plupart des expérimentations, le sentiment de présence est mesuré pendant ou suivant l'exécution d'une tâche dans un environnement virtuel.

Les méthodes de mesure de la présence peuvent être divisées en deux catégories : les mesures subjectives, incluant les questionnaires et les entrevues [20], et les mesures objectives, incluant les observations comportementales et les mesures psychophysiologiques [37]. Les questionnaires permettent de faire une évaluation subjective d'éléments connus pour affecter la présence [20]. Les entrevues consistent en une évaluation subjective ouverte, donnent généralement des résultats semblables aux questionnaires et sont très rarement utilisées [20]. Les analyses comportementales sont une mesure objective permettant d'évaluer à quel point les utilisateur.rice.s d'un environnement virtuel se comportent comme dans un environnement réel [20]. Finalement, les mesures psychophysiologiques sont utilisées pour évaluer objectivement l'activité cognitive et les réponses émotionnelles [38] et ainsi déterminer si la réponse physiologique dans l'environnement virtuel correspond à celle qui est attendue [20].

Mesures par questionnaire

Des revues de littérature ont établi que le sentiment de présence était majoritairement évalué à l'aide de questionnaires subjectifs [7, 13, 20]. Selon Gonçalves et al. [7], 44% des articles évaluaient le sentiment de présence uniquement de manière subjective, 41% l'évaluaient en utilisant des méthodes subjectives et objectives et dans seulement 15% des cas, le sentiment de présence faisait l'objet d'une évaluation uniquement objective. Parmi les méthodes subjectives

utilisées, 99 % étaient des questionnaires [13].

L'utilisation prépondérante des questionnaires serait due principalement à leur simplicité d'utilisation. Les questionnaires de présence sont généralement faciles à appliquer et à interpréter [13,20,39,40]. Étant disponibles en version papier, les questionnaires sont bon marché, faciles à administrer et applicables, quel que soit l'environnement virtuel utilisé [40]. Ils ne requièrent pas de longue préparation préalable, de compétences spécialisées ou d'instruments scientifiques particuliers comme c'est le cas pour les méthodologies objectives [40].

La complexité de la mesure de la présence réside dans le fait qu'il faille évaluer un état de conscience interne subjectif introduit par les sens [12] lié à la manière dont les utilisateur.rice.s perçoivent le réalisme d'un environnement virtuel [7]. Ayant la flexibilité de poser des questions sur des aspects spécifiques qui ont trait à des notions importantes liées au sentiment de présence, les questionnaires permettent théoriquement d'évaluer le sentiment de présence [7]. Possédant une fiabilité avérée et une validité apparente [20,39], les questionnaires se sont imposés comme mesure de base du construit de présence sur lesquels reposent les corrélations établies entre le concept de présence et les autres mesures telles que les mesures comportementales ou psychophysiologiques [7].

Or, divers questionnaires ont été développés pour évaluer le sentiment de présence malgré l'absence générale d'une définition standard du concept [40]. De plus, lors de l'utilisation de questionnaires, il est compliqué d'évaluer certaines notions importantes pouvant affecter le sentiment de présence, telles que les informations préalables liées aux expériences et à l'attente des participant.e.s par rapport à la tâche [20]. Ainsi, de nombreuses études font état de résultats hétérogènes, ce qui serait symptomatique d'une faible validité de construit des questionnaires [41]. Les questionnaires n'ont pas tous le même niveau de détail et sont affectés par la conception du sentiment de présence qu'en ont leurs auteurs [20]. De plus, les notions soutenant le sentiment de présence ont surtout été définies et étudiées sur la base des questionnaires, biaisant potentiellement les conclusions sur les variables influençant le concept [42]. Ainsi, les résultats des études utilisant différents questionnaires peuvent être difficiles à comparer en raison des différentes conceptualisations du sentiment de présence [40].

Un autre problème en lien avec les questionnaires est qu'il est uniquement possible de comparer les résultats provenant de systèmes utilisant des contingences sensorimotrices semblables, qui évoquent les mêmes qualia [19]. Par exemple, pour comparer une expérience virtuelle et une expérience réelle à l'aide de questionnaires, il faut que les paramètres de l'expérience, comme les modalités d'interaction, soient à peu près semblables à la réalité. Sinon, il peut arriver que dans certaines situations les questionnaires ne soient pas en mesure de faire la distinction entre une expérience virtuelle et une expérience réelle [43].

De plus, les méthodes de mesure par questionnaires peuvent interférer avec la tâche. Il est généralement conseillé de les administrer après une expérience virtuelle pour éviter de briser le sentiment de présence [40]. Toutefois, compléter un questionnaire dans ou hors d'un environnement virtuel n'a pas d'effet significatif sur la performance lorsque l'expérimentation est reprise [44].

L'utilisation de questionnaires pour évaluer la présence demeure cependant recommandée étant donné leur capacité à saisir au moins en partie le concept de présence [45]. Leurs résultats ne doivent néanmoins pas être considérés comme un absolu et doivent être complétés avec d'autres types de données objectives telles que des mesures psychophysiologiques ou comportementales [45]. Le couplage des questionnaires aux mesures objectives pour mesurer la présence permet notamment de produire des résultats plus fiables permettant de soutenir des hypothèses faites sur l'expérience virtuelle, d'identifier les biais subjectifs ou les limites de la méthodologie appliquée [7].

Mesures comportementales

Les mesures comportementales permettent de déterminer si les utilisateur.rice.s d'un environnement virtuel s'y comportent de la même manière que dans des conditions similaires à un environnement réel [20]. Par exemple, une mesure de la stabilité posturale des utilisateur.rice.s d'un environnement virtuel a été utilisée pour évaluer le sentiment de présence, la posture étant considérée moins stable lorsque l'illusion d'emplacement est moins forte [46]. Pour que cette méthode soit applicable, un environnement virtuel doit présenter des caractéristiques qui déclenchent les réactions corporelles analysées [20].

Ainsi, le principal problème des mesures comportementales est qu'elles ne sont pas généralisables [13]. Le développement d'une mesure comportementale valide peut être complexe et elles ne sont utiles que dans des expériences connues pour affecter le comportement des utilisateur.rice.s d'une manière particulière [13]. Ceci explique pourquoi les mesures comportementales sont peu utilisées pour étudier le sentiment de présence malgré la contribution substantielle qu'elles peuvent apporter à l'analyse du contexte d'une expérience virtuelle [20].

Mesures psychophysiologiques

Le sentiment de présence peut aussi être étudié objectivement en associant l'activité physiologique avec l'effet de la présentation de stimuli particuliers [40]. Cela permet d'étudier des phénomènes qui seraient complexes à évaluer avec d'autres méthodes d'évaluation [40]. Les mesures psychophysiologiques incluent l'activité cardiovasculaire, l'activité oculaire, l'ac-

tivité cérébrale, l'activité électrodermale et la température de la peau comme principales mesures liées au sentiment de présence [13, 40]. Les mesures psychophysiologiques présentent l'avantage de fournir une évaluation continue, quasi synchrone et objective du sentiment de présence [40].

Toutefois, les méthodes psychophysiologiques présentent des problèmes méthodologiques. Leurs indications ne sont pas généralisables dans tous les environnements virtuels [41]. Certains indicateurs psychophysiologiques ne sont affectés que par des comportements précis et ayant un niveau d'intensité relativement élevé [13, 20]. Plus précisément, les mesures psychophysiologiques pourraient être insensibles à des environnements virtuels qui n'occasionnent pas de comportements physiologiquement excitants [41]. L'utilisation de mesures psychophysiologiques dans des environnements avec des niveaux d'excitation modérés ou dans le but d'étudier des variations individuelles pourrait ne pas être appropriée pour l'étude du sentiment de présence [41].

De plus, les mesures psychophysiologiques sont difficiles à utiliser seules dans l'état actuel de la recherche en raison de l'absence d'indices physiologiques du sentiment de présence vérifiés et facilement reproductibles qui soient fortement corrélés à d'autres mesures de présence [40]. L'utilisation de mesures psychophysiologiques pour mesurer le sentiment de présence peut être contestée, car de nombreuses variables non contrôlées peuvent les influencer. Il peut donc être complexe d'établir un lien direct entre les réponses physiologiques des utilisateur.rice.s d'un environnement virtuel et leur sentiment de présence [13]. Les phénomènes influençant les mesures psychophysiologiques sont variés et incluent la frustration, l'anxiété [38], le cybermalaise [41, 47, 48], le stress et la charge cognitive [49]. On constate donc que les mesures psychophysiologiques ne permettent pas une évaluation directe du sentiment de présence. Elles mesurent plutôt ses conséquences. Ceci pose des défis pour leur évaluation et leur interprétation, car elles sont considérées comme des mesures distale dont les conclusions reposent sur la démonstration de l'hypothèse nulle.

En somme, des doutes sont émis quant au potentiel de généralisation des mesures psychophysiologiques [13]. Cependant, étant donné que différentes mesures psychophysiologiques sont sensibles à différents phénomènes, on espère tout de même être en mesure de développer une batterie de mesures psychophysiologiques permettant d'évaluer le sentiment de présence de manière généralisable [50].

Critique générale de la mesure de la présence

On remarque que la mesure du sentiment de présence est problématique. Les questionnaires et les mesures psychophysiologiques et comportementales sont critiqués sur le plan de leur

validité comme mesure du construit, leur sensibilité ou leur diagnosticité [19]. De plus, on s'aperçoit qu'il n'y a pas de manières valides et robustes de la mesurer objectivement [37].

Il est généralement complexe de faire des comparaisons entre les résultats de différentes recherches [37]. Cette complexité réside notamment dans le fait que plusieurs méthodes différentes de mesure du sentiment de présence sont utilisées sans standards. La plupart des mesures sont basées sur des concepts sous-jacents différents, mesurant à divers degrés la construction mentale du sentiment de présence ou les différentes propriétés du système [20]. Les différentes méthodes d'évaluation du sentiment de présence ne permettent pas d'obtenir une mesure uniforme du concept [20]. Cela cause une évaluation incohérente du sentiment de présence n'offrant qu'un aperçu limité des impacts du sentiment de présence sur une expérience virtuelle [20].

Étant un quale, donc sujet à une interprétation subjective et variable pour chaque individu, il est particulièrement complexe de faire une évaluation directe du sentiment de présence [19]. Une mesure idéale et valide du sentiment de présence devrait posséder cinq caractéristiques. Elle devrait être liée directement au sentiment de présence, elle devrait être fiable en ayant une répétabilité éprouvée, être sensible aux variations, être non intrusive pour éviter toute dégradation involontaire du sentiment de présence et être pratique, c'est-à-dire portable, peu coûteuse et facile à apprendre et à appliquer [13].

Pour évaluer le réalisme d'une expérience virtuelle, il n'est pas seulement nécessaire de prendre en compte la présence perçue de l'environnement, mais de la comparer avec la présence dans un environnement réel [10]. La similarité de la réponse comportementale entre le monde réel et un environnement virtuel devrait servir d'index objectif de présence [39]. Il devrait notamment y avoir une correspondance des caractéristiques objectives, comme les réponses physiologiques et comportementales, entre un environnement virtuel et son équivalent réel [9].

1.2.4 Variables influençant le sentiment de présence

La complexité et les difficultés liées à l'évaluation du sentiment de présence permettent de constater qu'il est encore nécessaire de déployer des efforts afin d'améliorer les méthodologies permettant de mesurer ce concept. Comme il en a été traité dans les sections précédentes, lorsque les utilisateur.rice.s font l'expérience d'un environnement virtuel, plusieurs variables ou phénomènes peuvent moduler la manière dont l'information est traitée et résulter en une variation du sentiment de présence ressentie. Ainsi, cette section propose d'analyser les variables qui ont une influence sur le sentiment de présence afin de cibler les concepts pouvant contribuer à l'amélioration de son évaluation.

Relation entre modalités sensorielles et présence

Le modèle de traitement de l'information pour la présence en environnement virtuel [25] a mis en évidence l'importance des stimuli et leur rôle par rapport à l'immersion. Les stimuli constituent la source d'information de laquelle découlent les traitements implicite et explicite entraînant éventuellement la formation d'un sentiment de présence. Les stimuli sont des variables importantes pour le sentiment de présence, ils représentent une composante objective qui peut être directement contrôlée. Pour être accessibles à l'expérience humaine, les stimuli sont traités par les modalités sensorielles.

Une modalité sensorielle fait référence à l'activité d'un organe sensoriel particulier qui sert à capter les stimuli de manière à ce qu'ils soient interprétés par le cerveau [51]. Ainsi, la capacité d'un système à fournir des stimuli appropriés constitue un déterminant de l'immersion ; décrite comme l'étendue et la précision de la simulation des modalités sensorielles par un système [20]. Cette définition comporte deux éléments explicitant le rôle utilitaire de l'immersion dans le développement du sentiment de présence. D'abord, pour favoriser le sentiment de présence, il est préférable de stimuler plusieurs modalités sensorielles. Ensuite, il faut que ces modalités sensorielles soient stimulées convenablement, de manière à ce que les utilisateur.rice.s d'un environnement virtuel soient en mesure d'effectuer des actions et d'atteindre leurs buts. Ainsi, il semble pertinent de se demander quel nombre de modalités sensorielles il est nécessaire de simuler et à quel degré la simulation des modalités sensorielles doit être précise.

Généralement, plus on simule de sens adéquatement, plus le sentiment de présence est grand. Des environnements virtuels utilisant à la fois la vision, l'audition et des retours haptiques ont permis de constater une augmentation significative du sentiment de présence, lorsque comparés aux mêmes environnements n'utilisant qu'une ou deux de ces modalités sensorielles [52–54]. Dans la vie réelle, notre cerveau est constamment engagé dans le traitement multimodal de l'environnement qui l'entoure. La coordination de deux ou plusieurs modalités sensorielles facilite la sélection des caractéristiques pertinentes parmi les stimuli non pertinents [54]. Par exemple, certaines caractéristiques sonores, comme l'habileté à localiser un son, ont une influence sur le comportement visuel humain, facilitant l'identification d'éléments dans l'environnement [55]. Ainsi, la combinaison de stimuli sensoriels suscite de manière générale un plus grand sentiment de présence, car elle aide les utilisateur.rice.s d'environnements virtuels à se représenter le monde duquel ils font l'expérience [54].

On indique que la simulation adéquate des sens est aussi nécessaire pour favoriser un sentiment de présence. Les systèmes offrant des propriétés sensorielles plus similaires aux propriétés humaines sont généralement associés à un plus important sentiment de présence et à une

performance plus réaliste [36]. Par exemple, une étude ayant comparé trois visiocasques selon leur fidélité d'affichage a montré qu'un visiocasque possédant une meilleure fidélité d'affichage permettait une meilleure compréhension de l'espace et un meilleur apprentissage du contenu présenté dans l'environnement virtuel [56].

Pour avoir une meilleure compréhension de l'impact de la précision de la simulation des modalités sensorielles sur l'immersion, il est nécessaire de les décomposer en différentes dimensions. Chaque modalité sensorielle est constituée de plusieurs dimensions, appelées *qualités sensorielles*, qui représentent l'information tirée de l'exposition à un stimulus [51]. Par exemple, l'acuité visuelle, désignant le plus fin niveau de détail qu'il est possible de discerner, et la perception des couleurs sont des qualités sensorielles associées à la vision. Ainsi, la précision de la simulation d'une modalité sensorielle fait référence à la finesse de la simulation de cette dernière par rapport aux limites humaines. La simulation des qualités sensorielles a divers degrés d'impact sur le sentiment de présence et le réalisme visuel d'un environnement virtuel [36].

Pour la modalité visuelle, les qualités sensorielles sont associées à la qualité des images présentées dans l'environnement virtuel, la vision binoculaire et l'aspect dynamique de la vision. Parmi les qualités sensorielles associées à la qualité des images présentées dans l'environnement virtuel, on retrouve l'acuité visuelle, le nombre de couleurs et la sensibilité au contraste et à la luminosité [19, 36, 57, 58]. Les qualités visuelles liées à la vision binoculaire, soit le fait d'observer une scène avec deux yeux, sont la vision stéréoscopique et la taille du champ de vision [19, 36, 52, 58]. Finalement, la fréquence de présentation des images est la qualité visuelle qui a trait à la simulation de l'aspect dynamique de la vision permettant de créer une sensation de mouvement [19, 36, 58].

La qualité sonore d'un environnement virtuel est une autre modalité sensorielle dont la précision de la simulation a un impact sur le sentiment de présence ressenti [52]. Elle est caractérisée par la présence ou l'absence de sons ambiants ou diégétiques [36] et la possibilité de les localiser dans l'espace virtuel [59].

Il a aussi été démontré que la proprioception, simulée par le degré auquel le corps est suivi dans l'environnement virtuel et la latence de ce suivi a aussi un impact sur la présence [60].

En somme, la plupart des études incluses dans cette revue suggèrent que le nombre de modalités sensorielles simulées et la précision de leur simulation jouent un rôle déterminant dans le sentiment de présence. Toutefois, il semble que la relation entre le sentiment de présence et ces deux caractéristiques de l'immersion n'est pas linéaire. Le fait d'augmenter le nombre de modalités sensorielles simulées dans un environnement virtuel ou d'améliorer la précision de leur simulation ne cause pas nécessairement en un sentiment de présence plus élevé.

Par exemple, l'ajout de stimulations sensorielles incongrues dans le contexte d'une expérience virtuelle, comme l'ajout de musique dans un environnement qui n'en nécessite pas, peut n'avoir aucun effet significatif sur la présence [61]. De plus, une stimulation sensorielle inadéquate, comme une contradiction entre un stimulus visuel et un mouvement corporel [62] ou un conflit de vergence-accommodation [62], peut créer de l'incompatibilité sensorielle qui est liée à un phénomène appelé "cybermalaise" [39].

Le cybermalaise a généralement une influence négative sur le sentiment de présence [39]. Il s'agit d'une sorte de malaise du mouvement induit visuellement [39]. Il se manifeste généralement sous la forme de fatigue oculaire, de vertige ou de nausée [63]. Les cybermalaises sont considérés comme une réponse normale à des phénomènes physiologiques inhabituels et leur intensité est extrêmement variable entre les individus [64]. Ils sont caractérisés par une activation du système nerveux sympathique et une désactivation du système parasympathique [47].

Ainsi, on remarque que l'immersion est un facteur objectif du sentiment de présence qui relève d'une certaine complexité. Il est important de porter une attention particulière aux différentes modalités sensorielles simulées et à leurs interactions avec d'autres modalités pour éviter d'affecter négativement le sentiment de présence.

Généralement, les propriétés immersives d'un environnement virtuel ont un effet modéré à élevé sur le sentiment de présence [36]. Cependant, toutes les propriétés immersives n'ont pas le même niveau d'impact sur le sentiment de présence. La simulation de certaines propriétés immersives a un apport plus important que d'autres [36]. Notamment la taille du champ de vision, la vision stéréoscopique, la fréquence de présentation des images et le niveau de suivi, soit le nombre et la précision avec laquelle les membres des utilisateur.rice.s réels.les sont reportés dans l'environnement virtuel [36].

Idéalement, les propriétés immersives d'un environnement virtuel seraient calquées sur les valeurs de la sensation humaine. Cependant, étant donné la complexité de la simulation des différentes modalités sensorielles, il serait plus pragmatique de viser l'atteinte d'une fidélité fonctionnelle qu'une fidélité physique parfaite [36]. D'ailleurs, les utilisateur.rice.s d'environnements virtuels ont tendance a rapporter un haut niveau de présence lorsqu'ils considèrent la simulation de leur sensation est suffisante [52]. Cependant, on ne sait pas à quel point il est nécessaire de simuler chaque modalité sensorielle de manière réaliste pour atteindre un niveau de présence considéré suffisant. L'état actuel de la littérature scientifique ne permet pas de cibler les valeurs adéquates des propriétés immersives permettant une fidélité fonctionnelle.

Influence du traitement explicite de l'information sur la présence

Le modèle de traitement de l'information pour la présence en environnement virtuel [25] indique qu'à la suite de l'intégration multisensorielle des stimuli, l'information est traitée de manière explicite. Le traitement explicite de l'information consiste en un déterminant difficilement mesurable du sentiment de présence à cause de sa subjectivité. C'est lors de cette étape que les utilisateur.rice.s établissent la crédibilité d'un environnement virtuel. Ainsi, on considère qu'il est important d'identifier des manières d'évaluer le traitement explicite de l'information effectué par les utilisateur.rice.s d'environnement virtuel.

Les éléments liés au traitement explicite de l'information peuvent être associés au facteur de cohérence du sentiment de présence. La cohérence est composée de facteurs extrinsèques et intrinsèques aux utilisateur.rice.s d'un environnement virtuel. Les éléments extrinsèques correspondent au résultat de l'intégration multisensorielle des stimuli, tandis que les éléments intrinsèques correspondent aux attentes et aux expériences antérieures des utilisateur.rice.s [65]. Ainsi, la cohérence est liée à la manière dont le résultat de l'expérience de sitmuli générés dans un environnement virtuel sont en mesure de rencontrer les attentes de ses utilisateur.rice.s.

Jusqu'à un certain point, il est possible de contrôler objectivement certains éléments de la cohérence d'un environnement virtuel qui résultent de l'intégration multisensorielle de l'information. Il est possible de contrôler objectivement ces éléments de la cohérence en agissant sur leur disponibilité dans un environnement virtuel. Parmi ces éléments on retrouve notamment la plausibilité externe et la plausibilité interne [31].

La plausibilité externe fait référence au fait qu'un environnement virtuel soit consistant par rapport aux interprétations qu'ont les utilisateur.rice.s des règles du monde réel [31]. Pour avoir une plausibilité externe, un environnement virtuel doit produire des événements virtuels de la même manière qu'ils se seraient produits dans le monde réel. Ainsi, un environnement qui respecte les lois de la physique évoque généralement un plus grande sentiment de présence qu'un environnement qui ne les respecte pas [52]. De la même manière, le réalisme de l'illumination aura un impact sur le sentiment de présence [9].

Une manière de contrevenir à la plausibilité externe sans nécessairement affecter négativement le sentiment de présence résultant est d'établir une plausibilité interne. La plausibilité interne fait référence au fait qu'un environnement virtuel soit consistant par rapport aux interprétations qu'ont les utilisateur.rice.s des règles qu'un environnement virtuel a lui-même mises en place [31]. Pour avoir une plausibilité interne, un environnement virtuel doit être en mesure de répondre aux attentes qu'il a lui-même créé [31]. Grâce à la plausibilité interne,

il est possible de maintenir la cohérence d'un environnement virtuel qui violerait certaines règles du monde réel. Ainsi, un environnement virtuel qui respecte les règlements qui ont été explicitement énoncés généralement évoque un plus grand sentiment de présence qu'un environnement qui ne les respecte pas [52].

Cependant, la manipulation de la plausibilité d'un environnement virtuel est aussi fortement influencée par des éléments subjectifs liés à l'attention, aux attentes et aux expériences antérieures des utilisateur.rice.s. Ainsi, à cause de ces phénomènes, les manipulations de la plausibilité interne et externe affecteront le sentiment de présence des utilisateur.rice.s d'un environnement virtuel de manière variable [66].

Les éléments qui permettent la plausibilité d'un environnement virtuel ne sont pas les mêmes pour tous les utilisateur.rice.s [30]. Certains peuvent vouer de l'importance à certaines interactions, alors que d'autres peuvent ne pas voir l'intérêt de ces éléments dans l'environnement virtuel [30]. Ainsi, des différences individuelles au niveau des traits de personnalité affectent le sentiment de présence [39].

Il est donc important de trouver des indicateurs permettant de représenter les attentes des utilisateur.rice.s d'un environnement virtuel de manière à déterminer s'ils auront tendance à faire l'expérience subjective du sentiment de présence. Une manière de mesurer les attentes dans un environnement virtuel est d'analyser la réponse des utilisateur.rice.s aux événements qui s'y produisent. La réponse des utilisateur.rice.s à un environnement virtuel est considérée comme étant composée de plusieurs niveaux. Elle inclut des réponses psychophysiologiques, des réflexes comportementaux inconscients, des réponses comportementales volitives, des réponses émotionnelles et des réponses cognitives de haut niveau [9]. L'évaluation de ces variables permettrait notamment de déterminer à quel point les utilisateur.rice.s d'un environnement virtuel agissent comme dans la réalité.

La section 1.2.3 Méthodes d'évaluation du sentiment de présence a déjà montré le lien entre les réponses psychophysiologiques et comportementales et le sentiment de présence. Or, on ne connaît généralement pas la nature de la relation entre les facteurs de la présence, comme la cohérence et l'immersion. De plus, les mesures d'autres variables pourraient être d'intérêt pour dresser un meilleur portrait cognitif et émotionnel des utilisateur.rice.s d'un environnement virtuel. Ce portrait permettrait de mieux juger de la manière dont un environnement virtuel répond aux attentes d'un environnement virtuel.

De manière à étudier les réponses cognitives de haut niveau pour évaluer le sentiment de présence, il est possible d'étudier des indicateurs de performance cognitive comme la charge cognitive et la conscience de la situation.

La charge cognitive est liée à la présence puisqu'elle peut affecter le traitement de l'information, comme l'attention ou la perception visuelle [67]. Plusieurs études ont établi une corrélation significative entre charge cognitive et sentiment de présence [10, 37, 68, 69]. Ainsi, à l'instar des mesures comportementales et psychophysiologiques, la charge cognitive pourrait être utilisée pour évaluer si un scénario virtuel implique des fonctions cognitives à un niveau attendu. Similairement, la conscience de la situation est un autre indicateur de performance cognitive qui pourrait être utilisé pour évaluer l'implication de fonctions cognitives dans un scénario virtuel et donc déterminer le sentiment de présence. Comme le sentiment de présence, les mesures de conscience de la situation sont affectées par des facteurs comme les ressources attentionnelles, la compréhension de la situation, le type d'information véhiculée ou certains facteurs technologiques [70]. La conscience de la situation est une évaluation des processus décisionnels qui permettent à un individu d'atteindre ses buts. Ceci constitue un autre lien entre la conscience de la situation et le sentiment de présence, puisque ce dernier est augmenté lorsque les utilisateur.rice.s d'un environnement virtuel, sont aptes à atteindre leur but dans un scénario virtuel [27].

Toutefois le lien entre les indicateurs de performance cognitive comme la charge cognitive ou la conscience de la situation n'a pas été exploré en profondeur. Notamment, on ne sait pas par quels facteurs du sentiment de présence ils sont influencés.

1.3 Plan de la thèse

Cette thèse est organisée en huit chapitres. Le chapitre 1 a permis d'introduire le concept de sentiment de présence, d'en comprendre l'importance et surtout d'évaluer les zones de la littérature scientifique qui pourraient bénéficier d'un ajout de connaissance.

Le chapitre 2 présente une revue de la littérature scientifique permettant de mieux comprendre l'état des connaissances sur les éléments spécifiques liés à la problématique abordée. D'abord, on y aborde le sujet de la modalité visuelle dans le but de mieux identifier les qualités sensorielles qui la constituent et la relation de cette modalité avec le sentiment de présence. Par la suite, on établit l'état des connaissances du lien entre le sentiment de présence et la charge cognitive, les mesures psychophysiologiques et la conscience de la situation.

Le chapitre 3 définit les objectifs de recherche et présente généralement la méthodologie utilisée pour les atteindre.

Les trois chapitres suivants sont constitués des trois articles réalisés dans le cadre de cette thèse. Le chapitre 4 présente une expérience réalisée afin de mesurer le seuil fonctionnel de trois qualités visuelles, la sensibilité au contraste, la taille du champ de vision et l'acuité

visuelle. Le chapitre 5 présente une expérience faisant usage des seuils fonctionnels de manière à manipuler objectivement l'immersion et constater l'effet de cette manipulation sur des indicateurs psychophysiologiques et sur la charge cognitive. Le chapitre 6 présente l'effet de la même manipulation de l'immersion sur la conscience de la situation.

Le chapitre 7 propose une discussion générale portant sur les résultats obtenus dans les trois articles, leurs limites, leur portée et de futures perspectives de recherche. Finalement, le chapitre 8 présente la conclusion de cette thèse.

CHAPITRE 2 REVUE DE LITTÉRATURE

Ce chapitre établit l'état de la littérature scientifique liée à la problématique de recherche à laquelle s'intéresse cette thèse. On y présente trois modalités visuelles, leur pertinence pour cette recherche et leur lien avec le sentiment de présence. Ensuite, on y précise l'état des connaissances pour des indicateurs cognitifs et physiologiques du sentiment de présence, soit la charge cognitive, de l'activité oculaire, de l'activité cardiovasculaire et de la conscience de la situation, ainsi que leur lien avec le sentiment de présence.

2.1 Qualités visuelles

On peut faire la distinction entre deux phénomènes qui modulent l'expérience humaine en lien avec l'immersion : la sensation et la perception. La sensation est un intrant d'information capté directement par les organes sensoriels, elle constitue l'habileté à détecter un stimulus et de potentiellement transformer cette détection en une expérience subjective [51]. La perception consiste en un traitement de ces informations permettant une représentation mentale des intrants sensoriels, soit l'acte de donner une signification à une sensation détectée [51]. Ainsi, les systèmes supportant des environnements virtuels contrôlent les sensations et perceptions humaines de manière à créer un environnement artificiel dans lequel des utilisateur.rice.s peuvent agir.

La modalité visuelle profite d'un biais culturel et social ayant favorisé son étude et son utilisation [71]. Ainsi, la vision est la principale modalité sensorielle reproduite par les systèmes de réalité étendue servant à générer des environnements virtuels. Comme l'étude de l'ensemble des qualités visuelles dépasse la portée de cette thèse, celle-ci se concentre sur trois qualités visuelles, soit la sensibilité au contraste, la taille du champ de vision et l'acuité visuelle. Nous avons ciblé ces trois qualités visuelles pour diverses raisons. Le lien entre la sensibilité au contraste et le sentiment de présence a été peu exploré de manière appliquée, cette thèse vise donc à faire la lumière sur la relation entre ces deux concepts. A contrario, la taille du champ de vision a été identifiée comme l'une des qualités visuelles ayant la contribution la plus importante pour le sentiment de présence [36], il est donc jugé pertinent de l'étudier. De la même manière, l'acuité visuelle s'est avérée importante pour le sentiment de présence, justifiant l'intérêt qui lui est porté.

Aucun seuil fonctionnel n'a été établi pour ces trois qualités visuelles. Ainsi, on ne connaît pas les valeurs constituant une limite perceptuelle, soit une valeur en dessous de laquelle

la perception doit être considérée comme fonctionnellement dégradée. Par conséquent, cette section de la revue de littérature vise à définir et préciser différents attributs du champ de vision, de l'acuité visuelle et de la sensibilité au contraste, notamment leur seuil perceptuel. Ensuite, on tente d'y spécifier les paramètres nécessaires pour que le système visuel humain soit fonctionnel en environnement virtuel utilisant un visiocasque. Lorsque possible, cette section établit le lien entre les trois qualités visuelles ciblées et le sentiment de présence.

2.1.1 Champ de vision

L'expérience visuelle est une reconstruction du monde basée sur deux images distinctes provenant de la rétine de chaque œil. La portion de l'espace que les yeux sont en mesure de capter lorsqu'ils sont immobiles est appelée champ de vision [72].

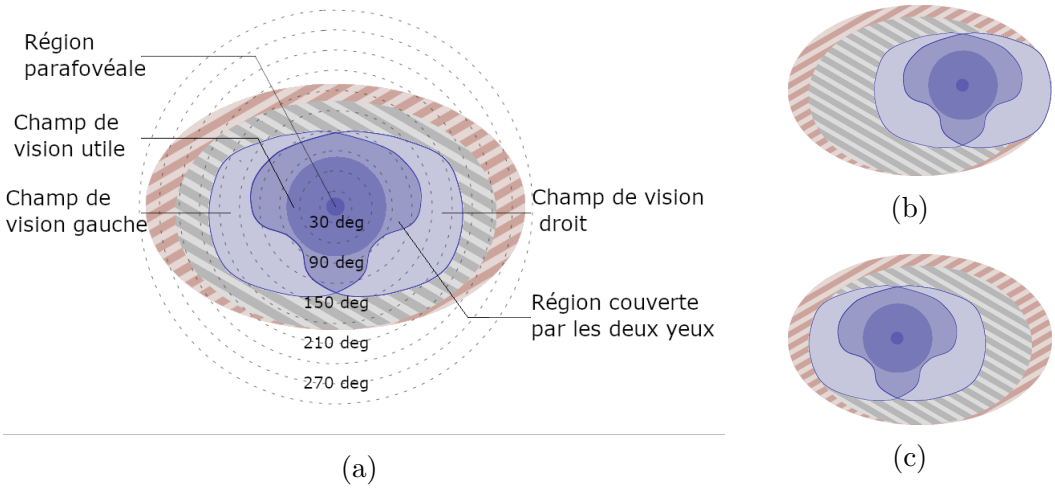


FIGURE 2.1 Représentation schématique de différentes portions du champ de vision

Note. À la figure a, les sections bleues remplies uniformément représentent la partie de l'espace qui est couverte par les yeux lorsqu'ils sont fixés sur un point, les régions rayées indiquent les parties du champ visuel qui peuvent être couvertes en déplaçant les yeux ; la figure b montre le champ de vision lorsque les yeux sont posés sur la limite droite du champ visuel ; la figure c montre le champ de vision lorsque les yeux sont posés sur la limite gauche de la région fovéale fixe.

La figure 2.1 montre une représentation des différentes portions du champ de vision. De manière autonome, un œil fixe a un champ de vision horizontal d'environ 150 [73] degrés et un champ de vision vertical de 120 à 140 degrés [51, 58, 73, 74]. Binoculairement, le champ de vision horizontal couvert lorsque les yeux sont fixes est plus grand. Lorsqu'on prend en compte les deux yeux, le système visuel humain couvre un champ horizontal d'environ 180 à 200 degrés [6, 51, 58, 74] dont 100 à 120 degrés sont couverts par les deux yeux en même

temps [6, 51, 74].

Dans un visiocasque, il est important de prendre en compte le fait que les yeux peuvent se déplacer dans le champ visuel présenté. Lorsque la tête est fixe, les yeux, en se déplaçant, peuvent couvrir une zone de 290 degrés horizontalement et 190 degrés verticalement [74]. Ces valeurs peuvent être considérées comme le seuil perceptuel du champ de vision. Toutefois, il est seulement possible de maintenir le regard dans la région fovéale fixe qui correspond à la zone comprise dans 40 à 50 degrés de la région centrale du champ de vision [75]. À titre de comparaison, le champ de vision horizontal de la plupart des visiocasques actuels se situe entre 90 et 110 degrés. Cette taille de champ de vision rend difficile les tâches où des individus doivent détecter des objets situés dans leur vision périphérique [6].

La littérature scientifique actuelle a montré que des valeurs plus proches de la taille totale du champ de vision évoquent une sentiment de présence plus élevée lorsque les facteurs confondants sont contrôlés [9,36,76,77]. Cette relation est d'ailleurs reflétée dans les mesures de performance [78]. Toutefois, certaines études comparant les effets de différentes tailles de champ de vision sur le sentiment de présence n'ont montré aucune différence significative entre les différentes tailles étudiées [52,79]. D'autres études, impliquant une illusion de mouvement dans l'environnement virtuel, ont montré que l'augmentation du champ visuel était plus propice à créer un conflit visuovestibulaire [39, 80] pouvant causer un cybermalaise [81, 82]. Puisque le cybermalaise est généralement négativement corrélé avec le sentiment de présence, une augmentation du champ de vision peut impliquer une diminution du sentiment de présence.

La partie du champ de vision utilisée par un observateur dépend de la tâche effectuée [83]. La performance visuelle n'est pas égale dans l'ensemble du champ de vision. Une augmentation de l'excentricité par rapport à la fovéa affecte la perception du mouvement [6,84,85], l'acuité visuelle [86–88], la sensibilité à la lumière et au contraste [88,89] et la perception des couleurs [85, 89–91]. Ces différences dans les performances de qualités visuelles à travers le champ de vision peuvent être responsables du fait que des fonctions différentes sont attribuées à différentes portions du champ visuel.

Par exemple, le concept de champ de vision utile décrit la fonction de la partie centrale du champ visuel. Le champ de vision utile est défini comme la surface autour d'un point de fixation à l'intérieur de laquelle les informations peuvent être perçues et traitées lors d'une tâche visuelle [92]. La taille du champ de vision utile varie en fonction de la tâche exécutée [93]. Généralement, elle est estimée à un diamètre de 30 à 40 degrés autour du point central du regard [94]. Toutefois, les études ayant évalué l'étendue du champ de vision utile ont tendance à négliger la contribution des informations périphériques pouvant être acquises [94]. En fait,

le champ de vision utile pourrait couvrir un diamètre s'étendant jusqu'à 80 degrés autour du point de fixation [95].

En résumé, on remarque que la taille du champ de vision est un facteur de l'immersion qui est associé au sentiment de présence. On peut identifier un seuil perceptuel autour de 290 degrés horizontalement et 190 degrés verticalement représentant la taille du champ de vision maximal qu'il est possible de couvrir lorsque la tête est fixe. L'état actuel de la littérature scientifique ne permet toutefois pas d'établir un seuil fonctionnel. Pour des tâches qui nécessitent une acquisition de données liées à des éléments présents dans la vision périphérique, la taille du champ de vision de l'actuelle génération de visiocasque est théoriquement trop petite et rend ce genre de tâche ardue. Toutefois, en se basant sur le concept de champ de vision utile, la taille du champ de vision offerte par les visiocasques apparaît suffisante. Donc, plus de recherche est nécessaire afin d'établir un seuil fonctionnel.

2.1.2 Acuité visuelle

L'acuité visuelle désigne le niveau de détail le plus fin qui peut être détecté ou identifié par l'œil [96]. Elle est physiologiquement déterminée par l'espacement entre les photorécepteurs sur la rétine [51]. Pour la mesurer de manière pertinente en environnement virtuel, on utilise généralement la notion d'acuité de résolution ou de minimum séparable, faisant référence à la plus petite séparation angulaire entre deux objets voisins qui peut être détectée [51].

L'acuité de résolution dépend des conditions de l'environnement. Pour une acuité maximale, on suppose une image stable affichée à l'intérieur de 0,5 degré du centre fovéal, avec un contraste maximal et des conditions d'illumination photopiques [62]. Dans ces conditions, l'œil humain a une acuité visuelle permettant de distinguer deux points entre 0,5 arcmin et 2 arcmin avec une moyenne de 1 arcmin [58, 73, 83, 97, 98].

L'acuité visuelle avec un visiocasque est moindre lorsqu'on la compare objectivement à l'acuité réelle [99, 100]. Ceci est dû à une densité de pixel trop faible pour reproduire l'acuité visuelle réelle [100]. Théoriquement, l'acuité visuelle limitée ne permet pas l'exécution en environnement virtuel immersif de toutes les tâches qu'il serait possible de faire dans un environnement réel [100]. La résolution angulaire de 4 à 6 arcmin de la plupart des visiocasques rend les objets difficiles à discerner à des distances où ils seraient clairement identifiés dans la vie réelle [6].

La relation entre la résolution et le réalisme perçue d'un environnement virtuel n'est reflétée qu'en partie dans la mesure du sentiment de présence [101]. Une plus grande résolution est liée à une augmentation significative du sentiment de présence [57, 77]. Cependant, ces

résultats ont été obtenus avec une grande différence entre les conditions de haute et de basse résolution, et il semble qu'une plus petite différence entre les conditions ne donne pas de résultats significatifs [102]. Les humains voient la plupart du monde à travers leur vision périphérique. Celle-ci ne permet pas de percevoir les images aussi nettement que la région fovéale, ce qui expliquerait pourquoi on ne trouve pas toujours de différence significative entre deux conditions de résolutions [103].

En fait, l'acuité visuelle n'est pas constante dans l'ensemble de la rétine. Elle est maximale dans la fovéa, région centrale couvrant jusqu'à environ 2 degrés d'excentricité, et diminue avec une augmentation de l'excentricité. Pour des stimuli fixes, l'acuité serait environ de 2,8 arcmin à 8 degrés d'excentricité, de 4,2 arcmin à 11 degrés et de 7,0 arcmin à 13 degrés [86]. L'œil a une sensibilité de 12 arcmin à une excentricité de 25 degrés de la fovéa [87]. Pour certains stimuli, un espacement jusqu'à 15 arcmin ne peut être discriminé à 30 degrés d'excentricité [85].

Les utilisateur.rice.s d'environnement virtuel ont tendance à surévaluer subjectivement l'acuité visuelle [58]. Ceci pourrait être dû à l'anticrénelage qui a pour effet d'améliorer le confort visuel [83] et de mitiger une partie de la perte de perception due à une mauvaise résolution [104], améliorant notamment la perception des distances [105]. Toutefois, cette technique ne permet pas de compenser le manque d'information qui nécessite une haute résolution pour être perçue [83].

En somme, on remarque qu'un lien entre acuité visuelle et sentiment de présence a été établi. Dans des conditions de vision optimales, le seuil perceptuel de cette qualité visuelle est de 1 arcmin. Toutefois, l'état actuel de la littérature scientifique ne permet pas d'identifier précisément un seuil fonctionnel.

2.1.3 Intensité lumineuse et contraste

Lorsqu'on traite du système visuel, l'intensité lumineuse correspond aux nombres de photons entrant en contact avec les photorécepteurs tapissant la rétine [51]. La sensibilité à l'intensité lumineuse est un phénomène qui varie dans le temps et qui dépend des conditions d'éclairage antérieures [106]. En raison d'un encodage effectué par les circuits neuronaux, l'œil humain est sensible aux différences d'intensité lumineuses plutôt qu'à l'intensité lumineuse moyenne [51]. Ce mécanisme explique pourquoi la vision est plus sensible aux contrastes d'intensité lumineuse entre les différentes aires de la rétine plutôt qu'à l'intensité lumineuse générale [51]. Ainsi, pour prendre en compte la variation temporelle et aborder cette qualité visuelle d'une manière qui permette l'identification d'un seuil perceptuel indépendant du contexte, le concept de sensibilité au contraste est employé plutôt que celui d'intensité lumineuse.

Le contraste correspond à l'amplitude de la variation de la luminance dans un stimulus normalisé par rapport à la luminance de l'arrière-plan, soit une quantification des différences d'intensité lumineuse d'une image [96]. Donc, la sensibilité au contraste est équivalente à la plus petite différence de luminance requise pour identifier une cible et est dépendante de la fréquence spatiale [96]. Elle est généralement exprimée comme le rapport entre la luminance faible et la luminance élevée d'une image. La sensibilité au contraste dépend de la fréquence spatiale [96], des conditions d'éclairage [107], et de la taille du champ de vision [72].

Pour la sensibilité au contraste, il existe un seuil perceptuel de 1 % qui est indépendant des conditions visuelles [108]. Toutefois, dans des conditions optimales, le contraste minimum qui peut être perçu dans la fovéa est de 0,2 % [88]. Pour un visiocasque avec un champ de vision de 110 degrés et une résolution équivalente à 4,6 arcmin, cette valeur passe à 0,45 % [107].

La sensibilité au contraste en périphérie est plus basse que la sensibilité au contraste fovéale [88]. Pour un stimulus crénelé, elle serait comprise entre 19 et 23 % à 30 degrés d'excentricité [88].

La sensibilité au contraste des utilisateur.rice.s de visiocasque est moins bonne que dans un environnement réel [99]. La diminution de la sensibilité au contraste serait due à une gamme limitée de luminances affichées [109], d'aberrations chromatiques et d'une résolution insuffisante [110]. Néanmoins, on peut raisonnablement s'attendre à ce que la gamme de contrastes offerte par les visiocasques soit suffisante pour que les utilisateur.rice.s se sentent présents dans un environnement virtuel.

Cette revue de littérature n'a pas été en mesure d'identifier des études ciblant l'effet de divers niveaux de contraste sur la présence ou la performance dans un environnement virtuel. Les études ayant établie son importance pour le sentiment de présence se sont basées sur des notions théoriques [19], ou sur des expérimentations évaluant la qualité des images dans leur ensemble [36].

Une étude a évalué l'effet de différents niveaux de contraste sur le réalisme perçu des images [111]. Cependant, comme les résultats rapportés dépendent de l'image utilisée, ils ne peuvent pas être utilisés comme une indication directe pour établir un seuil fonctionnel.

En somme, on note qu'un lien entre sentiment de présence et sensibilité au contraste a été établi théoriquement. Le seuil perceptuel de cette qualité visuelle est de 0,2 % dans des conditions visuelles idéales. Toutefois, l'état des connaissances est insuffisant pour identifier une valeur pouvant servir de seuil fonctionnel de la sensibilité au contraste.

Ainsi, on connaît les seuils perceptuels, soit les valeurs maximales pour lesquelles le système visuel humain est en mesure de percevoir. Cependant, comme le démontre l'utilisation actuelle

de la technologie de réalité étendue, l'humain n'est pas constamment en train de percevoir au maximum de ses capacités. Il est donc pertinent de se demander quelles sont les valeurs minimales, soit les seuils fonctionnels nécessaires pour permettre à l'humain de percevoir dans un environnement virtuel.

2.2 Charge cognitive

Cette section présente le concept de charge cognitive, les méthodes utilisées pour le mesurer, ainsi que son lien avec le sentiment de présence.

La charge cognitive est définie comme l'expérience phénoménologique d'individus qui affecte leurs comportements subséquents, soit leur performance et leur réponse physiologique à une situation [112]. Les attributs contribuant à l'expérience de la charge cognitive varient selon les tâches et les individus, puisque la charge cognitive n'est pas définie de manière unique par les qualités objectives de la tâche, mais reflète aussi la réponse des individus à la tâche [112].

Autrement dit, la charge cognitive fait référence au niveau de ressources attentionnelles requises pour répondre à des critères de performance objectifs et subjectifs, qui peut être modulé par les exigences de la tâche, les expériences antérieures des individus et d'autres facteurs externes [113].

Il est possible de distinguer trois types de charge cognitive : la charge intrinsèque, la charge extrinsèque et la charge utile. La charge intrinsèque est définie comme la charge induite par la nature des éléments traités, par exemple, la difficulté d'une tâche [114]. La charge extrinsèque est définie comme les facteurs externes à l'individu comme l'organisation du travail ou la pression due au temps [114]. Finalement, la charge utile est définie comme la charge nécessaire à une personne pour l'exécution de ses processus cognitifs, par exemple, accéder à la mémoire de travail, former des schémas cognitifs et automatiser leur application [115]. La charge utile serait un témoin de l'expertise puisqu'elle évalue la capacité à réaliser des tâches en appliquant des stratégies spécifiques qui permet de réduire le coût de traitement de l'information [116]. Ces trois types de charge cognitive sont additifs et constituent la charge cognitive totale [114].

L'environnement dans lequel une tâche est effectuée a un impact considérable sur la charge cognitive [117]. L'environnement inclut les caractéristiques des éléments utilisés dans l'accomplissement d'une tâche, les attributs de l'environnement et la présence d'autres personnes [117]. L'environnement affecte à la fois les individus effectuant une tâche et la nature de la tâche [117]. Cette affirmation est supportée par la théorie de la cognition distribuée qui suggère que les processus cognitifs sont distribués dans l'environnement physique de travail [117].

2.2.1 Mesure de la charge cognitive

La mesure de la charge cognitive permet d'estimer d'où proviennent les augmentations de demande dans une tâche menant à des performances jugées inacceptables. Elle peut ainsi être utilisée pour évaluer la désirabilité d'un système, surtout lorsque d'autres mesures ne permettent pas de faire un choix éclairé [118].

On classe les techniques de mesure de la charge cognitive en trois catégories : les mesures subjectives, les mesures de performances et les mesures psychophysiologiques [118].

Les mesures subjectives sont des mesures de la charge moyenne ou de la charge totale, dépendant du moment de la mesure [114]. La charge moyenne est l'intensité moyenne de la charge lors de l'exécution d'une tâche [119]. Tandis que la charge totale est la charge cognitive pour une procédure complète englobant plusieurs tâches [119]. Les mesures subjectives de la charge cognitive peuvent être divisées en deux catégories, les mesures unidimensionnelles et les mesures multidimensionnelles. Les mesures unidimensionnelles sont généralement plus faciles et rapides à utiliser [120] à défaut d'être moins sensibles et d'avoir moins de diagnosticités que les échelles multidimensionnelles [118, 121]. L'une des échelles multidimensionnelles les plus utilisées est le questionnaire NASA-TLX, dont le score représente le niveau de charge cognitive d'une personne [112].

Les mesures de performance sont considérées comme un indicateur de la charge maximale [114], soit la valeur maximale que peut atteindre la charge instantanée lorsqu'une tâche est exécutée [119]. Le niveau de performance tend généralement à diminuer lorsque la charge cognitive est trop élevée [117]. Cependant, les mesures de performances ne peuvent être considérées comme suffisantes à elles seules, puisqu'elles ne permettent pas de juger adéquatement de l'effort mental requis par une tâche [118]. De plus, elles manquent de rigueur scientifique pour l'interprétation de leurs résultats [118].

Les mesures psychophysiologiques consistent en une évaluation objective et continue de l'état mental d'une personne [118]. Elles seraient des représentations de la tendance dynamique de la charge cognitive et permettraient donc d'estimer la charge instantanée [114], soit la quantité de ressources mentales disponibles à un point précis dans le temps [119]. Toutefois, elles produisent des résultats ambigus et difficiles à interpréter si elles sont utilisées seules, car elles peuvent être attribuées à plusieurs phénomènes, comme une augmentation de la charge cognitive, mais aussi du stress [121]. Parmi les mesures psychophysiologiques liées à la charge cognitive, on retrouve notamment l'analyse de l'activité cardiovasculaire, de l'activité oculaire et l'analyse respiratoire [121].

On recommande l'utilisation de plusieurs types mesures pour pouvoir poser des hypothèses

sur les différences relatives à la charge cognitive utilisée [121, 122]. Par exemple, le fait de coupler une évaluation subjective de la charge cognitive à une mesure de la variabilité de la fréquence cardiaque, permet d'évaluer les ressources cognitives restantes et donc d'évaluer la charge cognitive effective [115].

2.2.2 Évaluation de la charge cognitive en environnement virtuel

Des études antérieures ont montré un lien entre le sentiment de présence et la charge cognitive [37, 68, 69]. Toutefois, les résultats rapportés reportent des corrélations contradictoires. Des expériences portant sur une tâche de neutralisation de mines terrestres [37] et une tâche de manoeuvre d'infanterie militaire [68] ont indiqué une corrélation négative significative entre la charge cognitive et la présence. Cependant, dans une tâche de lancer franc au basket-ball, la charge cognitive était significativement corrélée positivement avec la présence [69]. Ainsi, la charge cognitive ne doit pas être traitée comme un indicateur de présence absolu.

La charge cognitive devrait plutôt être utilisée pour évaluer si deux scénarios font intervenir des fonctions cognitives de manière similaire et avec la même intensité [10]. De cette façon, la charge cognitive peut être utilisée pour comparer les performances d'utilisateur.rice.s dans des environnements faisant appel à différents qualia. Cette méthode a par exemple été utilisée dans une comparaison réelle et virtuelle d'interactions avec des robots [123], pour comparer divers simulateurs de conduite automobile [10, 124] et pour comparer l'expérience de jeux vidéos sur console et en réalité virtuelle immersive [125]. On remarque que l'utilisation de la charge cognitive pour évaluer le sentiment de présence permet notamment d'évaluer le réalisme des environnements virtuels sur le plan cognitif en diagnostiquant les différences entre environnements réels et virtuels. Toutefois, l'effet d'une manipulation des facteurs du sentiment de présence, soit l'immersion et la cohérence, ne semble pas avoir été étudié. Ainsi, il n'est pas encore possible de déterminer de manière claire la relation entre la charge cognitive et les facteurs du sentiment de présence.

2.3 Mesures psychophysiologiques

Les mesures psychophysiologiques représentent une opérationnalisation des états internes d'une personne, dont la qualité peut varier d'une mesure à l'autre, et entre différents états [38]. En somme, les mesures psychophysiologiques peuvent être utilisées pour évaluer objectivement l'activité cognitive et les réponses émotionnelles d'individus lors de leurs interactions [38]. Les mesures psychophysiologiques présentent l'avantage de fournir une évaluation continue, relativement non invasive, quasi synchrone et objective de plusieurs phénomènes

dont le sentiment de présence [40]. De plus, les mesures peuvent être standardisées et plus facilement comparées entre différentes études. Elles nécessitent un échantillon relativement petit pour évaluer différents phénomènes cognitifs [113].

Toutefois, les mesures psychophysiologiques sont affectées par les caractéristiques de la tâche et les différences individuelles [113]. De plus, une des difficultés majeures liées à l'utilisation de ces mesures réside dans la complexité de l'inférence psychophysiologique [38]. En effet, plusieurs phénomènes peuvent impacter les mesures psychophysiologiques, notamment des états cognitifs comme le stress, la charge cognitive ou la somnolence [49], des réponses émotionnelles comme la frustration et l'anxiété [38] ou des états cognitifs et émotionnels plus complexes comme le cybermalaise et le sentiment de présence [39, 47, 48, 126]. De plus, le fait qu'elles constituent des mesures intermédiaires de la présence est problématique vis-à-vis de leur évaluation et de leur interprétation étant donnée qu'elles dépendent de la démonstration de l'hypothèse nulle.

Dans cette thèse, on s'intéresse à deux phénomènes physiologiques pouvant servir d'indicateurs du sentiment de présence, soit l'activité oculaire et l'activité cardiovasculaire.

2.3.1 Activité oculaire

L'activité oculaire fait référence aux différents événements oculaires donnant un aperçu de l'état cognitif et émotif d'une personne [127]. Les différents types d'activité oculaire sont généralement associés à des facteurs intrinsèques à l'individu comme sa charge cognitive et à des facteurs extrinsèques tels que la difficulté de la tâche [127]. L'évaluation de l'activité oculaire inclue notamment l'étude des événements oculomoteurs [127], la mesure et l'analyse des mouvements visuels [127, 128] et les mesures pupillaires [127, 129].

Parmi les mesures de l'activité oculaire, la mesure de l'activité pupillaire est liée au sentiment de présence [13]. L'activité pupillaire permet de mesurer le sentiment de présence étant donnée sa capacité à mesurer la distribution de l'attention [130] et d'explorer l'état cognitif et émotif d'individus [131].

Les réponses pupillaires agissent à titre d'indicateur de la capacité de traitement de l'information plutôt que de la demande absolue [132]. En fait, la taille de la pupille reflète les activités liées à l'effort cognitif, l'attention et la réponse aux stimuli émotionnels [129]. Plus précisément, des phénomènes cognitifs comme la somnolence cause une diminution du diamètre de la pupille [49, 133, 134] et d'autres comme une hausse de la charge cognitive cause une augmentation du diamètre de la pupille [49, 133]. En fait, la réponse pupillaire constituerait l'un des indicateurs physiologiques les plus sensibles‘à l'augmentation de la charge

cognitive [135]. La réponse pupillaire peut être mesurée par l'index d'activité pupillaire basse et élevée (LHIPA) [136].

Toutefois, plusieurs phénomènes peuvent moduler les réponses pupillaires. Les réponses émotionnelles interagissent notamment avec les phénomènes cognitifs. Ainsi, des stimulations associées à des émotions négatives de forte intensité rendent la dilatation des pupilles particulièrement sensible à la charge cognitive alors que des stimulations associées à des émotions positives de forte intensité la rendent moins sensible à la charge cognitive [129]. De plus, la dilatation de la pupille est affectée par des phénomènes physiologiques comme les mouvements oculaires et la position relative du regard [49]. Les mesures de la pupille sont aussi affectées par des phénomènes extrinsèques aux individus comme les conditions d'éclairage et la distorsion de la pupille lorsqu'elle se situe hors de l'axe de mesures des caméras de suivi oculaire [137]. De plus, les réponses pupillaires sont plus faibles chez les participant.e.s ayant une plus grande expertise [132]. Si les tâches sont trop difficiles pour les participant.e.s, la dilatation des pupilles peut diminuer [132].

2.3.2 Activité cardiovasculaire

Les réponses cardiovasculaires sont régulées par les systèmes nerveux sympathique et parasympathique et sont donc sensibles aux événements psychophysiologiques [138]. Les moyens de quantifier les réponses cardiovasculaires comprennent, sans s'y limiter, les mesures de base de la fréquence cardiaque et les mesures liées à la variabilité de la fréquence cardiaque.

Fréquence cardiaque

Les mesures de base de la fréquence cardiaque sont caractérisées par la relation entre le temps et les battements cardiaques. Elles consistent en une analyse du complexe QRS correspondant aux différentes étapes de dépolarisation et de contractions des ventricules du cœur. La fréquence cardiaque représente le nombre de battements du cœur pendant un temps donné [139].

Des études ont démontré une relation entre la fréquence cardiaque et le sentiment de présence. Toutefois, il existe des discordances dans les résultats présentés. Dans un environnement virtuel où des personnes étaient exposées à une situation anxiogène, la fréquence cardiaque était corrélée positivement avec le sentiment de présence [140]. Au contraire, dans un environnement virtuel représentant un simulateur de vol, la fréquence cardiaque était négativement corrélée au sentiment de présence [141]. Ainsi, bien qu'elle soit en mesure de refléter la force du sentiment de présence dans certains cas [41], la réaction des personnes à un environnement virtuel est souvent imprédictible [141]. La fréquence cardiaque devrait plutôt être

utilisée pour comparer les réactions physiologiques et cognitives entre environnements [10]. La fréquence cardiaque a par exemple été utilisée pour comparer un environnement virtuel utilisant un visiocasque et un autre sans visiocasque [142]. De cette manière, il est possible de déterminer si les réactions psychophysiologiques correspondent à celles qui sont attendues.

De plus, les mesures de la fréquence cardiaque sont liées à d'autres concepts qui ont une influence sur le sentiment de présence. Les mesures de la fréquence cardiaque ont été validées pour discriminer la charge cognitive dans divers domaines de recherche et sont donc recommandées pour cette utilisation [113]. Par exemple, il a été démontré que l'augmentation de la fréquence cardiaque est associée à une augmentation constante de la charge cognitive pendant les tâches de conduite en situation réelle et en simulateur [10, 49, 143]. Néanmoins, les mesures de base de la fréquence cardiaque sont généralement omises dans les analyses, ce qui peut être problématique en raison de leur capacité à distinguer des conditions de charge de travail spécifiques [143]. Elles semblent notamment présenter de nets avantages par rapport aux mesures de la variabilité de la fréquence cardiaque en ce qui concerne la sensibilité à la charge cognitive pour une tâche dont la charge cognitive est relativement faible [143].

Il a aussi été montré que la fréquence était sensible à la manipulation du facteur cohérence du sentiment de présence [52]. Cependant il n'est pas clair si la fréquence cardiaque est seulement liée à la cohérence d'un environnement virtuel, ou si l'immersion peut aussi l'impacter [52].

Variabilité de la fréquence cardiaque

La variabilité de la fréquence cardiaque est décrite comme la variation des intervalles de temps entre des battements cardiaques adjacents [144]. C'est un indicateur psychophysiologique qui est fonction des ressources cognitives disponibles d'une personne [115]. Les méthodes standard de mesure de la variabilité de la fréquence cardiaque comprennent les méthodes du domaine temporel, les méthodes du domaine fréquentiel et les méthodes non linéaires [145]. Les méthodes temporelles et fréquentielles sont des moyens complémentaires pour caractériser le même ensemble de variances. Les méthodes du domaine fréquentiel sont agrégées dans le temps et les méthodes du domaine temporel sont agrégées dans les fréquences [138].

De la même manière que la fréquence cardiaque, la variabilité de la fréquence cardiaque est un indicateur du sentiment de présence [40]. La variabilité de la fréquence cardiaque a d'ailleurs été utilisée pour faire la distinction entre le sentiment de présence évoquée par plusieurs scènes d'un environnement virtuel [146]. La variabilité de la fréquence cardiaque devrait donc être utilisée pour comparer les réactions physiologiques et cognitives engendrées par divers environnements virtuels [39]. Toutefois, la variabilité de la fréquence cardiaque est affectée par plusieurs phénomènes, il est donc requis d'être prudent avec les conclusions

tirées de son utilisation. La variabilité du rythme cardiaque augmente avec la somnolence, la fatigue et le désengagement et diminue avec la charge cognitive, le stress et la vigilance [49]. La variabilité de la fréquence cardiaque peut aussi varier au fil du temps. Il a notamment été démontré que le désengagement ou la démotivation lors de l'exécution d'une tâche avait pour effet d'augmenter la variabilité de la fréquence cardiaque [49].

Les mesures de la variabilité de la fréquence cardiaque sont aussi des indicateurs des ressources cognitives disponibles [115] et ont été validées pour discriminer la charge cognitive dans divers domaines de recherche. Elles sont donc recommandées pour cette utilisation [113]. Toutefois, dans les environnements virtuels, elles sont sensibles à des phénomènes comme le cybermalaise [147, 148]. De plus, l'utilisation de mesures de la variabilité de la fréquence cardiaque requiert un environnement imposant une charge cognitive relativement élevée. Les métriques de la variabilité de la fréquence cardiaque ont été utilisées pour évaluer la charge cognitive lors de tâches de conduite automobile. D'après Mehler et al. [143], la moyenne quadratique des intervalles successifs (RMSSD) et la distribution des basses fréquences (LF) peuvent effectuer une différenciation statistiquement significative, tandis que le écart-type des intervalles (SDNN), la proportion des intervalles successifs différant de plus de 50 ms (pNN50), la distribution des hautes fréquences (HF) et le rapport des distributions des hautes et basses fréquences (LF/HF) n'ont pas réussi à différencier les conditions. Stuiver et al. [149] ont utilisé la LF et la HF pour différencier les conditions lors d'une tâche de conduite automobile. Les deux étaient plus faibles en cas de charge cognitive plus élevée, mais ne différaient pas significativement entre les conditions. Shakouri et al. [150] ont rapporté une différenciation significative des conditions de charge cognitive telles que mesurées par NASA-TLX, tandis que la RMSSD, la LF, la HF et le LF/HF n'ont pas donné de résultats significatifs. Enfin, dans un contexte aéronautique, Alaimo et al. [151] ont indiqué que les mesures de la variabilité de la fréquence cardiaque étaient cohérentes avec les résultats obtenus avec NASA-TLX. Leurs résultats n'étaient pas significatifs, mais les auteurs ont signalé que leurs conditions avec un LF/HF plus élevé, une RMSSD plus faible et un SDNN correspondaient à un score NASA-TLX plus élevé.

En somme, on remarque que des travaux précédents ont montré une relation entre la variabilité de la fréquence cardiaque et le sentiment de présence. Toutefois, ceux-ci n'ont pas exploré le lien existant avec ses facteurs, soit l'immersion ou la cohérence.

2.4 Conscience de la situation

Cette section présente le concept de conscience de la situation, les méthodes utilisées pour le mesurer, ainsi que son lien avec le sentiment de présence.

La conscience de la situation est définie comme la perception des éléments présents dans un environnement circonscrit dans un volume temporel et spatial, la compréhension de leur signification et la projection de leur état dans un futur rapproché [152]. La conscience de situation permet de quantifier les processus cognitifs d'individus de manière à évaluer leur réponse comportementale dans un environnement dynamique. C'est un concept visant à décrire les éléments menant à une prise de décision permettant d'expliquer la performance d'une personne lors de l'exécution d'actions [152]. Ce concept peut être considéré comme le processus permettant de devenir conscient des éléments dans un environnement, comme le produit de cette prise de conscience ou comme une combinaison des deux [153].

La figure 2.2 présente le modèle de la conscience de la situation de Endsley [152]. Celui-ci représente la relation entre la conscience de la situation, la prise de décision et la performance résultant de cette prise de décision [152]. Selon ce modèle, la conscience de la situation est composée de trois niveaux ascendants. Le premier niveau correspond à la perception du statut, des attributs et de la dynamique des éléments présents dans l'environnement [152]. Le deuxième niveau est la compréhension de la situation basée sur une synthèse des différents éléments perçus au niveau 1 [152]. Le troisième niveau est la capacité d'anticiper les actions des éléments de l'environnement dans un futur rapproché en se référant aux informations provenant des niveaux précédents [152].

La conscience de la situation a un effet modéré sur la performance [154] et est donc considérée comme un déterminant de la performance lors de l'exécution d'une tâche [153]. Une mauvaise conscience de la situation augmente les chances de commettre des erreurs de performances [155]. Toutefois, une bonne conscience de la situation ne garantit pas un succès dans l'exécution d'une tâche, comme une mauvaise conscience de la situation ne garantit pas une mauvaise performance [153]. L'inconsistance de la conscience de la situation comme déterminant de la performance serait due au fait qu'elle est sujette à beaucoup de différences individuelles [155].

En fait, la conscience de la situation est influencée par plusieurs éléments dont des facteurs individuels, des attributs de la tâche exécutée et des facteurs influencés à la fois par la tâche et l'individu l'exécutant. Les facteurs individuels incluent notamment l'âge [155], les attributs cognitifs, les mécanismes de traitement de l'information, l'expérience et la formation. Les attributs de la tâche exécutée sont par exemple la complexité et le niveau d'automatisation de son interface utilisateur [152]. Finalement, les facteurs influencés à la fois par la tâche et l'individu l'exécutant consistent en des phénomènes comme le stress ou la charge cognitive [153, 156].

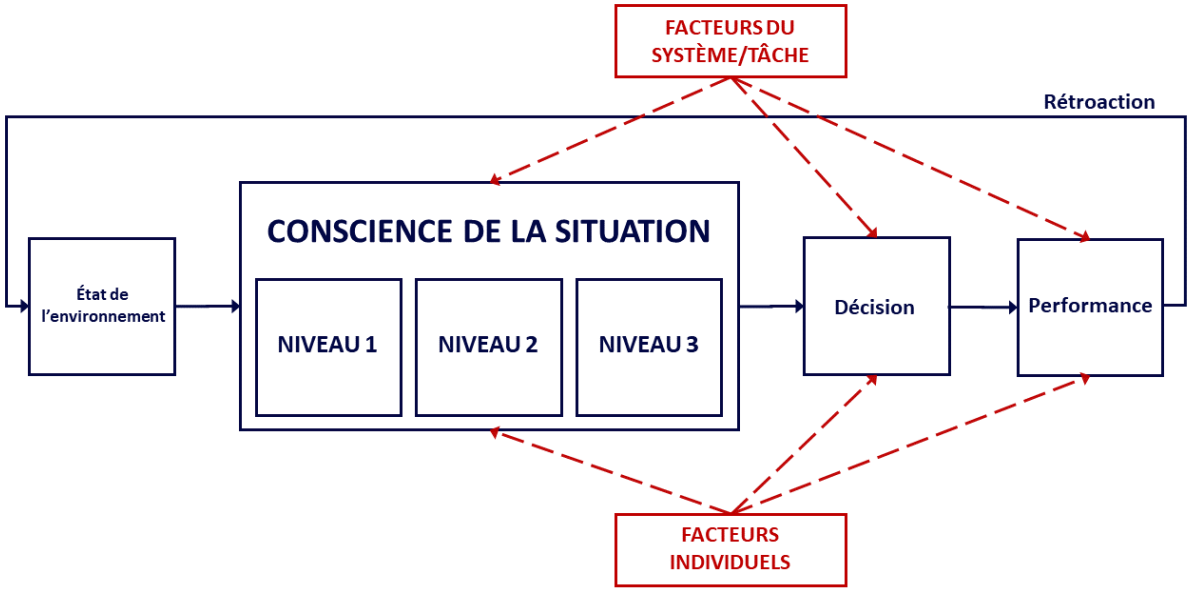


FIGURE 2.2 Modèle de la conscience de la situation

2.4.1 Mesure de la conscience de la situation

Deux catégories de mesures existent pour évaluer la conscience de la situation, soit les mesures indirectes et les mesures directes [156].

Les mesures indirectes consistent à inférer un niveau de conscience de la situation à partir des résultats des performances ou de traces laissées lors de l'exécution d'une tâche [157], comme l'analyse de protocoles verbaux ou l'évaluation de mouvements oculaires [153]. Les mesures indirectes sont utiles pour l'évaluation de la conscience de la situation lorsqu'une tâche est clairement définie et qu'elle possède des métriques de performance valides et sensibles [158]. Autrement, elles sont peu efficaces et peuvent être perturbées lorsque les sujets n'agissent pas comme prévu [157]. De plus, les mesures indirectes manquent de sensibilité et de diagnosticité pour évaluer la conscience de la situation [155]. Elles permettent plutôt de juger des processus sous-jacents de la conscience de la situation comme la perception d'informations pertinentes et l'attention portée à celles-ci [155].

Les mesures directes visent à évaluer les différents aspects de la conscience de la situation de manière subjective ou objective [156]. Les mesures subjectives se présentent généralement sous la forme de questionnaires utilisant des échelles de Likert remplis par les personnes dont on veut mesurer la conscience de la situation [159]. Elles sont souvent critiquées sur le plan de leur validité et de leur sensibilité [160]. Elles consisteraient plutôt en une évaluation de la

confiance des opérateurs en leur propre conscience de la situation [159].

Les méthodes directes et objectives d'évaluation de la conscience de la situation consistent à questionner les individus sur des éléments liés à leur conscience de la situation lors de l'exécution d'une tâche [159]. Les questionnaires utilisés lors de ces évaluations se basent sur des analyses de but permettant d'établir l'ensemble des éléments pertinents à la réussite de la tâche [161]. Les réponses récoltées sont comparées à la situation réelle et la proportion de bonnes réponses permet d'établir un score reflétant la qualité de la conscience de la situation [159]. Les méthodes objectives sont généralement plus compliquées à mettre en œuvre que les méthodes subjectives, mais elles offrent une meilleure diagnosticité, une meilleure sensibilité [155] et une plus grande validité des résultats [159]. Parmi les méthodes objectives de mesure de la conscience de la situation, la méthode SAGAT est l'une des plus répandues et validées [160]. SAGAT est déterministe, linéaire, significativement corrélée à la performance à une tâche [153] et moins sujette au biais d'échantillonnage que d'autres méthodes semblables [157].

2.4.2 Conscience de la situation et sentiment de présence

La conscience de la situation semble être une bonne candidate pour mesurer le sentiment de présence, car elle permet d'évaluer la similarité des réponses cognitives et comportementales entre deux environnements de manière objective. Théoriquement, les mesures de conscience de la situation sont sensibles à plusieurs éléments déterminants du sentiment de présence. Cependant, peu d'études appliquées ont fait l'évaluation de la relation entre la conscience de la situation et le sentiment de présence.

Les mesures de la conscience de la situation permettent d'évaluer les facteurs qui contribuent au sentiment de présence tels que la demande et l'offre de ressources attentionnelles, la compréhension de la situation, les types d'information utilisée et les facteurs technologiques [70]. De plus, une évaluation directe et objective de la conscience de la situation semble se rapporter à la modélisation de la présence basée sur les intentions et les actions de Riva et al. [27]. Cette modélisation indique que pour étudier le sentiment de présence, il est nécessaire d'intégrer une perspective cognitive en analysant comment les actions sont planifiées et contrôlées en réponse aux conditions environnementales et aux motivations des individus. Ainsi, pour évaluer le sentiment de présence, il est nécessaire d'étudier les activités exécutées, soit les ensembles d'actions effectuées de manière consciente et planifiée afin d'atteindre un but fixé. Pour ce faire, on recommande de décomposer l'activité de l'utilisateur.rice en ses différents composants [27]. Cette méthode d'évaluation est analogue à celle appliquée pour les mesures directes et objectives de la conscience de la situation.

Conceptuellement, la conscience de la situation semble sensible aux deux facteurs du senti-

ment de présence : l'immersion et la cohérence.

L'immersion est une évaluation directe de la façon dont un système simule la sensation humaine [36]. En évaluant la perception — l'acte de donner une signification à une sensation détectée [51] — la conscience de la situation consiste en une évaluation du résultat de l'immersion. Par exemple, l'utilisation d'un visiocasque dans un environnement virtuel implique un champ de vision plus petit, qui est connu pour affecter la conscience de la situation [162] et diminuer l'immersion. Un champ de vision de petite taille a un impact sur la quantité d'informations visuelles que l'on peut percevoir. L'immersion mesure les informations qui peuvent potentiellement être perçues, et la conscience de la situation mesure les informations qui sont effectivement perçues.

La cohérence correspond à l'ensemble des circonstances raisonnables évoquées par un scénario virtuel. Une circonstance raisonnable est un état de fait qui va de soi compte tenu des connaissances préalables d'une personne [65]. Il semble que la mesure de la conscience de la situation permettrait de déterminer, au moins partiellement, si un environnement virtuel fournit des circonstances objectivement raisonnables, c'est-à-dire si la situation présentée va de soi, compte tenu des connaissances préalables des personnes en faisant l'expérience. En effet, la conscience de la situation permet de faire l'évaluation de la capacité d'un individu à percevoir, comprendre et anticiper les éléments d'une situation menant à une prise de décision. Ainsi, puisque la variabilité de la capacité à prendre une décision appropriée est influencée en partie par les compétences et l'expérience d'une personne [152], qui sont directement liées à ses connaissances antérieures, la conscience de la situation offre une indication des connaissances préalables d'une personne dans un contexte donné.

Au niveau de l'application, peu d'études ont proposé d'utiliser les mesures de la conscience de la situation comme évaluation du sentiment de présence. Laptaned [163] et Jung et al. [164] ont trouvé une corrélation positive entre la conscience de la situation et le sentiment de présence avec une taille d'effet supérieure à 0,20, qui est la taille d'effet pratique minimale requise pour la recherche comportementale [165]. Cependant, la corrélation de Laptaned [163] provient d'une mesure subjective de la conscience de la situation. Cette étude ne permet donc pas de conclure sur la relation entre la conscience de la situation et le sentiment de présence telle qu'évaluée par une mesure objective de la conscience de la situation. Jung et al. [164] ont utilisé un dispositif de jeu vidéo portable lors de l'exécution de leur expérience. Ainsi, il n'est pas évident que leurs résultats peuvent être appliqués à un environnement de réalité virtuelle immersive.

Riley et al. [37] ont constaté que le sentiment de présence et la conscience de la situation sont négativement corrélées. Néanmoins, ils soupçonnent que l'interface de contrôle utilisée

dans leur expérimentation ait pu fausser leurs résultats. Leur interface exigeait de changer de point de vue pour développer la conscience situationnelle et aurait pu affecter l'engagement des participant.e.s dans leurs tâches. Ils affirment que cela aurait réduit les ressources attentionnelles disponibles pour la présence et aurait pu expliquer la diminution du score du questionnaire de présence et l'augmentation de la conscience situationnelle. Néanmoins, les auteurs ont suggéré que la mesure de la conscience de la situation pouvait servir de mesure objective de la présence.

Une étude a évalué la conscience de la situation, la charge cognitive et le sentiment de présence dans deux environnements virtuels différents [166]. Leur but était de vérifier si leur système utilisant un visiocasque permettait de former des individus au pilotage d'un aéronef de manière équivalente à celui utilisant un écran de projection. Ils rapportent un haut niveau pour les trois indicateurs dans les deux environnements. Toutefois, le système utilisant un visiocasque permettait un niveau un peu plus élevé de conscience de la situation, une plus grande sentiment de présence et un niveau de charge cognitive inférieur. Cette évaluation a notamment permis d'effectuer une comparaison impliquant deux systèmes utilisant des qualia différents.

Enfin, Ommerli [2] a évalué la relation entre la conscience de la situation et l'immersion, mais son modèle de présence repose sur une base théorique différente de celle utilisée dans cette thèse. Cette étude a utilisé le modèle de Mutterlein [167] selon lequel le terme immersion a une signification semblable à la définition du sentiment de présence utilisée par cette thèse [15]. L'étude a aussi utilisé le concept de téléprésence pour faire référence à l'illusion de lieu et le concept d'interactivité pour définir la capacité d'interagir avec un environnement, ce qui, selon la conception de Skarbez et al. [15], est lié à la fois à l'illusion de lieu et à l'illusion de plausibilité. Ommerli [2] a trouvé une relation directe et positive entre la téléprésence et les niveaux 1 et 2 de la conscience de la situation. De plus, cette recherche a établi que l'interactivité affecte directement le niveau 1 de la conscience de la situation et affecte indirectement le niveau 2 de la conscience de la situation [2].

En résumé, la plupart des études précédentes ont montré une corrélation positive entre le sentiment de présence et la conscience de la situation. Cependant, les expériences démontrant cette corrélation n'ont pas été réalisées dans un environnement virtuel immersif ou sont basées sur une théorie différente de la présence. Ainsi, certains doutes subsistent quant à la validité de l'utilisation de mesures de la conscience de la situation pour évaluer le sentiment de présence.

CHAPITRE 3 DÉMARCHE DU TRAVAIL DE RECHERCHE

D'abord, ce chapitre présente la problématique de recherche mise en lumière par l'introduction et la revue de littérature. On y établit ensuite l'objectif principal et les objectifs spécifiques poursuivis dans cette thèse. Puis, les sections suivantes offrent un aperçu de la méthodologie appliquée pour atteindre ces objectifs et des précisions sur les différentes méthodes de mesure utilisées.

3.1 Synthèse de la problématique de recherche

Les chapitres précédents ont permis de définir le sentiment de présence comme étant un phénomène issu de traitements implicites et explicites de l'information reflétant le réalisme perçu d'une expérience virtuelle. L'introduction a mis en évidence l'existence de problèmes liés à la mesure du sentiment de présence. Ces derniers sont principalement dus au fait que la mesure de ce concept repose essentiellement sur une évaluation subjective par questionnaire d'un phénomène supporté par des définitions manquant de standardisation. Ainsi, il n'est pas clair si les mesures actuelles du sentiment de présence permettent bel et bien d'évaluer le concept de manière valide. Or, des pistes de solutions pourraient être explorées pour améliorer la maîtrise du sentiment de présence et être en mesure d'obtenir un meilleur jugement du réalisme perçu d'un environnement virtuel.

Jusqu'à ce point, cette thèse a fait ressortir le fait que l'immersion est une caractéristique objective du sentiment de présence en environnement virtuel déterminée par la précision et l'étendue de la simulation des modalités sensorielles par un système. Généralement, la technologie actuelle ne permet pas de reproduire les seuils perceptuels humains, soit la valeur limite d'une qualité sensorielle pour laquelle les humains sont en mesure de percevoir. Cependant, même si l'atteinte d'une fidélité physique était théoriquement idéale pour maximiser le sentiment de présence prodiguée par les propriétés immersives d'un environnement virtuel, des études précédentes indiquent qu'une fidélité fonctionnelle serait suffisante afin de permettre aux utilisateur.rice.s de percevoir une expérience virtuelle comme étant réaliste. Toutefois, comme la revue de littérature l'a fait ressortir, aucune valeur de qualité sensorielle n'a été établie permettant de connaître le seuil fonctionnel des différentes modalités sensorielles. En d'autres mots, il n'est actuellement pas possible de définir une limite perceptuelle en dessous de laquelle une valeur de qualité sensorielle doit être considérée comme fonctionnellement dégradée.

Ensuite, bien que plusieurs variables aient été identifiées comme étant liées au sentiment de présence en environnement virtuel, leur relation avec les facteurs de la présence, soit l'immersion et la cohérence, est généralement inconnue. Nous ne connaissons donc pas la nature de la relation entre les facteurs de l'immersion, la charge cognitive, les mesures psychophysiologiques et la conscience de la situation.

3.2 Objectifs de recherche

La problématique de recherche a fait ressortir la nécessité d'explorer des méthodes d'évaluation du sentiment de présence permettant d'apporter plus de validité et d'objectivité à la mesure du concept. Ainsi, cette thèse a pour objectif principal d'effectuer une variation objective et non arbitraire de l'immersion en environnement virtuel afin d'évaluer ses impacts sur des indicateurs cognitifs et physiologiques du sentiment de présence.

Quatre objectifs spécifiques ont été définis de manière à atteindre cet objectif principal.

Le premier est relatif à la manipulation objective de l'immersion et consiste à mesurer le seuil fonctionnel de qualités visuelles, défini comme la limite perceptuelle en dessous de laquelle une valeur de qualité sensorielle est considérée comme fonctionnellement dégradée. Donc, le premier objectif spécifique de cette thèse consiste à :

O1 Identifier les seuils fonctionnels de trois qualités visuelles : la sensibilité au contraste, la taille du champ de vision et l'acuité visuelle dans un environnement virtuel utilisant un visiocasque.

L'atteinte de cet objectif permet d'effectuer une manipulation objective de l'immersion dont les différents niveaux sont établis sur une base fonctionnelle et non arbitraire. L'utilisation de cette méthode de manipulation de l'immersion permet d'évaluer son impact sur des éléments liés au sentiment de présence. Ainsi, les trois autres objectifs spécifiques de cette thèse sont :

O2 Évaluer l'impact d'une manipulation objective de l'immersion sur l'activité cardiovasculaire et l'activité oculaire de personnes engagées dans l'exécution d'une tâche dynamique dans un environnement virtuel.

O3 Évaluer l'impact d'une manipulation objective de l'immersion sur la charge cognitive de personnes engagées dans l'exécution d'une tâche dynamique dans un environnement virtuel.

O4 Évaluer l'impact d'une manipulation objective de l'immersion sur la conscience de la situation de personnes engagées dans l'exécution d'une tâche dynamique dans un environnement virtuel.

3.3 Démarche de recherche

De manière à atteindre ces objectifs de recherche, cette thèse fait état de deux expériences réalisées en environnement virtuel. Ces expériences sont rapportées dans trois articles de journaux publiés ou soumis pour publication dans la revue PRESENCE : Virtual and Augmented Reality.

3.3.1 Expérience 1 : Identification de seuils fonctionnels

La première expérience est décrite dans le premier article. Celui-ci introduit le concept de seuil fonctionnel. Dans cet article nous avons mesuré à quel point les valeurs de la sensibilité au contraste, la taille du champ de vision et l'acuité visuelle peuvent être dégradées sans qu'un.e utilisateur.rice d'un environnement virtuel ne soit en mesure de percevoir de dégradation de ses qualités sensorielles.

Pour identifier les trois seuils fonctionnels, nous avons appliqué une méthodologie provenant du domaine de la psychophysique : une version modifiée de la méthode des limites. Une à la fois, la sensibilité au contraste, la taille du champ de vision et l'acuité visuelle étaient progressivement dégradées pendant qu'un.e participant.e effectuait une tâche de sélection multidirectionnelle. 30 participant.e.s ont pris part à l'expérience. Leur but était d'identifier le moment auquel la dégradation de l'une de leurs qualités visuelles était perceptible tout en performant au meilleur de leur capacité à la tâche de sélection multidirectionnelle.

(a) Perspective initiale des participant.e.s (b) Murs ouest et nord (c) Murs est et sud

FIGURE 3.1 Environnement virtuel de l'expérience 1 selon trois perspectives

Note. a montre la perspective initiale des participant.e.s ; b montre les murs ouest et nord vus du coin sud-est de la pièce ; c montre les murs est et sud vus du coin nord-ouest de la pièce.

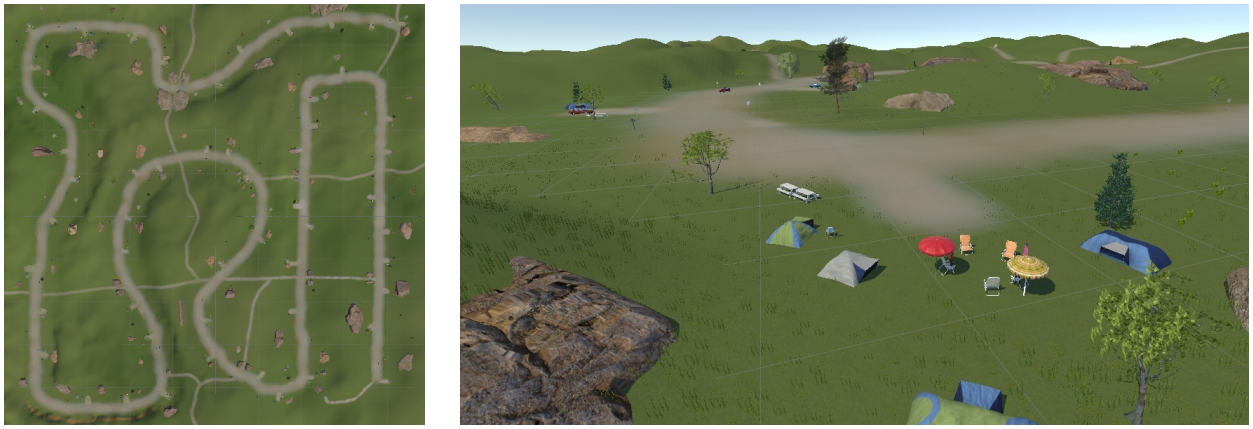
L'environnement virtuel dans lequel l'expérience a été réalisée est présenté à la figure 3.1. Il a été conçu de manière à présenter un ensemble diversifié de couleurs, de contraste, et de détails visuels dans l'ensemble du champ de vision des participant.e.s.

3.3.2 Expérience 2 : Évaluation de l'impact d'une manipulation de l'immersion sur des indicateurs physiologiques et cognitifs

La seconde expérience a fait l'objet de deux articles. Dans cette expérience, nous avons cherché à analyser l'effet d'une manipulation de l'immersion sur la charge cognitive, l'activité oculaire, l'activité cardiovasculaire et la conscience de la situation. Pour manipuler l'immersion, une à la fois, la sensibilité au contraste, la taille du champ de vision et l'acuité visuelle ont été fixées à des valeurs supérieures, égales ou inférieures au seuil fonctionnel identifié lors de la première expérience. La figure 3.2 offre un aperçu de l'environnement virtuel utilisé lors de la seconde expérience.

Pour cette expérience, 60 participant.e.s ont effectué une tâche de conduite automobile dans un environnement virtuel. Chaque participant.e effectuait trois scénarios de conduite. Un des trois scénarios consistait en un scénario de référence effectué alors que toutes les qualités visuelles étaient fixées à la valeur supérieure au seuil fonctionnel (aucune dégradation). Les deux autres étaient effectués alors que l'une des qualités visuelles étudiées était fixée au seuil fonctionnel ou à une valeur inférieure au seuil fonctionnel. Lors de chaque scénario, nous avons mesuré le sentiment de présence, la charge cognitive, l'activité oculaire, l'activité cardiovasculaire et la conscience de la situation.

La prochaine section offre une revue des méthodes utilisées pour mesurer les différents concepts présentés dans cette thèse.



(a) Carte du parcours (b) Image de l'environnement virtuel

FIGURE 3.2 Environnement virtuel de la seconde expérience

3.3.3 Instruments de mesure

Cette section présente des considérations méthodologiques liées aux différentes mesures utilisées lors des expérimentations réalisées dans le cadre de cette thèse.

Questionnaire de présence

Dans la seconde expérience réalisée dans le cadre de cette thèse, le sentiment de présence a été évaluée de manière subjective.

L'évaluation subjective du sentiment de présence a été réalisée à l'aide d'une version papier du questionnaire de présence du laboratoire de cyberpsychologie de l'Université du Québec en Outaouais [168]. Il s'agit d'un questionnaire de 24 items utilisant 19 items du *Presence Questionnaire* (PQ) version 2 [169] et 5 items liés à l'audio et à l'haptique. Ce questionnaire correspond à une version révisée du PQ de Witmer & Singer version 3.

Le PQ [169] est l'un des questionnaires de mesure du sentiment de présence le plus utilisé [20] et le plus court [13]. Il présente notamment l'avantage de ne pas être corrélé à une utilisation spécifique [20]. Toutefois, ce questionnaire est basé sur une conception différente de la présence que celle présentée dans cette thèse, considérant par exemple l'immersion comme une caractéristique subjective de la présence plutôt qu'une caractéristique objective [170]. Or, des études ont démontrées que le PQ était sensible aux variations de l'immersion et de la cohérence, réconciliant cet instrument de mesure avec la conception de la présence présentée dans cet ouvrage [52, 171]. L'illusion du lieu et l'illusion de plausibilité ont à peu près la même importance en ce qui concerne les scores au PQ [52]. Ainsi, son utilisation répandue, sa fiabilité et sa validité apparente assurent un certain niveau de comparabilité entre les études.

NASA-TLX

Le questionnaire NASA-TLX a été utilisé dans la seconde expérience réalisée dans le cadre de cette thèse.

Ce questionnaire est une méthode subjective de mesure de la charge cognitive utilisant une échelle d'évaluation multidimensionnelle pondérée à six facteurs [112]. Le fait qu'une pondération soit effectuée permet de tenir compte de la variabilité entre les sujets [112]. Parmi les six facteurs, trois sont liés à la tâche, soit la demande physique, la demande mentale et la demande temporelle et permettent de mesurer la charge extrinsèque [115]. Les trois autres facteurs sont liés à l'individu effectuant la tâche, soit son jugement de sa performance, l'ef-

fort nécessaire pour effectuer la tâche et son niveau de frustration [112]. Le jugement de la performance et du niveau de frustration sont des mesures de la charge intrinsèque, alors que l'effort nécessaire pour effectuer la tâche est à la fois une mesure de la charge intrinsèque et de la charge utile [116].

NASA-TLX possède une forte corrélation avec la performance, a une bonne diagnosticité et est relativement facile à utiliser lorsque comparée à d'autres méthodes subjectives multidimensionnelles d'évaluation de la charge cognitive [172].

Tableau 3.1 Distribution des scores du questionnaire NASA-TLX

min.	**10%**	**20%**	**25%**	**30%**	**40%**	**50%**	**60%**	**70%**	**75%**	**80%**	**90%**	**max.**
6.21	26.08	33.00	36.77	39.45	45.00	49.93	53.97	58.00	60.00	62.00	68.00	88.50

Le tableau 3.1 est tiré d'une méta-analyse présentant la distribution des scores obtenus après à l'exécution de tâches évaluées par NASA-TLX [173]. Pour les tâches de conduite de véhicule, on indique les valeurs présentées dans le tableau 3.2.

Tableau 3.2 Distribution des scores du questionnaire NASA-TLX pour une tâche de conduite de véhicule

min.	**25%**	**50%**	**75%**	**max.**
15.00	28.05	41.52	51.73	68.50

Indice d'activité pupillaire faible et élevée

L'indice d'activité pupillaire faible et élevée (LHIPA) a été utilisé dans la seconde expérience réalisée dans le cadre de cette thèse.

Le LHIPA est une mesure en temps réel de la charge cognitive [127] basée sur une analyse d'ondelettes qui utilise le rapport basse fréquence/haute fréquence reflétant les changements dans les composantes parasympathique et sympathique du système nerveux autonome [136]. Il s'agit d'une mesure inspirée de l'indice d'activité cognitive (ICA) et de l'indice de l'activité pupillaire (IPA) [136]. Des résultats empiriques montrent que le rapport des basses et des hautes fréquences de l'oscillation pupillaire exprimé par le LHIPA est robuste dans sa sensibilité à la mesure de la charge cognitive et varie en fonction de la difficulté d'une tâche [136]. Le LHIPA est plus efficace pour détecter une augmentation de la charge cognitive que des mesures comme l'IPA et l'ICA [127, 136]. De plus, l'effet de la distorsion de la pupille lorsqu'elle

est hors de l'axe de mesures des caméras de suivi oculaire est négligeable, ce qui consiste en un avantage par rapport aux autres mesures semblables [136].

Mesure de l'activité cardiaque

Lors de la seconde expérience, plusieurs mesures de l'activité cardiaque ont été utilisées. Les mesures que l'on retrouve dans cette étude sont divisées en trois catégories. La première catégorie comporte une mesure de la fréquence de base, soit la moyenne de la fréquence cardiaque. La deuxième catégorie comporte des mesures de la variabilité de la fréquence cardiaque du domaine temporel, soit le RMSSD, le SDNN et le pNN50. La troisième catégorie comporte des mesures de la variabilité de la fréquence cardiaque du domaine fréquentiel, soit la distribution des LF, la distribution des HF et le LF/HF. Les métriques de variabilité de la fréquence cardiaque ont été choisies en se basant sur les résultats des articles ayant étudié l'effet de la variabilité de la fréquence cardiaque lors de tâches de conduite présentées dans la revue de littérature.

Comme indiqué dans la revue de littérature, la fréquence cardiaque de base est une réponse physiologique sensible aux états cognitifs comme la charge cognitive [49]. Elle possède une sensibilité supérieure aux mesures de la variabilité de la fréquence cardiaque pour des conditions expérimentales ou la charge cognitive est faible [143].

Les mesures de la variabilité de la fréquence cardiaque des domaines temporels et fréquentiels, telles que RMSSD, SDNN, pNN50, la distribution des LF et la distribution des HF augmentent avec la somnolence, la fatigue et le désengagement et diminuent avec la charge de travail, le stress et la vigilance [49,174]. Le résultat du LF/HF est inverse. Un faible LF/HF reflète une dominance du système nerveux parasympathique et un LF/HF élevé indique une dominance du système nerveux sympathique [174]. De nombreux facteurs influencent les mesures de la variabilité de la fréquence cardiaque, il est donc important de connaître le contexte de l'enregistrement et des sujets. La variabilité de la fréquence cardiaque est influencée par la posture, l'âge, le sexe, la fréquence cardiaque, le type d'activité [174], l'état de santé [138], et la respiration [121, 174].

Pour tenir compte de ces facteurs et éviter les changements de fréquences cardiaques d'origine physiologique, certaines procédures méthodologiques doivent être suivies. Les mouvements du corps doivent être réduits de manière à limiter les impacts de changements respiratoires dus à la position du corps [49]. De plus, un nettoyage des données cardiaques doit être effectué pour éliminer le bruit causé par l'instrument de mesure et les rythmes anormaux normalement présents dans ce genre de mesure [49, 175]. Il est recommandé d'appliquer des interpolations pour appliquer ces corrections [175].

Les mesures d'activité cardiaques peuvent être affectées par des effets compensatoires [138]. Ainsi, pour mesurer des phénomènes psychophysiologiques comme la charge cognitive en évitant l'influence de ces effets compensatoires, il est recommandé d'utiliser des mesures à très court terme de la fréquence cardiaque [138, 149]. La variabilité de la fréquence cardiaque à très court terme représente la relation entre les branches sympathique et parasympathique [174]. Pour les mesures à très court terme, l'activité cardiaque doit être mesurée pendant un minimum de 120 secondes pour les mesures de fréquence cardiaque de base et pendant un minimum de 60 secondes pour la variabilité de la fréquence cardiaque [174].

Technique d'évaluation globale de la conscience de la situation

La conscience de la situation a été mesurée avec la technique d'évaluation globale de la conscience de la situation (SAGAT) dans le cadre de la deuxième expérience de cette thèse.

SAGAT est une méthode de mesure directe et objective de la conscience de la situation ayant été rigoureusement validée [160]. Elle est déterministe, linéaire et significativement corrélée à la performance à une tâche [153]. L'utilisation de cette technique de mesure n'a pas d'impact significatif sur la tâche évaluée [160]. De plus, elle est sensible aux changements de conditions expérimentales et moins sujette au biais d'échantillonnage que d'autres méthodes semblables [157]. Dans l'ensemble, elle consiste à questionner une personne effectuant une tâche sur des éléments liés à sa conscience de la situation.

Le développement de questionnaire de mesure de la conscience de la situation requiert une évaluation attentive des exigences liées aux trois niveaux de la conscience de la situation d'une personne [160]. Celle-ci s'effectue par une analyse de buts qui permet d'identifier les éléments menant aux prises de décision dans l'environnement étudié [161]. À partir des buts des personnes effectuant la tâche ciblée par l'analyse de la conscience de la situation, il est possible d'établir les questions appropriées pour en évaluer les différents niveaux. Le questionnaire utilisé lors de la seconde expérience est basé sur des analyses de but développées pour évaluer la conscience de la situation de conducteurs automobiles [176, 177]. Les résultats de ces dernières ont été adaptés pour que les questions reflètent les éléments pertinents de la tâche exécutée par les personnes participant à l'expérience.

Une fois que les questions relatives aux trois niveaux de conscience de la situation ont été déterminées, la méthode SAGAT requiert que la personne effectuant la tâche analysée soit questionnée lors de son exécution. Pour pouvoir questionner sans interrompre la personne dans l'exécution de sa tâche, la tâche est gelée à des moments aléatoires et les éléments rattachés à la tâche pouvant offrir de l'information relative à la conscience de la situation sont rendus inaccessibles. Dans le cas de la deuxième expérimentation, lorsque l'environnement

était mis sur pause, les vitres de la voiture devenaient opaques pour couper les participant.e.s des informations disponibles à l'extérieur du véhicule et les informations disponibles sur le tableau de bord n'étaient plus affichées.

Après la collecte d'information, les réponses des participant.e.s aux questions sont comparées aux données de la situation réelle. Cela permet d'établir un score indiquant combien les participant.e.s sont conscient.e.s de ce qui se passe dans l'environnement lors de l'exécution d'une tâche [153]. Le score de la conscience de la situation correspond à la proportion de bonnes réponses fournies par la personne ayant effectué la tâche. On recommande d'avoir environ 60 réponses par question posée pour être en mesure d'évaluer correctement la conscience de la situation [160].

CHAPITRE 4 ARTICLE 1 : IDENTIFICATION OF VISUAL FUNCTIONAL THRESHOLD FOR IMMERSION ASSESSMENT IN VIRTUAL REALITY

Cet article a été publié le 26 juillet 2022 dans PRESENCE : Virtual and Augmented Reality 2020 ; 29 1–22. doi : https ://doi.org/10.1162/pres_a_00354

Hébert-Lavoie, M. (Polytechnique Montréal), Doyon-Poulin P. (Polytechnique Montréal), Ozell, B. (Polytechnique Montréal)

4.1 Abstract

We consider that to objectively measure immersion, one needs to assess how each sensory quality is reproduced in a virtual environment. In this perspective, we introduce the concept of functional threshold which corresponds to the value at which a sensory quality can be degraded without being noticed by the user of a virtual environment. We suggest that the perceived realism of a virtual experience can potentially be evoked for sensory qualities values ranging from the perceptual threshold to the functional threshold. Thus, the identification of functional thresholds values allows us to constrain immersion. To lay the foundation for the identification of functional thresholds, we applied a modified version of the method of limits. We measured the value at which 30 participants were able to identify the degradation of their field of view (FOV), visual acuity, and contrast sensitivity while executing a multi-directional selection test. This enabled us to identify functional perceptual thresholds of 96.6 degrees for FOV, 12.2 arcmin for visual acuity, and 25.6% for contrast sensitivity.

4.2 Introduction

Immersive virtual reality simulates human sensory and perceptual processes through mechanical and computational means. Its use suggests multiple applications, offering opportunities to transform actual training and design methods [178,179]. To be valid and reliable, such usage requires a virtual environment to be equivalent to its real-world counterpart. Considered a measure of the realism of a virtual experience [2, 9, 10, 15, 19, 23], the concept of presence could be used to evaluate such correspondence [10].

Presence is a quale, it is subjective and its interpretation is variable for each individual [19]. It is composed of two principal aspects; coherence and immersion [171]. Coherence refers to the ability of the environment to act as the user expects [15]. Immersion refers to the

extensiveness of the simulation of sensory modalities by a system [15, 19].

Currently, the assessment of presence is done mostly by indirect methods such as subjective questionnaires [10, 21, 43], but also with physiological and behavioral measures [10, 37]. These methods are criticized for their lack of sensitivity, diagnosticity, and their validity as a measure of the construct [19, 180]. The subjective character of presence renders its direct measurement complex [19, 37]. However, because of its objective nature [19], evaluating immersion by quantifying each sensory modality could represent a way to circumvent a part of this complexity.

To do so we propose an approach similar to Perroud et al. [58], but more holistically; to assess how all sensory modalities are reproduced in a virtual environment. This requires measuring the sensory qualities of each sensory modality, i.e., the distinctive attributes that enable a sensory modality to carry out its function [51]. In that regard, we've identified the sensory qualities of vision. These are visual accommodation, eye movements, stereopsis, perception of movement, color perception, FOV, visual acuity, and light and contrast sensitivity. This paper intends to lay the foundation for the determination of a range of sensory quality values, bounded by perceptual and functional thresholds, in which immersion can occur.

4.2.1 Thresholds of Perception

Perceptual thresholds are the limiting sensory quality values for which humans can perceive, e.g., for visual acuity, it may be the smallest object that can be discriminated by the human eye. Within the framework of immersion, perceptual thresholds can be considered as the values to be reached by a system to maximize presence. However, at the moment, creating an immersive system able to reach these values represents a technical challenge that has yet to be solved.

Still, concepts such as useful FOV (UFOV) [94] imply that humans do not always perform at the limit of their perceptual capacities. This makes it relevant to evaluate to what extent a sensory quality can be degraded without being noticeable, i.e., identifying functional thresholds.

By using functional thresholds, it would be possible to constrain immersion functionally, between a maximal bound, defined by the perceptual thresholds, and a minimal bound, defined by the functional thresholds; thus, knowing between which perceptual values a feeling of presence, could potentially be evoked. Unlike most means of measuring presence, bounding perception addresses only its immersion component. However, it is an objective characterization that considers the technology used and that could be done prospectively without any

user intervention.

Functional thresholds have not been established for any sensory qualities. Therefore, the objective of this study is to determine three functional thresholds related to vision. In that regard, we measured the value for which a degradation was perceived for FOV, visual acuity, and contrast sensitivity which have undeniable impacts on presence. This was done in a virtual environment while executing a task where the attention is not entirely devoted to the discrimination of these phenomena. This will allow us to propose a framework for the study of immersion.

Thus, we first report on studies linked to the FOV, visual acuity, and light and contrast sensitivity. Then, we detail the methodology applied to evaluate their respective functional threshold in an immersive virtual reality. Thereafter, we present the measured functional thresholds, discuss their meaning in the context of the study and their foreseeable implications.

4.3 Related Work

This section presents the state of knowledge related to the three visual qualities that are targeted by this study. For each quality, we describe the notions that are useful to understand their role toward the visual system, determine their perceptual threshold, and highlight relevant information for the identification of functional thresholds.

4.3.1 Field of View

The visual experience is a reconstruction of the world based on two distinct images from the retina of each eye. The portion of space that the eyes can capture when motionless is called the field of view [51].

Figure 4.1 shows a representation of different sections of the FOV. Binocularly, when the eyes are fixed on a point, the human visual system covers a horizontal FOV of approximately 180 to 200 degrees [6, 58] of which 100 to 120 degrees are covered by both eyes [6, 51, 74], and a vertical FOV of 120 to 140 degrees [51, 73, 74].

In head-mounted displays (HMD), it is important to consider the fact that the eyes can move in the presented visual field. When the head is fixed, the eyes, by moving, can cover an area of 290 degrees horizontally and 190 degrees vertically [74]. These values could serve as a perceptual threshold for FOV. Also, it is important to note that it is possible to sustain the eye gaze in only 40 to 50 degrees in the central region of the FOV which is called the fixed

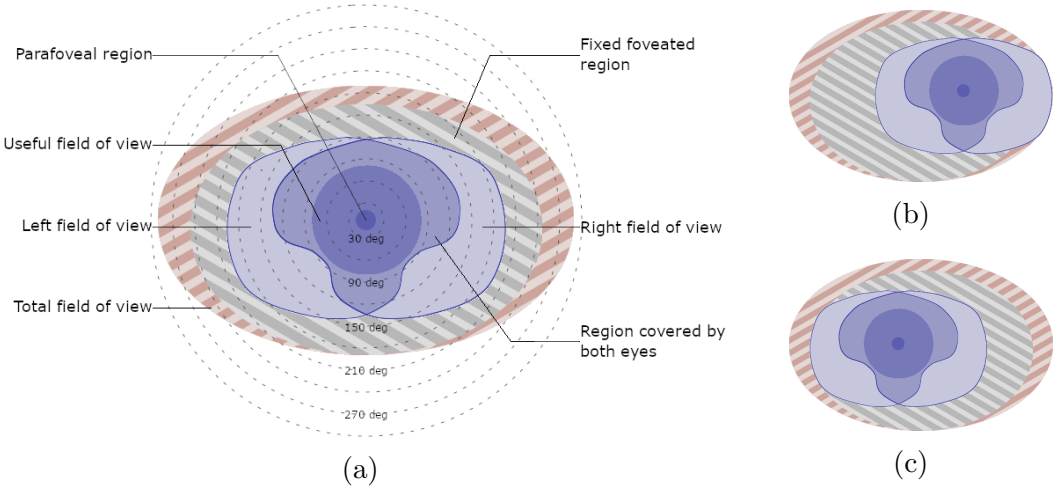


Figure 4.1 Schematic representation of different sections of the FOV.

Note. in a the blank blue sections represent the portion of space that is captured by the eyes when fixed on a point, the stripped regions indicate the parts of the FOV that can be covered by moving the eyes; b shows the eyes looking at the right limit of the FOV; c shows the eyes looking at the left limit of the fixed foveated region.

foveated region [75].

In comparison, the horizontal FOV of most current generation HMDs is between 90 and 110 degrees. We know that such FOV size makes any task where the observer is detecting objects in peripheral vision difficult to perform [6]. Research that evaluated the FOV size in a virtual environment tends to demonstrate that higher presence is evoked by values closer to the total FOV size [9, 76, 77]. This relation is also reflected in performance measure [78]. However, since the portion of the FOV used by an observer depends on the task performed [83] these relations do not allow us to pinpoint a FOV size that would be sufficient for general use.

Visual performance is not equal across the entire FOV. Increasing eccentricity from the fovea affects motion perception [6,84,85], visual acuity [86–88], light and contrast sensitivity [88,89], and color perception [85, 89–91]. These differences in visual qualities performance across the FOV may be responsible for the fact that different functions are attributed to different portions of the visual field.

For instance, the UFOV describes the function of the central part of the visual field. It is defined as the surface around the point of fixation inside which information can be perceived and processed during a visual task [92]. This definition reveals a similarity with our conception of a functional threshold of FOV. The UFOV size varies depending on the executed task [93]. Its range is often estimated to cover a diameter of 30 to 40 degrees from the central

gaze point [94]. Yet, methods assessing the extent of the UFOV tend to neglect the contribution of peripheral information that can be acquired [94], and it could cover a diameter of up to 80 degrees [95]. Thus, the state of knowledge on UFOV does not enable us to clearly establish a functional threshold of FOV.

In summary, the data available does not allow us to establish a functional threshold of FOV. For tasks that necessitate peripheral vision information acquisition, the FOV size offered by the current generation of HMD is theoretically too small and makes it harder to perform the task. However, according to the definition of UFOV, we could have expected such size to be sufficient. Hence, more research is needed to establish a functional threshold.

4.3.2 Visual Acuity

Visual acuity refers to the finest level of detail that can be detected or identified [96]. It is physiologically determined by the spacing between photoreceptors on the retina [51]. In a virtual environment, the notion of resolution acuity is generally used when measuring visual acuity. It is defined as the smallest angular separation between two neighboring objects that can be detected [51].

Visual acuity depends on the environment's visual conditions. Reaching the perceptual threshold of visual acuity requires a stable image displayed within 0,5 degrees of the center of the fovea, maximum contrast, and photopic conditions [62]. Under these conditions, the human eye has a resolution acuity between 0,5 and 2 arcmin with an average of 1 arcmin. [58, 73, 83, 97, 98].

The actual state of knowledge does not enable us to pinpoint a functional threshold for visual acuity. We know that the angular resolution of most current HMD is 4 to 6 arcmin. This is due to a pixel density too low to reproduce maximal visual acuity [100]. Theoretically, this resolution implies a lack of information that makes objects difficult to discern at distances where they would be clearly identified in real life [6, 83]. The relation between resolution and perceived realism of a virtual environment is reflected only to some extent in the measure of presence [101]. An increased resolution was shown to significantly increase the reported presence [57, 77]. However, these results were obtained with a great difference between the high and low-resolution conditions, and it appears that a smaller difference between conditions does not yield significant results [102]. This might indicate that a functional threshold can be reached with the actual technology.

The inability to see a difference between two resolution conditions could be due to the fact that humans see most of the world through their peripheral vision that cannot render images

as sharply as the foveal region [103]. Visual acuity is not constant throughout the retina and decreases with increasing eccentricity. For a fixed stimulus, the resolution acuity at 8 degrees of eccentricity of the fovea is 2.8 arcmin, dropping at 7.0 arcmin at 13 degrees [86]. At an eccentricity of 25 degrees, it is 12 arcmin [87], and at 30 degrees a spacing of up to 15 arcmin cannot be discriminated [85].

Briefly, we were unable to establish a functional threshold for visual acuity. However, previous research indicates that current HMD technology resolution acuity is sufficient to reach a visual acuity functional threshold. We would expect such value to correspond to the visual acuity in peripheral vision.

4.3.3 Light and Contrast Sensitivity

Light sensitivity corresponds to the number of photons absorbed by the cones and rods lining the retina. Due to neural encoding, the human eye is sensitive to differences in the intensity of regions of the receptive field rather than the average intensity [51]. Light sensitivity is a time-varying phenomenon that depends on previous lighting conditions [106]. To take into account the time-varying factor and address this visual quality in a way that enables the identification of a context-independent perceptual threshold, the concept of contrast sensitivity is employed. Thus, we will use the terms "contrast sensitivity" when referring to this phenomenon onwards.

Contrast sensitivity is equivalent to the smallest difference in luminance required to identify a target [96]. It is usually expressed as the ratio between the low and high luminance of an image. Contrast sensitivity depends on spatial frequency [96], illumination conditions [107], and FOV size [72].

There is a robust contrast threshold of 1% which is independent of the visual conditions [108]. Under optimal conditions, however, the minimum contrast that can be perceived in the fovea, the perceptual threshold, is 0.2 % [88]. For an HMD with a FOV of 110 degrees and a resolution equivalent to 4.6 arcmin this value increases to 0.45% [107].

Contrast sensitivity in the periphery is decreased compared to the foveal contrast sensitivity [88]. For an aliased stimulus, it is reported to range from 19% to 23% at 30 degrees of eccentricity [88].

Yet, we are unable to identify a functional threshold. We know that the contrast sensitivity of observers using an HMD is poorer than in a real environment [99]. Due to a limited range of displayed luminance [109], chromatic aberrations, and insufficient resolution, it has been concluded that VR technology is not able to reproduce standard contrast sensitivity

testing [110]. Still, we can reasonably expect the contrast range offered in current HMDs to be sufficient to let its user feel present in a virtual environment.

Also, we were unable to identify studies that targeted the effect of varying levels of contrast on presence or performance in a virtual environment. One study evaluated the effect of contrast on the perceived realism of desktop computer images, identifying a threshold for negative contrast offset of $0.6678(V(x,y) - 0.5) + 0.5$, where $V(x,y)$ is the value channel of the HSV color model, i.e., the pixel luminance [111]. However, since their results were reported as a contrast offset, which is image-dependent, it can't be used as a direct indication to establish a functional threshold.

Again, the body of knowledge was insufficient to identify a functional threshold value for contrast sensitivity. As for visual acuity, we would expect to find a functional threshold of contrast sensitivity that corresponds to values in the peripheral region since humans see most of the world through this area [103].

4.3.4 Research Objective

Following this review of related work, we can see that previous studies focused on establishing the function of the different visual qualities, and identifying perceptual thresholds. We also understand that no functional thresholds have been established for the three visual qualities targeted by this study. For each, some information is available but it is not sufficient to pinpoint them. Hence, the objective of this paper is to measure the functional thresholds for FOV, visual acuity, and contrast sensitivity in a virtual environment with an HMD.

4.4 Method

The methodology outlined below aims to identify thresholds at which a user of a virtual environment can perceive a degradation of its FOV size, visual acuity, and contrast sensitivity. The value at which a degradation is perceived is considered a functional threshold of the degraded quality for this specific environment. Generally, the methodology consisted of separately degrading these three visual qualities while the participant was performing a multi-directional selection test. Such a task was used so the participant's attention was not directed exclusively toward the identification of the degradation of the visual quality. The experiment was carried out in virtual reality using a Unity implementation of a multi-directional selection test.

4.4.1 Participants

The participants recruited were volunteers over 18 years of age, had no uncorrected vision problems, and did not have a known prevalence of being subject to cybersickness. The research project was reviewed by Polytechnique Montréal's research ethics board and met the current research ethics standards on ethical conduct for research involving humans.

30 participants (13 females, 17 males) aged between 22 and 63 years old ($M = 30.33$) took part in the experiment. All participants executed the experiment with their dominan t hand, 26 participants were right-handed and 4 left-handed. During the experiment, 11 participants wore glasses, 3 contact lenses and 16 did not need any device to correct their vision. Participants estimated their total experience time with virtual reality HMDs and had a median experience of 1 hour ($min = 0$, $M = 5.8$, $max = 100$).

4.4.2 Materials

HTC Vive Pro Eye

The participants used the HTC Vive Pro Eye HMD to complete the task. The HTC Vive Pro Eye is a virtual reality headset that allows the immersion of a participant in a virtual environment. It has a dual OLED display with a diagonal of 3.5 inches allowing a 110 degrees FOV, has a refresh rate of 90 Hz, and a resolution of 1440 x 1600 pixels per eye, for a maximum resolution acuity of 4.6 arcmin according to Maxwell et al. [100]. The headset generates a luminance between 0.04 and 130 cd/m^2 and occupies the AdobeRGB colorspace [181]. It is equipped with an audio system that was removed for the experiment. The controllers used in the experiment are those supplied with the device. The experiment's computer used an NVIDIA GeForce RTX 2080 SUPER GPU and an Intel Core i9-9900K CPU with 64GB RAM.

Virtual Environment

The virtual environment was created with Unity version 2019.3. As shown in figure 4.2, it consisted of a 6m x 4m x 6m room with a column on the northern wall. The floor had a concrete texture, the northern and western walls had a white paint texture, and the eastern and southern walls, as well as the ceiling, had a newly installed gypsum texture. A network of dark grey pipes extending to the northern wall and rectangular light fixtures covered the ceiling. On the northern side were plants and boxes of various colors. In front of the western wall were an industrial sink, a watering can, and yellow metal lockers. The eastern wall had

(a) Participant's initial view (b) West and North wall (c) East and South wall

Figure 4.2 The virtual environment from three different points of view.

*Note.*a shows the participant's initial view; b shows the West and North wall from the South-East corner of the room; and, c shows the East and South wall from the North-West corner of the room.

a dark gray door and the southern wall behind the participant was empty. The purpose of presenting many objects in the virtual environment was to diversify the range of colors, contrast, and level of visual detail across the participant's FOV.

4.4.3 Procedure

Multi-Directional Selection Test

As part of the experiment, participants performed several multi-directional selection tests, also referred to as Fitts' test. The objective of using a multi-directional selection test was to direct participants' attention toward specific objects and ensure that they were not only attentive to the degradation of their visual qualities.

A multi-directional selection test is based on Fitts' law which models human performance, reflecting the speed and accuracy of a selection task [182]. The overall performance of the test is captured by the throughput index which is measured in bits per second [183]. A higher throughput indicates a better efficiency to the task [183].

We based our experimental task on the Fitts' law evaluation procedure, standardized for a 2D performance evaluation [184]. More precisely, the task consisted of sequentially selecting 15 targets of size W arranged in a circle with a distance D between two selections. This was repeated for nine different combinations of target size and distance, each defined by a difficulty index (ID). The ID ranged from 0.85 to 3.17, which is in line with Soukoreff et al. [183] recommendations. All tasks were performed with targets located 0.5 m in front of the participant. The ID included three W of 0.05, 0.075, and 0.100 m, equivalent to visual angles of 343.5 arcmin, 514.7 arcmin, and 685.3 arcmin respectively. The D between

targets were 0.080, 0.240, and 0.400 m, corresponding to visual angles of 548.9 arcmin, 1619.5 arcmin, and 2616.2 arcmin. Figure 4.3 shows different task IDs as presented in the virtual environment.

(a) (b) (c) (d)

Figure 4.3 Representation of different multi-directional selection test IDs.

Note. a shows an ID of 1.05 corresponding to a W of 0.075 m and a D of 0.08 m; b shows an ID of 2.07 corresponding to a W of 0.075 m and a D of 0.24 m; c shows an ID of 2.32 corresponding to a W of 0.1 m and a D of 0.4 m; and, d shows an ID of 2.54 corresponding to a W of 0.05 m and a D of 0.24 m.

At the beginning of a trial, the 15 targets were displayed, disappearing after their respective selection. The target to select was opaque and yellow to distinguish it from the others which were transparent and navy blue. The first illuminated target was at the right side of the circle. After it was selected, the next target, located at about 180 degrees on the opposite side of the circle, became yellow and opaque. Selections then continued clockwise in a back and forth motion until all targets were selected. The first selected target was not included in the data analysis to minimize the impact of task initiation.

To make a selection, the participant had to place the tip of the controller in the target and pull the trigger. To help with the controller positioning, a red sphere was located at the tip of the virtual representation of the HTC Vive controller. Feedback confirming that the participant had made a selection was also provided, the sphere would turn green when the trigger was pulled.

It took a participant between 1 minute and 1 minute 30 seconds to complete a single multi-directional selection test.

Functional Threshold's Psychophysical Evaluation

We evaluated the functional thresholds using an adaptation of the method of limits [185] that incorporated a multi-directional selection test used to direct the observer's attention. The participants had to detect a visual degradation of the virtual environment.

The method of limits is based on the fact that there is a greater likelihood that a stimulus will be consciously detected if its intensity is greater [185]. It is a judgment task where participants are presented with a stimulus of incrementally varied intensity and need to tell if they can detect it or not.

Typically, the method of limits consists of two procedures: the ascending limit method and the descending limit method. The ascending limit method consists of increasing the intensity of a stimulus until it is detected by the participant. The descending limit method consists of decreasing the intensity of a stimulus until it is no longer detected by the participant. Experimenters then compute the combination of thresholds from both methods executed a multitude of times by many participants resulting in a distribution of perceptual thresholds [185].

The main modification we brought to the method was to only increment the stimulus intensity in one direction; degrading the visual qualities. This way, the detection probability increases with each increment because the degradation becomes more obvious, even if the participants are engaged in a task. We believe that incrementing the stimulus intensity in the other way would not have yielded appropriate results. In a normal use case, by doing a conscious effort, observers using the HTC Vive Pro Eye set at its maximal perceptual values can detect that their visual qualities are degraded. This is because the technology is not yet able to reproduce the perceptual thresholds. As a consequence, since it is expected to be difficult for participants to judge when the image has reached the maximal perceptual value that the HMD can offer, we anticipated that most participants would not have been apt to tell that they weren't able to detect the degradation of their visual qualities.

Thus, the size of the FOV, the screen resolution, and the contrast offset value were all separately decreased during the experiment. The participants had to indicate when they were able to perceive a degradation of either their FOV, visual acuity, or contrast sensitivity and, if possible, identify which quality was degraded, without disrupting the multi-directional selection test being performed.

We calibrated our degradation rates and final visual quality values in pre-test sessions. Each degradation was calibrated so that an increment was imperceptible. This avoided underestimating the thresholds since a variation with a large disparity could have erroneously led to earlier detection. We calibrated our final values by degrading the visual qualities until they reached a value judged detectable at a 100% rate. We made sure that this value was reached after around 2 minutes, giving each participant enough time to complete the multi-directional selection test while the degradation occurred.

Also, to identify functional thresholds, representing lower bound values of perception, we

made sure that participants were naive to the degradation to avoid an overestimation of the threshold. In that regard, each visual quality degradation was only carried out once per participant. We also made sure that they were unaware of which visual quality would be degraded or if a degradation would happen.

The following sections present the method that we used to achieve the degradation of the three visual qualities targeted by this study. Note that the visual quality values were measured according to the system values and then converted to generally interpretable values, e.g., the visual acuity was measured in resolution and then expressed in arcmin to ensure their communicability.

Field of View

(a) (b) (c)

Figure 4.4 Representation of FOV degradation.

Note. a shows the environment without any degradation, at an FOV of 110 degrees; b shows the FOV degraded to 102.4 degrees; and, c shows the FOV degraded to 50.6 degrees.

We reduced the 110-degree FOV offered by the HTC Vive Pro Eye with the *postprocessing* package from Unity by using the vignette function. Figure 4.4 shows an example of such degradation.

Measurement and Degradation Rate As the vignette function does not allow for a directly convertible value of the FOV in degrees, we measured it after the experiment. To do so, observers had to look at a point in the center of the screen and move their extended right arm holding the controller, away from the center toward the right side of their FOV. At the exact moment when the observers weren't able to see the controller, they pressed its trigger, recording the angle between the controller and the FOV's center. The total FOV corresponds to the recorded angle between the border of the vignette at the recorded participant's functional threshold value and the center of its FOV multiplied by 2. For each

functional threshold, this measure was repeated three times. The results presented in this paper consist of the mean of those three measurements.

In the HMD, the vignette wasn't visible until it reached -0.7 arbitrary unit (AU) after 19 seconds. Thus, the FOV varied according to the following equations :

$$FOV_{deg} = \begin{cases} 110, & \text{if } 0 < t < 19, \\ 126.76 - 0.88t, & \text{otherwise.} \end{cases} \tag{4.1}$$

Where FOV_{deg} is the size of the field of view in degrees and t is the elapsed time since the beginning of the degradation in seconds.

Visual Acuity

(a) (b) (c)

Figure 4.5 Representation of visual acuity degradation.

Note. a shows the environment without any degradation, at a visual acuity of 4.6 arcmin; b shows the visual acuity degraded to 15.5 arcmin; and, c shows the visual acuity degraded to 50.7 arcmin.

To reduce the visual acuity, we used the *SteamVR* package to virtually reduce the resolution of the presented image by a parameterizable factor. This resulted in the presentation to the observer of what appears to be larger and larger pixels. Figure 4.5 shows an example of such degradation.

Measurement and Degradation Rate When the resolution parameter is set to 1, the system offers the maximum possible resolution with the equipment, in this case, 2880 x 1600 or a resolution acuity of 4.6 arcmin. The visual acuity varied according to the following equation :

$$Res_{deg} = 2880 - 37.43t. \tag{4.2}$$

Where Res_{deg} is the horizontal resolution of the screen and t is the elapsed time since the beginning of the degradation in seconds. The vertical resolution was reduced by the same factor.

To convert the screen resolution in arcmin, we used the following equation, which is based on Maxwell et al. [100]

$$VA_{arcmin} = 120 \arctan \frac{2Lens \tan \frac{V}{2}}{Res_{deg} Width}. \tag{4.3}$$

Where VA_{arcmin} is the visual acuity in arcmin, and Res_{deg} is the horizontal pixel resolution, with $Lens$, the horizontal lens' size, here equal to 59.47 mm, V, the horizontal visual angle, equal to 110 degrees, and $Width$ the horizontal screen' size, equal to 88.9 mm.

Contrast Sensitivity

(a) (b) (c)

Figure 4.6 Representation of contrast degradation.

Note. a shows the environment without any degradation, at a contrast of 31.2%; b shows the contrast degraded to 23.5 %; and, c shows the contrast degraded to 14.7 %.

We achieved contrast degradation by reducing the extreme luminance levels presented in the image. This was done with the *postprocessing* package from Unity that allows the contrast of the pixels presented to the user to be offset, varying between 100 AU and -100 AU. For this experiment, the initial contrast value was set to 0 and could be degraded up to -100 AU. At -100 AU the pixels all have the same dark gray color. Figure 4.6 shows an example of such degradation.

Measurement and Degradation Rate The contrast varied according to the following equation:

$$C_{deg} = -0.8t. \tag{4.4}$$

Where C_{deg} is the contrast and t is the elapsed time since the beginning of the degradation in seconds.

The recommended measurement for natural contrast stimuli is RMS contrast [108]. It is expressed as the ratio of the standard deviation and the mean luminance of the image. For this environment, the contrast presented as a Unity parameter value can be converted in RMS contrast using the following equation.

$$C_{RMS} = 0.2882 + 0.0011(19.29 + C_{deg}) + 3.786 \times 10^{-7}(19.29 + C_{deg})^3. \tag{4.5}$$

This equation was obtained with a polynomial regression of order 3 with an $R^2 = 0.9983$. We calculated the RMS contrast of images of the virtual environment corresponding to each recorded functional contrast sensitivity threshold value to compute the regression.

4.4.4 Experimental Procedure

One participant at a time took part in the experiment. After reviewing and signing the information and consent form, the participants sat in a chair positioned so that they could move their arms to the limits of their kinesthetic sphere without hitting surrounding objects. The experimenter then explained to the participants that they would have to perform four multi-directional selection tests.

To limit the possible bias linked to the participant's expectation to see a degradation, they were told that they would not necessarily see the degradation of each visual quality. They were given examples such as: they might see two times the degradation of contrast and two times no degradation or that they might see three times the degradation of FOV and once the degradation of visual acuity.

Participants were asked to put on the HMD and were given a controller before being asked to follow the HTC Vive Pro Eye eye-tracking camera calibration procedure to adjust their interpupillary distance. It should be noted that at this point, all visual qualities were set at their maximum value. The participants then performed a practice run of the multi-directional selection test for which no data was recorded. They were reminded that they had to do the task continuously, without stopping, until they had selected all targets.

Before each multi-directional selection test, the participants were reminded that they had to focus on the multi-directional selection test, executing it as quickly and accurately as possible and that there may or may not be any degradation. They were also prompted to say out loud when they perceived a degradation if they perceived any.

The participants then executed four multi-directional selection tests. The moment when a participant identified a degradation was recorded. The participants had to select all targets for the task to be completed. In the case in which a participant did not notice the degradation, the visual quality value recorded was the one at the end of the last target selection. Once a multi-directional selection test was completed, the environment was returned to the initial visual condition.

4.4.5 Experimental Design

Each participant performed a total of four multi-directional selection tests, one in which no degradation occurred and three in which either the FOV size, visual acuity, or contrast sensitivity was degraded. To control for a possible bias due to the presentation order the conditions were balanced according to a Latin square experimental design.

The dependent variables were the value for which an observer was able to perceive a degradation of its visual quality, i.e., the functional thresholds. Each experiment yielded three measures of functional threshold, one for each visual quality.

During a multi-directional selection test, the presentation of each ID was randomized between trials to control for bias due to participant expectations. Throughout the experiment, the participants made a total of 540 selections (15 targets x 9 IDs x 4 degradation conditions).

4.4.6 Data Analysis

A psychometric function was computed for each degraded visual quality. This function provides a statistical estimate of the functional threshold distribution from which we estimated the point of subjective equality (PSE) and the just noticeable difference (JND).

The PSE is defined as the degradation value that will be detected by 50% of the observers. It is considered as a degradation level for which the observer will judge the environment with an equal likelihood to be degraded or not [186]. The JND is the additional amount of degradation needed to increase an observer's detection rate from 50% to 75%. It is considered to reflect the observer's sensitivity to a degradation difference [187].

Each psychometric function was modeled as a 2-parameter cumulative Weibull distribution. Their parameters were estimated with the *fitdist* function from the *fitdistrplus* library in R using the maximum likelihood estimator method. The 95% confidence intervals (CI) of the distribution were estimated with the *confint* function from the same library using likelihood profiling.

Between-subject analyses were performed to take into account the potential effect of age, gender, dominant hand, visual correction devices (glasses and contact lenses), experience in virtual reality, and order of presentation on the measured functional threshold. We compared the mean degradation value of FOV, visual acuity, and contrast sensitivity for the previously enumerated variables. For gender (female or male), dominant hand (left or right), correction devices (wearing or not wearing), and experience in VR (less than 2 hours or more than 2 hours), we performed an independent-samples t-test when the data met the assumptions for a parametric test and a Mann-Withney U test otherwise. For age (less than 30 y.o., between 30 and 40, or more than 40) and order (first, second, third, or fourth) we performed a one-way analysis of variance (ANOVA) when the data met the assumptions for a parametric test, and Kruskal-Wallis test otherwise.

Within-subject analyses were performed to compare the results of the multi-directional selection test and verify the potential effect of the measured functional threshold on the task. We compared the mean throughput for order (first, second, third, or fourth) and visual degradation condition (FOV, visual acuity, contrast sensitivity, or no degradation), performing a one-way repeated measures ANOVA when the data met the assumptions for a parametric test, and a Friedman test otherwise.

Between and within-subject analyses, significance was evaluated with an $\alpha = 0.05$.

4.5 Results

This section first presents the results for the three measured functional thresholds. For each, it reports the PSE, JND, the scale and shape parameters of the estimated Weibull distribution along with the parameters of the confidence intervals. Then, the significant results of the between-subject analyses and within-subject analyses are reported.

4.5.1 Field of View

The PSE for FOV is 96.6 degrees with a 95% CI [92.3, 100.3] and the JND is 6.9 degrees, as illustrated in figure 4.7a. The computed cumulative Weibull distribution has a shape parameter $\beta = 11.86$ and a scale parameter $\eta = 99.59$. The lower bound of the 95% confidence interval has a shape $\beta = 8.15$ and a scale $\eta = 96.50$, and the higher bound has a shape $\beta = 15.58$ and a scale $\eta = 102.69$. Figures 4.7b to 4.7d show respectively a representation of what was seen by an observer's left eye at the maximum value allowed by the system, the PSE, and the minimal value recorded during the experiment.

A Kruskal-Wallis test revealed a significant effect of age on the detection of FOV degradation

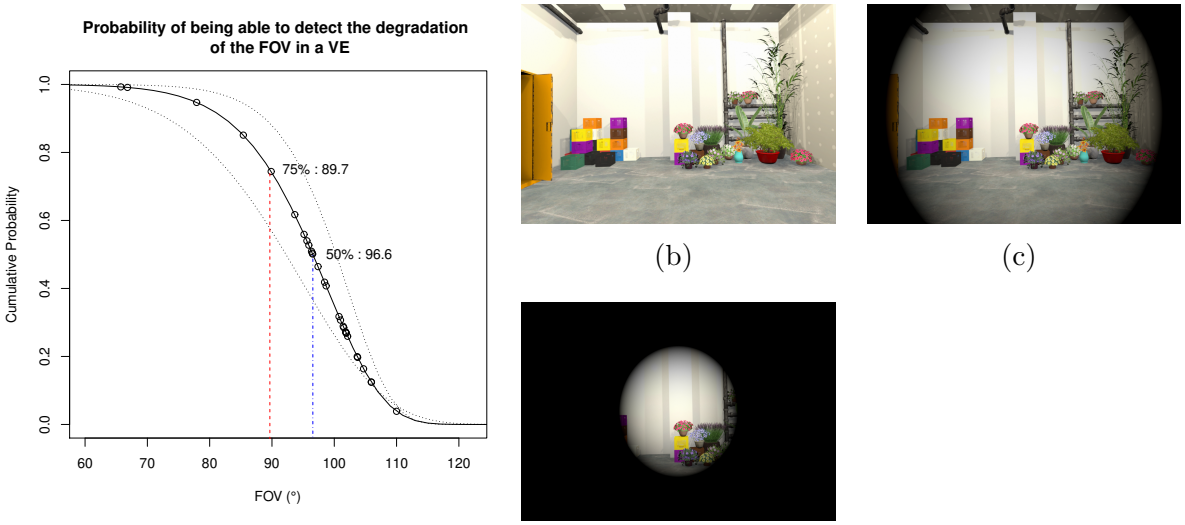


Figure 4.7 Results of FOV degradation.

*Note.*a Psychometric function showing the probability of a participant to be able to detect the degradation of its FOV in a virtual environment; b maximum recorded value; c PSE; d minimum recorded value.

$\chi^2(2) = 7.624$, $p < 0.05$. Post hoc pairwise comparisons using a Mann-Withney U test: $U = 66$, $n1 = 23$, $n2 = 3$, $p < 0.05$, showed that there was a significant difference between participants aged younger than 30 years old ($M = 97.34$ degrees, $SD = 8.98$ degrees) and older than 40 years old ($M = 67.59$ degrees, $SD = 17.43$ degrees).

Also, a Mann-Withney U test revealed that wearing a visual correction device had a significant effect on the detection of FOV degradation: $U = 2.037, n1 = 14, n2 = 16, p < 0.05$, participants wearing visual correction device ($M = 89.57$ degrees, $SD = 16.76$ degrees) detected the FOV degradation at a lower threshold than the participants not wearing them ($M = 99.56$ degrees, $SD = 7.01$ degrees). Further analysis did not reveal any significant differences between glasses and contact lenses wearer.

A Kruskal-Wallis test revealed no significant difference of the identified functional FOV threshold while comparing the order of execution of the multi-directional selection test $\chi^2(3) = 1.6376, p > 0.10$.

4.5.2 Visual Acuity

The PSE for visual acuity is 12.2 arcmin with a 95% CI [10.9, 14.2] and the JND is 3.3 arcmin as illustrated in figure 4.8a. The computed cumulative Weibull distribution has a

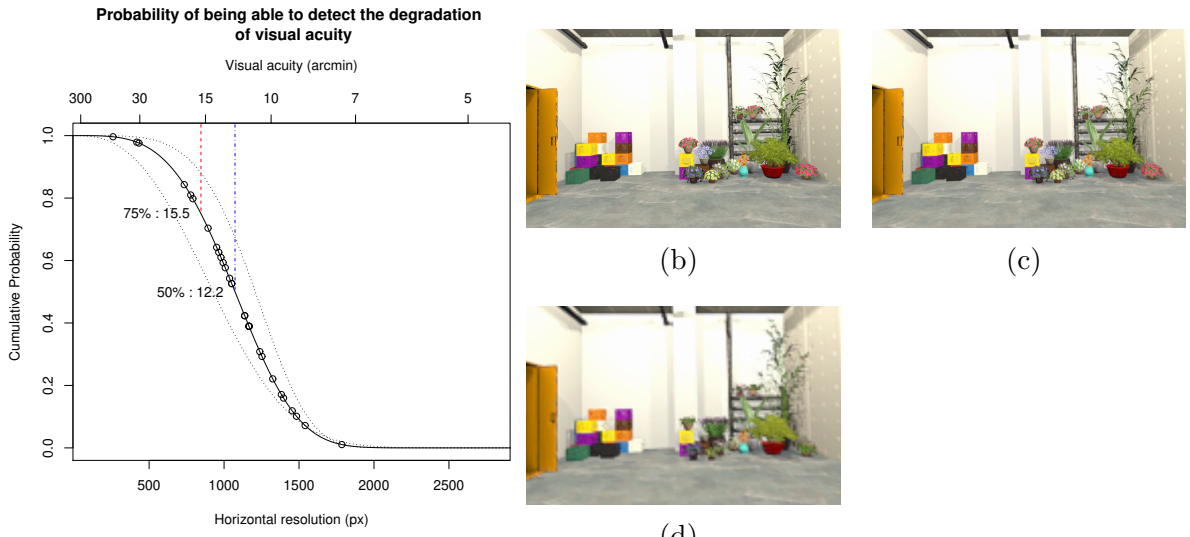


Figure 4.8 Results of visual acuity degradation.

Note. a Psychometric function showing the probability of a participant to be able to detect the degradation of its visual acuity in a virtual environment; b maximum recorded value; c PSE; d minimum recorded value.

shape parameter $\beta = 3.69$ and a scale parameter $\eta = 1185.04$. The lower bound of the 95% confidence interval has a shape $\beta = 2.64$ and a scale $\eta = 1064.96$, and the higher bound has a shape $\beta = 4.75$ and a scale $\eta = 1305.12$. Figure 4.8b to 4.8d show a representation of what was seen by an observer's left eye at the maximum value allowed by the system, the PSE, and the minimal value recorded during the experiment. We reported no significant between-subject analyses for this visual quality.

Moreover, a one-way ANOVA revealed no significant difference of the identified functional visual acuity threshold while comparing the order of execution of the multi-directional selection test $F(3, 26) = 0.359$, $p > 0.10$.

4.5.3 Contrast Sensitivity

The PSE for contrast sensitivity is 25.6% with a 95% CI [24.4, 26.6] and the JND is 2.1% as illustrated in figure 4.9a. The computed cumulative Weibull distribution has a shape parameter $\beta = 3.38$ and a scale parameter $\eta = 49.07$. The lower bound of the 95% confidence interval has a shape $\beta = 2.46$ and a scale $\eta = 43.57$, and the higher bound has a shape $\beta = 4.30$ and a scale $\eta = 54.56$. Figure 4.9b to 4.9d show a representation of what was seen by an observer's left eye for the maximum value allowed by the system, the PSE, and the

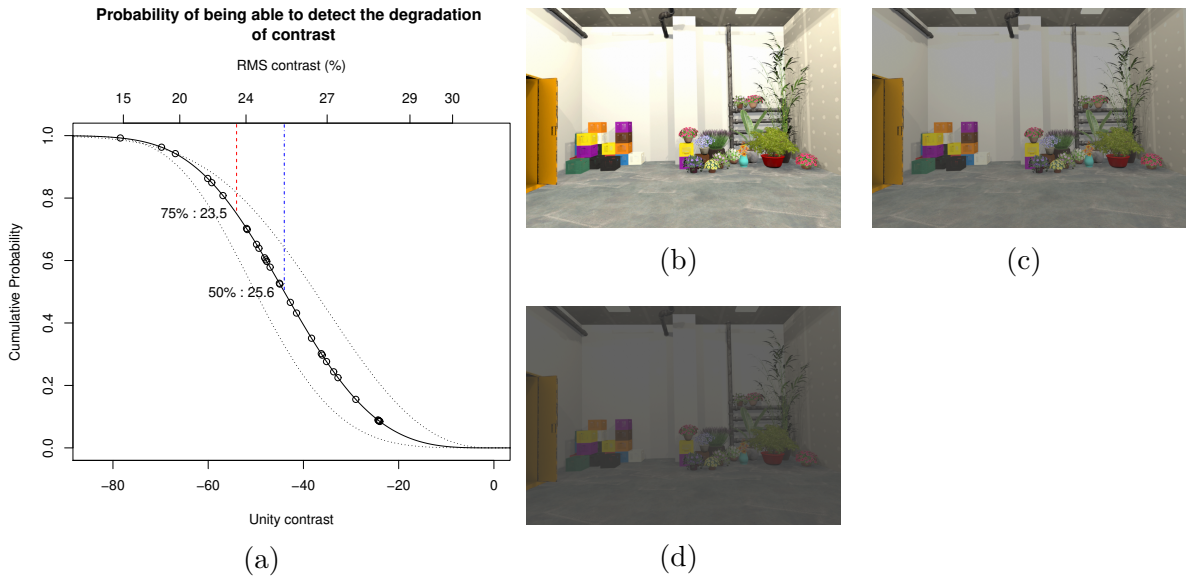


Figure 4.9 Results of contrast sensitivity degradation

Note. a Psychometric function showing the probability of a participant to be able to detect the degradation of its contrast sensitivity in a virtual environment; b maximum recorded value; c PSE; d minimum recorded value.

minimal value recorded during the experiment.

A two sample t-test showed that there was a significant effect of gender on the detection of contrast degradation $t(28) = -2.075$, $p < 0.05$, women ($M = 23{,}7\%$, $SD = 3{,}9\%$) detected the contrast degradation at a lower threshold than men ($M = 26{,}1\%$, $SD = 1{,}8\%$).

A one-way ANOVA revealed no significant difference of the identified functional contrast sensitivity threshold while comparing the order of execution of the multi-directional selection test $F(3, 26) = 0.2118$, $p > 0.10$.

4.5.4 Multi-Directional Selection Test Throughput Analysis

Table 4.1 Summary of post hoc pairwise comparison of task execution order

Comparison Group	Mean Difference	Sig.	Effect Size	DF
First and Second	-0.27	< 0.05	-2.768	29
First and Third	-0.41	< 0.001	-4.384	29
First and Fourth	-0.33	< 0.05	-2.828	29

A one-way repeated measures ANOVA showed a significant order effect on the throughput $F(3, 87) = 6.698$, $p < 0.001$. A post hoc comparison with 6 paired sample t-tests corrected with Holm-Bonferroni method to counteract multiple comparisons problem, was used to identify the differences between the conditions. As shown in table 4.1, the task executed first had a significantly lower throughput ($M = 4.68$ bps, $SD = 0.68$ bps) than the tasks performed second ($M = 4.96$ bps, $SD = 0.51$ bps), third ($M = 5.09$ bps, $SD = 0.64$ bps), or fourth (M = 5.01 bps, $SD = 0.62$ bps). Significant differences were only found between the first and the other order rank of execution of the task. Other comparisons revealed no significant differences.

Table 4.2 Summary of post hoc pairwise comparison of throughput between visual degradation conditions

Comparison Group	Mean Difference	Sig.	Effect Size	DF
FOV size and no degradation	-0.16	> 0.1	-1.563	29
Visual acuity and no degradation	0.09	> 0.1	0.9771	29
Contrast sensitivity and no degradation	-0.23	> 0.1	-2.114	29
Visual acuity and FOV size	0.25	< 0.05	2.825	29
Visual acuity and contrast sensitivity	0.32	< 0.01	3.373	29
FOV size and contrast sensitivity	0.07	> 0.1	0.612	29

Also, a one-way repeated measures ANOVA showed a significant effect of the degradation of visual qualities on the throughput $F(3, 87) = 4.2739$, $p < 0.01$. However, a post hoc comparison with 6 paired sample t-tests corrected with the Holm-Bonferroni method showed no difference between a degraded visual quality's throughput and the reference condition's throughput with no degradation. As shown in table 4.2 significant differences in throughput were only found between the visual acuity degradation and, respectively, the FOV size degradation, and the contrast sensitivity degradation. The visual acuity degradation condition had a significantly higher throughput ($M = 5.10$ bps, $SD = 0.56$ bps) than the FOV size condition ($M = 4.85$ bps, $SD = 0.61$ bps) and the contrast sensitivity condition ($M = 4.78$ bps, $SD = 0.71$ bps).

4.6 Discussion

The goal of this experiment was to measure the functional thresholds of FOV, visual acuity, and contrast sensitivity. These thresholds represent the values below which virtual environ-

ment's users can perceive a degradation of their visual qualities while their attention is not directed toward the discrimination of the degradation. In this section we analyze the results obtained following the experiment, first focusing on each targeted visual quality, and then in a more general manner.

4.6.1 Functional Thresholds

To determine the functional thresholds we used the PSE measured using a 2-parameter cumulative Weibull function based on the experimental data collected from 30 participants for each visual quality degradation. Since functional thresholds represent a lower bound of perception, we chose this conservative value to assess them.

In this section, we provide a critical analysis of the results obtained. For each targeted visual quality, we examine the relation between the measured functional threshold to the information identified in the related work and we discuss our study limitations.

Field of View

For FOV, we identified a functional threshold of 96.6 degrees. It represents a degradation of 13.4 degrees from the initial setting of 110 degrees. This functional threshold is smaller than the FOV covered by both eyes when an observer looks straight ahead [6, 74].

This result seems to corroborate the statement that methods assessing the UFOV neglect the contribution of peripheral information [94]. Using the common UFOV value of 40 degrees and adding 50 degrees to include the fixed foveated region, the region where an observer can sustain its eye gaze, we would expect a maximal value of 90 degrees for the FOV in which information can be perceived and processed. Yet, our measured FOV functional threshold is larger and seems to correspond to the UFOV of 80 degrees from the adapted UFOV of Itoh et al. [95]. When taking into account the fixed foveated region this value spans up to 130 degrees.

We suggest that the difference between the value of 130 degrees and our measured functional threshold of 96.6 degrees is because the portion of the FOV that is used by an observer depends on the task performed [83]. During the multi-directional selection test, the maximum portion of the FOV covered by the targets is estimated at 55 degrees which is encompassed in the measured functional threshold value. We consider that the measured value discussed in this section should be interpreted as the FOV functional threshold when the elements related to the task can all be included in the visual field of the observer. However, our experiment does not allow us to issue recommendations on the relation of the size of an object in the FOV

and the FOV functional threshold. Further experiments should be carried out to determine to what extent this functional threshold can be applied.

Some factors could have biased our functional threshold assessment of FOV. These factors essentially consist of a glitch in the virtual environment and the effect of age and visual correction devices on the identified degradation value.

In figure 4.7a we can observe that the confidence interval is relatively large and becomes narrower around a FOV size of 100 degrees. This may be due to a glitch that occurred when the vignette parameter was dynamically reduced, causing a lag in the update of the pixels near the border of the vignette. This glitch was only apparent when the participant's head moved. One participant described it as what one sees when its head emerges from the water while wearing swimming goggles. In total, six participants reported seeing this glitch, but it wasn't clear if it was before or after perceiving the degradation, and most of them mentioned that they reported the degradation because they saw the tunneling vision effect. Therefore, the functional FOV threshold measured may be overestimated and a lower threshold could be expected for the same task.

Also, our analysis revealed a significant effect of age and visual correction devices on the measured FOV functional threshold.

Participants older than 40 years old ($n = 3$) identified a significantly lower FOV functional threshold than participants aged younger than 30 years old ($n = 23$). Since our population of participants older than 40 years old is relatively low, we assume that more participants would be necessary to make a reliable conclusion on the effect of age.

Finally, participants wearing visual correction devices identified lower FOV functional thresholds. We hypothesize that this effect may be due to the glasses' frame that could have veiled the degradation, yielding a possible underestimation of the measured functional threshold value of FOV.

Visual Acuity

For visual acuity, we identified a functional threshold of 12.2 arcmin. The range of the 95% CI [10.9-14.2] corresponds to the visual acuity at 20 to 30 degrees of eccentricity from the fovea [87, 88]. This visual acuity functional threshold seems to reflect the fact that most of the perception is done through the peripheral vision field [103].

At the value of 12.2 arcmin, the smallest object relevant to the multi-directional selection test was still clearly visible, as the smallest target width was 343.5 arcmin. The degradation of visual acuity mostly made the object in the background difficult to discern. Most participants

measured the extent of this degradation only at the end of the task. At that point, they were able to take the time to look at the environment in general and not only focus on the targets.

Contrast Sensitivity

For contrast sensitivity, we identified a functional threshold of 25.6% RMS contrast. The identified RMS contrast was 5.6% less than the initial situation of 31.2 %. This functional threshold is close to the contrast sensitivity threshold of 19% to 23% at 30 degrees of eccentricity from the fovea [88]. It is also slightly higher than the threshold of 23.3% that we found when applying Dolhasz et al. [111] contrast offset threshold equation to our context. The difference may be the corollary of the poorer contrast sensitivity of observers using an HMD [99]. Yet, the value we measured seems to also indicate that perception is done through the peripheral vision field [103].

Three participants were not able to identify the contrast degradation during the experiment. However, they reported that the degradation was obvious either at the moment when they finished the task, or just after the environment was set back to its original state. We suggest two explanations for this phenomenon. First, we assume that they might have perceived the degradation and weren't just confident enough in their judgment to report it at this moment. Second, this inability to identify the degradation may be due to the degradation rate established for the experiment, which was too slow to reach a value sufficiently low to be detected in the allotted time.

We found a significant effect of gender on the detection of contrast degradation. In our experiment, women ($n = 13$) were less sensitive to contrast degradation than men ($n = 17$). According to our debriefing interviews, we believe that our male population had more video gaming experience. Hence, they were more likely to have already practiced tasks similar to the one executed during the experiment, making them better at shifting their attention between the multi-directional selection test and the contrast discrimination.

4.6.2 Multi-Directional Selection Test

In this section, we discuss the impact of the multi-directional selection test on our experiment. We examine the practice effect identified with a one-way repeated measures ANOVA, then we analyze the effect of the task on the degradation detection.

A one-way repeated measures ANOVA revealed an order effect on the throughput that can be attributed to the participants' lack of practice. The first multi-directional selection test

performed by the participants had a significantly lower throughput than the three others. As Soukoreff et al. [183] recommend, we asked our participants to do a practice test that was not recorded to familiarize themselves with the multi-directional selection test. They had the option to do as many practices run as they needed to feel comfortable with the task and every participant chose to do only one. As reflected by this analysis, participants should have completed two practice tests to eliminate the practice effect on the throughput. However, the throughput was not the variable of interest in this study.

To verify that this order effect did not impact the identification of the degradation, we performed a between-subject analysis of every visual quality degradation value. These tests revealed no significant difference when comparing the order of execution of a degradation. Thus, we cannot assume that the order in which the visual qualities were degraded affected the detection of the degradation

Also, to determine the effect of the multi-directional selection test on the degradation condition, we compared the throughput of each degradation condition. A one-way repeated measures ANOVA revealed no difference in performance between a task with degradation and without degradation. We suggest that this indicates that the multi-directional selection test was well suited for the identification of a functional threshold.

Surprisingly, we found that the throughput was significantly higher for the visual acuity condition than the FOV and the contrast sensitivity conditions. When looking at the mean difference in table 4.2, we see that the visual acuity is the only degradation to have a positive mean difference with the no degradation condition. We would have expected the degradation to affect all the throughput scores in the same way. Since the visual environment was degraded, we expected that all throughput scores would be less than the no degradation condition.

We suggest that the loss in visual acuity may have enhanced the precision of the participants in the multi-directional selection test. As the boundary between the target and the background became less clear, the participant had to make selections more toward the center to keep up with the instructions related to the multi-directional selection test, i.e., to select targets as quickly and as accurately as possible. We think that the degradation of FOV size and contrast sensitivity did not allow for such a shift in the participant's behavior. Thus, the significant differences between these conditions could be explained by the fact that the visual degradation affected the participant just enough to have a significantly different behavior between visual acuity degradation and FOV size or contrast sensitivity, but not enough to affect significantly these conditions when compared to the no degradation condition.

4.6.3 Implication of Functional Thresholds

The functional thresholds identified in this paper could serve as constraining values for the assessment of immersion in a virtual environment. We believe that for a virtual experience to be potentially considered realist, it must be able to simulate each visual quality with values comprised in the range bounded by the perceptual threshold as the upper limit and the functional threshold as the lower limit. This kind of evaluation consists in an objective way to characterize the level of immersion of a virtual experience provided by a specific technology. Such evaluation presents the advantage that it can be used to predict the level of immersion of a system without the intervention of any user. The values for the visual qualities targeted by this study are included in table 4.3.

Yet, these functional thresholds need to be used carefully regarding the limitation of our experiment.

We first assume that functional thresholds are task-dependent. Different task parameters could have yielded different thresholds, e.g., small characters in a reading task would have necessitated a greater visual acuity. Thus, we argue that when such task parameters are known, they should constitute the functional threshold for the said task.

The identified functional threshold values may also have been influenced by the device used for the study. Because of technological constraints, it is not yet possible to simulate the human's perceptual thresholds. Thus, some functional thresholds targeted by this study might have been underestimated since the starting perceptual value might have already been under the functional threshold. For example, using a device with a larger FOV base value might have yielded a larger functional threshold.

Additionally, our sample size prevents us from drawing conclusions that can be generalized about the identified functional thresholds. Especially the fact that we suspect an effect of age on our identified threshold, but don't have enough statistical power to draw reliable conclusions on that variable. Future work should address a larger and more varied population to resolve this issue.

Further, although our participants could move their head freely, the executed task was done in one specific location and direction, implying a static condition. Since non-static conditions are common in virtual environments, e.g., navigation, it is important to note their implications vis-a-vis the three functional thresholds addressed in this paper. In non-static conditions, the FOV functional threshold value could be greater. This is because visual performances are not equal across the entire FOV, as peripheral vision is more sensitive to movements. Similarly, the visual acuity and contrast sensitivity functional thresholds could be lower for

Table 4.3 Perceptual and functional thresholds of FOV, visual acuity and contrast sensitivity

Visual quality	**Measure**	**Perc. threshold**	**Funct. threshold**
FOV	FOV size	290 degrees horizontally, 190 degrees vertically [74]	96.6 degrees vertically and horizontally
Visual acuity	Resolution acuity	1 arcmin [98]	12.2 arcmin
Contrast sensitivity	RMS Contrast	0.2% [88]	25.6%

Note. The FOV sizes presented in table 4.3 take into account the eye movements.

such conditions, since the performance for the identification of these thresholds are better when the image is fixed and at the center of the fovea.

Moreover, in this study, a visual quality functional threshold was measured while the other visual qualities were kept at their maximum possible value. However, as mentioned in the related work section, visual qualities are not independent of each other. Using different visual quality values could result in the measurement of different functional thresholds. For example, a low-resolution value combined with a low contrast value might yield a lower functional threshold.

Nevertheless, we believe that using a multi-directional selection test to measure the visual qualities allowed us to estimate a limit that can be used in diverse situations since no particular sensitivity to the tested visual qualities was required by the task. The FOV size encompassed the entire task elements and the targets that needed to be selected were large enough and clearly distinct from other objects in the scene to be seen in any degradation condition. Therefore, we suggest that the measured values constitute general indications that can be used when the perception of a task is not limited by known values. We think that by using a conservative approach, i.e., assessing the functional thresholds with the identified PSE, we were able to present valid values of the concept of functional thresholds.

To measure the extent of our assessment, the results obtained in this paper should be evaluated with other tasks that require no particular sensitivity to the tested visual qualities. Also, additional experiments including a diverse sample of participants, a variety of devices, and distinct visual quality values combination would need to be conducted.

At the moment, more research is needed to characterize immersion using functional thresh-

olds. Only for vision, we still need to identify the functional threshold of visual accommodation, eye movements, stereopsis, perception of movement, and color perception. Similar work would also be required for other sensory modalities, i.e., audition, haptic, etc. However, such effort would not only benefit the measurement of immersion, it would also enhance our comprehension of human sensation and perception.

4.7 Conclusion

In this paper, we introduced the notion of functional threshold which constitutes a limit below which perception should be considered functionally degraded. Thus, below these values, a virtual environment seems unlikely to be appropriate to carry out equivalent real-world tasks.

We measured the functional thresholds for three visual qualities. We've assessed that for real-world tasks executed in a virtual environment a FOV of 96.6 degrees, a visual acuity of 12.2 arcmin, and a contrast of 25.6% constitute limiting values.

We support the idea that values farther from the perceptual thresholds than the functional thresholds will surely be considered as degraded by an observer, while values closer might not.

These results can be used as constraints to establish an objective measure of immersion for the emergence of presence in a virtual environment. In a future study, we will use them as a reference to evaluate the impact of the extensiveness of visual qualities simulation on presence in a virtual environment. Moreover, they will be used to make recommendations on the technological requirements for a virtual environment. For this purpose, the functional threshold reported in this paper shall be considered as minimal values to be used so observers are unable to detect with certitude that their visual qualities are degraded while in a virtual environment.

4.8 Acknowledgments

This work was supported by Mitacs, the Consortium for Aerospace Research and Innovation in Canada (CARIC), the Consortium de recherche et d'innovation en aérospatiale au Québec (CRIAQ) and by its partners under GRANT CARIC-CRIAQ MDO-1649-TRL4+. We would also like to express our sincerest gratitude to Antoine Martin for his technical contribution to software development.

CHAPITRE 5 ARTICLE 2 : THE RELATIONSHIP BETWEEN IMMERSION AND PSYCHOPHYSIOLOGICAL INDICATORS

Cet article a été soumis dans PRESENCE : Virtual and Augmented Reality le 3 février 2023. Ce chapitre consiste en une version corrigée suite à l'évaluation des relecteurs d'une première version de l'article.

Hébert-Lavoie, M. (Polytechnique Montréal), Doyon-Poulin P. (Polytechnique Montréal), Ozell, B. (Polytechnique Montréal)

5.1 Abstract

Psychophysiological indicators have been subject to a growing interest in *presence* research. However, despite the generated interest, the nature of the relationship between psychophysiological indicators and the factors of presence is still undetermined. We do not know how psychophysiological indicators are affected by immersion and coherence, the two factors of presence. To study this relationship, we objectively manipulated immersion by altering three visual qualities. The visual qualities were set to values over, at, or under their functional threshold, defined as a perceptual boundary under which a sensory quality value should be considered functionally degraded. Sixty participants executed a driving task in a virtual environment under different visual conditions while we measured their presence, mental workload, cardiovascular response, and eye response. Our results indicate that contrast sensitivity and visual acuity degradation under the functional threshold decreased presence. These immersion conditions yielded significantly different psychophysiological indicators results than the condition without degradation. Also, the degradation of visual acuity under the functional threshold increased our participants' mental workload. During our experiment, we observed that visual qualities under the functional threshold introduced unreasonable circumstances which modify a user's expectations and behavior in a virtual environment, resulting in decreased coherence. Since we only observed an effect of immersion degradation on our measured variables when it introduced unreasonable circumstances, we suggest that mental workload and psychophysiological indicators measure the coherence of a virtual environment. Also, we consider that immersion conditions under the functional threshold directly affect the coherence of a virtual environment.

5.2 Introduction

In a previous experiment, we introduced the concept of functional threshold, an objective indicator of the realism of a virtual experience [188]. A *functional threshold* is a sensory quality value beyond which perception is considered functionally degraded. It is a notion that complements the perceptual threshold, i.e., a limiting sensory quality value for which humans can perceive. Functional thresholds do not represent a fixed value indicating when a task should not be considered realistic. Instead, we should interpret them as such: values farther from the perceptual thresholds than the functional thresholds are more likely to be deemed degraded by an observer, while values closer might not.

Immersion is defined as the extensiveness of the simulation of sensory modalities by a system [20]. It is a factor of the concept of presence [19]. It is possible to constrain immersion between a maximal and a minimal bound using functional thresholds. The perceptual thresholds represent the maximal bounds, and the functional thresholds represent the minimal bounds. These constraints define the perceptual values between which presence could be evoked. Functional thresholds bring objectivity and meaning to the manipulation of immersion. Rather than setting their experiment at arbitrarily chosen values of low or high immersion, researchers can use them as reference points.

Since functional thresholds allow the manipulation of immersion in an objective and meaningful way, we used them to verify the effect of variables sensitive to presence. By manipulating immersion, we hope to get insights into how a variable is related to the feeling of presence. Psychophysiological indicators are of interest to researchers in the field of presence as they are considered an objective assessment of presence [50]. However, while their sensitivity has been demonstrated, their relation to presence must be clarified. They have been associated with both increases and decreases in presence [40]. Thus, we suspect they might have been wrongfully used as a holistic measure of the concept.

We have previously measured the functional thresholds for contrast sensitivity, size of the field of view (FOV), and visual acuity [188]. In the present study, we manipulated those visual qualities while participants executed a driving task to investigate the effect of immersion on psychophysiological indicators. This method of manipulating presence will allow us to analyze immersion's impact on psychophysiological indicators. The goal is to understand how they can be used as an objective way to quantify presence in a virtual environment. Hence, this paper aims to verify the effect of an objective degradation of immersion – a factor of presence – on psychophysiological indicators.

To contextualize the topic of this study, we will first discuss matters related to the concept

of presence in a virtual environment, detail the role of the sensory modalities targeted by the study toward presence, and present the psychophysiological indicators and their relation to presence. Afterward, we will describe the methodology applied to reach our objectives and present its results. Finally, we will discuss the implication of psychophysiological indicators on immersion and conclude on their impact on presence.

5.3 Related Work

5.3.1 Presence and Related Concepts

Presence qualifies and quantifies a user's feeling of a virtual environment [13]. Higher presence contributes to better performance, engagement, and skill learning [12, 37].

Skarbez et al. [15] define the concept of presence in a virtual environment as the "perceived realness of a mediated or virtual experience." As shown in figure 5.1, the subjective experience of presence is formed by two illusions: the plausibility illusion and the place illusion. The plausibility illusion refers to the illusion that the events that occur in a virtual environment are happening for real. The place illusion refers to the quale of being in a real place [19]. These illusions are regulated by two aspects of a virtual environment: coherence and immersion. Coherence contributes to the plausibility illusion, and immersion contributes to the place

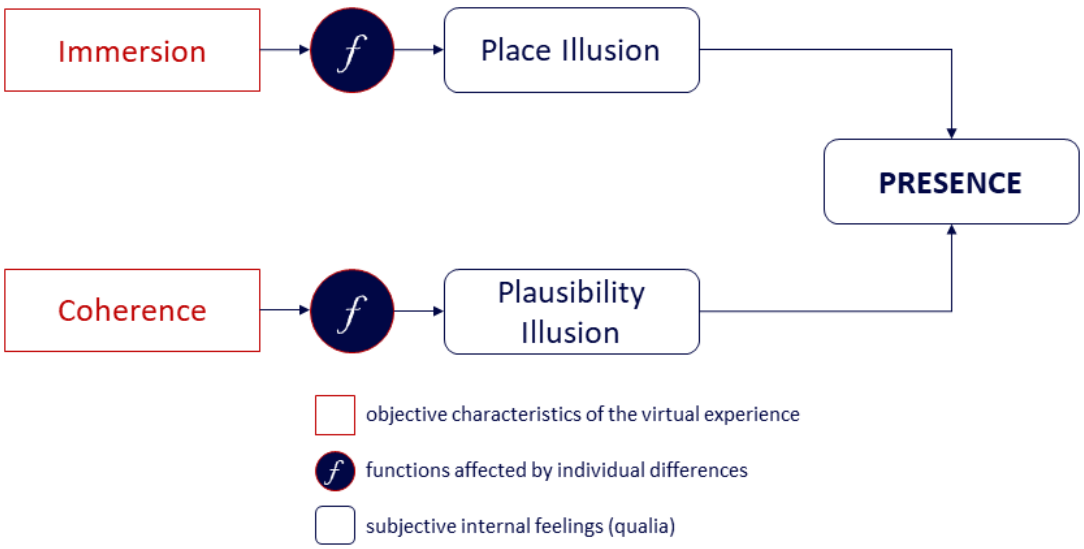


Figure 5.1 Relationship between presence concepts

Note. This figure is based on Skarbez et al. [15] proposed relationship between presence concepts.

illusion. Coherence and immersion define the limits in which the plausibility and the place illusion can be evoked [15].

Coherence refers to the aspects of a virtual environment that contributes to the plausibility illusion. It is defined as the ability of the environment to act as the user expects it to act [15]. Coherence represents the set of reasonable circumstances that a scenario can evoke without introducing unreasonable circumstances. A reasonable circumstance is a state of a virtual scenario that is evident from prior knowledge of the individual experiencing it. Reasonable circumstances are different for every individual [30].

The place illusion is part of the concept of immersion. Immersion represents the set of valid actions supported by a virtual environment system [19]. It is an objective characteristic of a virtual environment designating the extent and the precision of the simulated sensory modalities [20]. For example, to evaluate immersion provided by the visual modality, one would need to assess its qualities, namely visual accommodation, eye movements, stereopsis, perception of movement, size of the field of view, visual acuity, light and contrast sensitivity, and color perception [188]. Immersion provides the boundaries within which the place illusion can occur. Thus, the place illusion depends on the system used to evoke it [19].

Functional Thresholds as Indicators of Immersion

Immersion can be quantified by measuring the qualities composing each sensory modality simulated by a system [36]. A virtual environment simulating more relevant sensory modalities has been shown to increase presence [54]. Still, this kind of evaluation is not meaningful in itself, as there is no clear information on the outcome of a given level of immersion.

To better indicate the expected effect of the value of sensory qualities, we have introduced the notion of "functional threshold" [188]. This notion is used as a perceptual boundary. It is defined as a sensory quality value below which perception should be considered functionally degraded. The other sensory boundary is the perceptual threshold, defined as the limiting sensory values resulting in perception. Hence, we expect that a virtual environment has a greater probability of evoking the place illusion if it involves sensory quality values between the functional and perceptual threshold [188].

Thus, ideally, technologies that simulate the perceptual threshold of every sensory quality would be used to maximize presence. However, simulating the perceptual threshold of every sensory quality has yet to be achievable. In fact, it is believed that functional fidelity would be more attainable than perfect physical fidelity [36]. It is difficult to identify exact thresholds of immersion since their assessment is relative rather than absolute [36]. Still,

functional thresholds give an estimation of the sensory quality values that are necessary to evoke presence.

Measuring immersion through the assessment of sensory quality values has several advantages that resemble a valid ideal measure of presence defined by Souza et al. [13]. It is relevant, with a direct link to presence; it is reliable and can be applied to any virtual environment; it allows the objective quantification of variation; it is non-intrusive and inexpensive, with the possibility to be applied asynchronously without the need for participants. Moreover, it is an objective characterization that considers the system used and that can be done prospectively without any user intervention. Its most prominent flaws are that it measures immersion rather than the whole concept of presence and that data related to functional and perceptual thresholds can be difficult to gather for the time being.

The following paragraphs present the main features of the three visual qualities for which a functional threshold has been previously measured [188]: contrast sensitivity, size of the field of view, and visual acuity.

Contrast Sensitivity The concept of contrast sensitivity is a means of taking into account light sensitivity without being affected by the time-varying nature of this visual quality [51]. It enables the use of a context-independent measure of the phenomenon. Contrast sensitivity is equivalent to the slightest difference in luminance required to identify a target [96]. When dealing with natural contrast stimuli, the recommended measurement is the root mean square (RMS) contrast expressed as the ratio of the standard deviation and the mean luminance of the image [108]. Contrast sensitivity is a visual quality. Manipulating it implies an impact on immersion and should theoretically affect presence. However, we could not find any paper addressing the relationship between contrast sensitivity and presence.

One's contrast sensitivity is dependent on visual conditions. When optimal, the perceptual threshold, i.e., the minimum contrast that can be perceived in the fovea, is 0.2 % [88]. Furthermore, the functional threshold corresponds to 25.6% RMS contrast with a 95% confidence interval (CI) [24.4, 26.6] [188].

Size of the Field of View The FOV corresponds to the portion of space that the eyes can capture when motionless [51]. Generally, higher presence is evoked by values closer to the total FOV size [9, 76, 77]. Larger FOV size also increases workload [189] and improves task performance [78]. However, in virtual environments with vection, causing a visual-vestibular cue conflict [39,80], a large FOV is a significant factor in cybersickness [81,82], a phenomenon generally negatively correlated with presence [39, 190].

When considering the amplitude in which the eyes can move while the head is fixed, to encompass the whole range of possible values that equipment such as a head-mounted display (HMD) should have, the perceptual threshold of the FOV is 290 degrees horizontally and 190 degrees vertically [74]. In comparison, the functional threshold of FOV is 96.6 degrees both horizontally and vertically [188].

Visual Acuity Visual acuity refers to the finest detail that can be detected or identified [96]. Generally, the notion of resolution acuity is used to measure visual acuity in a virtual environment. Resolution acuity corresponds to the smallest angular separation between two neighboring objects that can be detected [51].

Under optimal conditions [62], the human eye has a resolution acuity of 1 arcmin in average [58,73,83,97,98]. The identified functional threshold for visual acuity when the head is stable is 12.2 arcmin [188].

Since the angular resolution of most current HMDs is 4 to 6 arcmin, the pixel density is too low to reproduce maximal visual acuity [100]. Thus, objects are difficult to discern at distances where they would be clearly seen in real life [6, 83], such as writing on road signs. The lack of realism concerning visual acuity is reflected only to some extent in presence measurement [101]. An increased screen resolution significantly increases the reported presence when there is a great difference between the tested resolution condition [57, 77]. However, smaller differences between conditions do not yield significant results [102].

In summary, changing the values of visual qualities allows for manipulating immersion objectively and consistently. The contrast values that a virtual environment should offer to evoke the place illusion should be between 0.2% and 25.6%, the FOV between 96.6 degrees both horizontally and vertically, and 290 degrees horizontally and 190 vertically, and the resolution values should be between 1 arcmin and 12.2 arcmin. However, the presented functional thresholds were obtained for a task where the head was stable, and the participants only had to focus on one element of the virtual environment [188]. We do not know if these values would still be valid in a different context.

5.3.2 Psychophysiological Indicators

In recent years, there has been a growing interest in using psychophysiological indicators to measure presence [50]. Some psychophysiological indicators, like basic heart rate have been shown to be related to the coherence factor of presence [65]. Psychophysiological indicators include cardiovascular response metrics, and eye response metrics, as the most widely used measures across varied research domains [113].

Psychophysiological indicators offer the benefits of being continuous, relatively non-invasive, real-time, and, to some extent, an objective assessment of presence [40,50]. Also, contrary to subjective self-reports of presence, psychophysiological indicators are believed to be matched between a virtual environment and a real environment [9]. Thus, helping investigate phenomena that are difficult to capture using self-report methods and providing a more accurate analysis of the state of the use of a virtual environment.

The preferred use of psychophysiological indicators in evaluating presence is to assess related factors [40]. Identifying the concepts related to presence should provide a way to improve the validity, sensitivity, and diagnosticity of its measurement by psychophysiological indicators. Generally, psychophysiological indicators are used to objectively evaluate cognitive activity, emotional responses, and cognitive-emotional states [38]. Cardiovascular and eye responses have been used to evaluate phenomena such as frustration, anxiety [38], cybersickness [47,48], stress, and workload [49] which have been shown to affect presence [13,37,41,68,69,126,191, 192].

Among the many concepts evaluated by psychophysiological indicators, mental workload appears particularly relevant for presence assessment. Mental workload represents the cost of mental resources necessary to accomplish a task [112, 113, 116]. It is relevant to every presence evaluation context, unlike emotional responses or cognitive-emotional states, which might not have enough intensity to apply to every presence assessment situation.

Previous studies showing a link between presence and mental workload [37, 68, 69] reported inconsistent results. Experiments studying a landmine neutralizing task [37] and an army room clearing task [68] indicated a significant negative correlation between mental workload and presence. However, in a free throw basketball task, mental workload was significantly positively correlated with presence [69]. Thus, mental workload should not be treated as an absolute presence indicator. Instead, it should be used to evaluate if two scenarios involve cognitive functions similarly with the same intensity [10].

Psychophysiological indicators represent the dynamic trend of mental workload and thus make it possible to estimate the instantaneous load [114], i.e., the number of mental resources available at a specific point in time [119]. While most psychophysiological indicators were found to discriminate changes in mental workload, they were not universally valid in all task scenarios. Psychophysiological responses caused by mental workload are highly scenario dependent, are affected by several task characteristics and individual differences [113], and, as previously pointed out, can also be attributed to several other phenomena. Hence, the use of multiple types of mental workload measures, such as self-report measures and psychophysiological indicators, is recommended to capture the concept. Various mental workload measures

allow researchers to hypothesize about the factors contributing to the total load [122] and to obtain a more accurate assessment of mental workload [121].

Cardiovascular Response Metrics

There is a relation between cardiovascular responses and mental workload [149]. Cardiovascular responses – the heart's reaction and its related systems to psychophysiological events – are regulated by the sympathetic and parasympathetic nervous system [138]. Cardiovascular measures are thought to be sensitive to mental workload since sympathetic nerves take control of cardiac activity when individuals are under a high mental workload causing distinguishable variation in the cardiovascular response [113].

Cardiovascular response measures have been validated in discriminating mental workload across varied research domains and thus are recommended for such use [113]. Means of quantifying cardiovascular responses include basic heart rate measurements and heart rate variability (HRV) measures.

Basic heart rate refers to the relation between time and heartbeats. An increase in heart rate is associated with a consistent increase in mental workload during real-world and simulator driving tasks [10, 49, 143].

HRV is described as the change in the time intervals between adjacent heartbeats [144]. Since many factors influence HRV measures, an awareness of the recording context and subjects is essential. HRV is influenced by posture, age, sex, heart rate, type of activity [174], health status [138], and information regarding respiration [121, 174].

Generally, HRV metrics, such as RMSSD, SDNN, pNN50, LF power, and HF power increase with drowsiness, fatigue, and disengagement and decrease with workload, stress, and vigilance [49, 174]. LF/HF does not have the same dynamic toward these phenomena. A low LF/HF reflects parasympathetic dominance, which occurs when individuals engage in tend-and-befriend behaviors. A high LF/HF indicates sympathetic dominance, which is linked to fight-or-flight behaviors [174].

HRV metrics were used to evaluate mental workload in car driving tasks. According to Mehler et al. [143], RMSSD and LF power were reported to provide statistically significant differentiation, while SDNN, pNN50, HF power, and the LF/HF metrics all failed to differentiate conditions. Stuiver et al. [149] used LF and HF power to differentiate between conditions. Both were lower with a higher mental workload but did not differ significantly between conditions. Shakouri et al. [150] reported significant differentiation of mental workload conditions as measured by NASA-TLX, while RMSSD, LF, HF, and LF/HF did not

yield significant results. Finally, in an aviation context, Alaimo et al. [151] indicated that HRV metrics were consistent with the results obtained with NASA-TLX. Their results were not significant, but they reported that their conditions with a higher LF/HF, lower RMSSD, and SDNN corresponded to a higher NASA-TLX score.

Past studies have obtained mixed results on the effectiveness of HRV metrics as a measure of mental workload in driving or driving-like context. The insensitivity of the HRV measures is supposedly due to an insufficient mental or emotional demand [150] or to a confound resulting from a continuing rise in heart rate caused by the initial reaction of the individual to the studied condition and a decrease in heart rate due to regulation effect [149].

Eye Response Metrics

Eye response metrics can capture mental workload changes induced by visually demanding tasks [113]. Among the many types of eye response metrics [127], this paper will focus on a pupil size-related metric, i.e., the low-frequency/high-frequency index of pupillary activity (LHIPA).

Studies have revealed that pupillary activity is correlated with mental workload [127] and drowsiness [134]. An increase in pupil diameter is reliably associated with an increase in mental workload, while a decrease is associated with drowsiness [49]. Pupil diameter measures are widely used to assess users' mental workload when interacting with human-computer interfaces [127].

The LHIPA is a wavelet-based algorithm inspired by the Index of Cognitive Activity (ICA) and Index of Pupillary activity (IPA). It is sensitive to variations of mental workload manipulated by task difficulty. Like in cardiovascular response analysis, the low frequency/high frequency (LF/HF) ratio reveals changes in components of the autonomic nervous system. Decreases in pupil diameter indicate parasympathetic activation or sympathetic inhibition. Increases in pupil diameter are related to sympathetic excitation or parasympathetic inhibition [136].

The LHIPA is more effective at detecting load than the similar IPA [127]. It also presents the advantage of yielding significant results in working memory capacity and is negligibly affected by small angles off-axis distortion of the pupil diameter. However, it appears that it is not sufficiently sensitive to distinguish between some levels of task difficulty [136].

5.3.3 Research Objectives and Hypothesis

In a previous experiment, we measured the functional thresholds for contrast sensitivity, size of the field of view, and visual acuity [188]. Functional thresholds provide a way to meaningfully manipulate immersion, indicating under which visual quality value a virtual environment should be considered degraded. It provides a way to analyze the effect of an objective degradation of immersion with more diagnosticity. Thus, this paper aims to study the effect of a manipulation of immersion on presence, mental workload, cardiovascular response, and eye response.

Immersion has been related to presence. Thus, we expect to observe a decrease in presence by varying the value of visual qualities under their respective functional threshold.

Basic heart rate has been shown to be related to the coherence factor of presence [65]. However, there is no evidence establishing if it also pertains to immersion. Moreover, we have found no studies about the relationship between other psychophysiological indicators like HRV metrics and LHIPA and factors of presence. Because of the added perceptual load generated by degrading visual qualities, we expect the psychophysiological indicators to be sensitive to an immersion's manipulation under the functional threshold.

Since psychophysiological indicators like cardiovascular response and eye response are associated with mental workload variations, studying the impact of immersion degradation on psychophysiological indicators also represents an opportunity to study its impact on mental workload. While a link between mental workload and presence has been shown, to our knowledge, the relationship between mental workload and presence factors has yet to be explored. It is unclear if a degradation in immersion affects the mental workload of a virtual environment user.

To this end, this paper reports the results of an experiment where we measured the presence, mental workload, cardiovascular response, and eye response of participants driving a car in a virtual environment under various visual quality levels set over, under, or at their respective functional thresholds.

5.4 Method

5.4.1 Participants

This research project was reviewed by Polytechnique Montréal's research ethics board and met the current research ethics standards on ethical conduct for research involving humans (CER-2122-35-D). Participants recruited were required to possess a valid driver's license

(learner, probationary, or full license). They were all volunteers over 18 years of age, had no uncorrected vision problems, and did not have a known prevalence of being subject to cybersickness. Also, to ensure the validity of the cardiac measurements, participants had not consumed alcohol or caffeinated beverages and had not engaged in strenuous exercise for at least 12 hours before the experiment.

We recruited 67 participants (25 females, 42 males) for the experiment. Of these, 7 could not complete the experience due to cybersickness (6 females, 1 male). The following analyses exclude their data. Participants estimated their total experience time with virtual reality HMDs with a median experience of 2 hours (min = 0 hour, M = 32.5 hours, max = 1000 hours).

5.4.2 Apparatus

The experiment's computer used an NVIDIA GeForce RTX 2080 SUPER GPU and an Intel Core i9-9900K CPU with 64GB RAM.

The participants used the HTC Vive Pro Eye HMD to complete the driving task. This HMD has a dual OLED display with a diagonal of 3.5 inches allowing a 110 degrees FOV, a refresh rate of 90 Hz, and a resolution of 1440 x 1600 pixels per eye for maximum resolution acuity of 4.6 arcmin. It generates a luminance between 0.04 and 130 cd/m^2 and occupies the AdobeRGB colorspace [181]. We removed the HMD headphones for the experiment.

To drive the car during the experiment, participants used the Flashfire Suzuka Wheel 900A steering wheel and pedal assembly. The kit has a steering wheel that allows 900 degrees of adjustable rotation. It has twenty action buttons, including two handles behind the steering wheel. The gas and a brake pedal were positioned on the ground in front of our participants.

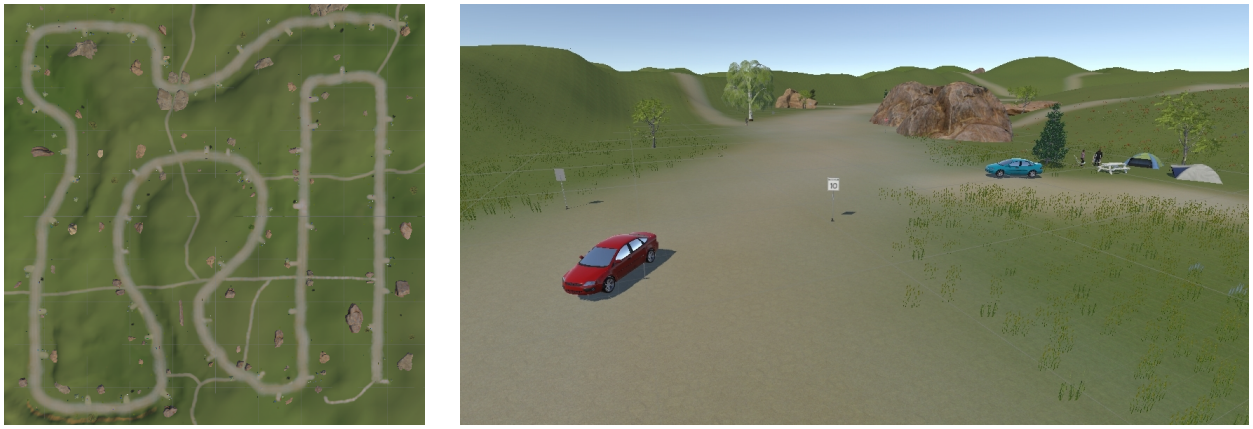
In addition, the participants used the HTC Vive Pro Eye controllers. They were strapped to their forearms to allow their hands to be tracked during the experiment while being able to touch the steering wheel without hindrance.

For the cardiovascular data gathering, we used the Polar H10 (Polar H10; www.polar.com/ca-en). It is a heart rate monitor with an adjustable strap that can be attached to a participant's chest. The device allows a heart rate measurement with a sampling frequency of 1000 Hz. It has been reported to have a high correlation and agreement level with a golden standard reference device [193]. Its Bluetooth connection enabled the transfer of heart rate data during the experiment. These data were used to evaluate the mental workload of our participants through the analysis of their cardiovascular response. We transferred the ECG raw data in a .csv file using the Elite HRV smartphone application, giving us access to the time elapsed

between two successive R-waves, called the RR intervals.

We used the HTC Vive Pro Eye embedded eye-tracker for the eye response analysis. It has a 120 Hz gaze data output frequency and an accuracy of 0.5–1.1 degrees within 20 degrees of the center of the FOV.

Virtual Environment



(a) Map (b) Screen shot of the virtual environment

Figure 5.2 Virtual environment used for the driving task.

We created the virtual environment with Unity version 2020.3.6f1. As shown in figure 5.2a, it consisted of a sizeable hilly terrain scattered with tall grass and flowers divided by a large road composed of straight lines, slight curves, and sharper turns. The 4 km long road mainly crossed footpaths and was flanked by trees, large rocks, parked cars, camping tents, chairs, and tables.

The virtual environment included the following elements to create a dynamic situation. Road signs such as stop signs, turn signs, and speed limit signs were placed on both sides of the road (see figure 5.3). The speed limits indicated a value between 10 km/h and 30 km/h. Since a high navigation speed may increase cybersickness [81], a low speed was used to mitigate the chances that our participants develop cybersickness symptoms.

During the experiment, participants could encounter a total of 27 cars on the road. A simple algorithm controlled these cars, allowing them to follow a predetermined route and avoid obstacles that may have been encountered. They were moving along the route at speeds varying between 10 and 35 km/h. They were not necessarily respecting the speed limit. Moreover, thirty-three pedestrians were walking or running along the road, sometimes crossing it, at speeds varying between 3 and 10 km/h. Other animated human characters

enriched the environment. They were not considered pedestrians as they were motionless and positioned farther from the road. They were mainly standing up or seated and performed what appeared to be discussion animations.

(a) Speed limit sign (b) Stop sign (c) Right turn sign

Figure 5.3 Road signs in the virtual environment.

5.4.3 Self-Report Measures

To evaluate the presence and mental workload of our participants during the experiment, we used the presence questionnaire (PQ) and NASA-TLX, respectively. Both are empirically validated self-report measurement methods.

Presence Questionnaire

We measured presence with a paper version of the UQO Cyberpsychology Lab revised presence questionnaire (PQ) [168], corresponding to the Witmer & Singer PQ version 3. It is a 24-item questionnaire based on Witmer et al. [169] PQ using 19 items of the PQ version 2 [169] and 5 items related to audio and haptics. The PQ [169] is one of the most widely used [20] and shortest questionnaires [13] to measure presence. It presents the advantages of not being correlated with any specific use [20] and being sensitive to increases in immersion and coherence [52, 171]. Place illusion and plausibility illusion are approximately of equal importance in regard to scores on the PQ [52]. Thus, its widespread use, reliability, and face validity ensure a certain level of comparability between studies.

NASA-TLX

We used NASA-TLX to measure our participants' mental workload subjectively. NASA-TLX is a subjective method of measuring mental workload using a multi-dimensional rating scale requiring operator judgment of several psychological and task-related variables [112]. It comprises a total of six factors. Three are task-related, namely physical demand, mental demand, and temporal demand, and three are operator-related, namely the individual's judgment of their performance, the effort required to perform the task, and the individual's level of frustration [112]. Physical, mental, and temporal demands are determined by the situation's complexity and seem to be indicators of both intrinsic and extraneous workload factors [115]. Effort is believed to be an indicator of germane load [116].

A review of 36 published NASA-TLX scores [173] allows for comparing mental workload. For the task of driving a car, the statistical distribution range found was 15.00 (min), 28.05 (25th percentile), 41.52 (50th percentile), 51.73 (75th percentile), and 68.50 (max). When considering any task, the following values were obtained: 6.21 (min.), 36.77 (25th percentile), 49.93 (50th percentile), 60.00 (75th percentile), and 88.50 (max). Thus, driving tasks impose a relatively lower mental workload than general tasks.

The questionnaire was implemented in our virtual environment. Each question appeared in a separate pane. Participants had to read the question and validate that they understood the scale, e.g., from very low to very high or from perfect to failure. Then, they would rate each factor of their mental workload using their HTC Vive controller to move the cursor on the scale to a value between 0 and 20. For our analysis, we use the weighted NASA-TLX score.

5.4.4 Procedure

Driving Task

During the experiment, we asked our participants to perform a driving task in the previously described virtual environment. It was required that they respect the Highway Code at all times. Each participant had to complete three scenarios during which they were exposed to either one of seven visual conditions. Three scenarios were used to prevent the participants from being able to guess the events that would occur during their driving task. The different scenarios corresponded to three different starting positions on the road. Each participant did all three different scenarios in random order.

The seven visual conditions include a condition where either the contrast sensitivity, the size of the FOV, or the visual acuity is set at the functional threshold or a minimum value

and a condition without degradation. The functional thresholds and minimum value were determined in a previous study [188]. The functional thresholds correspond to the results of that study, while the minimum value is the smallest functional threshold recorded in that study. The seven presented visual conditions experienced by the participants are shown in figure 5.4, 5.5, and 5.6 along with their respective values. Figure 5.4a, 5.5a, and 5.6a are all the same image. Every participant saw two degradation conditions and performed a scenario without degradation. All participants saw the condition without degradation as it was used as a baseline. Participants experienced only one of the seven studied visual conditions per scenario. For example, if a participant experienced a degradation of their FOV, they had their visual acuity and contrast sensitivity set to the maximum possible value. Each participant did not see all degradation conditions because such an experiment would have been too lengthy, which would have caused possible bias due to fatigue. The conditions were balanced according to a randomized block design.

We measured the participants' situational awareness using freeze probes (see Appendix A.) during the driving task. However, the present paper will not discuss the situational awareness-related results. Each scenario consisted of three trials ending with a freeze. The experimenter triggered the freezes at pre-selected times, occurring randomly every 3 to 6 minutes and 30 seconds while the driving task was executed, and the cardiovascular response and eye response data were recorded. During a freeze, the simulation was stopped, and the participants had to answer questions about their situational awareness. After the first and second freeze, participants resumed driving under the visual conditions displayed before the freeze. It took between 15 to 35 minutes to complete a scenario.

In summary, each participant performed three driving scenarios for which data was analyzed. All participants did the baseline scenario with no degradation and two degraded scenarios.

(a) (b) (c)

Figure 5.4 Contrast sensitivity degradation conditions

Note. a shows the environment without any degradation, at a contrast of 31.2%; b shows the contrast degraded to 25.6 %; c shows the contrast degraded to 14.7 %.

(a) (b) (c)

Figure 5.5 FOV size degradation conditions

*Note.*a shows the environment without any degradation, at a FOV of 110 degrees; b shows the FOV to 96.6 degrees; c shows the FOV degraded to 50.6 degrees.

(a) (b) (c)

Figure 5.6 Visual acuity degradation conditions

Note. a shows the environment without any degradation, at a visual acuity of 4.6 arcmin; b shows the visual acuity degraded to 12.2 arcmin; and, c shows the visual acuity degraded to 22.8 arcmin.

Experimental Procedure

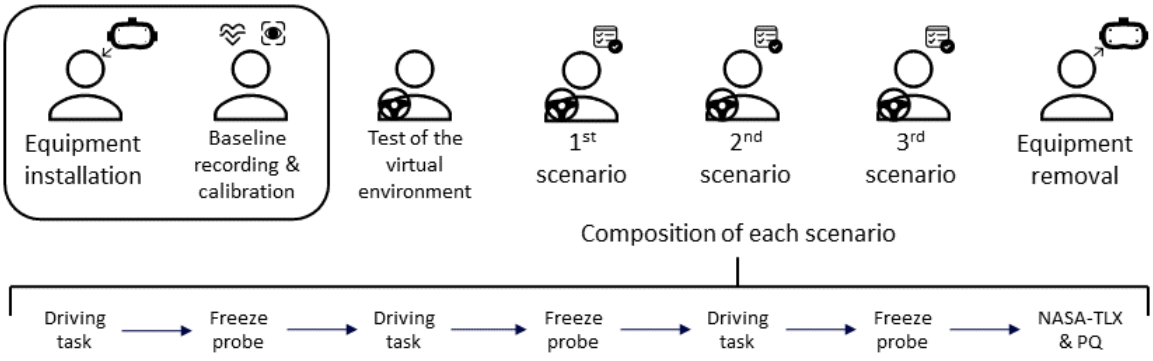


Figure 5.7 Overview of the experimental procedure

Figure 5.7 presents an overview of the experimental procedure. Upon arrival at the experi-

ment site, we provided the participants with a Polar H10 heart rate monitor and explained how to wear it appropriately. After correctly positioning the device, we asked them to sit comfortably in the driver's seat and to minimize body movements to prevent adding noise to the heart rate recording. We then proceeded to a baseline recording of their resting heart rate, reminding them that they had to stay calm, breathe normally and be as motionless as possible. During the recording, we explained the study procedures and gave the instructions to follow in case of cybersickness, i.e., to remove the HMD and wait until the sensation subsides. We made sure that they were able to identify cybersickness symptoms. We also asked them to stop the experiment if they were feeling cybersickness symptoms, as they were a source of confound in our experiment. Finally, we told them that they could stop the experiment if they felt uncomfortable for any other reason.

At the end of the 5 minutes necessary to record their resting heart rate baseline, we asked the participants to put on the HMD and follow the HTC Vive Pro Eye eye-tracking camera calibration procedure. The participants then performed a practice drive, allowing them to get accustomed to the car and virtual environment parameters and evaluate if they were prone to suffer from cybersickness. At this point, all visual qualities were set at their maximum value. At the end of the practice, we did a freeze probe, specifying that the questions seen in practice would not necessarily be the ones being asked during the experiment. After answering all probes, they were prompted to complete the NASA-TLX questionnaire in the virtual environment. Afterward, they removed the HMD and weighed each NASA-TLX factor for the driving task, and filled out the Presence Questionnaire (PQ) [169].

The three driving scenarios were then executed. At the end of each scenario, the participants were questioned to make sure that they did not feel cybersickness symptoms. They were then asked to complete the NASA-TLX questionnaire, take off the HMD and fill out the PQ.

5.4.5 Data Analysis

Each participant experienced two of the six possible degradation condition scenarios and a scenario without degradation, completing the PQ and the NASA-TLX at the end of each scenario. Thus, the questionnaires were completed twenty times ($n = 20$) for each degradation condition and sixty times ($n = 60$) for the condition without degradation. In the following section, we provide descriptive statistics of the questionnaires as they are not the main focus of the experiment but brings an interesting contribution to the discussion. The presented descriptive statistics are the mean, standard deviation, min, max, and quartiles.

For the cardiovascular and eye response analysis, since each scenario comprised three trials, we gathered 60 measurements ($n = 60$) for each degradation condition and 180 measurements

($n = 180$) for the condition without degradation. Since our data were unbalanced and our participants had been subjected to repeated measurements, we used linear mixed models for our analysis. Our linear mixed models were fitted with the lme4 R package. We used the restricted maximum likelihood (REML) as an estimation method since it can give an unbiased estimate of the variance parameters given the number of participants ($n = 60$) that took part in the experiment. We assumed an unstructured covariance matrix of random factors. Satterthwaite's method was used to compute the degrees of freedom for the t-tests and p-values. Every linear mixed model analysis significance was evaluated with an $\alpha = 0.05$. Each model Intra-Class Correlation (ICC) was evaluated with the R sjstats package, using the adjusted value. Marginal (R^2m) and conditional (R^2c) R squared values were evaluated with the R MuMIn package.

Cardiovascular Response Analysis The collected RR intervals were processed using Kubios HRV Standard 3.5.0 analysis package. The default settings were used to compute the basic heart rate values and the HRV statistics. Artifacts were corrected by interpolation with the Kubios HRV Standard beat correction threshold algorithm. We only analyzed the periods with less than 1% of their beats corrected. For each trial, we analyzed a fixed period of 2 minutes and 30 seconds, starting 30 seconds after the beginning of a trial.

To analyze our participants' basic HR and HRV, we investigated the effect of the degradation of contrast sensitivity, size of the field of view, and visual acuity on mean heart rate, HRV time-domain, and HRV frequency-domain metrics. The time-domain metrics were SDNN (ms), RMSSD (ms), pNN50 (%), and the frequency-domain metrics were LF (ms^2), HF (ms^2), LF/HF. The frequency-domain metrics were obtained based on a fast Fourier transform spectrum. Our cardiovascular metrics were log-transformed when necessary to obtain normally distributed data and residuals. Each cardiovascular response model was built with the degradation level as a fixed factor and a random intercept relating to each participant. As it is essential to consider subject-related factors [174], we introduced age and gender as fixed effects in every model and HR in every HRV model.

Eye Response Analysis The LHIPA was computed with an implementation of Duchowski et al. [136] algorithm. Eye-related data was obtained using the VIVE SRanipal SDK, which provided the position and pupil diameter of each eye during the experiment. To compute the LHIPA, the acquired data was preprocessed by removing data collected 200 ms prior to the start of a blink and 200 ms following the end of a blink. To analyze the eye response of our participants, we investigated the effect of the degradation of contrast sensitivity, size of the field of view, and visual acuity on the LHIPA metric.

5.5 Results

5.5.1 Presence Questionnaire

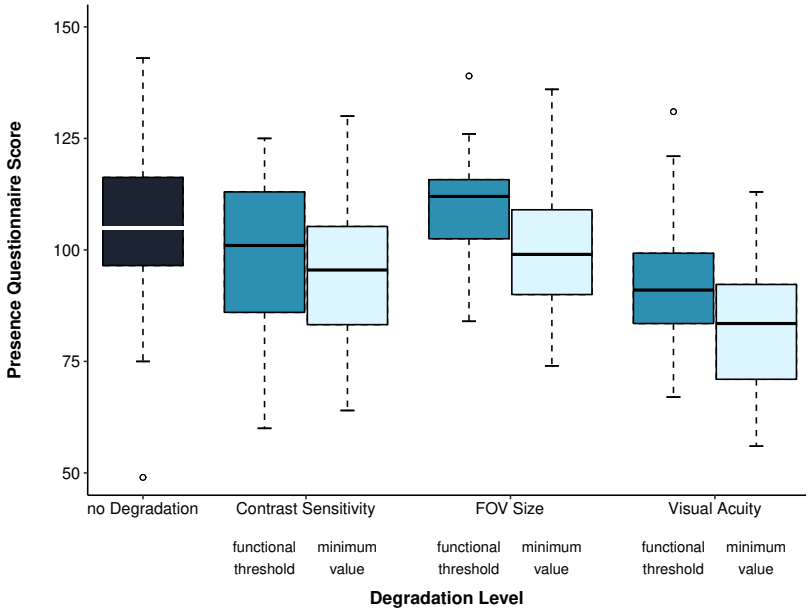


Figure 5.8 Presence questionnaire score for each degradation condition

Note. From left to right, the boxplots represent the PQ score for the following conditions: no degradation, contrast sensitivity at the functional threshold and the minimum value, size of the FOV at the functional threshold and the minimum value, and visual acuity at the functional threshold and the minimum value.

Figure 5.8 shows the results of the PQ score for each experimental condition. Table 5.1 shows the min., 1st quantile, median, 3rd quantile, max., mean, and standard deviation values of the PQ score for each degradation condition.

The results show that the PQ score median is higher for the condition without degradation, except when compared to the size of the FOV at the functional threshold. Otherwise, when comparing each degradation with each other, the PQ score median is higher for the degradation of the FOV than the other degradation conditions. Furthermore, we can see that the visual acuity conditions have the lowest presence values among the studied degradation conditions.

Table 5.1 Descriptive statistics of Presence Questionnaire score for each degradation condition

	Min	1st quantile	Median M (SD)	3rd quantile	Max
No degradation	49.00	96.50	105.00 105.2 (16.45)	116.25	143.00
Contrast sensitivity functional threshold	60.00	86.00	101.00 99.15 (17.14)	113.00	125.00
Contrast sensitivity min.	64.00	83.25	95.50 96.45 (17.72)	105.25	130.00
FOV functional threshold	84.00	102.50	112.00 109.20 (12.98)	115.75	139.00
FOV min.	74.00	90.00	99.00 100.15 (14.25)	109.00	136.00
Visual acuity functional threshold	67.00	83.50	91.00 92.40 (16.00)	99.25	131.00
Visual acuity min.	56.00	71.00	83.50 82.15 (14.74)	92.25	113.00

Note. The upper part of each line shows the min., 1st quantile, median, 3rd quantile, and max. value of the PQ score for each degradation condition, the lower part shows the mean and the standard deviation.

5.5.2 NASA-TLX

Figure 5.9 shows the results of the weighted NASA-TLX score for each degradation condition. Table 5.2 shows the min., first quartile, median, third quartile, max., mean, and standard deviation values of the NASA-TLX score for each degradation condition.

The NASA-TLX score median is the lowest for the condition without degradation. Generally, the NASA-TLX score median is similar for every degradation condition except for the visual acuity degradation condition. The functional threshold condition for visual acuity has a lower median NASA-TLX score than the minimal value condition.

5.5.3 Cardiovascular Response

Basic Heart Rate

As shown in table 5.3, variables that significantly impacted heart rate were the size of the field of view at the functional threshold and the age of our participants. For the condition with the size of the FOV at the functional threshold, we recorded a significantly lower mean HR ($M = 71.94$, $SD = 12.41$) compared to the condition without degradation ($M = 77.96$, $SD = 11.61$), $t(476.25) = -2.92$, $p = 0.004$. The estimate for this variable had a value of

Table 5.2 Descriptive Statistics of NASA-TLX Score for Each Degradation Condition

	Min	**1st quartile**	**Median** **M (SD)**	**3rd quartile**	**Max**
No degradation	2.11	24.48	35.32 36.84 (17.25)	49.00	83.55
Contrast sensitivity functional threshold	3.72	26.66	38.26 39.18 (19.27)	52.29	78.33
Contrast sensitivity min.	21.83	28.72	35.40 40.71 (15.74)	50.79	72.34
FOV functional threshold	13.13	31.08	39.15 42.26 (17.72)	52.16	91.42
FOV min.	5.38	20.52	36.81 36.04 (18.93)	51.17	66.69
Visual acuity functional threshold	18.48	34.55	46.13 45.38 (15.29)	56.33	84.01
Visual acuity min.	13.94	40.29	56.86 54.74 (20.26)	66.94	85.74

Note. The upper part of each line shows the min., first quartile, median, third quartile, and max. value of the NASA-TLX score for each degradation condition, the lower part shows the mean and the standard deviation.

Table 5.3 Results of the Linear Mixed Model for Mean HR

	Mean HR			
	Estimate	**df**	**t-value**	**sig.**
(Intercept)	**4.5008** **4.3558** **4.6458**	57.1434	60.2520	**<0.0001**
Contrast sensitivity functional threshold	0.0074 -0.0056 0.0203	476.2658	1.1110	0.2672
Contrast sensitivity min.	-0.0118 -0.0248 0.0011	476.3829	-1.7870	0.0745
FOV functional threshold	**-0.0198** **-0.0330** **-0.0066**	476.2504	-2.9150	**0.0037**
FOV min.	0.0014 -0.0115 0.0144	476.2566	0.2160	0.8290
Visual acuity functional threshold	-0.0070 -0.0196 0.0057	476.2470	-1.0780	0.2814
Visual acuity min.	0.0020 -0.0110 0.0149	476.3544	0.3000	0.7646
Gender	0.0125 -0.0609 0.0859	57.0273	0.3310	0.7419
Age	**-0.0062** **-0.0108** **-0.0016**	57.0456	-2.6020	**0.0118**
	ICC	**R2m**	**R2c**	
	0.9166	0.1016	0.9250	

Note. The upper part of each line of the table shows the estimate, degrees of freedom, t statistic, and significance of each variable. The lower part shows the 95% confidence interval of the estimate, with the lower limit presented on the left and the upper limit presented on the right. The last line shows the Intra-Class Correlation (ICC), marginal R squared (R^2m), and conditional R squared (R^2c) for the model.

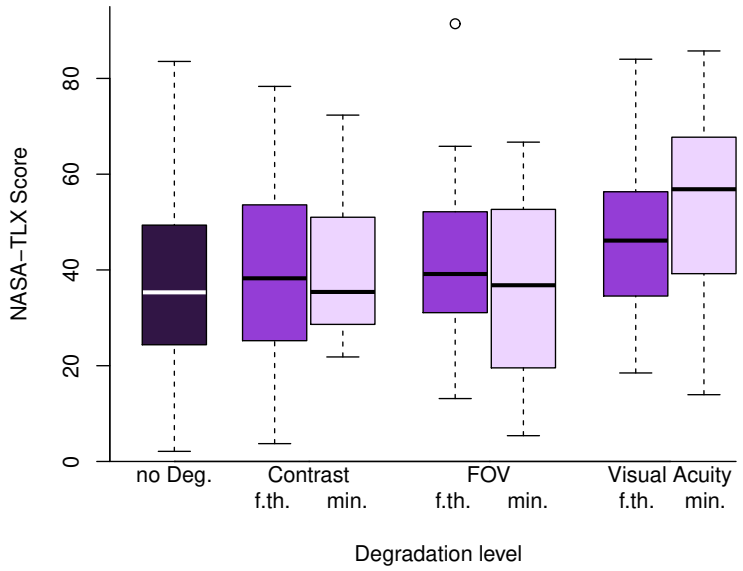


Figure 5.9 NASA-TLX Score for Each Degradation Condition

Note. From left to right, the boxplots represent the NASA-TLX score for the following conditions: no degradation, contrast sensitivity at the functional threshold and the minimum value, size of the FOV at the functional threshold and the minimum value, and visual acuity at the functional threshold and the minimum value.

-0.0198 with a 95% CI [-0.0330, -0.0066]. We also found that the mean HR was significantly lower for older participants.

HRV Metrics

Table 5.4 shows the results of the linear mixed models for RMSSD, SDNN, and pNN50. For these three models, we can see a significant negative impact of age and mean HR on the three dependent variables.

For the RMSSD model, apart from age and mean HR, the variables that significantly impacted RMSSD are the size of the FOV at the functional threshold and the minimum value. For the condition with the size of the FOV at the functional threshold ($M = 38.00$, $SD = 26.89$), we recorded a significantly higher RMSSD compared to the condition without degradation ($M = 30.90$, $SD = 17.27$), $t(480.27) = 2.42$, $p = 0.02$. At the minimum value ($M = 36.58$, $SD = 22.77$), we also recorded a significantly higher RMSSD than the condition without degradation, $t(478.46) = 2.26$, $p = 0.02$. In addition, the estimate for the FOV at the functional threshold had a value of 0.0616 with a 95% CI [0.0119, 0.1112], and the estimate for the FOV at the minimum value had an estimate of 0.0557 with a 95% CI [0.0077, 0.1041].

Table 5.4 Results of the Linear Mixed Models for RMSSD, SDNN, and pNN50

	RMSSD				SDNN				pNN50			
	Estimate	**df**	**t-value**	**sig.**	**Estimate**	**df**	**t-value**	**sig.**	**Estimate**	**df**	**t-value**	**sig.**
(Intercept)	**7.1553**	141.0709	28.1200	**<0.0001**	**6.3037**	141.6749	25.0990	**<0.0001**	**10.5612**	138.6079	14.1020	**<0.0001**
	6.6631 7.6472				**5.8177 6.7888**				**9.1139 12.0137**			
Contrast sensitivity	0.0149	478.6043	0.6030	0.5467	0.0131	483.8708	0.4330	0.6653	-0.0949	433.9610	-1.0740	0.2833
functional threshold	-0.0333 0.0630				-0.0460 0.0721				-0.2672 0.0772			
Contrast sensitivity	-0.0396	480.1875	-1.5980	0.1108	**-0.0688**	486.0708	-2.2710	**0.0236**	**-0.1913**	435.5707	-2.0990	**0.0364**
min.	-0.0878 0.0088				**-0.1277 -0.0094**				**-0.3684 -0.0132**			
FOV	**0.0616**	480.2774	2.4190	**0.0159**	0.0224	485.9566	0.7180	0.4732	0.1119	436.7386	1.2230	0.2218
functional threshold	**0.0119 0.1112**				-0.0385 0.0830				-0.0667 0.2897			
FOV	**0.0557**	478.4639	2.2550	**0.0246**	**0.0640**	483.6649	2.1160	**0.0348**	**0.1874**	433.7299	2.0790	**0.0382**
min.	**0.0077 0.1041**				**0.0053 0.1233**				**0.0128 0.3646**			
Visual acuity	0.0345	478.9524	1.4260	0.1545	0.0080	484.2965	0.2690	0.7881	0.1687	434.8552	1.9170	0.0559
functional threshold	-0.0128 0.0816				-0.0501 0.0655				-0.0036 0.3395			
Visual acuity	0.0273	478.7507	1.1070	0.2689	0.0412	484.1982	1.3620	0.1737	-0.0272	436.2862	-0.3070	0.7592
min.	-0.0209 0.0754				-0.0178 0.1001				-0.2002 0.1452			
Gender	-0.1587	56.8992	-1.7290	0.0893	-0.0841	56.7301	-1.0650	0.2914	-0.3326	55.0750	-1.4650	0.1487
	-0.3364 0.0191				-0.2366 0.0686				-0.7713 0.1062			
Age	**-0.0214**	59.0977	-3.6690	**0.0005**	**-0.0187**	59.6280	-3.7020	**0.0005**	**-0.0542**	58.8473	-3.7110	**0.0005**
	-0.0327 -0.0101				**-0.0285 -0.0090**				**-0.0824 -0.0260**			
Mean heart rate	**-0.0406**	456.9145	-20.3480	**<0.0001**	**-0.0269**	297.2933	-12.2160	**<0.0001**	**-0.0920**	264.8345	-13.6300	**<0.0001**
	-0.0445 -0.0367				**-0.0312 -0.0226**				**-0.1052 -0.0790**			
	ICC	***R²m***	***R²c***		***ICC***	***R²m***	***R²c***		***ICC***	***R²m***	***R²c***	
	0.8208	0.6071	0.9296		0.6862	0.4438	0.8255		0.6953	0.5062	0.8495	

Note. The table shows the result for the linear mixed model of the RMSSD, SDNN, and pNN50, respectively. The upper part of each line of the table shows the estimate, degrees of freedom, *tstatistic*, and significance of each variable. The lower part shows the 95% confidence interval of the estimate, with the lower limit presented on the left and the upper limit presented on the right. For each model, the last line shows the Intra-Class Correlation (ICC), marginal R squared (R^2m), and conditional R squared (R^2c) for the model.

For the SDNN model, other than age and mean HR, the variables that significantly impacted SDNN were the contrast sensitivity and the FOV at the minimum value. For the condition with contrast sensitivity at the minimum value ($M = 42.83$, $SD = 21.95$), we recorded an SDNN significantly different from the condition without degradation ($M = 41.94$, $SD = 18.03$), $t(486.07) = -2.27$, $p = 0.02$. At the FOV at the minimum value ($M = 45.33$, $SD = 18.40$), the SDNN was also significantly different $t(483.66) = 2.12$, $p = 0.03$. The estimate for the contrast sensitivity at the minimum value is -0.0688 with a 95% CI [-0.1277, -0.0094], and the estimate for the FOV at the functional threshold is 0.0640 with a 95% CI [0.0053, 0.1233].

Furthermore, the variables that significantly impacted pNN50, other than age and mean HR, were the contrast sensitivity and the FOV at the minimum value. For the condition with contrast sensitivity at the minimum value ($M = 12.91$, $SD = 17.21$), we recorded a pNN50 significantly different from the condition without degradation ($M = 11.37$, $SD = 13.41$), $t(435.57) = -2.10$, $p = 0.04$. For the FOV at the minimum value ($M = 15.52$, $SD = 18.31$), we recorded a pNN50 also significantly different from the condition without degradation, $t(433.73) = 2.08$, $p = 0.04$. The estimate for the contrast sensitivity at the minimum value

is -0.1913 with a 95% CI [-0.3684, -0.0132], and the estimate for the FOV at the functional threshold is 0.1874 with a 95% CI [0.0128, 0.3646].

Table 5.5 Results of the Linear Mixed Model for LF, HF, and LF/HF

	LF				HF				LF/HF			
	Estimate	**df**	**t-value**	**sig.**	**Estimate**	**df**	**t-value**	**sig.**	**Estimate**	**df**	**t-value**	**sig.**
(Intercept)	**11.5729**	120.1059	18.2740	**<0.0001**	**13.2196**	142.9317	18.8670	**<0.0001**	-1.245581	132.8444	-1.9500	0.0533
	10.3485 12.7944				**11.8612 14.5739**				-2.4789 -0.0129			
Contrast sensitivity functional threshold	-0.0764	492.1198	-0.8230	0.4107	0.0478	482.7294	0.5820	0.5609	-0.115543	487.6107	-1.3650	0.1729
	-0.2576 0.1043				-0.1125 0.2076				-0.2805 0.0496			
Contrast sensitivity min.	**-0.2121**	494.6341	-2.2830	**0.0229**	-0.1016	484.8681	-1.2360	0.2171	-0.118851	489.9969	-1.4020	0.1615
	-0.3922 -0.0297				-0.2615 0.0592				-0.2839 0.0466			
FOV functional threshold	-0.0429	494.0567	-0.4490	0.6533	**0.2312**	484.7978	2.7370	**0.0064**	**-0.239459**	489.6905	-2.7490	**0.0062**
	-0.2297 0.1427				**0.0663 0.3957**				**-0.4092 -0.0695**			
FOV min.	0.0878	491.8600	0.9460	0.3445	0.0579	482.5311	0.7060	0.4807	0.020212	487.3779	0.2390	0.8113
	-0.0924 0.2696				-0.1011 0.2192				-0.1462 0.1842			
Visual acuity functional threshold	-0.0972	492.4890	-1.0700	0.2852	-0.0062	483.1533	-0.0770	0.9387	-0.077994	488.0258	-0.9410	0.3471
	-0.2753 0.0791				-0.1636 0.1498				-0.2391 0.0841			
Visual acuity min.	0.0322	492.7276	0.3470	0.7285	0.1099	483.0264	1.3400	0.1809	-0.074602	488.0609	-0.8820	0.3782
	-0.1485 0.2131				-0.0502 0.2696				-0.2392 0.0907			
Gender	-0.0691	56.6643	-0.3880	0.6994	-0.3066	56.4678	-1.3680	0.1768	0.252862	56.8999	1.3340	0.1875
	-0.4129 0.2747				-0.7400 0.1270				-0.1132 0.6189			
Age	**-0.0394**	59.7014	-3.4200	**0.0011**	**-0.0498**	59.2863	-3.4720	**0.0010**	0.007928	59.9482	0.6500	0.5182
	-0.0616 -0.0172				**-0.0776 -0.0221**				-0.0156 0.0315			
Mean heart rate	**-0.0456**	181.5217	-7.7450	**<0.0001**	**-0.0792**	318.1893	-13.0420	**<0.0001**	**0.029208**	233.3170	5.0480	**<0.0001**
	-0.0569 -0.0342				**-0.0909 -0.0673**				**0.0180 0.0404**			
	ICC	***R²m***	***R²c***		***ICC***	***R²m***	***R²c***		***ICC***	***R²m***	***R²c***	
	0.5281	0.2737	0.6573		0.7070	0.4703	0.8448		0.6114	0.1546	0.6715	

Note. The table shows the result for the linear mixed model of the LF, HF, and LF/HF, respectively. The upper part of each line of the table shows the estimate, degrees of freedom, *tstatistic*, and significance of each variable. The lower part shows the 95% confidence interval of the estimate, with the lower limit presented on the left and the upper limit presented on the right. For each model, the last line shows the Intra-Class Correlation (ICC), marginal R squared (R^2m), and conditional R squared (R^2c) for the model.

Table 5.5 shows the results of the linear mixed models for LF, HF, and LF/HF. For these three models, we can see a significant negative impact of age and mean HR on LF and HF, and a significant positive impact on LF/HF.

For the LF model, contrast sensitivity at the minimum value was the only model variable with a significant impact on LF apart from age and mean HR. For contrast sensitivity at the minimum value ($M = 1430.86$, $SD = 1531.68$), we recorded a significantly lower LF compared to the condition without degradation ($M = 1541.29$, $SD = 1530.10$), $t(494.63) = -2.28$, $p = 0.02$. The estimate for contrast sensitivity at the minimum value had an estimate of -0.2121 with a 95% CI [-0.3922, -0.0297].

For the HF model, other than age and mean HR, the FOV at the functional threshold was the only model variable with a significant impact on HF. For FOV at the functional threshold ($M = 927.03$, $SD = 1852.05$), we recorded an HF significantly higher from the condition without degradation ($M = 454.00$, $SD = 651.89$), $t(484.80) = 2.74$, $p = 0.006$. The estimate

for the FOV at the functional threshold is 0.2312 with a 95% CI [0.0663, 0.3957].

For the LF/HF model, other than mean HR, the FOV at the functional threshold was the only variable with a significant impact on LF/HF. For FOV at the functional threshold ($M = 3.71$, $SD = 2.72$), we recorded an LF/HF significantly lower from the condition without degradation ($M = 6.27$, $SD = 5.83$), $t(489.69) = -2.75$, $p = 0.006$. The estimate for the FOV at the functional threshold is -0.2395 with a 95% CI [-0.4092, -0.0695].

5.5.4 Eye Response

Table 5.6 Results of the Linear Mixed Model for LHIPA

	LHIPA			
	Estimate	**df**	**t-value**	**sig.**
(Intercept)	**1.6370**	89.2800	50.9550	**<0.0001**
	1.5736 **1.6996**			
Contrast sensitivity functional threshold	0.0228	473.6000	0.5550	0.5794
	-0.0573 0.1030			
Contrast sensitivity min.	0.0227	474.1000	0.5400	0.5892
	-0.0595 0.1046			
FOV functional threshold	0.0204	471.9000	0.5090	0.6113
	-0.0582 0.0986			
FOV min.	0.0663	475.1000	1.5380	0.1248
	-0.0178 0.1506			
Visual acuity functional threshold	0.0007	473.5000	0.0180	0.9859
	-0.0794 0.0813			
Visual acuity min.	**0.1008**	472.9000	2.4590	**0.0143**
	0.0209 **0.1810**			
	ICC	***R²m***	***R²c***	
	0.3943	0.0109	0.4009	

Note. The table shows the result for the linear mixed model of the LHIPA. The upper part of each line of the table shows the estimate, degrees of freedom, *tstatistic*, and significance of each variable. The lower part shows the 95% confidence interval of the estimate, with the lower limit presented on the left and the upper limit presented on the right. For each model, the last line shows the Intra-Class Correlation (ICC), marginal R squared (R^2m), and conditional R squared (R^2c) for the model.

Table 5.6 shows the linear mixed model results for the LHIPA. Visual acuity at the minimum value was the only model variable with a significant impact on this model. For visual acuity at the minimum value ($M = 1.76$, $SD = 0.32$), we recorded a significantly higher LHIPA compared to the condition without degradation ($M = 1.63$, $SD = 0.31$), $t(472.90) = 2.46$, $p = 0.01$. In addition, the estimate for the visual acuity at the minimum value had an estimate of 0.1008 with a 95% CI [0.0209, 0.1810].

5.6 Discussion

The objective of this paper was to study the effect of a manipulation of immersion on presence, mental workload, and psychophysiological indicators. For this purpose, we degraded three visual qualities while our participants executed a driving task in a virtual environment. During the driving task, we recorded their cardiac and eye response and asked them to fill out the PQ and NASA-TLX questionnaires.

5.6.1 Effect of an Objective Degradation of Immersion on the PQ

In this section, we first assessed how each degradation affected presence by evaluating the PQ score. It is important to note that due to a limited amount of observation, $n = 20$ per degradation condition, we limited our statistical analyzes of the PQ to descriptive statistics.

The PQ score resulting from the degradation of contrast sensitivity at the functional threshold was close to the score obtained without degradation. For the contrast sensitivity at the minimum value, the impact of the degradation appears to be consequential as the PQ score is lower than both previously discussed values. While more work is needed to determine if these effects are significant, the PQ seems sensitive to immersion variation of contrast sensitivity under the functional threshold.

For the degradation of FOV, our evaluation indicates a higher PQ score at the functional threshold than without degradation. Also, the PQ score obtained for the size of the FOV at the minimum value is close to the score of the condition without degradation. Results from previous studies indicated that a higher presence is evoked by values closer to the total FOV size [9, 76, 77]. Thus, we would have expected to measure a decrease in the PQ score for the conditions where the FOV size is smaller than the condition without degradation. We suspect this result was linked to an unconsciously lower cybersickness due to the reduced FOV and will discuss the matter in greater detail in the following sections.

The scores obtained for the visual acuity at the functional threshold and the minimum value were lower than the condition without degradation. We expected these results since we believe that the visual acuity at the functional threshold for the present experiment differs from the experiment in which it was measured [188]. In this previous experiment, we found a functional threshold for a task that engaged no particular visual acuity since the objects involved in the interactions were large. Thus, it was postulated that functional thresholds are context-dependent and that the task executed by the participants in a virtual environment should modulate the functional threshold. In the present experiment, since the task goal required reading characters at a distance, the functional threshold should have been set at

a level allowing the accomplishment of this task without difficulty. However, the technology at our disposal did not provide a sufficient resolution. Still, the PQ score appeared sensitive to immersion variation of visual acuity under the functional threshold.

5.6.2 Effect of an Objective Degradation of Immersion on NASA-TLX

An assessment of mental workload was done subjectively using the NASA-TLX. However, our limited number of observations restricted our statistical analysis of the questionnaire to descriptive statistics.

The NASA-TLX score of the condition without degradation had a median of 35.32, which is lower than the first quartile of a global NASA-TLX analysis [173]. Hence, the task without degradation can be considered as imposing a low workload on our participants. Since psychophysiological measures are sensitive to arousing virtual environment [41], using a virtual environment with a low mental workload allowed us to mitigate the confounds in our investigation on the effect of degradation of immersion on psychophysiological measures. Thus, our methodology of objectively varying immersion helped us gain insights into presence and its related concept.

For the contrast sensitivity degradation at the functional threshold and the minimum value, the NASA-TLX score medians were similar to the median of the condition without degradation. Likewise, the NASA-TLX scores for the size of the FOV at the functional threshold and the minimum value appear equivalent to the condition without degradation.

However, for the visual acuity at the functional threshold, the NASA-TLX score median was higher than for the condition without degradation. Furthermore, it was even higher for visual acuity at the minimum value than at the functional threshold and, consequently, the condition without degradation. The NASA-TLX score indicates a higher mental workload for visual acuity in the two degraded conditions than for the condition without degradation.

As previously explained, the results regarding the visual acuity degradation NASA-TLX score were expected, mainly because we believe that the visual acuity at the functional threshold for the present experiment differs from the experiment in which it was measured [188]. To further investigate the effect of visual acuity at the functional threshold, it could be interesting to compare its impact on the NASA-TLX score with higher visual acuity measures using an HMD with a higher resolution.

5.6.3 Effect of an Objective Degradation of Immersion on Psychophysiological Indicators

There were no significant differences between the contrast sensitivity degradation at the functional threshold and the condition without degradation for cardiac and eye response metrics. Thus, our experiment does not allow us to suspect an effect of degradation of the contrast sensitivity at this value on the studied psychophysiological indicators. However, as shown in table 5.4 and 5.5, there was a significant but relatively low effect on the degradation of contrast sensitivity at the minimum value on SDNN, pNN50, and LF. The three HRV metrics indicate a lower score in that degradation condition. In this condition, we could assume a higher mental workload, stress, or vigilance due to the low contrast rendering many elements hard to distinguish during the driving task. Participants had to be more focused to be able to reach their goals in the virtual environment.

The linear mixed models of mean HR, RMSSD, HF, and LF/HF revealed a significant difference between the measures of FOV at the functional threshold and the condition without degradation. Mean HR and LF/HF were lower for the condition where the size of the FOV was set at the functional threshold than for the condition without degradation. Additionally, RMSSD and HF were higher for the condition with the FOV at the functional threshold than for the condition without degradation. The same effect was observed for the condition with the size of the FOV at the minimum value. RMSSD, SDNN, and pNN50 were significantly different in the condition with the FOV minimum value than in the condition without degradation. The minimum value of FOV degradation had a relatively low effect on the significant HRV metrics. These results were unexpected as we anticipated that a smaller FOV, implying a lower immersion, would affect the cardiovascular response indicating a higher mental workload, stress, or vigilance. We assumed that a reduced FOV would force our participants to make more head movements, spending more mental resources gathering the information needed to accomplish their tasks.

A first explanation would link these effects on psychophysiological indicators to an unconscious decrease in cybersickness. It is known that a virtual environment like ours, with camera displacement while the body stays static, can cause this kind of discomfort [80] due to visual-vestibular cue conflict [39]. This fact is evidenced by the number of participants that had to withdraw from our experience because of nausea or some informal comments at the end of the experiment regarding slight dizziness. Hence, we suppose that our participants were unconsciously experiencing cybersickness, even with the measure put in place to mitigate this phenomenon. The psychophysiological indicators were significantly affected in the conditions where the FOV was degraded because a smaller FOV imposed a lower

cybersickness. Alternatively, we could explain these results because there are less data to acquire and process with a smaller FOV causing a decrease in mental workload. We suppose so because, when compared to the Grier [173] NASA-TLX analysis, our driving task imposed a low workload. The condition without degradation has a median score of 35.32, which is lower than the 50th percentile of a driving task. Thus, the information required to carry out the task might have been readily available even with a reduced FOV.

Visual acuity at the functional threshold does not affect any variable significantly. Thus, our experiment does not allow us to suspect an effect of the degradation of the visual acuity at the functional threshold on psychophysiological indicators. The condition with the visual acuity at the minimum value was the LHIPA linear model's only significant variable, showing a relatively low effect of visual acuity at the minimum value on the measure. A low visual acuity makes many elements, such as speed limits, indistinguishable. As for the condition of contrast sensitivity at the minimum value, participants had to be more focused and vigilant in this condition to reach their goals in the virtual environment. Hence, this effect could imply a higher mental workload at the minimum value for this condition, as supported by the NASA-TLX score.

Not all psychophysiological indicators appear sensitive to the same visual quality degradation. Our data imply that SDNN, pNN50, and LF are sensitive to contrast sensitivity under the functional threshold. We cannot confirm if the degradation of FOV affects any psychophysiological measure since the effect found in mean HR, RMSSD, SDNN, pNN50, HF, and LF/HF seems to be confounded. However, they may help assess the effect of FOV reduction. Furthermore, our results indicate that LHIPA is sensitive to a virtual environment with a visual acuity under the functional threshold.

Analysis of linear mixed models

For basic heart rate, the residual variability was explained by the participants' random intercept with an ICC of 0.92, indicating excellent reliability between our raters. The ICC of HRV metrics, ranging from 0.53 to 0.82, indicated moderate to good reliability between our participants. Hence, for these models, the variation was mainly due to differences between participants. On the other hand, the LHIPA model had an ICC of 0.39, indicating poor reliability. For this model, the variation between participants was less important than the variation in each participant measurement. Thus, a greater number of measurements for each participant would have helped reduce the uncertainty of the model.

Otherwise, it appears that the inclusion of random effects of participants was relevant. This is shown by the differences between the R^2m, which estimates the fraction of the variance

explained by the fixed effects, and R^2c, which estimates the fraction explained by the fixed and random effects. Since R^2c is constantly greater than R^2m, including random effects improved our models' accuracy.

In summary, the linear mixed models used in this paper enabled us to deepen our comprehension of the influence of visual quality degradation on immersion. They helped understand the context in which the psychophysiological measures can be applied.

5.6.4 Theoretical Implications

Our results regarding the degradation of contrast sensitivity and visual acuity indicated that these immersion conditions significantly affected our measured psychophysiological indicators only when they were at a value under the functional threshold. These visual conditions seemed to decrease presence, although we could not evaluate their significance due to a lack of statistical power.

Still, we do not think this presence decrease is only due to the immersion factor's degradation. Instead, we argue that these visual conditions affected the coherence of our environment, the other factor of presence. Coherence is defined as the ability of a virtual environment to act as a user expects it to act [15]. We introduced unreasonable circumstances while manipulating immersion by degrading contrast sensitivity and visual acuity under the functional threshold. These unreasonable circumstances changed how our participants would typically respond and thus their expectations toward the virtual environment. For instance, when the visual acuity in the virtual environment was at the minimum value, the resolution was so low that our participants had to change their strategy to acquire relevant information. For example, they had to delay the moment they perceived the information or change their positioning to perceive it appropriately. Thus, we suspect that the unreasonable circumstances introduced by the degradation of visual qualities under the functional threshold were responsible for the psychophysiological indicators' significant results. The introduction of unreasonable circumstances implies that the resulting coherence degradation was responsible for these significant results. This assertion is in agreement with Skarbez [65] results. Skarbez indicated that inconsistent behavior of a virtual environment affected their participants' heart rates. While the basic heart rate was not significantly different for our contrast sensitivity and visual acuity conditions, we observed a significant effect on the other studied psychophysiological indicators for which variations are associated with similar phenomena, such as HRV metrics and LHIPA. Thus, we consider that an immersion's degradation does not affect psychophysiological indicators unless they introduce unreasonable circumstances. While more work is needed to confirm such affirmation, we suggest that psychophysiological

indicators could constitute an objective measure of coherence. Although we suspect a similar effect of the size of the FOV on our psychophysiological metrics, we could not show it with our experiment as other factors affected our measures.

Consequently, since it measures a factor responsible for variations in psychophysiological indicators, we suggest that mental workload assessed by the NASA-TLX is also related to the coherence of a virtual environment. In our experiment, the visual acuity's degradation conditions seemed to impose a higher mental workload as subjectively evaluated with the NASA-TLX. However, we would need more statistical power to evaluate the significance of the difference between the NASA-TLX score of the condition without degradation and its score for the visual acuity's degradation conditions. Still, our visual acuity's degradation results allow us to argue that subjectively assessed mental workload is also affected by unreasonable circumstances. Moreover, we suspect we did not observe a similar effect in the contrast sensitivity condition under the functional threshold because participants felt they could still perform adequately. Some participants in the contrast sensitivity at the minimum value condition mentioned that they felt they were driving in fog. Their ability to create reasonable circumstances explaining the degradation helped them adapt their behavior without subjectively impacting their mental workload. Thus, we suggest that subjectively assessed mental workload also evaluates the coherence of a virtual environment.

Moreover, the points discussed in this section have implications that affect the way presence is modeled. As shown by the arrow linking "Immersion" and "Coherence" in figure 5.10, we propose to add a relationship between immersion and coherence to have a better representation of the way that immersion affects presence as previously described.

5.6.5 Limitations

The primary outcome of this paper is the suggestion that psychophysiological indicators, such as cardiovascular and eye response, are only sensitive to considerable degradation of immersion. Since they have been proven to affect presence, we have proposed that it was a measure of coherence rather than immersion. However, this affirmation must be taken carefully since we have yet to test all combinations. Other psychophysiological indicators might have unsuspected effects on our tested sensory quality. Moreover, the tested cardiovascular and eye metrics might significantly affect other sensory qualities.

As shown by the related work, cardiovascular response and eye response are related to variations in mental workload. However, this link is not direct and other phenomena such as cybersickness, and anxiety can also cause these indicators to vary. The complexity of psychophysiological inferences [38] represent considerable limitations in the usage of psychophys-

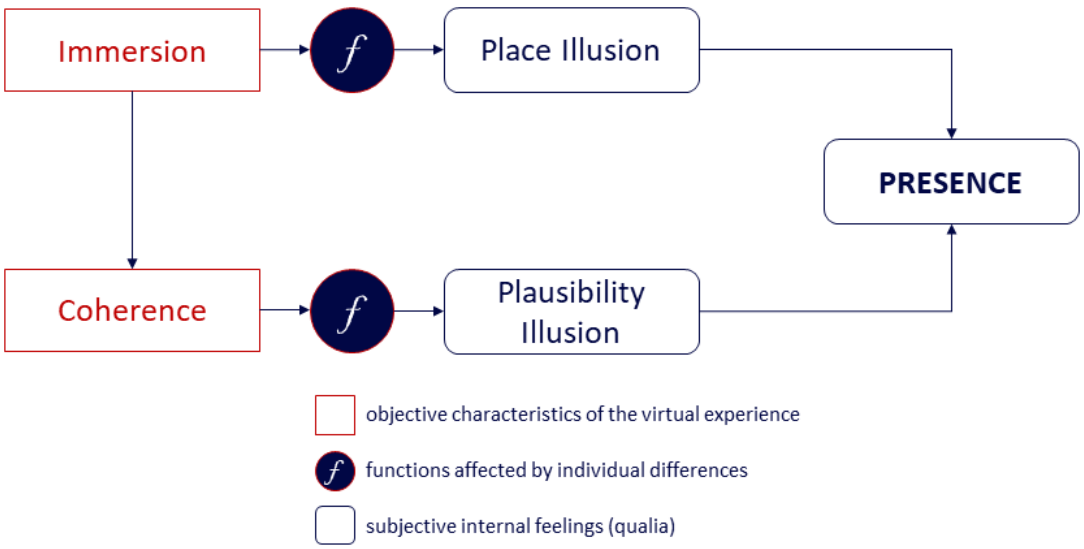


Figure 5.10 Revised Relationship Between Presence Concepts

Note. This figure is based on Skarbez et al. [15]'s proposed relationship between presence concepts.

iological indicators as a measure of presence.

A limitation of this study is the potential impact of luminance on the computation of LHIPA, as it was not controlled for during the manipulation of visual quality. This suggests that the results could have been affected by changes in pupil size due to variation in luminance. However, it should be noted that LHIPA has been shown to be more resilient to variations in luminance compared to other similar pupillary activity measures [136]. Future studies should take into consideration the potential impact of luminance on LHIPA and aim to control for this variable.

Cybersickness could have confounded our results. As previously discussed, a virtual environment with vection like the one we used, can cause cybersickness [82]. We believe that cybersickness might have had a minor effect on participants that completed the experience, but it was not important enough to be consciously acknowledged during its execution. Moreover, we do not believe that this phenomenon influenced conditions other than those where we manipulated the size of the FOV. Our participants received a clear directive to stop the experience if they felt cybersickness symptoms. Thus, other than for the degradation of the FOV, we do not suspect that this phenomenon affected our results in a way that would radically change our conclusions.

Also, since the chosen metrics were selected for a driving task, they might not be appropriate

for every virtual environment activity. For example, to study mental workload, HRV metrics are only suited for tasks where participants are seated or mostly immobile [174]. Thus, researchers must carefully choose the psychophysiological indicators they use and be aware of the context in which they will be applied while studying presence in a virtual environment.

Moreover, we do not imply that mental workload can completely capture the concept of coherence. Instead, we believe that cardiovascular and eye response indicators are sensitive to some variation in coherence and can be a helpful diagnostic tool to assess presence in a virtual environment.

5.7 Conclusion

In this experiment, we objectively manipulated immersion by setting a virtual environment contrast sensitivity, size of the FOV, or visual acuity values over, at, or under the functional threshold. The functional thresholds are sensory quality values under which perception is considered degraded. They represent a meaningful value to set visual qualities in a virtual environment since they allow using a reference value for immersion study. By using meaningful visual quality values rather than arbitrarily chosen ones, it is possible to better grasp the effect of immersion degradation in a virtual environment.

We found that this manipulation of immersion affected presence. Setting the contrast sensitivity and visual acuity under the functional threshold decreased the presence measured by the PQ. However, we found that manipulating the size of the FOV increased presence. We suspect this result is due to a reduction of cybersickness caused by an attenuation of the visual-vestibular cue conflict. Concerning the NASA-TLX results, we found that only the manipulation of visual acuity increased the subjectively reported mental workload. Other immersion manipulation seemed not to affect the NASA-TLX score. Moreover, we found that contrast sensitivity and visual acuity set at a value under the functional threshold significantly affected cardiovascular and eye response. We noticed that when sensory qualities are set under the functional threshold, they introduce unreasonable circumstances that affect the users' expectations of a virtual environment. Thus, we suggest that psychophysiological indicators and subjectively assessed mental workload are not a measure of immersion but could instead be used to assess the coherence factor of presence objectively. Also, we consider that immersion directly impacts coherence when sensory quality values are set under the functional threshold.

5.8 Acknowledgments

This work was supported by Mitacs, the Consortium for Aerospace Research and Innovation in Canada (CARIC), the Consortium de recherche et d'innovation en aérospatiale au Québec (CRIAQ) and by its partners under GRANT CARIC-CRIAQ MDO-1649-TRL4+.

Appendix A.
SAGAT Probes Used in the Virtual Environment

Table 5.6 SAGAT probes and answers for a driving task

SA Level		Probe	Answers
1	1.	Please indicate your current speed.	A. 0-4 km/h B. 5-9 km/h C. 10-14 km/h D. 15-19 km/h E. 20-24 km/h F. 25-29 km/h G. 30-35 km/h H. Over 36 km/h
1	2.	How fast is the closest driver in front of you going?	A. No cars in front B. Slower than me C. About the same speed D. Faster than me
1	3.	Please indicate how close the pedestrians are to you.	A. No pedestrians in sight B. Within 5 second of travel time in front of me C. Between 6–15 seconds of travel time in front of me D. Between 16-30 seconds of travel time in front of me E. Just past pedestrians within the last 10 seconds (in back of me)
1	4.	What is the current color of the closest traffic sign in front of you?	A. None B. Yellow C. White D. Red
2	5.	Do you need to apply your brakes to comply with any traffic signs?	A. Yes B. No
2	6.	Are you currently exceeding the speed limit?	A. More than 10 km/h under the speed limit B. 6-10 km/h under the speed limit C. Within 5 km/h of speed limit D. 6-10 km/h over the speed limit E. More than 10 km/h over the speed limit
3	7.	How far until your next turn to stay on your planned route?	A. Within 50 m B. 50 to 100 m C. 100 to 250 m D. >250 m
3	8.	Will a driver pass you in the next 30 seconds?	A. No cars behind me B. No C. Yes

This table shows the probes used in the virtual environment to assess our participants situational awareness. They were based on SAGAT probes for driving tasks by Bolstad et al. [176] and Endsley [177].

CHAPITRE 6 ARTICLE 3 : THE INFLUENCE OF IMMERSION ON SITUATIONAL AWARENESS IN A VIRTUAL ENVIRONMENT

Cet article a été soumis dans PRESENCE : Virtual and Augmented Reality le 22 février 2023.

Hébert-Lavoie, M. (Polytechnique Montréal), Doyon-Poulin P. (Polytechnique Montréal), Ozell, B. (Polytechnique Montréal)

6.1 Abstract

Researchers have pointed out the need to find an alternative to subjective questionnaires to measure presence in a virtual environment. Situational awareness has been proposed to objectively measure the concept of presence. However, the link between situational awareness and specific factors of presence has not been established. To study this relationship, sixty participants executed a driving task in a virtual environment under different visual conditions while we measured their situational awareness with the situational awareness global assessment technique (SAGAT), and their presence with the presence questionnaire (PQ). During the driving task, we objectively and meaningfully manipulated immersion, a factor of presence, by varying our participants contrast sensitivity, size of the field of view, and visual acuity. The meaningful manipulation of presence also allowed us to evaluate the functional thresholds of the three aforementioned visual qualities for a driving task, which were previously measured from a multi-directional selection test. Our results indicated a significant positive correlation between SAGAT and PQ. They also showed that SAGAT was sensitive to an immersion's degradation and brought a good diagnosticity on the effect of an immersion's manipulation. Consequently, we conclude that it could represent an objective alternative to subjective questionnaires to measure presence in a virtual environment. Moreover, our assessment of the functional thresholds allowed us to confirm that they were context dependent. Our results indicated that only the contrast sensitivity functional threshold was valid in both a multi-directional selection test and a driving task.

6.2 Introduction

Presence qualifies and quantifies a virtual environment [13]. It is a quale described as the feeling that a virtual environment is becoming real for participants [9, 10, 23]. A virtual environment enabling high presence contributes to enhancing performance, engagement and

skill learning [12, 37], and is thus desirable.

One of the considerable challenges in our ability to produce virtual environments with a strong sense of presence is the difficulty of evaluating this concept. Measuring presence poses a serious challenge because of its highly subjective nature. Most studies rely on questionnaires to measure it [12, 13] although their flaws have been exposed frequently over the last 20 years [40, 42, 180].

To find alternatives to assess presence in a virtual environment, researchers proposed using the concept of situational awareness (SA) [2, 37, 163, 164]. It is an objective concept that describes the elements leading to human decision-making. While these previous studies have demonstrated a link between SA and presence, the relationship between SA and specific presence factors has yet to be established. Thus, we studied the effect of an objective manipulation of immersion – a factor of presence – on SA. To do so, we manipulated three visual qualities: the contrast sensitivity, the field of view (FOV) size, and the visual acuity of participants executing a driving task. During the driving task, the visual qualities were set to values over, at, or under their functional threshold, a perceptual boundary under which a sensory quality value should be considered functionally degraded [188].

This paper also discusses the validity of the functional thresholds measured in Hébert-Lavoie et al. [188]. Consequently, we will analyze the validity of functional thresholds of contrast sensitivity, FOV size, and visual acuity using the findings from our present study and the results from a previous investigation [194].

The following section reviews the work related to presence and SA. Then, we present the experimental methodology used to study the impact of the degradation of three visual qualities on SA. Subsequently, we present and discuss the results obtained from this experiment.

6.3 Related work

6.3.1 Presence

Presence is characterized by the perceived realism of a virtual experience and the sensation of being in a place other than one's physical location [22]. It is defined as the "perceived realness of a mediated or virtual experience" [15]. Presence is a perceptual illusion. As opposed to a cognitive illusion, knowing it is an illusion does not change how an individual responds to it [26]. It is an embodied evaluation [195] based on an efficient and congruent multisensory integration that leads to a subjective experience [25]. Two factors are predominately associated with presence: coherence and immersion.

Coherence refers to the ability of a virtual environment to act as the user expects it to act [15]. It represents the reasonable circumstances that a scenario can evoke without introducing unreasonable circumstances [65]. During a sequence of events, reasonable circumstances are the set of states of the virtual environment that are evident from prior knowledge. Reasonable circumstances are subjective and are different for each person [30].

Immersion designates the extent and precision of the sensory modalities simulated in a virtual environment [20]. Theoretically, to fully immerse an individual in a virtual environment, one must perfectly reproduce all its sensory qualities. For example, the visual modality includes the following sensory qualities: visual accommodation, eye movements, stereopsis, perception of movement, FOV size, visual acuity, light and contrast sensitivity, and color perception [188].

Coherence and immersion are not directly related to presence. Instead, these two factors provide the constraints in which perceptual illusions can arise [15]. Coherence contributes to the plausibility illusion, the illusion that the events that occur in a virtual environment are happening for real [19]. Immersion contributes to the place illusion, the subjective experience of being in a real place [19]

Objective and meaningful quantification of immersion

Since presence is a quale, it is not possible to objectively manipulate the whole concept. Coherence depends on highly subjective notions such as the prior knowledge of the individuals experiencing presence. Thus, for now, the only way to have objective control of presence is through immersion.

Immersion assessed through the extent and precision of the simulated sensory qualities provides many advantages. It objectively quantifies the variations and applies to any virtual environment. However, the concept alone does not grant an estimation of the extent and precision at which it is necessary to simulate the sensory qualities.

A naive approach can assume that every sensory quality needs to be reproduced at its perceptual threshold, i.e., the limiting sensory values resulting in perception. However, reaching perceptual thresholds in all environments is unnecessary since the ability to sense at the perceptual threshold requires a high level of attention [51] and optimal sensing conditions [62,88]. Moreover, for the time being, it is not technically possible to reproduce all human sensations at the perceptual threshold.

Thus, instead of reproducing perceptual thresholds, we propose working with functional thresholds, i.e., the minimum values necessary to evoke place illusion. Functional thresholds

are perceptual boundaries defined as a sensory quality value below which perception is considered functionally degraded [188]. A virtual environment is more likely to evoke place illusion — the result of immersion — if it involves sensory quality values between the functional and perceptual thresholds. Functional thresholds represent the reasonable sensory quality values in a virtual environment. We have previously evaluated the functional threshold for contrast sensitivity, FOV size, and visual acuity [188].

Contrast sensitivity is a context-independent way to consider light sensitivity [51]. It is equivalent to the smallest difference in luminance required to identify a target [96]. RMS contrast, an image's standard deviation and mean luminance ratio, represent a way to measure the contrast sensitivity of natural stimuli [108]. In optimal visual conditions, the perceptual threshold for contrast sensitivity is 0.2% [88], and the functional threshold is 25.6% [188].

FOV size corresponds to the portion of space that the eyes can capture [51]. For equipment such as HMD, the amplitude in which the eyes can move while the head is fixed defines the FOV size. By moving, the eyes can cover an area of 290 degrees horizontally and 190 degrees vertically [74], which is considered the perceptual threshold for this visual quality. In comparison, the functional threshold of FOV is 96.6 degrees both horizontally and vertically [188].

Visual acuity refers to the finest level of detail that can be detected or identified [96]. Under optimal conditions, the human eye has a resolution acuity of 1 arcmin on average [73, 97, 98]. The identified functional threshold for visual acuity when the head is stable is 12.2 arcmin [188].

We have previously evaluated the impact of the degradation of contrast sensitivity, FOV size, and visual acuity at and under the functional threshold on psychophysiological indicators of mental workload [194]. The results of this assessment indicated that setting these sensory qualities at values under the functional threshold affects the physiological response and imposes a higher mental workload on virtual environment users.

The concept of functional threshold is still recent. Since it represents a way to objectively and meaningfully manipulate immersion, its use is promising for studying presence in virtual environments. It enables researchers to know under which visual quality values perception shall be considered functionally degraded.

Measuring presence

Measuring presence is a challenge since it is a subjective internal awareness state introduced by the senses [12]. The subjectivity of the concept gave rise to several definitions and measures

of presence. The use of many measures based on different concepts restricted their comparability and gave limited insights into presence variations [20, 37]. Thus, results obtained with presence measures are heterogeneous, indicating a low validity of the concept [41].

Currently, the most common measure of presence is by questionnaire [13,20]. Because of their adaptability, enabling them to capture abstract concepts, questionnaires are deemed appropriate to measure the subjective concept of presence [40]. Questionnaires' main advantages are that they are generally easy to apply and interpret [13, 20] regardless of the environment used [40], are shown to be reliable [20], and have face validity [39].

However, despite their popularity, questionnaires possess many flaws. First, their scores are subject to participants' prior information, experience, and task expectations, representing a significant source of bias [13, 20, 40, 180]. These biases affect the measures' sensitivity, diagnosticity, and validity as a measure of the concept [19]. Moreover, there are a variety of questionnaires measuring presence. Due to the lack of a standard definition of the concept, they are generally arduous to compare [40] since not all questionnaires offer the same level of detail [20]. Furthermore, they cannot compare systems that do not evoke the same qualia [19]. For example, they cannot distinguish between virtual and real experiences [43]. Finally, the ability of questionnaires to predict task performance in virtual environments is inconclusive [40].

Hence, current presence measures are unreliable. There is a need to find a valid and objective measure of presence [42].

6.3.2 Situational awareness

The similarity of the behavioral response between the real world and a virtual environment could serve as an objective index of presence [39]. An objective way to quantify a person's cognitive process that could help assess their behavioral response is the concept of situational awareness (SA). SA describes the elements leading to human decision-making. It aims at explaining an individual's performance when executing actions. It is defined as the perception of the elements present in a temporally and spatially determined environment, the comprehension of their meaning, and the projection of their state in the near future [152]. SA can be considered as the process of becoming aware, as the product of awareness, or a combination of both [153].

Many elements influence SA, such as individual factors, attributes of the task or system, and factors related to the system. The individual factors include one's age [155], goals, expectations, cognitive attributes, and information processing mechanisms, which depend on

the individual's skills, experience, and training [152]. The attributes of the task or system are, for example, its complexity. Finally, factors related to the system include elements implying both the individual and the task performed, such as stress or mental workload [152].

As shown in figure 6.1 inspired by Endsley [152], SA is composed of three ascending levels [196]. Level 1 is the perception of the status, attributes, and dynamics of the elements in the environment. Level 2 is the understanding of the situation based on a synthesis of the different elements perceived at level 1. Level 3 is the ability to anticipate the actions of the elements of the environment in the near future by referring to the information from the previous levels.

In the case of car driving, a driver must be aware of road conditions, the presence of other cars or obstacles, and the rate of change of traffic in order to predict whether it is safe to change lanes or pass another car. SA assesses whether drivers can perceive and understand what is happening around them and whether they can guide their future actions based on their understanding of the current situation. A decrease in SA increases the driver's chances of making performance errors [155].

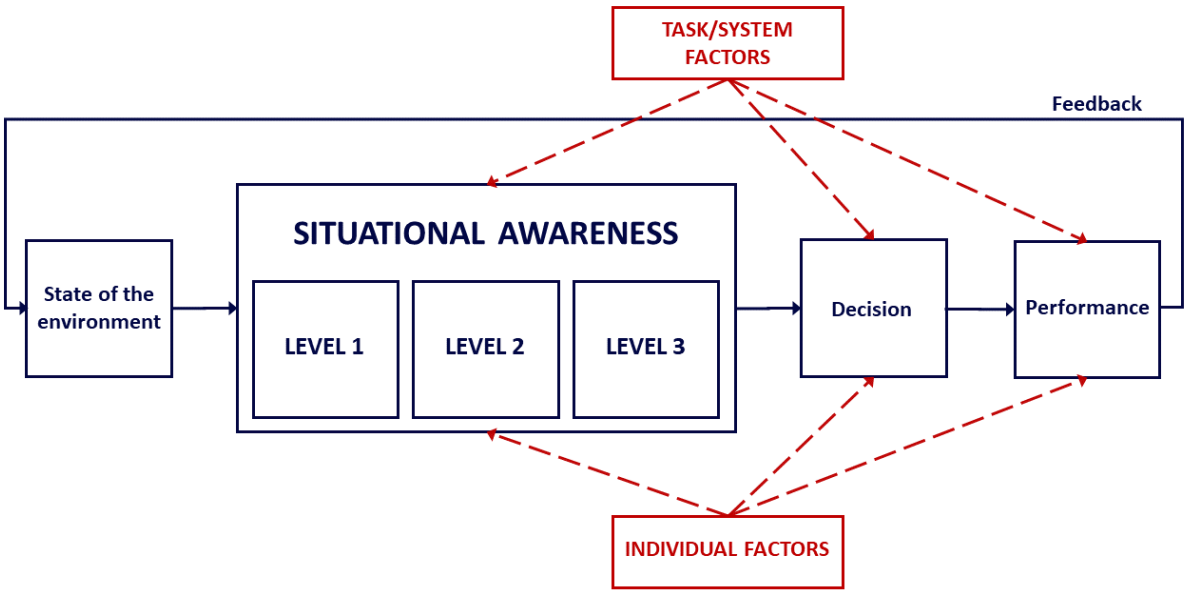


Figure 6.1 Situational awareness model

Situation Awareness Global Assessment Techniques

The Situation Awareness Global Assessment Technique (SAGAT) is an objective, direct measure of a person's SA. The technique consists of questioning individuals on elements related to their SA while they are executing a task.

Applying the SAGAT requires the task to be frozen at random times during its execution. During a freeze, the elements that can offer task-relevant information are inaccessible to the task performer. The freezes serve to ask the task performer to answer SAGAT probes without affecting their performances. SAGAT probes are based on SA requirements, the relevant information that enables a person to reach the three levels of SA. After executing the task, the answers to the SAGAT's probes are compared to the actual situation. The SA is evaluated based on the proportion of correct answers provided by the participant [161]. SAGAT produces a score of the participant's ability to recall items in the environment, their properties, and how they will act in the future [153].

Although complicated to implement, this method offers a good sensitivity, diagnosticity and validity of the results [153, 155, 159] while showing no significant evidence of disturbance due to the intrusiveness of the method [160]. SAGAT score is deterministic, linear, and positively correlated with task performance [153]. It is an objective measurement of the SA construct [156]. SAGAT is usually preferred to subjective measures since they rather tend to assess the individual's confidence in their SA rather than the concept itself [155].

Situational Awareness and Presence

We believe that SA could constitute a measure of presence. It enables researchers to evaluate the similarity of cognitive and behavioral response aspects, which could serve as an objective index of presence [39]. Moreover, SA measurements capture factors that contribute to presence such as attentional resources, understanding situation understanding, information types, and technological factors [70]

According to Riva et al. [27], presence is higher when individuals can reach their goals in a virtual environment. Using individuals' goals helps design a virtual environment in a presence-minded way [27]. They enable constraining a person's intentions in a reduced set of logical or physical alternatives. In this way, SA could act as a measure of presence since it evaluates a person's ability to gather and use the information necessary to achieve their goals.

Conceptually, SA appears to pertain to both factors of presence: immersion and coherence.

Immersion is a direct evaluation of how a system simulates human sensation [36]. By assessing

perception – the act of giving meaning to a detected sensation [51] – SA consists of a conscious evaluation of the result of immersion. For example, using a head-mounted display (HMD) in a virtual environment implies a smaller FOV which is known to affect SA [162] and decreases immersion. A small FOV size impacts the amount of visual information one can perceive. In short, immersion measures the information that can potentially be perceived, and SA measures the information that is effectively perceived.

Coherence corresponds to the set of reasonable circumstances evoked by a virtual scenario. A reasonable circumstance is a state of affairs that is self-evident, given prior knowledge [65]. In a way, SA is an evaluation of one's prior knowledge. A SA assessment evaluates an individual's decision-making process based on their capacity to perceive, understand, and anticipate a situation [152]. Part of the variability in the capacity to make an appropriate decision is influenced by a person's skills and experience [152], which are directly linked to their prior knowledge.

Few studies have proposed using SA measures as an assessment of presence. Laptaned [163] and Jung et al. [164] have found a positive correlation between SA and presence with an effect size greater than 0.20, which is the required minimum practical effect size for behavioral research [165]. However, Laptaned [163] found his correlation with a subjective measure of SA and Jung et al. [164] did their experiment with a portable game device which did not resort to immersive virtual reality.

Riley et al. [37] found that presence and SA are negatively correlated. Nevertheless, they suspect that their control interface might have biased their results. Their interface required shifting viewpoints to develop situational awareness and could have affected participants' engagement in their tasks. They argued that this would have reduced the attentional resources available for presence and could have accounted for the decrease in presence questionnaire score and the increase in situational awareness. Still, the authors suggested that the measure of situational awareness could serve as an objective measure of presence.

Finally, Ommerli [2] evaluated the relation between SA and immersion, but their model of presence stands on a different theoretical ground from the one used in this paper. They used the model of Mutterlein [167] in which the term immersion has a similar meaning to the definition of presence of Skarbez et al. [15]. They used the word telepresence while referring to place illusion and interactivity as the ability to interact with an environment, which seems related to both place illusion and plausibility illusion. They found a direct and positive relationship between telepresence and SA level 1, telepresence and SA level 2, and that interactivity affects SA level 1 and indirectly affects SA level 2 [2].

Most previous studies have shown a positive correlation between presence and SA. However,

experiments demonstrating this correlation have not been done in an immersive virtual environment or are based on a different theory of presence. Thus, some doubts subsist about the validity of this assertion. Moreover, the sensitivity of SA to immersion has never been tested.

6.3.3 Research Objectives

The first objective of the present research is to evaluate the effect of an objective and meaningful degradation of presence through a degradation of immersion on SA.

The second objective of the research presented in this paper is to evaluate the validity of the functional thresholds measured in Hébert-Lavoie et al. [188]. Functional thresholds are task-dependent. They have yet to be validated for tasks different from the one from which they have been measured. Hence, we want to assess the validity of the contrast sensitivity, FOV size, and visual acuity functional thresholds in a virtual environment. We consider that a driving task is a good way of verifying the extent of the measured functional thresholds since it relies primarily on the visual system.

6.4 Method

6.4.1 Participants

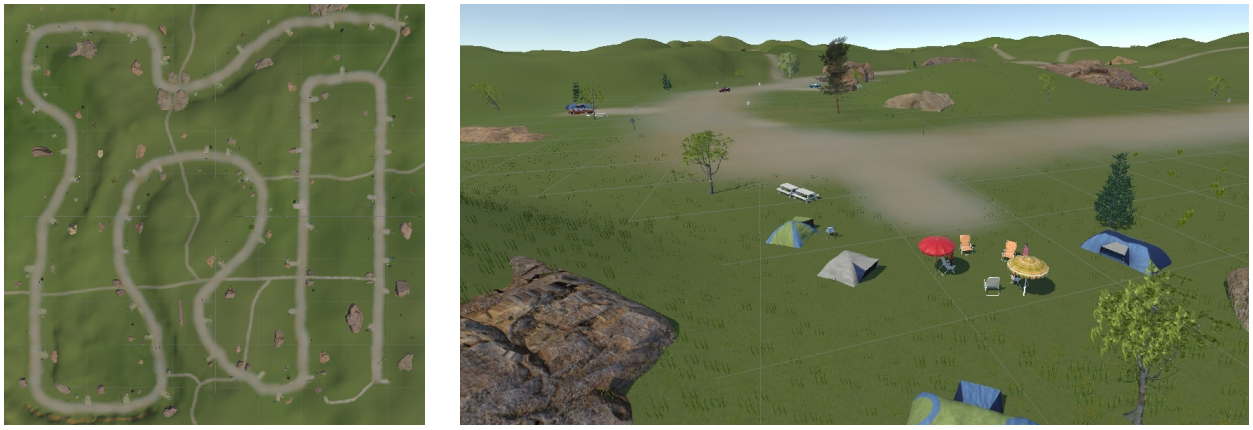
We recruited 67 voluntary participants (25 females, 42 males) for the experiment. 7 were unable to complete the experience due to cybersickness (6 females, 1 male). Thus, this paper analyzed the data of the 60 participants (19 females, 41 males) who completed the experiment. Participants were all over 18 years of age, had no uncorrected vision problems, did not have a known prevalence of being subject to cybersickness, and possessed a valid driver's license (learner, probationary or full license). To ensure the validity of the cardiac measurements, they had not consumed alcohol or caffeine and had not engaged in strenuous exercise for at least 12 hours before the experiment.

The experiment was reviewed by Polytechnique Montréal's research ethics board and met the current research ethics standards on ethical conduct for research involving humans (CER-2122-35-D).

6.4.2 Apparatus

We created a virtual environment with Unity version 2020.3.6f1. For the experiment, we used a computer with an NVIDIA GeForce RTX 2080 SUPER GPU and an Intel Core i9-9900K

CPU with 64GB RAM.



(a) Map (b) Screen shot of the virtual environment

Figure 6.2 Virtual environment used for the driving task.

Figure 6.2 shows a map and a screen shot of the virtual environment used for this experiment. The virtual environment was a hilly terrain with tall grass and flowers. A road composed of straight lines, slight curves, and sharp turns divided the terrain. The road, flanked by trees, large rocks, parked cars, camping tents, chairs, and tables, mostly crossed footpaths and spanned approximately 4 km. The virtual environment included many virtual objects to create a dynamic situation, allowing the measurement of the participants' SA. Road signs such as stop signs, turn signs, and speed limit signs were placed on both sides of the road. The speed limits indicated values between 10 km/h and 30 km/h. Twenty-seven cars were driving along the route at speeds varying between 10 km/h and 35 km/h. They were not necessarily respecting the speed limit. A simple algorithm controlled these cars, allowing them to follow a predetermined route and avoid obstacles that may have been encountered. Thirty-three pedestrians were walking or running along the road, sometimes crossing it, at speeds varying between 3 and 10 km/h. Other animated human characters were standing up or seated and performed discussion animations.

Participants used the HTC Vive Pro Eye head-mounted display (HMD) to complete the driving task. The HTC Vive Pro Eye has a dual OLED display with a diagonal of 3.5 inches allowing a 110-degree FOV, a refresh rate of 90 Hz, and a resolution of 1440 x 1600 pixels per eye for a maximum resolution acuity of 4.6 arcmin. It generates a luminance between 0.04 and 130 cd/m^2 and occupies the AdobeRGB colorspace [181]. The device is also equipped with an eye tracker with a 120 Hz gaze data output frequency and an accuracy of 0.5–1.1 degrees within 20 degrees of the center of the FOV. The HMD headphones were removed for the experiment.

To drive the virtual car, participants used the Flashfire Suzuka Wheel 900A steering wheel and pedal assembly. The steering wheel allows 900 degrees of rotation and has 20 action buttons, including two handles behind the steering wheel. The gas and a brake pedal assembly were positioned on the ground in front of the participants. To allow our participants' hands to be tracked during the experiment while being able to touch the steering wheel without hindrance, we strapped the HTC Vive Pro Eye controllers to their forearms.

Participants wore the Polar H10 (Polar H10; www.polar.com/ca-en) during the experiment. It is a heart-rate monitor with an adjustable strap that allows it to be attached to the participant's chest. These data were used to evaluate our participants' mental workload by analyzing their cardiovascular response in a previous study [194].

Questionnaires

Presence Questionnaire During the experiment, participants filled out a paper version of the UQO Cyberpsychology Lab revised presence questionnaire (PQ), corresponding to the Witmer & Singer PQ version 3. It is a 24-item questionnaire based on Witmer et al. [169] PQ using the 19 items of the PQ version 2 [169] and the 5 items related to audio and haptics. The PQ [169] is one of the most widely used [20] and shortest questionnaire [13]. It has been used in various studies, including general assessment of presence, simulations, social presence, play, or storytelling, without correlation to any specific use [20]. Its score increases according to an increase in immersion and coherence [52, 171]. Place illusion and plausibility illusion are equally important in scores on the PQ [52]. Thus, its widespread use, reliability, and face validity ensure a certain level of comparability between studies.

SAGAT probes Probes used in the virtual environment were based on SAGAT probes for driving tasks by Bolstad et al. [176] and Endsley [177] (see Table 6.1). They were selected to assess the three levels of SA of our participants.

Table 6.1 SAGAT probes and answers for a driving task

SA Level		**Probe**	**Answers**
1	1.	Please indicate your current speed.	A. 0-4 km/h B. 5-9 km/h C. 10-14 km/h D. 15-19 km/h E. 20-24 km/h F. 25-29 km/h G. 30-35 km/h H. Over 36 km/h
1	2.	How fast is the closest driver in front of you going?	A. No cars in front B. Slower than me C. About the same speed D. Faster than me
1	3.	Please indicate how close the pedestrians are to you.	A. No pedestrians in sight B. Within 5 second of travel time in front of me C. Between 6–15 seconds of travel time in front of me D. Between 16-30 seconds of travel time in front of me E. Just past pedestrians within the last 10 seconds (in back of me)
1	4.	What is current color of the closest traffic sign in front of you?	A. None B. Yellow C. White D. Red
2	5.	Do you need to apply your brakes to comply with any traffic signs?	A. Yes B. No
2	6.	Are you currently exceeding the speed limit?	A. More than 10 km/h under the speed limit B. 6-10 km/h under the speed limit C. Within 5 km/h of speed limit D. 6-10 km/h over the speed limit E. More than 10 km/h over the speed limit
3	7.	How far until your next turn to stay on your planned route?	A. Within 50 m B. 50 to 100 m C. 100 to 250 m D. >250 m
3	8.	Will a driver pass you in the next 30 seconds?	A. No cars behind me B. No C. Yes

6.4.3 Procedure

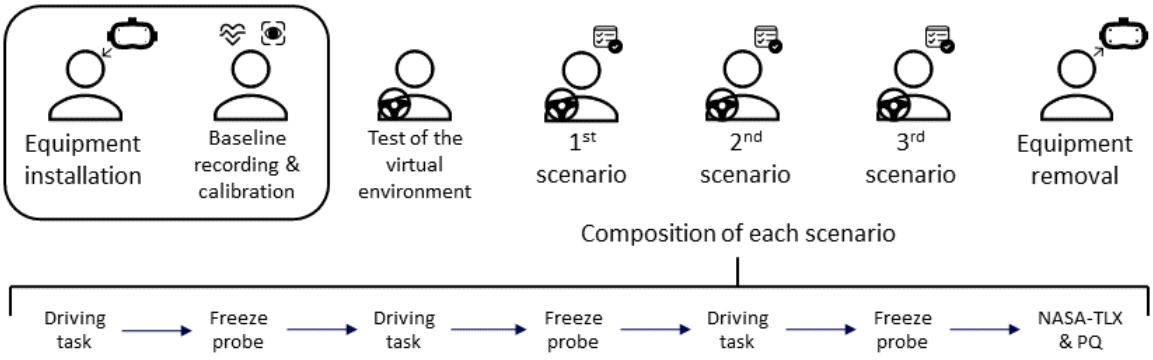


Figure 6.3 Overview of the experimental procedure

Figure 6.3 presents an overview of the experimental procedure. Upon arrival at the lab, we provided our participants with the Polar H10 heart-rate monitor and explained how to wear it appropriately. After correctly positioning the device, the participants sat in the driver's seat. We then proceeded to a 5-minutes baseline recording of their resting heart-rate. We explained the study procedures and gave the instructions to follow in case of cybersickness during the recording. After, we asked the participants to put on the HTC Vive Pro Eye and follow its eye-tracking camera calibration procedure.

Then, our participants performed a practice drive to get accustomed to the car and virtual environment parameters. All visual qualities were set at their maximum value. At the end of the practice, we did a freeze probe, specifying that the questions seen in practice would not necessarily be the ones asked during the experiment. After answering all probes, our participants completed the NASA-TLX questionnaire, for which the present paper will not discuss the results, and filled out the Presence Questionnaire (PQ) [169].

Then, our participants executed three driving scenarios. They had to respect the Highway Code at all times. We used three scenarios to prevent our participants from being able to guess the events that would occur during their driving task. The different scenarios corresponded to three different starting positions on the road. Each participant did all three different scenarios in random order. During each scenario, they were exposed to either of seven visual conditions shown in figure 6.4 along with their respective values.

This study targeted three visual qualities: contrast sensitivity, FOV size, and visual acuity. Participants either experienced no visual degradation or degradation of either one of the three visual qualities at the functional threshold or the minimum value totaling seven conditions.

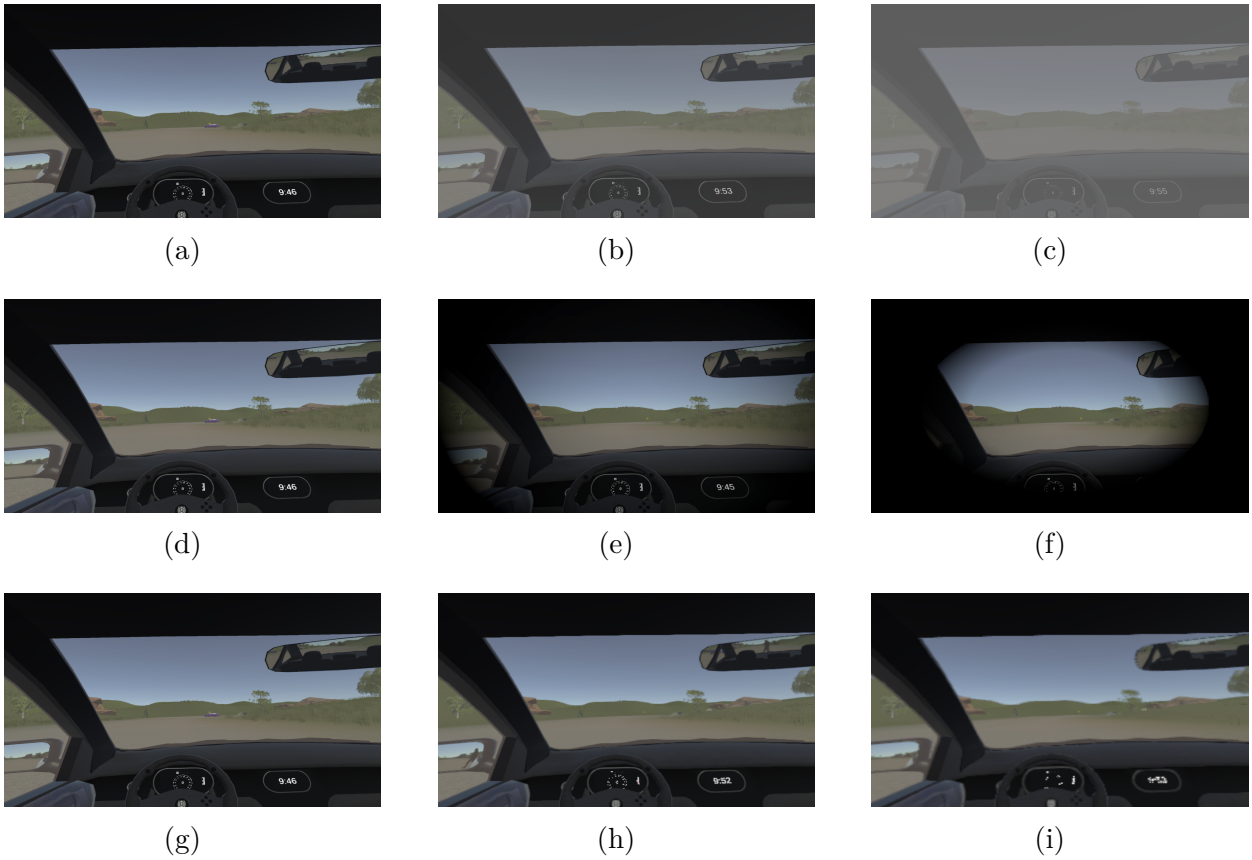


Figure 6.4 Virtual environment degradation conditions

Note. a shows the environment without any degradation, at a contrast of 31.2%; b shows the contrast degraded to 25.6 %; c shows the contrast degraded to 14.7 %; d shows the environment without any degradation, at a FOV of 110 degrees; e shows the FOV to 96.6 degrees; f shows the FOV degraded to 50.6 %; g shows the environment without any degradation, at a visual acuity of 4.6 arcmin; h shows the visual acuity degraded to 12.2 arcmin; and, i shows the visual acuity degraded to 22.8 arcmin. a, d, and g are the same image shown three times to evidence each degradation.

We determined the functional thresholds and minimum value based on a previous study [188]. The functional thresholds correspond to the results of that study, while the minimum value is the lowest functional threshold recorded in that study.

All participants experienced a scenario without degradation and two degradation conditions, meaning that each participant did not see all degradation conditions. Furthermore, they experienced only one of the seven studied visual conditions simultaneously. For example, participants experiencing the FOV size degradation had their contrast sensitivity and visual acuity set to the maximum possible value. We balanced the conditions using a Latin square experimental design to control for a presentation order bias.

Each scenario consisted of three trials ending with a freeze. We used the freeze to measure our participants' SA following the SAGAT method. During the driving task, the experimenter triggered the freezes at preselected times, occurring randomly every 3 to 6 minutes and 30 seconds. During a freeze, we set the visual qualities to their maximal values, the participant car windows became opaque, and the dashboard indications were hidden. Then, the SAGAT probes, presented in 6.1, were shown one by one in a random order to the participants who gave their answers out loud. After the first and second freeze, participants resumed driving under the visual conditions displayed before the freeze. After the third freeze, participants were asked to complete the NASA-TLX questionnaire and fill out the PQ. Participants took between 15 to 35 minutes to complete a scenario.

The experiment ended when participants completed the three scenarios. They then removed the equipment they were wearing and did a little debrief.

6.4.4 Data analysis

We gathered 60 answers ($n = 60$) for each degradation condition and 180 answers ($n = 180$) for the condition without degradation for each SAGAT probe. The SAGAT total score corresponds to the mean ratio of good answers to the SAGAT probes ranging from 0 (all answers were incorrect) to 1 (all answers were correct). The SA level 1 score corresponds to the participants' answers to probes 1 to 4, the SA level 2 score corresponds to the participants' answers to probes 5 and 6, and the SA level 2 score corresponds to the participants' answers to probes 7 and 8 found in table 6.1. We used linear mixed models for our analysis since we used repeated measurements and our data were unbalanced. We fitted our linear mixed model with the lme4 R package. We used the restricted maximum likelihood (REML) as an estimation method since it can give an unbiased estimate of the variance parameters given the number of participants ($n = 60$) that took part in the experiment. We assumed an unstructured covariance matrix of random factors. We used Satterthwaite method to compute the degrees of freedom for the t-tests and p-values. We evaluated every linear mixed model analyzes significance with an $\alpha = 0.05$. We assessed our model Intra-Class Correlation (ICC) with the R sjstats package using the adjusted value and the marginal (R^2m) and conditional (R^2c) R squared values with the R MuMIn package.

The PQ was completed 20 times ($n = 20$) for each degradation condition and 60 times ($n = 60$) for the condition without degradation. Thus, we provide descriptive statistics of this measurement in the following section. The presented statistics are the mean, standard deviation, min, max, and quartiles.

Finally, we evaluated the Pearson correlation coefficient between the SAGAT score and the

PQ score with an $\alpha = 0.05$.

6.5 Results

6.5.1 Situational Awareness

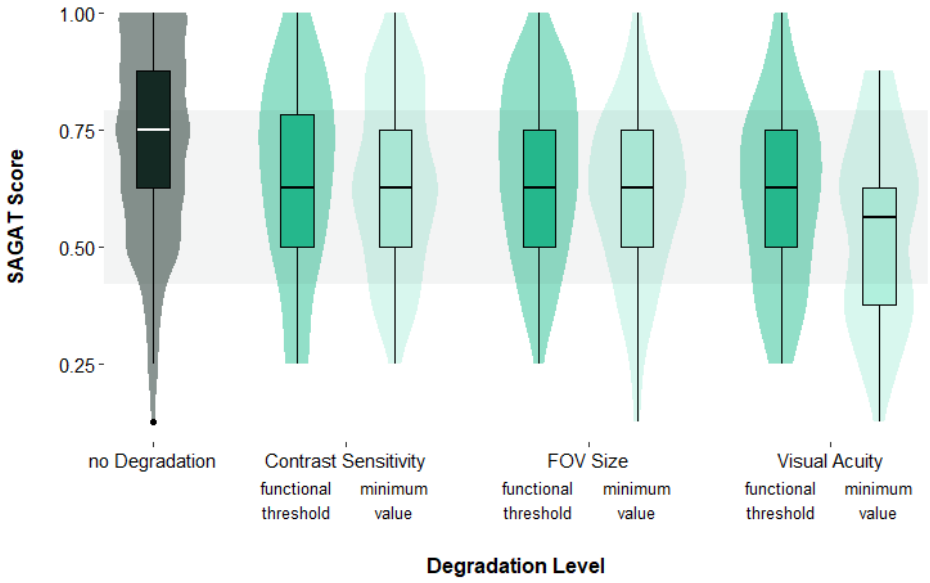


Figure 6.5 SAGAT score for each degradation condition

Note. The violin plots display a mirrored density plot of the SAGAT score for each degradation condition. They also show a boxplot which includes the median, two hinges, two whiskers, and individual outlying points.

Figure 6.5 shows the results of the SAGAT score for each degradation condition. Table 6.2 shows the min., 1st quartile, median, 3rd quartile, max., mean, and standard deviation values of the SAGAT score for each degradation condition.

The SAGAT score median was highest for the condition without degradation. Generally, the SAGAT score median was similar for every degradation condition except for the visual acuity degradation at the minimum value. In this condition, the SA was lower.

Table 6.3 shows the results of the linear mixed models for the SAGAT score. The variables that significantly impacted the SAGAT score were the FOV size at the minimum value, the visual acuity at the functional threshold, and the visual acuity at the minimum value. For the FOV size at the minimum value and the visual acuity at the functional threshold, the violin plots in figure 6.5 show a distribution of the SAGAT score around the median. However,

Table 6.2 Descriptive statistics of SAGAT score for each degradation condition

	Min	1st quartile	Median M (SD)	3rd quartile	Max
No degradation	0.1250	0.6250	0.7500 0.7250 (0.2103)	0.7250	1.0000
Contrast sensitivity functional threshold	0.2500	0.5000	0.6250 0.6563 (0.2066)	0.6563	1.0000
Contrast sensitivity min.	0.2500	0.5000	0.6250 0.6417 (0.1848)	.6417	1.0000
FOV functional threshold	0.2500	0.5000	0.6250 0.6645 (0.1802)	0.6645	1.0000
FOV min.	0.1250	0.5000	0.6250 0.6104 (0.1754)	0.6104	1.0000
Visual acuity functional threshold	0.2500	0.5000	0.6250 0.6052 (0.1894)	0.6052	1.0000
Visual acuity min.	0.1250	0.3750	0.5625 0.5375 (0.19433)	0.5375	0.8750

Note. The upper part of each line shows the min., 1st quartile, median, 3rd quartile and max. value of the SAGAT score for each degradation condition, the lower part shows the mean and the standard deviation.

Table 6.3 Results of the linear mixed model for SAGAT Score

	SAGAT Score			
	Estimate	**df**	**t-value**	**sig.**
(Intercept)	**0.6715** **0.6432** **0.6998**	251.5712	46.0120	**0.0000**
Contrast sensitivity functional threshold	-0.0140 -0.0682 0.0399	530.4412	-0.5040	0.6142
Contrast sensitivity min.	-0.0243 -0.0789 0.0295	530.4375	-0.8760	0.3813
FOV functional threshold	0.0022 -0.0535 0.0576	531.3714	0.0760	0.9395
FOV min.	**-0.0644** **-0.1184** **-0.0102**	530.6039	-2.3210	**0.0207**
Visual acuity functional threshold	**-0.0736** **-0.1265** **-0.0201**	531.4333	-2.7030	**0.0071**
Visual acuity min.	**-0.1387** **-0.1926** **-0.0844**	530.6653	-4.9960	**0.0000**
Driving experience	-0.0162 -0.0351 0.0027	59.6055	-1.6600	0.1021
	ICC	**R2m**	**R2c**	
	0.0561	0.0595	0.1122	

Note. The table shows the result for the linear mixed model of the global SAGAT Score. The upper part of each line of the table shows the estimate, degrees of freedom, *tstatistic* and significance of each variable. The lower part shows the 95% confidence interval of the estimate with the lower limit presented on the left and the upper limit presented on the right. For each model, the last line shows the Intra-Class Correlation (ICC), marginal R squared (R^2m) and conditional R squared (R^2c) for the model.

since less than 15% of the SAGAT score reached values higher or equal to 0.875, the difference between these conditions and the condition without degradation is significant.

For the condition with the FOV size at the minimum value ($M = 0.61$, $SD = 0.18$), we obtained a significantly lower SAGAT score compared to the condition without degradation ($M = 0.73$, $SD = 0.21$), $t(530.60) = -2.32$, $p = 0.0207$. The estimate for the FOV at the minimum value was -0.0644 with a 95% CI [-0.1184, -0.0102], indicating a lower SA in this condition.

For the condition with the visual acuity at the functional threshold ($M = 0.61$, $SD = 0.19$), we obtained a significantly lower SAGAT score compared to the condition without degradation ($M = 0.73$, $SD = 0.21$), $t(531.43) = -2.70$, $p = 0.0071$. The estimate for the visual acuity at the functional threshold was -0.0736 with a 95% CI [-0.1265, -0.0201], indicating a lower SA in this condition.

For the condition with the visual acuity at the minimum value ($M = 0.54$, $SD = 0.19$), we obtained a significantly lower SAGAT score compared to the condition without degradation ($M = 0.73$, $SD = 0.21$), $t(530.67) = -5.00$, $p < 0.0001$. The estimate for the visual acuity at the minimum value was -0.1387 with a 95% CI [-0.1926, -0.0844], indicating a lower SA in this condition.

Table 6.4 Results of the linear mixed model for SAGAT for each SA level

	SA Level 1				**SA Level 2**				**SA Level 3**			
	Estimate	**df**	**t-value**	**sig.**	**Estimate**	**df**	**t-value**	**sig.**	**Estimate**	**df**	**t-value**	**sig.**
(Intercept)	0.6458	336.2078	34.6700	0.0000	0.7917	226.0148	29.9750	0.0000	0.6028	241.7558	21.7690	0.0000
	0.6097 0.6820				0.7404 0.8429				0.5491 0.6565			
Contrast sensitivity functional threshold	-0.0333	515.1627	-0.8980	0.3696	0.0131	531.8259	0.2660	0.7910	-0.0036	531.0336	-0.0700	0.9440
	-0.1057 0.0389				-0.0826 0.1091				-0.1062 0.0978			
Contrast sensitivity min.	-0.0613	515.2704	-1.6500	0.0995	0.0196	531.8219	0.3990	0.6900	-0.0053	531.0285	-0.1020	0.9190
	-0.1342 0.0104				-0.0759 0.1159				-0.1087 0.0960			
FOV functional threshold	-0.0127	518.8464	-0.3320	0.7403	0.0574	531.9996	1.1350	0.2570	-0.0185	531.7190	-0.3440	0.7310
	-0.0870 0.0618				-0.0435 0.1554				-0.1227 0.0875			
FOV min.	-0.0599	515.7860	-1.6120	0.1077	-0.0844	531.8746	-1.7140	0.0870	-0.0560	531.1593	-1.0710	0.2850
	-0.1323 0.0123				-0.1804 0.0114				-0.1575 0.0468			
Visual acuity functional threshold	**-0.0973**	519.7698	-2.6650	**0.0079**	-0.0711	531.9982	-1.4720	0.1420	-0.0202	531.7539	-0.3940	0.6940
	-0.1678 -0.0256				-0.1649 0.0234				-0.1202 0.0799			
Visual acuity min.	**-0.1396**	516.0247	-3.7570	**0.0002**	**-0.2365**	531.8916	-4.8050	**0.0000**	-0.0378	531.2062	-0.7230	0.4700
	-0.2118 -0.0671				**-0.3320 -0.1400**				-0.1399 0.0641			
Driving experience	-0.0124	59.8837	-1.1140	0.2696	-0.0231	58.5788	-1.2660	0.2110	-0.0169	58.7334	-0.9060	0.3690
	-0.0336 0.0089				-0.0586 0.0123				-0.0532 0.0194			
	ICC	**R2m**	**R2c**		**ICC**	**R2m**	**R2c**		**ICC**	**R2m**	**R2c**	
	0.0045	0.0345	0.0388		0.0757	0.0595	0.1307		0.0622	0.0049	0.0668	

Note. The table shows the result for the linear mixed model of the SAGAT Score of SA Level 1, 2, and 3 respectively. The upper part of each line of the table shows the estimate, degrees of freedom, *tstatistic* and significance of each variable. The lower part shows the 95% confidence interval of the estimate with the lower limit presented on the left and the upper limit presented on the right. For each model, the last line shows the Intra-Class Correlation (ICC), marginal R squared (R^2m) and conditional R squared (R^2c) for the model.

Table 6.4 shows the results of the linear mixed models for the SAGAT Score of SA Levels 1, 2, and 3. Visual acuity at the functional threshold and the minimum value significantly impacted SA Level 1. Moreover, visual acuity at the minimum value significantly impacted SA Level 2. However, no variable included in the model significantly affected SA Level 3.

For the condition with visual acuity at the functional threshold ($M = 0.55$, $SD = 0.22$), we obtained a significantly lower SAGAT score for SA Level 1 compared to the condition without degradation ($M = 0.71$, $SD = 0.26$) , $t(519.77) = -2.67$, $p = 0.0079$. The estimate for the visual acuity at the functional threshold was -0.0973 with a 95% CI [-0.1678, -0.0256], indicating a lower Level 1 SA in this condition.

For the condition with visual acuity at the minimum value ($M = 0.51$, $SD = 0.26$), we obtained a significantly lower SAGAT score for SA Level 1 compared to the condition without degradation ($M = 0.71$, $SD = 0.26$) , $t(516.02) = -3.76$, $p = 0.0002$. The estimate for the visual acuity at the minimum value was -0.1396 with a 95% CI [-0.2118, -0.0671], indicating a lower Level 1 SA in this condition.

For the condition with the visual acuity at the minimum value ($M = 0.57$, $SD = 0.44$) we obtained a significantly lower SAGAT score for SA Level 2 compared to the condition without degradation ($M = 0.83$, $SD = 0.28$), $t(531.89) = -4.81$, $p < 0.0001$. The estimate for the visual acuity at the minimum value was -0.2365 with a 95% CI [-0.3320, -0.1400], indicating a lower Level 2 SA in this condition.

6.5.2 Presence Questionnaire

Figure 6.6 shows the results of the PQ score for each degradation condition. Table 6.5 shows the min., 1st quartile, median, 3rd quartile, max., mean, and standard deviation values of the PQ score for each degradation condition.

Our results show that the PQ score median was higher for the condition without degradation, indicating a higher presence, except when compared to the size of the FOV at the functional threshold. The FOV degraded at the functional threshold increased the PQ score compared to the condition without degradation.

When comparing each degradation pair, the PQ score median was higher for the degradation of the FOV than the other degradation conditions. Also, we can see that the visual acuity conditions had the lowest values among the studied degradation conditions.

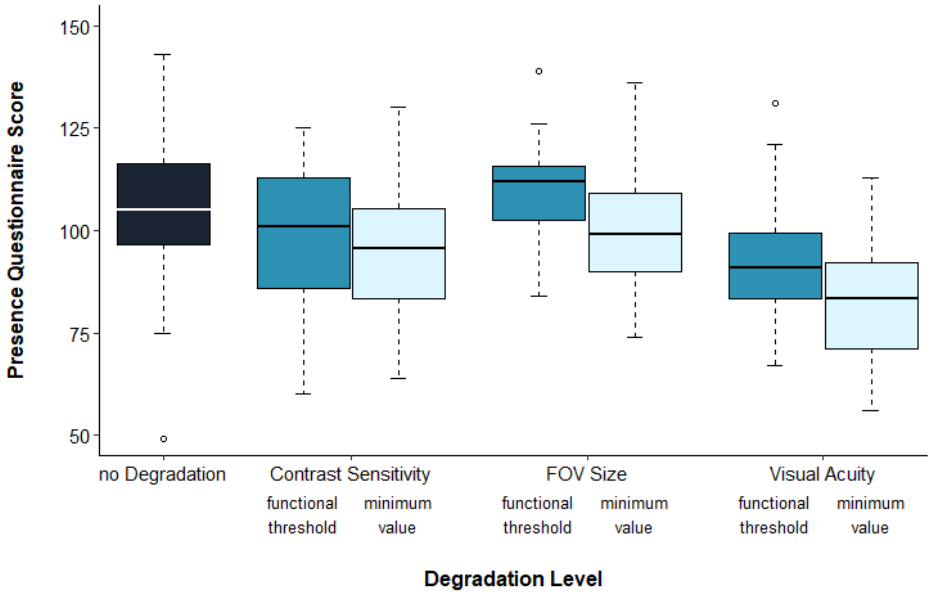


Figure 6.6 Presence questionnaire score for each degradation condition

Note. The presented boxplots include the median, two hinges, two whiskers, and individual outlying points.

Table 6.5 Descriptive statistics of Presence Questionnaire score for each degradation condition

	Min	1st quartile	Median M (SD)	3rd quartile	Max
No degradation	49.00	96.50	105.00 105.2 (16.45)	116.25	143.00
Contrast sensitivity functional threshold	60.00	86.00	101.00 99.15 (17.14)	113.00	125.00
Contrast sensitivity min.	64.00	83.25	95.50 96.45 (17.72)	105.25	130.00
FOV functional threshold	84.00	102.50	112.00 109.20 (12.98)	115.75	139.00
FOV min.	74.00	90.00	99.00 100.15 (14.25)	109.00	136.00
Visual acuity functional threshold	67.00	83.50	91.00 92.40 (16.00)	99.25	131.00
Visual acuity min.	56.00	71.00	83.50 82.15 (14.74)	92.25	113.00

Note. The upper part of each line shows the min., 1st quartile, median, 3rd quartile, and max. value of the PQ score for each degradation condition, the lower part shows the mean and the standard deviation.

6.5.3 Correlation between SA and PQ

Among our participants, the SAGAT score was slightly positively correlated to the PQ score, $r(178) = 0.20, p < 0.01$. The correlation is shown in figure 6.7.

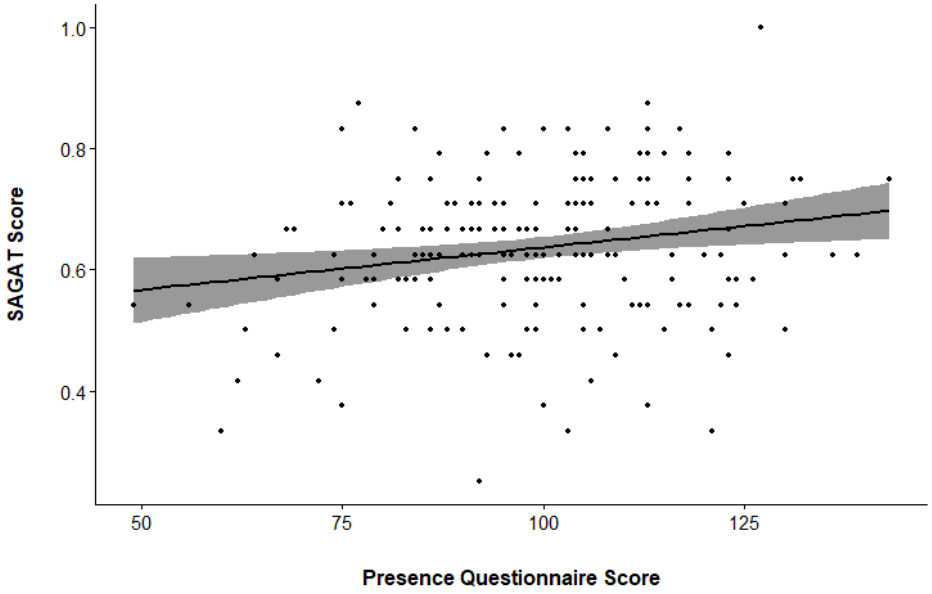


Figure 6.7 Correlation between the presence questionnaire score and the SAGAT score

6.6 Discussion

The goals of this research were twofold. First, we evaluated the effect of an objective and meaningful presence degradation through immersion on SA. To do so, we initially assessed the influence of the degradation of immersion on the SAGAT and the PQ scores separately. Then, we evaluated the relationship between the two concepts. Second, we used this experiment's results to assess the validity of the functional thresholds measured in Hébert-Lavoie et al. [188].

6.6.1 The effect of immersion's degradation on SAGAT scores

To evaluate the effect of an objective and meaningful degradation of immersion on SA, we degraded the environment by setting the contrast sensitivity, the FOV size, or the visual acuity at the functional threshold or a value under it. We then compared them to the condition without degradation by analyzing the results obtained from our four SAGAT linear mixed models. A significant difference between the condition without degradation and a

degradation condition indicated that the degradation impacted our participants' SA. In this section, we argue that SAGAT possesses two desirable qualities of a presence measurement method [13]. It has diagnosticity, and it is sensitive to an immersion's degradation.

We found a significant difference between the global SAGAT scores of the condition without degradation and the FOV size at the minimum value. In this condition, participants had to perform more head movements to acquire the information needed to complete their task. They had to change their behavior to be able to achieve their goals. However, when analyzing each SA level individually, we observed that the difference between SA level 1, 2, or 3 scores of the condition without degradation and the FOV size at the minimum value was not significant. The fact that this sensory quality significantly affected the global score but not the individual levels indicates that there is no SA level preponderantly responsible for our participants' loss of SA. Thus, the small size of the FOV altogether impacted our participants' ability to perceive, understand, and anticipate the elements related to their goals.

We found a significant difference between the global SAGAT scores of the condition without degradation and the visual acuity at the functional threshold. In this condition, the image presented to the participants was blurred, forcing them to change their behavior to be able to achieve the goals related to the driving task. Analyzing the significant difference between the SA level 1 scores of the condition without degradation and the visual acuity at the functional threshold brings more clarifications. It indicates that this immersion's degradation predominantly affected our participants' perception. Hence, we attribute the lower global SAGAT score of the visual acuity at the functional threshold to the fact that participants had to wait until they were closer to the driving-related elements, such as road signs or pedestrians, to be able to perceive them adequately.

Finally, we found a significant difference between the global SAGAT scores of the condition without degradation and the visual acuity at the minimum value. Since it was lower than the functional threshold, it imposed an evident behavior change. For example, participants had to place their heads inconveniently close to the car's dashboard to read their speed. They were also unable to read the characters on the road signs as they were too blurry. In this condition, participants had to abandon some of their driving-related goals because of their inability to perceive adequately. An analysis of the significant difference between the SA level 1 and 2 scores of the condition without degradation and the visual acuity at the minimum value explains this situation. Participants could not sufficiently perceive and comprehend the driving-related information in this immersion condition.

In summary, our results illustrated some advantages of SAGAT as a measure of presence. They suggest that the SAGAT score is sensitive to the impact of an immersion's degradation

if the affected sensory qualities are goal-related. In that case, participants tended to adapt their behavior to the new situation or abandoned their goals when they could not find an appropriate strategy. Also, our results offered diagnosticity by allowing the analysis of each SA level, clarifying the effects of the studied visual qualities' degradation on SA.

Analysis of linear mixed models

In this section, we analyze our linear mixed models' descriptive statistics to gauge the extent and scope of our findings.

As shown in table 6.3 and 6.4, the residual variability of the linear mixed models of SAGAT score, SA Level 1, SA Level 2, and SA Level 3 is explained by the participants' random intercept with an ICC between .005 and .08, indicating low reliability between our raters. The variation between participants was less substantial than the variation in each participant measurement, and a more considerable number of measurements for each participant would have helped reduce the uncertainty of the model.

The variability of SAGAT explains the poor reliability between our raters. During each trial of a SAGAT assessment, a question will have a different answer every time because the context is constantly changing. Thus, more data would have been necessary to have better rater reliability.

Moreover, the values of our models' R^2m, R^2c exposed the need for a more diverse source of information. The variance explained by the fixed effects (R^2m) was between .005 and .06, whereas the variance explained by the entire model (R^2c) was between .04 and .13 and was constantly higher than the R^2m. These low results indicate that a decrease in immersion and our participants' driving experience do not explain most of our results' variation. We expected these results as SA is a behavioral measure known to be affected by many factors that are difficult to capture.

Even though our R^2m and R^2c are low, the inclusion of random effects of participants was relevant. Our models' R^2c are constantly greater than our models' R^2m. Thus, including random effects improved our models' accuracy.

The linear mixed models used in this paper deepened our understanding of the influence of visual quality degradation on SA. The low descriptive statistics scores indicated the context in which the SAGAT measures can be applied to study presence. To enhance the SAGAT linear mixed models descriptive statistics, a high number of participants, sampling, and a diverse source of relevant goal-related information are necessary.

6.6.2 The effect of immersion degradation on the PQ score

To better understand how the degradation of contrast sensitivity, size of the FOV, and visual acuity impacted presence, we assessed how each degradation affected the PQ score.

As expected, the PQ score resulting from the degradation of contrast sensitivity at the functional threshold was close to the score obtained without degradation. For the contrast sensitivity at the minimum value, the impact of the degradation was consequential as the PQ score was lower than both previously discussed values. While more work is needed to determine if these effects are significant, our results appear to indicate that the PQ score is sensitive to immersion variation caused by a degradation of contrast sensitivity.

The results concerning the size of the FOV were unexpected. Our evaluation indicated a higher PQ score at the functional threshold than without degradation. Also, the PQ score obtained for the FOV size at the minimum value was similar to the condition without degradation. Results from previous studies indicated that values closer to the total FOV size evoked a higher presence [9, 76, 77]. Thus, we would have expected to find at least a more considerable impact on the PQ score of the FOV size at the minimum value. A virtual environment like ours where the camera moves while the body stays static can likely cause cybersickness [80]. While we did not directly measure this phenomenon, we noticed its impact on mental workload as reported in our previous paper [194]. Moreover, a few participants informally reported slight dizziness at the end of the experiment, and seven interrupted it due to nausea. Thus, we suspect we obtained this result because the reduced FOV lessened the cybersickness felt by the participants. Since we explicitly told our participants to interrupt the experiment if they were to feel cybersickness symptoms, we do not believe this ailment significantly impacted our results. However, we must acknowledge that the PQ score may not be sensitive to variations in FOV size or that we obtained these results because of a relatively small number of answers to the questionnaire.

The scores obtained for the visual acuity at the functional threshold and the minimum value were the lowest of all conditions. As discussed above, we expected such low results because we believe that the visual acuity at the functional threshold for the present experiment differs from the experiment in which it was measured [188]. To further investigate the effect of visual acuity at the functional threshold, it could be interesting to compare its impact on the PQ score with higher visual acuity values using an HMD with a better resolution. Overall, the results indicate a sensitivity of the PQ score to variation in immersion caused by a degradation of visual acuity.

Due to a limited number of observations, $n = 20$ per degradation condition, we limited our

statistical analyzes of the PQ to descriptive statistics. However, the obtained results offer an interesting perspective on the effect of the objective degradation of immersion. The analysis of the PQ results appears to reveal different tendencies than the SAGAT concerning the effect of reducing the FOV size on presence. While a decrease in FOV size seemed to have a positive or no effect on the PQ [194], it significantly decreased the SAGAT score when set at the minimum value. Otherwise, the presented results seem to corroborate Skarbez et al. [171] affirmation indicating that the PQ is sensitive to a variation in immersion. However, each visual quality does not affect the PQ with the same intensity. It is not surprising as the questionnaire does not address each visual quality directly. Thus, participants evaluate them subjectively based on their prior experiences and sensibility.

6.6.3 Situational Awareness and Presence

In the previous sections, we analyzed the effect of an immersion's degradation on the SAGAT and PQ scores individually. In this section, we assess the relationship between the two measures.

We found a low positive correlation between the SAGAT and PQ score, indicating a relationship between the two variables. Our correlation's effect size is just above 0.20, the required minimum practical effect size significance for behavioral research [165].

We explain the small correlation effect size by the difference of focus of the SAGAT and PQ in the experimental context. We did not expect a perfect agreement between the two measures. The PQ encompasses many subjectively assessed factors, such as involvement, sensory fidelity, and adaptation to the virtual environment [169]. SAGAT is more specific as it objectively measures part of a decision process; a person's ability to perceive, understand and anticipate elements related to their goals [152].

In the context of our experiment, the impact of FOV size's degradation illustrates the difference between the two measures and explains, at least partly, the low correlation. Our results suggest that a FOV reduction under the functional threshold decreases the SAGAT score. In contrast, due to fewer cybersickness symptoms, a FOV reduction at the functional threshold increased the PQ score, and a FOV reduction at the minimum value did not affect the PQ score. In our experiment, the manipulation did not affect the SAGAT score in the same way as the PQ score. The difference is due to the nature of SAGAT probes which did not focus on factors addressed by the PQ.

Previous studies reported a greater positive correlation between SA and presence than our experiment. Laptaned [163] found a low positive correlation $r = 0.373, p < 0.003$ between

SA and PQ score. However, this result is based on a subjective questionnaire of SA. Subjective questionnaires are not adapted for a precise context, as is SAGAT. They assess the individual's confidence in their SA [159], which may explain why the correlation found in Laptaned's [163] study is higher than ours. Jung et al. [164] found a highly positive correlation $r = 0.57, p = 0.001$ between the SAGAT and PQ score. However, they used a flight simulation game on a handheld console with a limited display size to carry out their experiment. All the relevant information was readily available to the participant as it was contained on the screen of the handheld console. Since Laptaned [163] and Jung et al. [164] did not manipulate immersion, coherence was the major contributor to presence variations, also explaining the difference with our results.

Moreover, Ommerli [2] used partial least square modeling to investigate the relationships between telepresence and SA levels 1 and 2. They did not evaluate presence but the analog concept of telepresence. They obtained an effect size $f = 0.10$ for telepresence and SA level 1 and $f = 0.13$ for telepresence and SA level 2. Their results go in our sense as they reported a relationship with a positive but small effect between telepresence and SA levels 1 and 2.

Since the basis of the measurements of SA and presence is different, we would not have expected the SA measure and a subjective presence questionnaire to be in complete agreement. Still, our results and the presented literature suggest that a measure of SA could be a candidate measure of presence. As Slater [42] pointed out, an alternative to presence questionnaires is needed since they may not provide a valid assessment of the concept because of the subjectivity of their basis. Hence, following this analysis, we believe that SA can be used to quantify presence in a virtual environment. A measure like SAGAT can fulfill the need for objectivity to effectively assess the "perceived realness of a mediated situation" [15]. As our results demonstrated, it is able to capture factors that contribute to presence such as demand and supply of attentional resources, and understanding of situation [70].

6.6.4 Functional thresholds

The functional thresholds used in this experiment have yet to be validated in a context different from a multi-directional selection test. Hence, the second objective of this paper was to assess the validity of the contrast sensitivity, FOV size, and visual acuity functional thresholds for a driving task.

The degradation values of the present experiment were chosen based on our previous study [188]. The functional thresholds used resulted from a psychophysical experiment where the contrast sensitivity, the FOV size, and the visual acuity were slowly degraded. At the same time, the participants executed a multi-directional selection test. The participants had to tell

the experimenter when they could perceive a degradation of either visual quality targeted by the experiment. The value of the functional threshold corresponded to the value detected by 50% of the participants. Also, we drew the minimum values used in the present experiment from this previous study [188]. We did not arbitrarily choose these values as they consisted of the minimum values measured in that previous experiment.

We consider a functional threshold valid if it does not significantly affect a factor leading to a decrease in presence. Thus, this study evaluated the validity of the contrast sensitivity, FOV size, and visual acuity functional thresholds based on their impact on SA and psychophysiological indicators. We based our evaluation on the results of this paper and those of a previous paper that evaluated the relationship between psychophysiological indicators and the sensory qualities targeted by this paper [194]. We considered a functional threshold invalid if its SAGAT scores were significantly lower than the condition without degradation or if the psychophysiological indicators differed significantly from the condition without degradation.

For the contrast sensitivity, the condition at the functional threshold was not significantly different from the condition without degradation, but so did the condition at the minimum value. We expected the contrast sensitivity at the minimum value to significantly impact the SAGAT scores since it was set at a value under the functional threshold. The absence of a significant difference at the minimum value could have led us to reevaluate the contrast sensitivity functional threshold. Even at this degradation value, SA – a presence-related factor – was not different from the condition without degradation.

We justify these non-significant results with our experiment driving parameters. A low contrast renders objects of a scene harder to distinguish. However, during the experiment, the speed limit was low, enabling the participants to take the time to get close to and read the speed signs. Moreover, we probed participants on SA elements perceivable even in low contrast conditions. The SA probes focused on the largest components of the scene, e.g., the road, the cars, the pedestrians, or on components for which the contrast was optimal, e.g., speed signs composed of black numbers on a white background, easing perception. Even if the contrast was low, it did not significantly affect how our participants attempted to achieve their goals. Thus, to use SAGAT to capture the effect of an immersion's degradation under the functional threshold, the studied sensory quality must be related to the goal of the participants.

However, the contrast sensitivity degradation at the minimum value was evident. Many participants reported feeling like they were driving through fog in that condition. Thus, it impacted the psychophysiological indicators measured in Hébert-Lavoie et al. [194]. Psy-

chophysiological phenomena such as mental workload were significantly different for the condition where the contrast sensitivity was set at the minimum value. Hence, since there was a significant difference in a factor related to presence at the minimum value, we consider the functional threshold of contrast sensitivity appropriate for this experiment.

Our results related to the validity of the FOV size at the functional threshold are inconclusive. We observed no significant difference between the SAGAT scores of the functional threshold and the condition without degradation. The SAGAT scores at the minimum value differed significantly from those without degradation. However, the psychophysiological indicators implied that phenomena such as mental workload decreased significantly due to a decrease of cybersickness in the condition of the functional and minimum value [194]. Reducing a visual-vestibular cue conflict might have been responsible for these significant effects. Thus, even though our SA results support the FOV size's functional threshold and show that SA could capture the effect of the FOV size's degradation under the functional threshold, the results of the psychophysiological indicators do not allow us to conclude on the validity of this functional threshold.

We expected the functional threshold of visual acuity for the present experiment to differ from the experiment in which it was measured [188]. This previous experiment measured the functional threshold while interacting with large objects, which could be seen easily at low visual acuity. In this previous paper, we postulated that the task executed by the participants should modulate the functional threshold. In the present experiment, since the task required reading characters at a distance, the functional threshold should have been set at a level allowing reading characters without difficulty. However, the technology at our disposal did not allow us to test such conditions [100]. Thus, the visual acuity functional threshold measured for a multi-directional selection test [188] does not apply to a driving task since the SAGAT scores, and the psychophysiological indicators [194] are significantly different from the condition without degradation.

Summarily, the contrast sensitivity functional threshold measured in a multi-directional selection test can be applied to a driving task. The results regarding the FOV size are inconclusive due to a confound in the psychophysiological indicators measurement. Finally, the visual acuity functional threshold is invalid in a driving context. Still, altogether, this experiment confirms that functional thresholds are context-dependent.

6.6.5 Limitations

The conclusions drawn from this paper are subject to some limitations.

We used a driving task which is not representative of all tasks that can be carried out in an immersive virtual environment. Our conclusions should hold in general, except perhaps for tasks that would be drastically different from the current task. Moreover, an evaluation with SAGAT is not appropriate for every task. A measure of SA must be taken in a dynamic environment to be sensitive to conditions variation. The probes must evaluate sensory qualities that are relevant to the task to be able to capture variation in presence.

The SAGAT probes used in the experiment may not reflect the immersion degradation, biasing the measure toward certain visual qualities. Hence, we may have seen an effect of FOV size at the minimum value on SA level 1 if the probes focused on that aspect. However, we based our SA questionnaires on Bolstad et al. [176] and Endsley [177] and on the specific context of our experiment to address the relevant elements of a driving task. Thus, even if the probes are biased toward a specific visual quality, they are still relevant in a presence research context. They allow assessing a person's decision process by accounting for the relevant elements.

Our experiment does not demonstrate that SA captures all factors of presence. Still, we show that SAGAT could assess a part of the concept of presence. It presents a much-needed alternative to subjective questionnaires. However, the results of our experiment do not lead us to believe that SA can be used as an independent and unique measurement of presence. Even though SA allows objectively quantifying a person's decision-making process, it is only a partial picture of the cognitive processes and elements giving rise to presence. For example, the PQ assessment of presence picked off the effect of cybersickness, for which our SA measurement was not sensitive. For now, an adequate assessment of presence needs diverse sources of information.

Also, we only targeted contrast sensitivity, size of the FOV and visual acuity, which represent only a small portion of the sensory qualities. More sensory qualities must be manipulated to be able to generalise our conclusions on the link between SA and presence.

Moreover, to have a better understanding of the effect of SA on presence, the coherence factor shall be manipulated. In this experiment, we chose to manipulate only immersion as the degradation was objective and the level of degradation were meaningful, helping us to reduce the biases in the measurement of our different conditions. However, a manipulation of coherence would benefit our understanding of the relationship between SA and presence.

The number of participants in the study limits our conclusions as well. A higher number of participants could have revealed significant numbers in the evaluation of the PQ score. Such evaluation would have enabled us to have a better understanding of the relation between the PQ and the SAGAT and could have strengthened our conclusion regarding the validity of

the studied functional thresholds.

Finally, following this analysis, we noticed that many factors of presence must be evaluated to assess the validity of a functional threshold. Hence, the conclusions on the validity of the functional thresholds for a driving task must be taken carefully, as they were only assessed by evaluating SA and psychophysiological indicators. To ensure the validity of these functional thresholds, other factors related to presence should be considered.

6.7 Conclusion

We studied the effect of an immersion's degradation on SA. To do so, we degraded our participants' contrast sensitivity, FOV size, or visual acuity while they executed a driving task. We then measured their SA with SAGAT.

Our results indicated that SAGAT was sensitive to an immersion's degradation. We demonstrated that this SA's measure was sensitive to an objective immersion's degradation when the degraded sensory qualities were related to the participants' goals. Also, using SAGAT provided diagnosticity, helping researchers identify the cause of presence variation. Moreover, we found a small but positive correlation between the SAGAT score and the PQ score that corroborates the relationship between SA and presence. Thus, SA could represent an objective alternative to subjective questionnaires to measure presence in a virtual environment.

We also assessed the validity for a driving task of the contrast sensitivity, FOV size, and visual acuity functional thresholds measured in a multi-directional selection test [194]. We concluded that only the contrast sensitivity functional threshold could be used in both contexts.

6.8 Acknowledgments

This work was supported by Mitacs, the Consortium for Aerospace Research and Innovation in Canada (CARIC), the Consortium de recherche et d'innovation en aérospatiale au Québec (CRIAQ) and by its partners under GRANT CARIC-CRIAQ MDO-1649-TRL4+.

CHAPITRE 7 DISCUSSION GÉNÉRALE

La mesure du sentiment de présence est complexe étant donné sa nature subjective et un relatif manque de validité des méthodologies utilisées pour l'étudier. Des méthodes de manipulation objective couplée à l'étude d'indicateurs pertinents du sentiment de présence devraient permettre d'apporter des solutions afin de combler les lacunes actuelles liées à l'évaluation de ce concept. Ainsi, l'objectif principal de cette thèse était d'évaluer les impacts d'une variation objective et non arbitraire de l'immersion en environnement virtuel sur des indicateurs cognitifs et physiologiques du sentiment de présence. Pour ce faire, nous avons réalisé deux expériences dont les résultats ont été présentés dans trois articles.

L'objectif du premier article (chapitre 4) était d'établir grâce à une méthodologie psychophysique des seuils fonctionnels pour trois qualités visuelles, soit la sensibilité au contraste, la taille du champ de vision et l'acuité visuelle. Grâce à ces seuils fonctionnels, il est possible de manipuler l'immersion d'une manière objective et non arbitraire. Les objectifs des deuxième et troisième articles étaient relatifs à l'effet d'une manipulation de l'immersion sur des variables liées au sentiment de présence identifiées dans la revue de littérature. Ainsi, le deuxième article (chapitre 5) avait pour objectif d'évaluer l'effet d'une manipulation objective et non arbitraire de l'immersion sur la charge cognitive, l'activité cardiovasculaire et l'activité oculaire. Finalement, le troisième article (chapitre 6) avait deux objectifs. Le premier était d'évaluer l'effet d'une manipulation objective et non arbitraire de l'immersion sur la conscience de la situation. Le second était d'évaluer la validité des seuils fonctionnels mesurés dans le premier article à la lumière des résultats des chapitres 5 et 6.

Les résultats de cette thèse permettent de proposer des pistes de solution pour établir une méthodologie d'étude valide et objective du sentiment de présence. En proposant le concept de seuil fonctionnel, cette thèse présente une méthode objective et non arbitraire pour manipuler l'immersion, un facteur important du sentiment de présence. Cette méthode de manipulation a permis d'étudier l'impact de l'immersion sur des indicateurs physiologiques et des mesures de performances cognitives liées à la présence. L'impact d'une manipulation de l'immersion, l'un des deux principaux facteurs du sentiment de présence, permet de montrer des relations entre une variable et le sentiment de présence. Ainsi, les résultats de cette thèse montrent que la charge cognitive, l'activité cardiovasculaire, l'activité oculaire et la conscience de la situation sont sensibles à des dégradations de l'immersion uniquement lorsque les paramètres sont fixés à une valeur inférieure au seuil fonctionnel. Les conditions d'immersion fixées à des valeurs inférieures au seuil fonctionnel introduisent des circonstances déraisonnables qui

modifient les attentes et les comportements des utilisateur.rice.s d'un environnement virtuel, résultant en une dégradation de la cohérence. Ainsi, cette thèse montre que les variables psychophysiologiques et les mesures de processus cognitifs évalués dans le cadre des chapitres précédents consistent en des mesures de la cohérence d'un environnement virtuel.

7.1 Seuils fonctionnels

Cette thèse a introduit le concept de seuil fonctionnel. Un seuil fonctionnel est défini comme la limite perceptuelle en dessous de laquelle une valeur de qualité sensorielle est considérée comme fonctionnellement dégradée. Il s'agit d'une mesure complémentaire au seuil perceptuel, qui consiste en une valeur limite d'une qualité sensorielle pour laquelle une personne est en mesure de percevoir. Lorsque les qualités sensorielles sont simulées à des valeurs supérieures au seuil fonctionnel pour une tâche donnée, un environnement virtuel a le potentiel d'évoquer un sentiment de présence suffisante pour permettre l'exécution d'une tâche de manière à ce qu'elle engendre les comportements attendus par les individus la réalisant. En somme, les seuils fonctionnels quantifient la fidélité fonctionnelle d'un environnement. Ils indiquent si les simulations sensorielles sont suffisantes pour qu'un environnement virtuel puisse supporter les buts qu'une personne souhaite atteindre.

Les études ayant analysé les effets de l'immersion en environnement virtuel le font généralement en variant le niveau d'immersion de manière arbitraire [9, 30, 59]. Elles proposent des niveaux élevé, modéré ou bas d'immersion sans bases théoriques pour justifier leur catégorisation. Les seuils fonctionnels présentent l'avantage d'établir une base qui permet une telle catégorisation. Ainsi, l'utilisation de seuils fonctionnels est non arbitraire, car elle permet la manipulation de l'immersion à des niveaux définis représentant une possibilité élevée d'être en mesure d'évoquer un sentiment de présence. De plus, l'utilisation de seuils fonctionnels pour mesurer le sentiment de présence est objective, car elle est basée sur la mesure des phénomènes physiques qui sont impliqués dans la sensation.

Les résultats du chapitre 4 indiquent un seuil fonctionnel de la sensibilité au contraste de 25,6 %, un seuil fonctionnel de la taille du champ de vision horizontal et vertical de 96,6 degrés et un seuil fonctionnel de l'acuité visuelle de 12,2 arcmin. Ces résultats ont été obtenus lors d'une tâche standardisée de sélection multidirectionnelle. Nous avions choisi d'imposer l'exécution de cette tâche lors de l'identification des seuils fonctionnels pour obtenir une mesure représentative d'une tâche générale exécutée en environnement virtuel. D'abord, elle servait à diriger l'attention des participant.e.s sur des éléments précis de l'environnement virtuel. Les participant.e.s n'étaient donc pas seulement attentifs à la dégradation de leurs qualités visuelles. De plus, la tâche de sélection multidirectionnelle n'impliquait pas de sensibilité par-

ticulière aux qualités visuelles évaluées. Les éléments pertinents pour la tâche étaient donc contrastés, ils avaient des couleurs vives, ce qui les mettait en évidence. Ils étaient placés au centre du champ de vision des participant.e.s et ils étaient d'une taille suffisamment importante pour les distinguer facilement même lorsque la résolution de l'écran était basse. De cette manière, les seuils perceptuels ont été mesurés dans un contexte permettant de ne pas les surestimer, puisque la nature de la tâche ne menait pas à une détection hâtive de la dégradation. La figure 7.1 offre une représentation schématique des seuils fonctionnels des trois qualités visuelles étudiées. Elle présente une représentation graphique des valeurs supérieures au seuil fonctionnel, des valeurs du seuil fonctionnel et des valeurs inférieures au seuil fonctionnel.

Toutefois, il a été postulé dans le chapitre 4 que les seuils fonctionnels dépendaient du contexte dans lequel était réalisée une tâche en environnement virtuel. Les seuils identifiés sont donc potentiellement appelés à être révisés pour s'adapter à une tâche particulière. Cependant, nous affirmons que les seuils perceptuels identifiés au chapitre 4 agissent en tant que plancher étant donné que les paramètres de la tâche choisie pour les évaluer n'impliquaient pas de sensibilité particulière aux qualités sensorielles évaluées. Ainsi, nous estimons que ces seuils fonctionnels peuvent être utilisés comme valeurs minimales que doit offrir une technologie

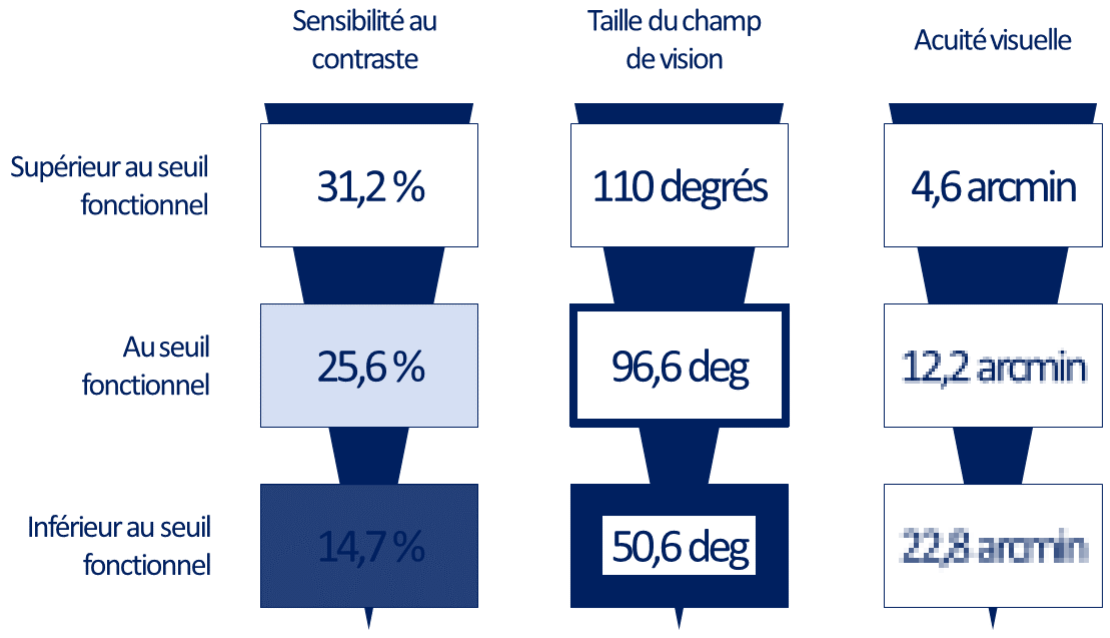


FIGURE 7.1 Représentation schématique des trois qualités visuelles à des valeurs autour du seuil fonctionnel

pour espérer évoquer un sentiment de présence.

Les chapitres 5 et 6 traitent d'une expérience visant à évaluer la validité des seuils fonctionnels mesurés dans un contexte sensoriel différent de celui où ils ont été mesurés. Lors de cette expérience, 60 participant.e.s ont effectué une tâche de conduite automobile en environnement virtuel. La tâche de conduite automobile a été choisie, car elle implique une sensibilité aux différentes qualités visuelles étudiées. Notamment, les contrastes entre certains éléments, comme les piétons et la route, étaient moins accentués que dans l'expérience précédente. Plusieurs informations d'intérêt, comme des piétons ou des panneaux de signalisation, se retrouvaient dans le champ de vision périphérique. De plus, un niveau de résolution élevé était nécessaire pour lire les caractères inscrits sur les panneaux de signalisation et les indications du tableau de bord.

Lors de notre évaluation, un seuil fonctionnel était considéré comme valide pour la tâche de conduite automobile si le fait de fixer une qualité sensorielle à cette valeur n'affectait pas significativement des variables liées au sentiment de présence et à la tâche étudiée. Ainsi, nous avons évalué la sensibilité de la charge cognitive, de l'activité cardiovasculaire, de l'activité oculaire et de la conscience de la situation aux seuils fonctionnels de la sensibilité au contraste, de la taille du champ de vision et de l'acuité visuelle. Comme indiqué dans le chapitre 6, les résultats obtenus suggèrent que seul le seuil fonctionnel de la sensibilité au contraste peut être considéré comme valide pour une tâche de conduite automobile. Compte tenu d'un conflit visuovestibulaire affectant plusieurs variables étudiées, nous n'avons pas été en mesure de conclure sur la validité de la taille du champ de vision. Finalement, nos résultats ont montré que le seuil fonctionnel déterminé pour l'acuité visuelle était invalide pour une tâche de conduite automobile puisqu'il affectait significativement la charge cognitive, l'activité physiologique et la conscience de la situation. En somme, ces résultats permettent d'évaluer la portée des seuils fonctionnels mesurés et confirment que les seuils fonctionnels sont dépendants de la tâche effectuée.

7.2 Impacts d'une manipulation de l'immersion sur des indicateurs de ressources attentionnelles

Dans le cadre de cette thèse, nous étions les premiers à étudier les effets d'une manipulation objective et non arbitraire de l'immersion sur des variables liées au sentiment de présence. Dans le chapitre 5, nous avons évalué l'impact d'une manipulation de l'immersion sur la charge cognitive, l'activité cardiovasculaire et l'activité oculaire. La charge cognitive a été mesurée à l'aide du questionnaire NASA-TLX, l'activité cardiovasculaire à l'aide de la fréquence de base moyenne et de la variabilité de la fréquence cardiaque et l'activité oculaire à

l'aide de l'indice LHIPA, utilisé pour la première fois dans le cadre d'une étude du sentiment de présence. Dans le chapitre 6, nous avons évalué l'impact d'une manipulation de l'immersion sur la conscience de la situation à l'aide de la méthode SAGAT. Ces indicateurs liés aux ressources attentionnelles des participant.e.s ont montré une sensibilité aux manipulations de l'immersion. Cependant, ils n'étaient pas sensibles aux mêmes phénomènes.

De manière à analyser les apports des différents indicateurs de présence évalués dans les articles précédents, cette section adopte une approche quantitative et qualitative. L'approche quantitative se base sur une évaluation des résultats statistiques obtenus dans les chapitres 5 et 6 pour déterminer les impacts significatifs des manipulations de l'immersion sur les différentes variables étudiées. Étant donné que la possibilité d'atteindre un but est un élément important du sentiment de présence et un déterminant de la cohérence d'un environnement virtuel [27], l'approche qualitative considère la manière dont les manipulations ont affecté les buts des participant.e.s.

Les résultats du questionnaire NASA-TLX et l'indicateur d'activité pupillaire LHIPA ont été impactés seulement par la manipulation de l'acuité visuelle. Les résultats du questionnaire NASA-TLX montrent une augmentation de la charge cognitive pour les conditions où l'acuité visuelle est fixée au seuil fonctionnel ou à un niveau inférieur au seuil fonctionnel par rapport à la condition sans dégradation. Les résultats du LHIPA indiquent une différence significative entre la condition sans dégradation et la condition où l'acuité visuelle est fixée à une valeur inférieure au seuil fonctionnel. La condition où l'acuité visuelle est fixée à une valeur inférieure au seuil fonctionnel engendre un LHIPA plus élevé, ce qui indique potentiellement une hausse de la charge cognitive [136]. Ces conditions d'immersion présentaient la dégradation la plus évidente. Il était possible pour les participant.e.s d'identifier que leur perception était altérée et il était évident que la manipulation de l'acuité visuelle affectait les différents buts en lien avec la tâche de conduite automobile. Ainsi, les résultats du questionnaire NASA-TLX et du LHIPA semblent être sensibles aux variations d'immersion qui affectent de manière évidente les buts poursuivis par une personne, causant une variation de la charge cognitive et donc la cohérence de l'environnement.

Les résultats liés à l'activité cardiovasculaire ont été impactés significativement par les manipulations de la sensibilité au contraste et de la taille du champ de vision.

Le SDNN, le pNN50 et la distribution des LF indiquent une baisse de la variabilité de la fréquence cardiaque pour la condition où la sensibilité au contraste est fixée à une valeur inférieure au seuil fonctionnel. Dans cette condition, les résultats des indicateurs SDNN et pNN50 reflètent une augmentation du stress, de la vigilance ou de la charge cognitive [49]. Par surcroît, le résultat de la distribution des LF offre des indications similaires en plus de

suggérer une augmentation du contrôle du système nerveux sympathique [138].

Dans la condition où la sensibilité au contraste était fixée à une valeur inférieure au seuil fonctionnel, les participant.e.s étaient tout de même en mesure de percevoir les informations nécessaires pour atteindre leurs buts de manière adéquate étant donné la conduite à basse vitesse. La manipulation de la sensibilité au contraste n'affectait pas les buts poursuivis par les participant.e.s de manière aussi évidente que la dégradation de l'acuité visuelle. Néanmoins, la manipulation de la sensibilité au contraste à une valeur inférieure au seuil fonctionnel imposait une perception dégradée évidente. Certain.e.s participant.e.s ont mentionné avoir l'impression de conduire dans du brouillard. Cette condition d'immersion, pourrait avoir imposé un état cognitif comme le stress [49] auquel les autres mesures utilisées ne sont pas particulièrement sensibles. Ainsi, des mesures de la variabilité de la fréquence cardiaque comme le SDNN, le pNN50 et la distribution des LF semblent être sensibles à une variation de la sensibilité au contraste engendrant un changement d'état cognitif. Ces variables peuvent être considérées comme des indicateurs d'une variation de la cohérence d'un environnement virtuel puisqu'ils font état d'un changement de la manière avec laquelle les participant.e.s appréhendent une situation.

Les résultats de la moyenne de la fréquence cardiaque de base, du RMSSD, de la distribution des HF et du LF/HF indiquent une différence significative entre la condition sans dégradation et la condition où la taille du champ de vision est fixée au seuil fonctionnel. Dans cette condition, la moyenne de la fréquence cardiaque est plus basse, ce qui peut être indicateur de plusieurs phénomènes dont une baisse du stress ou de la charge cognitive [113]. Similairement, le RMSSD indique une augmentation de la variabilité de la fréquence cardiaque qui peut être associée à une baisse du stress, de la vigilance ou de la charge cognitive [49]. De plus, la diminution du LF/HF et l'augmentation de la distribution des HF indiquent potentiellement une diminution de la charge cognitive [174] et une activation du contrôle du système nerveux parasympathique [138]. D'autre part, les résultats du RMSSD, du SDNN et du pNN50 indiquent une augmentation de la variabilité de la fréquence cardiaque pour la condition où la taille du champ de vision est fixée à une valeur inférieure au seuil fonctionnel. Cette augmentation de la variabilité de la fréquence cardiaque peut être associée à une baisse du stress, de la vigilance ou de la charge cognitive [49].

Dans les conditions où la taille du champ de vision était manipulée, les différences significatives avec la condition sans dégradation semblent liées à une diminution du cybermalaise. Le cybermalaise est lié à une augmentation du contrôle du système nerveux sympathique et une diminution du contrôle du système nerveux parasympathique [47]. Dans les conditions où la taille du champ de vision était manipulée, nous pouvons constater une augmentation

du contrôle du système nerveux parasympathique possiblement liée à une diminution des effets du cybermalaise sur l'activité cardiovasculaire. Bien que les participant.e.s à la seconde expérience déclaraient ne pas ressentir de cybermalaise, nous supposons que les effets d'un conflit visuovestibulaire peuvent malgré tout être observés [39, 80]. Ainsi, la moyenne de la fréquence cardiaque, le RMSSD, le SDNN, le pNN50 et la distribution des HF et le LF/HF semblent sensibles à une variation des conditions d'immersion lorsque celles-ci affectent le cybermalaise. Le cybermalaise a un impact sur les circonstances dans lesquelles se déroule une expérience virtuelle et influence donc la cohérence d'un environnement. Les mesures de l'activité cardiovasculaires sensibles au cybermalaise peuvent donc être considérées comme des indicateurs de la cohérence d'un environnement virtuel.

Aussi, il est possible de constater que les différentes mesures de la variabilité de la fréquence cardiaque ne sont pas toutes sensibles aux mêmes manipulations de l'immersion. Cette sensibilité variable peut paraître étonnante puisque ces indicateurs psychophysiologiques mesurent tous le même phénomène. Cependant, comme les résultats présentés ne sont pas contradictoires, nous estimons que les différences dans les indications des mesures de la variabilité de la fréquence cardiaque seraient dues à un contrôle à la fois du système nerveux sympathique et parasympathique. Le fait que certaines mesures sont plus sensibles au contrôle de l'un ou de l'autre expliquerait pourquoi elles n'indiquent pas toutes les mêmes différences significatives.

Les résultats du questionnaire SAGAT évaluant la conscience de la situation indiquent une sensibilité aux manipulations de la taille du champ de vision fixée à une valeur inférieure au seuil fonctionnel, et aux deux dégradations de l'acuité visuelle. Dans ces trois conditions de manipulations, les résultats du questionnaire SAGAT sont significativement inférieurs à la condition sans dégradation. La condition où la taille du champ de vision est fixée à une valeur inférieure au seuil fonctionnel a un impact sur le score global de SAGAT. Cette condition d'immersion implique un impact général sur la capacité des participant.e.s à percevoir, comprendre et anticiper la situation lorsque la taille de leur champ de vision est grandement réduite. La condition où l'acuité visuelle est fixée au seuil fonctionnel a un impact sur le score global de SAGAT, ainsi que sur le score de niveau 1 de SAGAT. L'acuité visuelle fixée au seuil fonctionnel implique un impact général sur la conscience de la situation des participant.e.s, mais aussi un impact particulier sur leur capacité à percevoir les éléments importants dans le cadre d'une tâche de conduite. La condition où l'acuité visuelle est fixée à une valeur inférieure au seuil fonctionnel a un impact sur le score global de SAGAT, ainsi que sur le score des niveaux 1 et 2 de SAGAT. La dégradation de l'acuité visuelle à une valeur inférieure au seuil fonctionnel implique donc une diminution générale de la conscience de la situation, en plus d'impacter particulièrement la capacité des participant.e.s à percevoir et à comprendre les éléments importants dans le cadre de la tâche de conduite automobile.

Les résultats du questionnaire SAGAT permettent d'identifier les conditions présentant des différences significatives par rapport à la condition sans dégradation. Le fait que ces conditions présentent des diminutions significatives de la conscience de la situation indique que les individus évoluant dans un environnement virtuel présentant ces conditions d'immersion ont une moins bonne capacité à percevoir, comprendre ou anticiper les éléments de la situation en lien avec les buts qu'ils souhaitent accomplir. Cette diminution de la conscience de la situation peut être associée à l'introduction de circonstances déraisonnables affectant la cohérence de la situation. Ainsi, on remarque que le score du questionnaire SAGAT est assez sensible pour identifier les environnements présentant un manque de cohérence.

Ces résultats permettent de préciser les informations qu'il est possible de retirer d'une évaluation du sentiment de présence basée sur une analyse de la charge cognitive et d'indicateurs psychophysiologiques. On remarque que les mesures de la charge cognitive, de l'activité oculaire, de l'activité cardiovasculaire et de la conscience de la situation sont complémentaires. Les mesures de la charge cognitives et les indicateurs psychophysiologiques permettent d'évaluer les ressources attentionnelles utilisées, tandis que la conscience de la situation indique si les ressources attentionnelles sont utilisées de manière à ce qu'une personne soit en mesure d'atteindre ses buts.

Bien que NASA-TLX, LHIPA, la moyenne de la fréquence cardiaque et les mesures de la variabilité de la fréquence cardiaque évaluent un même concept, soit la disponibilité des ressources attentionnelles, cette évaluation ne se fait pas de la même manière pour chaque mesure. Notre évaluation et les travaux de Riley et al. [37], Lackey et al. [68] et Ma et Kaber [69] montrent que NASA-TLX permet d'identifier des variations de la charge cognitive en environnement virtuel associées au but des individus avec diagnosticité et sensibilité. Selon les résultats du chapitre 5, le LHIPA permet moins de diagnosticité, mais offre une méthodologie permettant de faire une telle évaluation de manière instantanée. De plus, ces deux mesures ne semblent pas sensibles au potentiel cybermalaise observé dans nos conditions expérimentales.

Les mesures de l'activité cardiovasculaire ont montré une sensibilité à des phénomènes différents. Des mesures de la variabilité de la fréquence cardiaque comme le SDNN, le pNN50 et la distribution des LF sont sensibles à des phénomènes comme la diminution de la sensibilité au contraste. Ce phénomène n'a pas un impact directement relié à la tâche, mais il affecte tout de même la cohérence de l'environnement de manière évidente par rapport à la condition sans dégradation. La moyenne de la fréquence cardiaque, le RMSSD, le SDNN, le pNN50, la distribution des HF et le LF/HF ont démontré être sensible au conflit visuovestibulaire lié au cybermalaise. De la même manière que la manipulation de la sensibilité au contraste à une

valeur inférieure au seuil fonctionnel, les manipulations de la taille du champ de vision n'affectaient pas les buts des participant.e.s de manière directe et évidente. Ainsi, notre analyse permet de conclure que ces mesures peuvent être utilisées pour évaluer de manière sensible si la cohérence de l'environnement virtuel est affectée par le cybermalaise. Toutefois, malgré leurs bénéfices, les mesures de l'activité cardiovasculaire et particulièrement les mesures de la variabilité de la fréquence cardiaque possèdent un inconvénient limitant leur possibilité d'application. Comme révélé par notre revue de littérature, elles ne peuvent pas être utilisées pour évaluer tous les types de tâche, car elles sont affectées par plusieurs phénomènes. Pour éviter d'introduire des facteurs de confusion dans cette mesure, il est notamment nécessaire que la tâche implique des mouvements réduits du corps et n'induise pas de changements respiratoires [49].

Par rapport aux autres mesures étudiées, la mesure de la conscience de la situation à l'aide de la méthode SAGAT offre un point de vue alternatif sur l'utilisation des ressources attentionnelles. Plutôt que d'estimer la **quantité** de ressources attentionnelles utilisées, cette méthode indique **comment** elles sont utilisées. Ainsi SAGAT est en mesure d'évaluer avec une bonne diagnosticité des éléments liés à la cohérence d'un environnement virtuel d'une autre perspective que les mesures de la charge cognitive et les indicateurs psychophysiologiques. De plus, notre expérience semble montrer que SAGAT n'est pas sensible au cybermalaise lorsque ce dernier n'affecte pas directement le but des personnes impliquées dans une tâche en environnement virtuel.

En somme, les mesures utilisées dans le cadre de cette thèse sont pertinentes pour l'évaluation du sentiment de présence en environnement virtuel. Elles offrent différentes informations liées à la quantité de ressources attentionnelles utilisées et à la manière dont ces ressources sont utilisées. Ainsi, elles permettent de faire une évaluation comparative de la cohérence d'un environnement virtuel par rapport à un environnement de référence.

7.3 Implications théoriques

Cette section traite des implications théoriques des résultats de cette thèse. Dans un premier temps, elle discute de la relation entre le sentiment de présence et la conscience de la situation. Ensuite, elle pose une réflexion critique sur la méthode de manipulation du sentiment de présence utilisée lors de cette thèse. Finalement, elle traite de l'impact de l'immersion sur la cohérence d'un environnement virtuel.

La littérature scientifique avait déjà établi un lien entre le sentiment de présence et les mesures de charge cognitive, l'activité cardiovasculaire et l'activité oculaire. Toutefois, certains doutes

subsistaient par rapport au lien entre le sentiment de présence et la conscience de la situation. Une corrélation positive a été montrée par quelques études [2, 163, 164], cependant les bases théoriques relatives à la présence ou la conscience de la situation de ces dernières différaient de celles de cette thèse. D'un autre côté, Riley et al. [37] qui s'appuyaient sur les mêmes concepts théoriques que nous, avaient trouvé une corrélation négative entre la conscience de la situation et le sentiment de présence. Toutefois, ils affirmaient que ce résultat était dû à leur interface de contrôle et affirmaient s'attendre à trouver une corrélation positive entre les deux concepts. Ainsi, la corrélation entre le sentiment de présence et la conscience de la situation présentée dans le chapitre 6 a permis de confirmer que celle-ci est positive. Ce résultat permet de confirmer que le sentiment de présence est affecté par la manière dont les ressources attentionnelles sont utilisées de manière à atteindre un but. Lorsqu'un environnement virtuel permet à des individus d'utiliser leurs ressources attentionnelles de manière efficace leur permettant de compléter leurs objectifs, le sentiment de présence est plus élevée. La conscience de la situation est donc une variable qui permet d'évaluer la fidélité d'un environnement virtuel en fonction des buts.

Lors de cette thèse, nous avions pour objectif d'évaluer l'impact d'une manipulation de l'immersion sur des indicateurs de ressources attentionnelles liées au sentiment de présence. De cette manière, nous projetions être en mesure de préciser la relation entre ces variables et les facteurs du sentiment de présence. En poursuivant notre objectif, nous avons introduit une méthode de manipulation du sentiment de présence consistant à manipuler l'immersion en fixant des valeurs de qualités sensorielles à différents niveaux. Cette méthode de manipulation du sentiment de présence est i. objective, ii. non arbitraire, iii. présente l'avantage d'avoir une bonne reproductibilité et iv. offre des résultats qui peuvent être interprétés et comparés indépendamment du système utilisé pour générer l'environnement virtuel,

i. elle est objective, car elle se base sur la manière de générer les phénomènes physiques responsables de la sensation ;
ii. elle est non arbitraire lorsqu'on se base sur les seuils perceptuels pour déterminer les niveaux d'immersion fixés ;
iii. elle est facilement reproductible, étant donné que les niveaux d'immersion utilisés sont énoncés de manière explicite ;
iv. le fait de connaître les niveaux d'immersion fixés permet une meilleure comparaison des résultats provenant de différents systèmes, même s'ils ne font pas appel aux mêmes qualia.

Il est ainsi plus facile d'identifier les différences entre les études et évaluer la portée des résultats. En fait, la principale critique de cette méthode pour étudier le sentiment de présence

est qu'elle ne permet pas de manipuler l'ensemble du sentiment de présence, mais seulement l'immersion. Néanmoins, comme le sentiment de présence est composé de deux facteurs, soit l'immersion et la cohérence, il est possible de poser des hypothèses sur l'un des facteurs en faisant varier l'autre. De plus, cette méthode met en évidence l'importance de préciser les différentes valeurs de qualités sensorielles simulées de manière à contrôler les conditions d'immersion et leurs effets sur les différentes variables affectant la sentiment de présence.

Nous avons utilisé cette méthodologie pour étudier la relation entre différents indicateurs de ressources attentionnelles et les facteurs du sentiment de présence. En manipulant l'immersion lors de la mesure de la charge cognitive, l'activité cardiovasculaire, l'activité oculaire et la conscience de la situation et en analysant les résultats, nous avons conclu que ces variables étaient sensibles à la cohérence d'un environnement virtuel. Nous sommes parvenus à cette conclusion, car les différentes variables étudiées étaient sensibles aux manipulations d'immersion que lorsque celles-ci modifiaient la cohérence d'un environnement virtuel. NASA-TLX, LHIPA, le SDNN, le pNN50, la distribution des LF et le score de SAGAT ont été significativement affectées lorsque les conditions d'immersion étaient fixées à des valeurs inférieures au seuil fonctionnel. Dans ces conditions d'immersion, la perception des individus était affectée de manière évidente, entraînant des changements dans leurs comportements ou dans leurs états cognitifs. De plus, la moyenne de la fréquence cardiaque, le RMSSD, le SDNN, le pNN50, la distribution des HF et le LF/HF montrent un changement de l'état cognitif et physiologique des individus dans ces conditions par rapport à la condition de référence indiquant que la cohérence de l'environnement virtuel est affectée. L'utilisation de ces mesures permet de déterminer les états d'un scénario virtuel qui en changent la cohérence. Elles sont sensibles à divers phénomènes et permettent d'avoir une meilleure maîtrise du sentiment de présence. Elles permettent de juger si le niveau de ressources attentionnelles utilisées et la manière dont elles sont utilisées correspondent à ce à quoi on s'attend de l'environnement virtuel. Ainsi, comme cette thèse a démontré que les conditions d'immersion ont un impact sur la cohérence d'un environnement virtuel, nous proposons qu'un modèle du sentiment de présence doive tenir compte de cette relation tel qu'illustré à la figure 7.2.

7.4 Limites de la thèse

Cette thèse comporte quelques limites. D'abord, le type d'échantillon de participant.e.s ayant pris part aux expérimentations affecte le potentiel de généralisation des conclusions. Les deux expériences ont utilisé des échantillons de convenance pour le recrutement des participant.e.s. Ces échantillons ne sont probablement pas représentatifs de l'ensemble de la population et des études subséquentes avec des personnes ciblées pour leurs sensibilités ou leurs insensi-

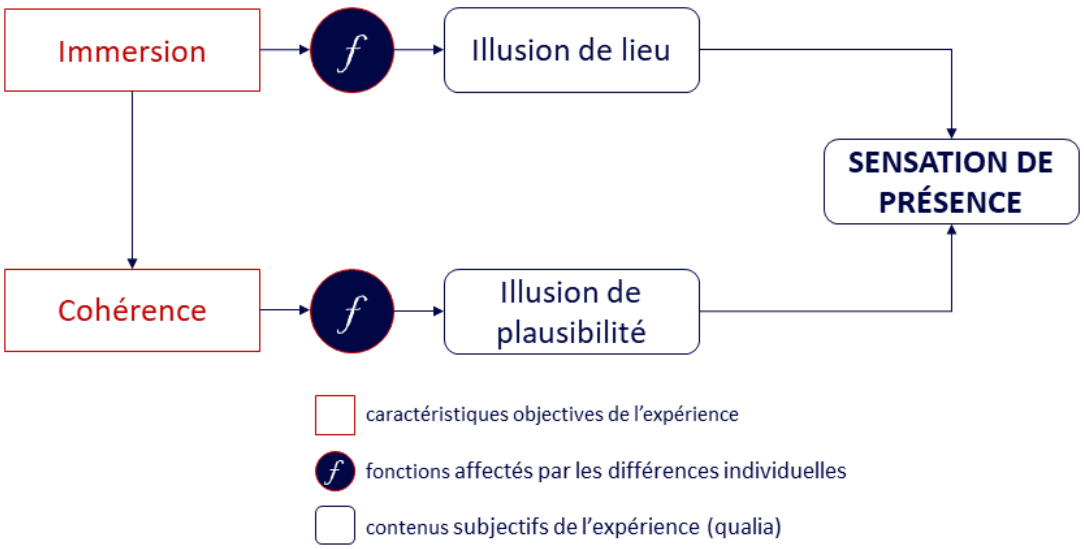


FIGURE 7.2 Relation révisée entre les concepts du sentiment de présence

Note. Cette figure se base sur les relations entre les concepts du sentiment de présence établie Skarbez et al. [15].

bilités aux phénomènes étudiés devraient permettre d'avoir une meilleure confiance dans le potentiel de généralisation des résultats présentés dans cette thèse. La taille des échantillons a aussi un impact sur le potentiel de généralisation des résultats. Les seuils fonctionnels ont été déterminés à partir d'un échantillon de 30 participant.e.s. La même mesure auprès d'une plus grande population pourrait être appelée à varier, néanmoins les indications des intervalles de confiance de chaque seuil permettent d'estimer que de telles variations n'impacteraient pas significativement le résultat des conclusions des articles suivants. En effet, les valeurs incluses dans les intervalles de confiance impliquent des conditions d'immersion semblables à celles utilisées dans cette thèse. Autrement, la taille de l'échantillon de la seconde expérimentation a notamment eu un impact sur la mesure de la charge cognitive et du sentiment de présence. Ces deux variables n'ont pas permis de récolter une quantité de mesures suffisamment importante pour juger de leur signification sur le plan statistique. Toutefois, les résultats des indicateurs psychophysiologiques et l'effet sur l'immersion ont tout de même permis de poser des hypothèses et tirer des conclusions raisonnables sur l'effet de la manipulation de ces variables.

Le fait que les seuils fonctionnels identifiés soient dépendants du contexte dans lequel ils ont été mesurés limite aussi la portée des conclusions de cette thèse. Il est nécessaire de connaître les paramètres associés à une tâche pour connaître les seuils fonctionnels qui s'y

appliquent. De plus, la validation des seuils fonctionnels est complexe. Dans cette thèse, nous avons validé les seuils étudiés en vérifiant s'ils impliquaient une dégradation significative des variables associées au sentiment de présence et à la tâche effectuée. Cependant, l'identification de l'ensemble des variables importantes permettant cette validation n'est pas aisée. Bien que les variables utilisées semblent donner un bon portrait des impacts des seuils fonctionnels sur le plan cognitif, d'autres variables pourraient potentiellement bonifier les conclusions tirées. Aussi, cette thèse n'a tenu compte de l'effet que de trois qualités visuelles et celles-ci ont été manipulées une à la fois. Or, les qualités sensorielles ne sont pas indépendantes les unes des autres. Il est donc possible que les paramètres utilisés lors de la première expérience aient influencé les résultats de celle-ci. Idéalement, il faudrait identifier le seuil fonctionnel de chaque qualité sensorielle afin d'évaluer leur importance par rapport au sentiment de présence.

Les conclusions sur la sensibilité des variables à la cohérence ont été tirées à la suite de l'évaluation d'une quantité limitée de conditions. Cette thèse a seulement fait l'étude de la manipulation de trois qualités sensorielles. Celles-ci ont été manipulées à deux niveaux, soit au seuil fonctionnel et à une valeur inférieure au seuil fonctionnel. La manipulation d'autres qualités sensorielles pourrait avoir des effets qui ne se sont pas révélés dans les expérimentations tenues dans le cadre de cette thèse.

Cette thèse propose des alternatives pour mesurer le sentiment de présence en environnement virtuel. Ces alternatives visent notamment à remplacer la mesure du sentiment de présence par questionnaire dont la validité de construit est remise en cause [42]. Cependant, contrairement aux questionnaires qui offrent un score du sentiment de présence, les alternatives proposées ne permettent qu'une évaluation comparative du sentiment de présence. Il est donc nécessaire d'avoir un environnement et une tâche de référence pour être en mesure de juger du sentiment de présence offerte avec les méthodes proposées dans cette thèse.

Malgré ses limites, cette thèse offre une perspective intéressante sur l'évaluation du sentiment de présence en environnement virtuel. Elle permet notamment d'accroître et d'enrichir la théorie qui y est rattachée.

CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS

Cette thèse a permis d'étoffer les méthodes d'étude du sentiment de présence en environnement virtuel.

D'abord, nous y avons établi la notion de seuil fonctionnel. Celle-ci détermine une limite perceptuelle en dessous de laquelle une valeur de qualité sensorielle est considérée comme fonctionnellement dégradée. Nous avons mesuré le seuil fonctionnel de la sensibilité au contraste, de la taille du champ de vision et de l'acuité visuelle pour une tâche de sélection multidirectionnelle en environnement virtuel. Lors de l'exécution de cette tâche, nous avons mesuré un seuil fonctionnel de 25.6 % pour la sensibilité au contraste, de 96.6 degrés pour la taille du champ de vision horizontal et vertical et de 12.2 arcmin pour l'acuité visuelle. Cette thèse a mis en évidence que les seuils fonctionnels sont dépendants du contexte dans lequel ils sont utilisés. Par exemple, parmi les trois mesures étudiées, uniquement la sensibilité au contraste est considérée comme valide pour une tâche semblable à la conduite automobile. Toutefois, comme les seuils fonctionnels identifiés dans cette thèse ont été mesurés lors d'une tâche ne requérant pas de sensibilité particulière aux qualités sensorielles étudiées, nous considérons qu'ils peuvent être utilisés comme valeur plancher. Un système utilisé pour générer un environnement virtuel devrait donc minimalement présenter ces valeurs pour augmenter le potentiel d'évoquer un sentiment de présence.

Ensuite, l'établissement de seuils fonctionnels aura permis de manipuler l'immersion de manière objective et non arbitraire. Grâce à cette méthode de manipulation, nous avons étudié l'effet d'une dégradation de l'immersion sur des indicateurs de ressources attentionnelles. Les ressources attentionnelles ciblées par cette thèse ont été mesurées avec le NASA-TLX, la moyenne de la fréquence cardiaque, des mesures de la variabilité de la fréquence cardiaque, le LHIPA et la méthode SAGAT. En environnement virtuel, le NASA-TLX, la moyenne de la fréquence cardiaque, les mesures de la variabilité de la fréquence cardiaque et le LHIPA offrent une estimation des ressources attentionnelles disponibles, tandis que la méthode SAGAT indique si les ressources attentionnelles sont utilisées en fonction des buts poursuivis par un individu. Ces variables étaient sensibles aux variations de l'immersion uniquement lorsque les qualités visuelles étaient fixées sous le seuil fonctionnel. Dans ces conditions, les valeurs de qualité visuelles affectaient de manière évidente les circonstances dans lesquelles se déroulait l'expérience. Ainsi, nous suggérons que les mesures de ressources attentionnelles nommées ci-dessus offrent des indications sur la cohérence d'un environnement virtuel. Cette relation implique une relation entre l'immersion et la cohérence dans la modélisation du sen-

timent de présence. Le niveau d'immersion aurait donc un impact direct sur la cohérence d'un environnement virtuel.

En somme, cette thèse aura permis d'approfondir les connaissances liées à la manipulation du sentiment de présence. Elle aura aussi permis de mieux définir la relation entre les facteurs du sentiment de présence et des indicateurs liés aux ressources attentionnelles des personnes prenant part à une expérience en environnement virtuel.

8.1 Perspectives futures

Cette thèse comporte quelques propositions pour l'évaluation du sentiment de présence en environnement virtuel qui mériteraient d'être approfondies.

Parmi ces propositions, nous ciblons notamment l'étude des seuils fonctionnels comme avenue prometteuse pour l'évaluation du sentiment de présence. Idéalement, nous estimons que de futures recherches devraient être consacrées à établir une valeur plancher de seuil fonctionnel pour l'ensemble des qualités sensorielles humaines. L'identification de ces seuils fonctionnels permettrait de déterminer si une technologie est suffisamment performante pour évoquer un sentiment de présence en environnement virtuel ou définir certaines contraintes qui empêchent ce phénomène de se manifester. Les seuils fonctionnels permettraient ainsi de faire des recommandations technologiques, notamment lors de la comparaison de systèmes.

Lors de cette thèse, nous n'avons fait varier l'immersion qu'à trois différents niveaux. Ainsi nous estimons pertinent de poursuivre la recherche en ce sens en faisant l'étude de plusieurs niveaux d'immersion. Cela permettrait d'avoir une meilleure idée de la sensibilité des utilisateur.rice.s aux seuils fonctionnels et de mieux caractériser les intervalles dans lesquels ils s'appliquent. Une analyse de multiples niveaux d'immersion permettrait aussi de mieux comprendre à quel point les différents indicateurs du sentiment de présence sont affectés par l'immersion. Ceci pourrait se faire en analysant les tailles d'effets et en vérifiant si celles-ci sont consistantes dans différents contextes. Les conclusions de ces analyses permettraient notamment d'évaluer la sensibilité des mesures à différentes manipulations de qualités sensorielles.

L'utilisation de seuils fonctionnels permettrait aussi de guider la recherche sur les aspects critiques d'un système servant à générer un environnement virtuel. Par exemple, en ce moment, on sait que certains cybermalaises sont dus au conflit visuovestibulaire [39, 80]. Ce conflit se manifeste, car le cerveau n'est pas en mesure d'intégrer les différents stimuli de manière cohérente. Le système visuel donne une information qui est en contradiction avec le système vestibulaire. Le fait de connaître les différentes qualités sensorielles qui modulent le cyber-

malaise permettrait d'établir les seuils fonctionnels nécessaires pour éviter que ce phénomène se produise. L'étude des seuils fonctionnels permettrait notamment d'identifier les éléments objectifs liés à la sensation qui affectent la présence en environnement virtuel.

Une autre avenue pouvant être explorée pour donner suite à cette thèse concerne l'analyse des seuils fonctionnels. Celles-ci devraient se faire sur plusieurs indicateurs liés au sentiment de présence. Nous avons déjà identifié des indicateurs liés à l'utilisation de ressources attentionnelles comme la charge cognitive ou la conscience de la situation comme mesures permettant d'évaluer la validité d'un seuil fonctionnel. Toutefois, nous sommes d'avis que plusieurs autres indicateurs sont pertinents, notamment des indicateurs liés aux états émotifs qui peuvent être engendrés dans un environnement virtuel. L'évaluation de l'état émotionnel des individus permettrait d'obtenir une meilleure évaluation de la fidélité d'un environnement virtuel. L'évaluation des seuils fonctionnels bénéficierait certainement de l'exploration de la dimension affective d'une expérience virtuelle. L'évaluation des états émotionnels permettrait notamment d'avoir un meilleur portrait des intentions des individus en environnement virtuel. Par exemple, être en mesure de déceler de l'anxiété dans un environnement virtuel qui ne devrait pas en générer permettrait de poser des hypothèses et d'avoir une meilleure compréhension des éléments affectant les intentions des individus et le sentiment de présence.

RÉFÉRENCES

[1] J. McFetrich, "A structured literature review on the use of high fidelity patient simulators for teaching in emergency medicine," *Emergency Medicine Journal*, vol. 23, n°. 7, p. 509–511, 2006.

[2] C. Ommerli, "Examining the effects of perceived telepresence, interactivity, and immersion on pilot situation awareness during a virtual reality flight exercise," Mémoire de maîtrise, Carleton University, 2020.

[3] J. I. Cross *et al.*, "Using extended reality in flight simulators : A literature review," *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[4] A. T. Lee, *Flight simulation : virtual environments in aviation.* Routledge, 2017.

[5] Office québécois de la langue française. (2021) réalité étendue. [En ligne]. Disponible : https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26558992

[6] B. Blissing, *Driving in virtual reality : Requirements for automotive research and development.* Linköping University Electronic Press, 2020, vol. 2085.

[7] G. Gonçalves *et al.*, "Systematic review on realism research methodologies on immersive virtual, augmented and mixed realities," *IEEE Access*, 2021.

[8] D. J. Harris *et al.*, "A framework for the testing and validation of simulated environments in experimentation and training," *Frontiers in Psychology*, vol. 11, p. 605, 2020.

[9] M. Slater, B. Spanlang et D. Corominas, "Simulating virtual environments within virtual environments as the basis for a psychophysics of presence," *ACM Transactions on Graphics*, vol. 29, n°. 4, p. 1–9, 2010.

[10] I. Milleville-Pennel et C. Charron, "Driving for real or on a fixed-base simulator : Is it so different ? An explorative study," *Presence*, vol. 24, n°. 1, p. 74–91, févr. 2015.

[11] A. Schnack, M. J. Wright et J. L. Holdershaw, "Immersive virtual reality technology in a three-dimensional virtual simulated store : Investigating telepresence and usability," *Food Research International*, janv. 2018. [En ligne]. Disponible : http://linkinghub.elsevier.com/retrieve/pii/S0963996918300280

[12] O. I. Caldas *et al.*, "Leading presence-based strategies to manipulate user experience in virtual reality environments," *Virtual Reality*, p. 1–12, 2022.

[13] V. Souza *et al.*, "Measuring presence in virtual environments : A survey," *ACM Computing Surveys (CSUR)*, vol. 54, n°. 8, p. 1–37, 2021.

[14] S. Jeon et W. Park, “Computing presence in nonfigurative virtual environment under color and motion stimulus,” dans *2016 Future Technologies Conference (FTC)*. IEEE, 2016, p. 1100–1106.

[15] R. Skarbez, F. P. Brooks, Jr. et M. C. Whitton, “A survey of presence and related concepts,” *ACM Computing Surveys*, vol. 50, n°. 6, p. 1–39, nov. 2017. [En ligne]. Disponible : http://dl.acm.org/citation.cfm?doid=3161158.3134301

[16] T. W. Schubert, “A new conception of spatial presence : Once again, with feeling,” *Communication Theory*, vol. 19, n°. 2, p. 161–187, 2009.

[17] J. R. Wilson, “Virtual environments and ergonomics : needs and opportunities,” *Ergonomics*, vol. 40, p. 1057–1077, oct. 1997. [En ligne]. Disponible : http://www.tandfonline.com/doi/abs/10.1080/001401397187603

[18] M. M. N. Diyana et D. R. A. Rambli, “Sense of presence on image based virtual reality determines user’s mental imagery,” dans *2012 IEEE Conference on Open Systems*. IEEE, 2012, p. 1–4.

[19] M. Slater, “Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments,” *Philosophical Transactions of the Royal Society B : Biological Sciences*, vol. 364, n°. 1535, p. 3549–3557, 2009.

[20] D. Hein, C. Mai et H. Hußmann, “The usage of presence measurements in research : a review,” dans *Proceedings of the International Society for Presence Research Annual Conference (Presence)*. The International Society for Presence Research Prague, 2018, p. 21–22.

[21] B. G. Witmer et M. J. Singer, “Measuring presence in virtual environments : A presence questionnaire,” *Presence : Teleoperators and Virtual Environments*, vol. 7, n°. 3, p. 225–240, juin 1998. [En ligne]. Disponible : http://www.mitpressjournals.org/doi/10.1162/105474698565686

[22] S. Weber, D. Weibel et F. W. Mast, “How to get there when you are there already ? defining presence in virtual reality and the importance of perceived realism,” *Frontiers in psychology*, p. 1538, 2021.

[23] C. Deniaud *et al.*, “The concept of "presence" used as a measure for ecological validity in driving simulators,” dans *Proceedings of the 2014 European Conference on Cognitive Ergonomics*. Vienna, Austria : ACM Press, 2014, p. 1–4.

[24] J. Höcherl *et al.*, “Human-robot assembly : Methodical design and assessment of an immersive virtual environment for real-world significance,” dans *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, vol. 1. IEEE, 2020, p. 549–556.

[25] M. Berthiaume *et al.*, "A novel integrated information processing model of presence," *Presence : Teleoperators and Virtual Environments*, vol. 27, n°. 4, p. 378–399, 2018.

[26] M. Slater, "Immersion and the illusion of presence in virtual reality," *British Journal of Psychology*, vol. 109, n°. 3, p. 431–433, 2018.

[27] G. Riva *et al.*, "Intention, action, self and other : An evolutionary model of presence," dans *Immersed in media*. Springer, 2015, p. 73–99.

[28] B. K. Szabó et A. Gilányi, "The notion of immersion in virtual reality literature and related sources," dans *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 2020, p. 000 371–000 378.

[29] S. B. Gilbert, "Perceived realism of virtual environments depends on authenticity," *Presence*, vol. 25, n°. 4, p. 322–324, 2016.

[30] R. Skarbez *et al.*, "A psychophysical experiment regarding components of the plausibility illusion," *IEEE transactions on visualization and computer graphics*, vol. 23, n°. 4, p. 1369–1378, 2017.

[31] R. Busselle et H. Bilandzic, "Fictionality and perceived realism in experiencing stories : A model of narrative comprehension and engagement," *Communication theory*, vol. 18, n°. 2, p. 255–280, 2008.

[32] J. Steuer, "Defining virtual reality : Dimensions determining telepresence," *Journal of communication*, vol. 42, n°. 4, p. 73–93, 1992.

[33] D. Alais, F. Newell et P. Mamassian, "Multisensory processing in review : from physiology to behaviour," *Seeing and perceiving*, vol. 23, n°. 1, p. 3–38, 2010.

[34] P. Bouvier, "The five pillars of presence : guidelines to reach presence," dans *Proceedings of the 11th annual international workshop on presence.*, 2008, p. 246–249.

[35] J. F. Hoorn, E. A. Konijn et G. C. Van der Veer, "Virtual reality : Do not augment realism, augment relevance," *Human-Computer Interaction : Overcoming Barriers*, vol. 4, n°. 1, p. 18–26, 2003.

[36] J. J. Cummings et J. N. Bailenson, "How immersive is enough ? a meta-analysis of the effect of immersive technology on user presence," *Media Psychology*, vol. 19, n°. 2, p. 272–309, 2016.

[37] J. M. Riley, D. B. Kaber et J. V. Draper, "Situation awareness and attention allocation measures for quantifying telepresence experiences in teleoperation," *Human Factors and Ergonomics in Manufacturing & Service Industries*, vol. 14, n°. 1, p. 51–67, 2004.

[38] S. H. Fairclough, "Fundamentals of physiological computing," *Interacting with computers*, vol. 21, n°. 1-2, p. 133–145, 2009.

[39] S. Weech, S. Kenny et M. Barnett-Cowan, "Presence and cybersickness in virtual reality are negatively related : a review," *Frontiers in psychology*, vol. 10, p. 158, 2019.

[40] S. Grassini et K. Laumann, "Questionnaire measures and physiological correlates of presence : A systematic review," *Frontiers in psychology*, vol. 11, p. 349, 2020.

[41] W. M. Felton et R. E. Jackson, "Presence : A review," *International Journal of Human–Computer Interaction*, vol. 38, n°. 1, p. 1–18, 2022.

[42] M. Slater, "How colorful was your day ? why questionnaires cannot assess presence in virtual environments," *Presence*, vol. 13, n°. 4, p. 484–493, 2004.

[43] M. Usoh *et al.*, "Using presence questionnaires in reality," *Presence*, vol. 9, n°. 5, p. 497–503, oct 2000. [En ligne]. Disponible : http://www.mitpressjournals.org/doi/10.1162/105474600566989

[44] S. Putze *et al.*, "Breaking the experience : Effects of questionnaires in vr user studies," dans *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, p. 1–15.

[45] M. Slater et M. Garau, "The use of questionnaire data in presence studies : do not seriously likert," *Presence : Teleoperators and Virtual Environments*, vol. 16, n°. 4, p. 447–456, 2007.

[46] H. Ohnishi, K. Mochizuki et Y. Y. Sugimoto, "Measuring the sense of reality produced by highly realistic surround display systems : Application of an arx model to observers' body-sway data," *Electronics and Communications in Japan (Part I : Communications)*, vol. 87, n°. 2, p. 46–58, 2004.

[47] Y. Y. Kim *et al.*, "Characteristic changes in the physiological components of cybersickness," *Psychophysiology*, vol. 42, n°. 5, p. 616–625, 2005.

[48] E. Chang, H. T. Kim et B. Yoo, "Predicting cybersickness based on user's gaze behaviors in hmd-based virtual reality," *Journal of Computational Design and Engineering*, vol. 8, n°. 2, p. 728–739, 2021.

[49] M. Lohani, B. R. Payne et D. L. Strayer, "A review of psychophysiological measures to assess cognitive states in real-world driving," *Frontiers in human neuroscience*, vol. 13, p. 57, 2019.

[50] J. Schirm, "Standardizing objective measures of presence in immersive virtual environments," dans *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, p. 1–4.

[51] J. M. Wolfe *et al.*, *Sensation & perception*, 5e éd. Sinauer Sunderland, MA, 2006.

[52] R. Skarbez, F. P. Brooks Jr et M. C. Whitton, “Immersion and coherence in a stressful virtual environment,” dans *Proceedings of the 24th acm symposium on virtual reality software and technology*, 2018, p. 1–11.

[53] G. Gonçalves *et al.*, “Impact of different sensory stimuli on presence in credible virtual environments,” *IEEE transactions on visualization and computer graphics*, vol. 26, n°. 11, p. 3231–3240, 2019.

[54] M. Marucci *et al.*, “The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence,” *Scientific Reports*, vol. 11, n°. 1, p. 1–15, 2021.

[55] G. Song, D. Pellerin et L. Granjon, “Different types of sounds influence gaze differently in videos,” *Journal of Eye Movement Research*, vol. 6, n°. 4, 2013.

[56] B. Rudolph *et al.*, “Investigating the effects of display fidelity of popular head-mounted displays on spatial updating and learning in virtual reality,” dans *International Symposium on Visual Computing*. Springer, 2020, p. 666–679.

[57] J. Hvass *et al.*, “Visual realism and presence in a virtual reality game,” dans *2017 3DTV Conference : The True Vision-Capture, Transmission and Display of 3D Video*. IEEE, 2017, p. 1–4.

[58] B. Perroud *et al.*, “Model of realism score for immersive VR systems,” *Transportation Research Part F : Traffic Psychology and Behaviour*, vol. 61, p. 238–251, févr. 2019, special TRF Issue : Driving Simulation.

[59] A. S. Azevedo, “3d sound enhanced presence in virtual environments,” *PhD thesis, Master’s thesis*, 2013.

[60] T. Muender *et al.*, “Haptic fidelity framework : Defining the factors of realistic haptic feedback for virtual reality,” dans *CHI Conference on Human Factors in Computing Systems*, 2022, p. 1–17.

[61] S. Bialkova et M. S. Van Gisbergen, “When sound modulates vision : Vr applications for art and entertainment,” dans *2017 IEEE 3rd Workshop on Everyday Virtual Reality (WEVR)*. IEEE, 2017, p. 1–6.

[62] J. Iskander, M. Hossny et S. Nahavandi, “A review on ocular biomechanic models for assessing visual fatigue in virtual reality,” *IEEE Access*, vol. 6, p. 19 345–19 361, 2018.

[63] “Cybermalaises – Laboratoire de Cyberpsychologie de l’UQO,” http://w3.uqo.ca/cyberpsy/index.php/cybermalaises/, accessed : 2020-02-20.

[64] L. Rebenitsch et C. Owen, “Review on cybersickness in applications and visual displays,” *Virtual Reality*, vol. 20, n°. 2, p. 101–125, juin 2016.

[65] R. T. Skarbez, “Plausibility illusion in virtual environments,” Thèse de doctorat, The University of North Carolina at Chapel Hill, 2016.

[66] M. Hofer *et al.*, “The role of plausibility in the experience of spatial presence in virtual environments,” *Frontiers in Virtual Reality*, p. 2, 2020.

[67] J. Sweller, “Cognitive load theory, learning difficulty, and instructional design,” *Learning and instruction*, vol. 4, n°. 4, p. 295–312, 1994.

[68] S. J. Lackey *et al.*, “The stress and workload of virtual reality training : the effects of presence, immersion and flow,” *Ergonomics*, vol. 59, n°. 8, p. 1060–1072, 2016.

[69] R. Ma et D. B. Kaber, “Presence, workload and performance effects of synthetic environment design factors,” *International Journal of Human-Computer Studies*, vol. 64, n°. 6, p. 541–552, 2006.

[70] R. S. Kalawsky *et al.*, “The validity of presence as a reliable human performance metric in immersive environments,” dans *3rd International Workshop on Presence*, 2000, p. 1–16.

[71] F. Hutmacher, “Why is there so much more research on vision than on any other sensory modality ?” *Frontiers in psychology*, vol. 10, p. 2246, 2019.

[72] B. Perroud *et al.*, “Application of the relative visual performance model in a virtual reality immersive system,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, n°. 10, p. 3128–3132, 2019.

[73] M. S. Elbamby *et al.*, “Toward low-latency and ultra-reliable virtual reality,” *IEEE Network*, vol. 32, n°. 2, p. 78–84, 2018.

[74] B. E. Riecke, H.-G. Nusseck et J. Schulte-Pelkum, “Selected technical and perceptual aspects of virtual reality displays,” Max Planck Institute for Biological Cybernetics, Tübingen, Germany, Rapport technique 154, 2006.

[75] B. C. Kress, “Digital optical elements and technologies (EDO19) : Applications to AR/VR/MR,” dans *Digital Optical Technologies 2019*, vol. 11062. International Society for Optics and Photonics, 2019, p. 1106222.

[76] J.-W. Lin *et al.*, “Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment,” dans *Proceedings IEEE Virtual Reality 2002*. IEEE, 2002, p. 164–171.

[77] H. B.-L. Duh *et al.*, “Effects of characteristics of image quality in an immersive environment,” *Presence : Teleoperators & Virtual Environments*, vol. 11, n°. 3, p. 324–332, 2002.

[78] E. D. Ragan *et al.*, “Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, n°. 7, p. 794–807, 2015.

[79] D. M. Johnson et J. E. Stewart, “Use of virtual environments for the acquisition of spatial knowledge : Comparison among different visual displays,” *Military Psychology*, vol. 11, n°. 2, p. 129–148, 1999.

[80] M. Melo *et al.*, “Impact of different role types and gender on presence and cybersickness in immersive virtual reality setups,” dans *2021 International Conference on Graphics and Interaction (ICGI)*. IEEE, 2021, p. 1–8.

[81] L. Rebenitsch et C. Owen, “Review on cybersickness in applications and visual displays,” *Virtual Reality*, vol. 20, n°. 2, p. 101–125, 2016.

[82] J. Teixeira et S. Palmisano, “Effects of dynamic field-of-view restriction on cybersickness and presence in hmd-based virtual reality,” *Virtual Reality*, vol. 25, n°. 2, p. 433–445, 2021.

[83] A. Kemeny, “Simulation and perception,” dans *Introduction de la Conférence sur la Simulation de Conduite*, 1999.

[84] B. Lidestam, L. Eriksson et O. Eriksson, “Speed perception affected by field of view : Energy-based versus rhythm-based processing,” *Transportation Research Part F : Traffic Psychology and Behaviour*, vol. 65, p. 227–241, 2019.

[85] D. Hoffman, Z. Meraz et E. Turner, “Limits of peripheral acuity and implications for VR system design,” *Journal of the Society for Information Display*, vol. 26, n°. 8, p. 483–495, 2018.

[86] M. Ogawa et K. Shidoji, “Relationship between perception of image resolution and peripheral visual field in stereoscopic images,” dans *Stereoscopic Displays and Applications XXII*, vol. 7863. International Society for Optics and Photonics, 2011, p. 78630S.

[87] A. Siekawa *et al.*, “Foveated ray tracing for VR headsets,” dans *International Conference on Multimedia Modeling*, vol. 11295, Lecture Notes in Computer Science. Springer, 2019, p. 106–117.

[88] L. N. Thibos, D. L. Still et A. Bradley, “Characterization of spatial aliasing and contrast sensitivity in peripheral vision,” *Vision research*, vol. 36, n°. 2, p. 249–258, 1996.

[89] A. Patney *et al.*, “Towards foveated rendering for gaze-tracked virtual reality,” *ACM Transactions on Graphics*, vol. 35, n°. 6, p. 1–12, 2016.

[90] A. T. Duchowski *et al.*, “On spatiochromatic visual sensitivity and peripheral color LOD management,” *ACM Transactions on Applied Perception (TAP)*, vol. 6, n°. 2, p. 1–18, 2009.

[91] T. Hansen, L. Pracejus et K. R. Gegenfurtner, "Color perception in the intermediate periphery of the visual field," *Journal of vision*, vol. 9, n°. 4, p. 26–26, 2009.

[92] S. Langlois, "ADAS HMI using peripheral vision," dans *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. New York, NY, USA : Association for Computing Machinery, 2013, p. 74–81.

[93] R. V. Ringer *et al.*, "Creating a new dynamic measure of the useful field of view using gaze-contingent displays," dans *Proceedings of the Symposium on Eye Tracking Research and Applications*, 2014, p. 59–66.

[94] B. Wolfe *et al.*, "More than the useful field : Considering peripheral vision in driving," *Applied ergonomics*, vol. 65, p. 316–325, 2017.

[95] N. Itoh, K. Sagawa et Y. Fukunaga, "Useful visual field at a homogeneous background for old and young subjects," *Gerontechnology*, vol. 8, n°. 1, p. 42–51, 2009.

[96] C. R. Bennett *et al.*, "The assessment of visual function and functional vision," dans *Seminars in Pediatric Neurology*, vol. 31. WB Saunders, 2019, p. 30–40.

[97] J. Kim *et al.*, "Foveated AR : dynamically-foveated augmented reality display," *ACM Transactions on Graphics*, vol. 38, n°. 4, p. 1–15, 2019.

[98] M. A. Livingston, "Quantification of visual capabilities using augmented reality displays," dans *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE, 2006, p. 3–12.

[99] D. Sproule *et al.*, "Characterization of visual acuity and contrast sensitivity using head-mounted displays in a virtual environment : A pilot study," dans *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 63, n°. 1. SAGE Publications Sage CA : Los Angeles, CA, 2019, p. 547–551.

[100] D. Maxwell *et al.*, "Evaluating the applicability of repurposed entertainment virtual reality devices for military training," *MODSIM World*, n°. 0028, p. 1–10, 2018.

[101] W. M. Felton, "Presence in virtual environments : Visual factors and measure convergence," Thèse de doctorat, University of Idaho, 2021.

[102] H. Q. Dinh *et al.*, "Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments," dans *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. IEEE, 1999, p. 222–228.

[103] K. M. Lee, "Why presence occurs : Evolutionary psychology, media equation, and presence," *Presence : Teleoperators & Virtual Environments*, vol. 13, n°. 4, p. 494–505, 2004.

[104] J. Ryu, N. Hashimoto et M. Sato, “Influence of resolution degradation on distance estimation in virtual space displaying static and dynamic image,” dans *2005 International Conference on Cyberworlds (CW'05)*. IEEE, 2005, p. 8–pp.

[105] K.-M. Jää-Aro et L. Kjelldahl, “Effects of image resolution on depth perception in stereo and nonstereo images,” dans *Stereoscopic Displays and Virtual Reality Systems IV*, vol. 3012. International Society for Optics and Photonics, 1997, p. 319–326.

[106] Z. Yan *et al.*, “Exploring eye adaptation in head-mounted display for energy efficient smartphone virtual reality,” dans *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications*, 2018, p. 13–18.

[107] P. G. Barten, “The effects of picture size and definition on perceived image quality,” *IEEE Transactions on Electron Devices*, vol. 36, n°. 9, p. 1865–1869, 1989.

[108] D. G. Pelli et P. Bex, “Measuring contrast sensitivity,” *Vision research*, vol. 90, p. 10–14, 2013.

[109] P. Ledda, A. Chalmers et H. Seetzen, “HDR displays : a validation against reality,” dans *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, vol. 3. IEEE, 2004, p. 2777–2782.

[110] G. Vivas-Mateos *et al.*, “Screen and virtual reality-based testing of contrast sensitivity,” dans *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society*. IEEE, 2020, p. 6054–6057.

[111] A. Dolhasz, I. Williams et M. Frutos-Pascual, “Measuring observer response to object-scene disparity in composites,” dans *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, 2016, p. 13–18.

[112] S. G. Hart et L. E. Staveland, “Development of NASA-TLX (Task Load Index) : Results of Empirical and Theoretical Research,” dans *Advances in Psychology*. Elsevier, 1988, vol. 52, p. 139–183.

[113] D. Tao *et al.*, “A systematic review of physiological measures of mental workload,” *International journal of environmental research and public health*, vol. 16, n°. 15, p. 2716, 2019.

[114] F. Paas *et al.*, “Cognitive Load Measurement as a Means to Advance Cognitive Load Theory,” *Educational Psychologist*, vol. 38, n°. 1, p. 63–71, mars 2003.

[115] E. Galy, M. Cariou et C. Mélan, “What is the relationship between mental workload factors and cognitive load types ?” *International Journal of Psychophysiology*, vol. 83, n°. 3, p. 269–275, mars 2012.

[116] E. Galy, J. Paxion et C. Berthelon, "Measuring mental workload with the nasa-tlx needs to examine each dimension rather than relying on the global score : an example with driving," *Ergonomics*, vol. 61, n°. 4, p. 517–527, 2018.

[117] H.-H. Choi, J. J. G. van Merriënboer et F. Paas, "Effects of the Physical Environment on Cognitive Load and Learning : Towards a New Model of Cognitive Load," *Educational Psychology Review*, vol. 26, n°. 2, p. 225–244, juin 2014.

[118] B. Cain, "A Review of the Mental Workload Literature," Defence Research and Development, Toronto, Ontario, Canada, Rapport technique, 2007.

[119] B. Xie et G. Salvendy, "Prediction of mental workload in single and multiple tasks environments," *International journal of cognitive ergonomics*, vol. 4, n°. 3, p. 213–242, 2000.

[120] A. H. Roscoe et G. A. Ellis, "A Subjective Rating Scale for Assessing Pilot Workload in Flight : A decade of Practical Use," Royal aerospace establishment, Farnborough, Hampshire, Technical Report TR90019, mars 1990.

[121] S. Miller, "Workload measures," National Advanced Driving Simulator, Iowa city, Rapport technique N01-006, 2001.

[122] A. Korbach, R. Brünken et B. Park, "Differentiating Different Types of Cognitive Load : A Comparison of Different Measures," *Educational Psychology Review*, vol. 30, n°. 2, p. 503–529, juin 2018.

[123] V. Villani, B. Capelli et L. Sabattini, "Use of Virtual Reality for the Evaluation of Human-Robot Interaction Systems in Complex Scenarios," dans *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, août 2018, p. 422–427.

[124] J. Mueller *et al.*, "Driving Simulator and Scenario Effects on Driver Response," *IIE Annual Conference. Proceedings*, p. 507–513, 2014.

[125] H. C. Lum *et al.*, "How Immersion, Presence, Emotion, & Workload Differ in Virtual Reality and Traditional Game Mediums," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 62, n°. 1, p. 1474–1478, sept. 2018.

[126] E. Nalivaiko *et al.*, "Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time," *Physiology & behavior*, vol. 151, p. 583–590, 2015.

[127] B. Mahanama *et al.*, "Eye movement and pupil measures : A review," *Frontiers in Computer Science*, vol. 3, p. 733531, 2022.

[128] V. Faure, R. Lobjois et N. Benguigui, "The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior," *Transportation research part F : traffic psychology and behaviour*, vol. 40, p. 78–90, 2016.

[129] K. Krejtz *et al.*, "Pupillary and microsaccadic responses to cognitive effort and emotional arousal during complex decision making," *Journal of eye movement research*, vol. 13, n°. 5, 2020.

[130] J. Laarni, N. Ravaja et T. Saari, "Using eye tracking and psychophysiological methods to study spatial presence," dans *Annual International Workshop on Presence, USA, 2003*, 2003.

[131] M. P. Huang et N. E. Alessi, "Presence as an emotional experience," dans *Medicine meets virtual reality*. Ios Press, 1999, p. 148–153.

[132] B. C. Fehringer, "Optimizing the usage of pupillary based indicators for cognitive workload," *Journal of Eye Movement Research*, vol. 14, n°. 2, 2021.

[133] Y. Morad *et al.*, "Pupillography as an objective indicator of fatigue," *Current eye research*, vol. 21, n°. 1, p. 535–542, 2000.

[134] B. J. Wilhelm *et al.*, "Objective and quantitative analysis of daytime sleepiness in physicians after night duties," *International Journal of Psychophysiology*, vol. 72, n°. 3, p. 307–313, 2009.

[135] M. I. Ahmad *et al.*, "A framework to estimate cognitive load using physiological data," *Personal and Ubiquitous Computing*, p. 1–15, 2020.

[136] A. T. Duchowski *et al.*, "The low/high index of pupillary activity," dans *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, p. 1–12.

[137] K. Krejtz *et al.*, "Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze," *PloS one*, vol. 13, n°. 9, p. e0203629, 2018.

[138] G. G. Berntson *et al.*, "Cardiovascular psychophysiology." dans *Handbook of psychophysiology, 4th ed.* Cambridge University Press, 2017, p. 183–216.

[139] B. De Maria *et al.*, "Concomitant evaluation of heart period and qt interval variability spectral markers to typify cardiac control in humans and rats," *Frontiers in Physiology*, vol. 10, p. 1478, 2019.

[140] M. Meehan *et al.*, "Physiological measures of presence in stressful virtual environments," *Acm transactions on graphics (tog)*, vol. 21, n°. 3, p. 645–652, 2002.

[141] B. Busscher *et al.*, "Physiological measures and self-report to evaluate neutral virtual reality worlds," *Journal of CyberTherapy and Rehabilitation*, vol. 4, n°. 1, p. 15–25, 2011.

[142] F. Pallavicini, A. Pepe et M. E. Minissi, “Gaming in virtual reality : What changes in terms of usability, emotional response and sense of presence compared to non-immersive video games ?” *Simulation & Gaming*, vol. 50, n°. 2, p. 136–159, 2019.

[143] B. Mehler, B. Reimer et Y. Wang, “A comparison of heart rate and heart rate variability indices in distinguishing single-task driving and driving under secondary cognitive workload,” dans *Proceedings of the sixth international driving symposium on human factors in driver assessment, training and vehicle design*, 2011, p. 590–597.

[144] R. McCraty et F. Shaffer, “Heart rate variability : new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk,” *Global advances in health and medicine*, vol. 4, n°. 1, p. 46–61, 2015.

[145] T. F. of the European Society of Cardiology, the North American Society of Pacing et Electrophysiology, “Heart rate variability : standards of measurement, physiological interpretation, and clinical use,” *Circulation*, vol. 93, n°. 5, p. 1043–1065, 1996.

[146] A. P. Anderson *et al.*, “Relaxation with immersive natural scenes presented using virtual reality,” *Aerospace medicine and human performance*, vol. 88, n°. 6, p. 520–526, 2017.

[147] A. H. X. Yang, N. Kasabov et Y. O. Cakmak, “Machine learning methods for the study of cybersickness : a systematic review,” *Brain Informatics*, vol. 9, n°. 1, p. 24, 2022.

[148] N. Setiowati, T. Wijayanto et F. Trapsilawati, “Identifying cybersickness when wearing a head-mounted display through heart rate variability data,” dans *IOP Conference Series : Materials Science and Engineering*, vol. 885, n°. 1. IOP Publishing, 2020, p. 012069.

[149] A. Stuiver *et al.*, “Short-term cardiovascular measures for driver support : Increasing sensitivity for detecting changes in mental workload,” *International Journal of Psychophysiology*, vol. 92, n°. 1, p. 35–41, 2014.

[150] M. Shakouri *et al.*, “Analysis of the sensitivity of heart rate variability and subjective workload measures in a driving simulator : the case of highway work zones,” *International journal of industrial ergonomics*, vol. 66, p. 136–145, 2018.

[151] A. Alaimo *et al.*, “Aircraft pilots workload analysis : heart rate variability objective measures and nasa-task load index subjective evaluation,” *Aerospace*, vol. 7, n°. 9, p. 137, 2020.

[152] M. R. Endsley, “Toward a Theory of Situation Awareness in Dynamic Systems,” *Human Factors : The Journal of the Human Factors and Ergonomics Society*, vol. 37, n°. 1, p. 32–64, mars 1995. [En ligne]. Disponible : http://journals.sagepub.com/doi/10.1518/001872095779049543

[153] P. M. Salmon *et al.*, "Measuring situation awareness in complex systems : Comparison of measures study," *International Journal of Industrial Ergonomics*, vol. 39, n°. 3, p. 490–500, 2009.

[154] J. Z. Bakdash *et al.*, "The validity of situation awareness for performance : a meta-analysis," *Theoretical Issues in Ergonomics Science*, p. 1–24, 2021.

[155] C. A. Bolstad *et al.*, "Measurement of situation awareness for automobile technologies of the future," *Performance metrics for assessing driver distraction : the quest for improved road safety*, p. 195–213, 2010.

[156] M. R. Endsley, *Designing for situation awareness : An approach to user-centered design*. CRC press, 2013.

[157] ——, "A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness : A Comparison of SAGAT and SPAM," *Human Factors : The Journal of the Human Factors and Ergonomics Society*, p. 001872081987537, sept. 2019.

[158] A. R. Pritchett, R. J. Hansman et E. N. Johnson, "Use of testable responses for performance-based measurement of situation awareness," *Situation awareness analysis and measurement*, p. 189–209, 2000.

[159] M. R. Endsley, "The Divergence of Objective and Subjective Situation Awareness : A Meta-Analysis," *Journal of Cognitive Engineering and Decision Making*, p. 20, sept. 2019.

[160] ——, "Direct Measurement of Situation Awareness : Validity and Use of SAGAT," dans *Situational Awareness*, 1er éd., E. Salas, édit. Routledge, 2000, p. 129–156. [En ligne]. Disponible : https://www.taylorfrancis.com/books/9781351548564/chapters/10.4324/9781315087924-9

[161] M. Endsley, "Situation Awareness in Aviation Systems," dans *Handbook of Aviation Human Factors, Second Edition*, J. Wise, V. Hopkin et D. Garland, édit. CRC Press, déc. 2009, p. 12–1–12–22.

[162] R. M. S. Clifford *et al.*, "The Effect of Immersive Displays on Situation Awareness in Virtual Environments for Aerial Firefighting Air Attack Supervisor Training," dans *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, mars 2018, p. 1–2.

[163] U. Laptaned, "Situation awareness in virtual environments : A theoretical model and investigation with different interface designs," *UTCC Engineering Research Papers*, 2006.

[164] D. Jung, S. Jo et R. Myung, "A study of relationships between situation awareness and presence that affect performance on a handheld game console," dans *Proceedings of the*

2008 International Conference on Advances in Computer Entertainment Technology, 2008, p. 240–243.

[165] C. J. Ferguson, "An effect size primer : a guide for clinicians and researchers." *Professional Psychology : Research and Practice*, vol. 40, n°. 5, p. 532–538, 2009.

[166] R. M. Clifford *et al.*, "The effects of multi-sensory aerial firefighting training in virtual reality on situational awareness, workload, and presence," dans *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2020, p. 93–100.

[167] J. Mütterlein, "The three pillars of virtual reality ? investigating the roles of immersion, presence, and interactivity," dans *Proceedings of the 51st Hawaii international conference on system sciences*, 2018.

[168] UQO Cyberpsychology Lab, "Presence questionnaire," Université du Québec en Outaouais, 2013. [En ligne]. Disponible : https://marketinginvolvement.files.wordpress.com/2013/12/pq-presence-questionnaire.pdf

[169] B. G. Witmer, C. J. Jerome et M. J. Singer, "The factor structure of the presence questionnaire," *Presence : Teleoperators & Virtual Environments*, vol. 14, n°. 3, p. 298–312, 2005.

[170] M. Slater *et al.*, "Measuring presence : A response to the witmer and singer presence questionnaire," *Presence : teleoperators and virtual environments*, vol. 8, n°. 5, p. 560–565, 1999.

[171] R. Skarbez, F. Brooks et M. Whitton, "Immersion and coherence : Research agenda and early results," *IEEE transactions on Visualization and Computer Graphics*, 2020.

[172] S. Rubio *et al.*, "Evaluation of Subjective Mental Workload : A Comparison of SWAT, NASA-TLX, and Workload Profile Methods," *Applied Psychology*, vol. 53, n°. 1, p. 61–86, janv. 2004.

[173] R. A. Grier, "How high is high ? a meta-analysis of nasa-tlx global workload scores," dans *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 59, n°. 1. SAGE Publications Sage CA : Los Angeles, CA, 2015, p. 1727–1731.

[174] F. Shaffer et J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in public health*, p. 258, 2017.

[175] M. Peltola, "Role of editing of rr intervals in the analysis of heart rate variability," *Frontiers in physiology*, vol. 3, p. 148, 2012.

[176] C. A. Bolstad, "Age-related factors affecting the perception of essential information during risky driving situations," dans *Human Performance, Situation Awareness, and Automation Conference, Savannah, GA*, 2000.

[177] M. Endsley, “Situation awareness measurement : How to measure situation awareness in individuals and teams,” *Human Factors and Ergonomics Society*, 2021.

[178] M. Oberhauser et D. Dreyer, “A virtual reality flight simulator for human factors engineering,” *Cognition, Technology & Work*, vol. 19, n°. 2, p. 263–277, 2017.

[179] M. Safi, J. Chung et P. Pradhan, “Review of augmented reality in aerospace industry,” *Aircraft Engineering and Aerospace Technology*, vol. 91, n°. 9, p. 1187–1194, 2019.

[180] H. J. Gardner et M. A. Martin, “Analyzing ordinal scales in studies of virtual environments : Likert or lump it !” *Presence : Teleoperators and Virtual Environments*, vol. 16, n°. 4, p. 439–446, 2007.

[181] O. Clausen *et al.*, “Towards predictive virtual prototyping : Color calibration of consumer VR HMDs,” dans *16th GI AR/VR Workshop*, 2019.

[182] I. S. MacKenzie, *Human-Computer Interaction : An Empirical Research Perspective*, 1er éd. San Francisco, CA, USA : Morgan Kaufmann, 2013.

[183] R. W. Soukoreff et I. S. MacKenzie, “Towards a standard for pointing device evaluation, perspectives on 27 years of fitts’ law research in hci,” *International journal of human-computer studies*, vol. 61, n°. 6, p. 751–789, 2004.

[184] International Organization for Standardization, “Ergonomics of human-system interaction — part 411 : Evaluation methods for the design of physical input devices,” International Organization for Standardization, Geneva, CH, Standard ISO/TS 9241-411 :2012, mai 2012.

[185] R. M. Herrick, “Psychophysical methodology : Comparisons within the method of limits,” *Perceptual and Motor Skills*, vol. 28, n°. 2, p. 503–514, 1969.

[186] K. Mania *et al.*, “Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity,” dans *Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, 2004, p. 39–47.

[187] S. R. Ellis *et al.*, “Generalizeability of latency detection in a variety of virtual environments,” dans *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 48, n°. 23. SAGE Publications Sage CA : Los Angeles, CA, 2004, p. 2632–2636.

[188] M. Hébert-Lavoie, P. Doyon-Poulin et B. Ozell, “Identification of visual functional thresholds for immersion assessment in virtual reality,” *PRESENCE : Virtual and Augmented Reality*, p. 1–44, 2022.

[189] K. Grinyer et R. J. Teather, “Effects of field of view on dynamic out-of-view target search in virtual reality,” dans *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2022, p. 139–148.

[190] R. Shi *et al.*, "Virtual reality sickness mitigation methods : A comparative study in a racing game," *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, vol. 4, n°. 1, p. 1–16, 2021.

[191] Y. Ling *et al.*, "A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders," *PloS one*, vol. 9, n°. 5, p. e96144, 2014.

[192] G. Riva *et al.*, "Affective interactions using virtual reality : the link between presence and emotions," *Cyberpsychology & behavior*, vol. 10, n°. 1, p. 45–56, 2007.

[193] M. Umair *et al.*, "Hrv and stress : a mixed-methods approach for comparison of wearable heart rate sensors for biofeedback," *IEEE Access*, vol. 9, p. 14 005–14 024, 2021.

[194] M. Hébert-Lavoie, P. Doyon-Poulin et B. Ozell, "The relationship between immersion and psychophysiological indicators," p. 1–53, 2022, manuscript submitted for publication.

[195] T. Willans, "Spatial presence in virtual worlds as a perceptual emotion : An expansion on cognitive feeling ?" dans *2012 Sixth International Conference on Complex, Intelligent, and Software Intensive Systems*. IEEE, 2012, p. 899–904.

[196] M. R. Endsley, "Situation awareness misconceptions and misunderstandings," *Journal of Cognitive Engineering and Decision Making*, vol. 9, n°. 1, p. 4–32, 2015.