

THESE DE DOCTORAT DE

L'UNIVERSITE D'ANGERS
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Informatique*

Par

Eulalie VERHULST

**Contribution de l'étude de l'interaction en environnement virtuel : intérêt
de la charge mentale**

Thèse présentée et soutenue à Angers, le 14 Décembre 2018

Unité de recherche : Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS - EA7315)

Rapporteurs avant soutenance :

Professeur Grigore Burdea, Université de Rutgers
Professeur Bernard N'Kaoua, Université de Bordeaux

Composition du Jury :

Président : Professeur Sébastien George, Le Mans Université

Professeur Dominique Bechmann, Université de Strasbourg
Indira Thouvenin Maître de conférences, Université de
Compiègne
Professeur Bernard N'Kaoua, Université de Bordeaux

Directeur de thèse
Paul Richard Maître de conférences, Université d'Angers

Co-directeur de thèse
Professeur Philippe Allain, Université d'Angers
Co-encadrant de thèse
Professeur Frédéric Banville, Université du Québec à Rimouski

Remerciements

Ce manuscrit est le fruit de trois années de travail qui ont pu aboutir grâce à l'aide et à la présence de mes encadrants de thèse. Je souhaite remercier mes directeurs Paul Richard et Philippe Allain ainsi que mon co-encadrant Frédéric Banville pour leur soutien au long de cette thèse.

Merci à Pierre Chauvet pour son investissement dans le développement de l'application pour traiter les données cardiaques. Ainsi qu'aux quatre étudiants qui ont travaillé sur le projet.

Je remercie les personnes qui ont accepté d'être rapporteur et membres du jury, à savoir : Grigore Burdea, Bernard N'Kaoua, Sébastien George, Dominique Bechmann et Indira Thouvenin. Merci de l'intérêt que vous portez à ces travaux, pour votre présence lors de la soutenance ainsi que pour la qualité des échanges que nous avons formulés.

Je remercie ma famille et mes amis.

Le mot de la fin devrait être destiné à ceux sans qui l'aboutissement de la thèse n'aurait pas été possible. Je remercie donc chacun des participants qui ont donné de leur temps en participant au protocole de thèse.

Table des matières

1	Introduction générale	17
2	Interaction en environnement virtuel	23
2.1	Introduction	23
2.2	Définition et classification des techniques d'interaction	25
2.2.1	La navigation en environnement virtuel	25
2.2.2	Sélection et manipulation d'objets virtuels	29
2.2.3	Classification des approches proposées	31
2.3	Evaluation des techniques d'interaction	33
2.3.1	Techniques de navigation	33
2.3.2	Techniques de sélection	35
2.3.3	Utilisabilité et heuristique	37
2.4	Analyses des techniques d'interaction	40
2.4.1	Peut-on privilégier une technique d'interaction ?	40
2.4.2	Choix des techniques d'interaction pour le VMT	41
2.4.3	Comparaison des techniques	45
2.5	Synthèse	55
3	La charge mentale en environnement virtuel	57
3.1	Introduction	57
3.2	Définition et approches théoriques	58
3.2.1	Définition	58
3.2.2	Théorie des ressources cognitives limitées (Kahneman, 1973)	59
3.2.3	Théorie de la charge cognitive (Paas & Van Merriënboer, 1994)	59
3.2.4	Cadre théorique proposé	60
3.3	Evaluation de la charge mentale	61
3.3.1	Les mesures subjectives	62

3.3.2	Les mesures physiologiques	64
3.3.3	Les mesures comportementales et les index de performance	66
3.3.4	Les mesures basées sur les données	67
3.4	Estimation de la charge mentale	69
3.4.1	La charge mentale associée à la tâche	69
3.4.2	La charge mentale associée aux facteurs environnementaux	72
3.4.3	La charge mentale associée à l'utilisateur	75
3.5	Synthèse	83
4	Evaluation de la performance lors d'IADL	85
4.1	Introduction	85
4.2	Sélection des variables pour la qualification de la performance	86
4.2.1	Sélection des objets d'étude	86
4.2.2	Comparaison de la performance à une même tâche en réel et en virtuel	92
4.2.3	Limites de l'approche	94
4.3	Variables utilisées dans les EVs lors d'IADL	96
4.3.1	Temps de réalisation	96
4.3.2	Types d'erreurs	97
4.3.3	Autres variables dépendantes	101
4.4	Interaction entre processus cognitif et utilisabilité de la technique	103
4.4.1	Problématique	103
4.4.2	Profil du participant	104
4.4.3	Intégration de variables liées à la technique d'interaction	106
4.5	Synthèse	108
5	Cas d'application : le <i>Virtual Multitasking Test</i>	109
5.1	Introduction	109
5.2	Spécificités du VMT	110
5.3	Implémentation des tâches dans le VMT	114
5.3.1	Contexte théorique lié aux tâches	114
5.3.2	Présentation des scénari	116
5.3.3	Enregistrement des variables	120
5.4	Implémentation des interactions dans le VMT	121
5.4.1	Approche proposée	121

5.4.2	Présentation des techniques d'interaction	122
5.4.3	Enregistrement des variables	124
5.5	Implémentation du rapport d'analyses	126
5.6	Synthèse	127
6	Méthode	129
6.1	Hypothèses	129
6.2	Outils pour le recueil de données	131
6.3	Procédure expérimentale	132
6.4	Dispositifs d'interaction	134
6.5	Participants à l'étude	136
7	Résultats	139
7.1	Présentation de la base de données	139
7.2	Elaboration des groupes d'individus	140
7.2.1	Expertise du participant	140
7.2.2	Elaboration du profil de charge mentale	146
7.3	Utilisation de la technique d'interaction	147
7.3.1	Répartition de l'utilisation des techniques d'interaction : Navigation . . .	147
7.3.2	Score au <i>Throughput</i>	149
7.4	Analyses des questionnaires	150
7.4.1	Score au NASA-TLX	150
7.4.2	Score au System Usability Scale	151
7.4.3	Score de présence	152
7.4.4	Score de cybermalaise	152
7.5	Analyse de la performance au VMT	154
7.5.1	Phase de familiarisation	154
7.5.2	Temps de réalisation du VMT	155
7.5.3	Score au VMT	155
7.5.4	Autres variables	157
7.5.5	Modèle de régression	159
7.6	Analyse des données cardiaques	162
8	Discussion	165
8.1	Retour sur les résultats	165

8.1.1	Modulation de la charge mentale en fonction de l'environnement virtuel .	165
8.1.2	Modulation de la charge mentale selon le profil de l'utilisateur	167
8.1.3	Modulation de la charge mentale en fonction de la difficulté de la tâche .	168
8.1.4	Impact de la charge mentale sur la performance	169
8.2	Discussion générale : à quoi est due la charge mentale?	170
8.3	Application à un contexte d'évaluation neuropsychologique	175
8.3.1	Ecologie de la tâche	175
8.3.2	Quelle technique d'interaction utiliser?	177
8.3.3	Quelles variables privilégier?	180
8.4	Limites	182
8.5	Synthèse	184
9	Conclusion et perspectives	187
9.1	Conclusions générales	187
9.2	Perspectives envisagées	189
	Bibliographie	190
A	Annexe : Journal d'information issu du VMT	209
B	Annexe : Questionnaire Général	211
C	Annexe : <i>System Usability Scale</i> - SUS	215
D	Annexe : <i>Presence Questionnaire</i> - PQ	217
E	Annexe : NASA-TLX	223
F	Annexe : <i>Simulator Sickness Questionnaire</i> - (SSQ)	225
G	Annexe : Recette de cuisine	227

Table des figures

2.1	Modèle des $3I^2$ de Fuchs et Moreau (2006).	24
2.2	Illustration d'une téléportation (Frommel, Sonntag, & Weber, 2017).	28
2.3	Illustration des six Degrés de liberté (DDL).	28
2.4	Illustration d'une sélection en 2D et en 3D.	29
2.5	Illustration de la technique du rayon.	30
2.6	Taxinomie des techniques de navigation tiré de Tan, Robertson, et Czerwinski (2001) et traduit en français.	34
2.7	Visualisation de la navigation (Chittaro, Ranon, & Ieronutti, 2006) avec le tracé suivi par tous les utilisateurs (a) et visualisation de la durée totale passée par endroits (b).	36
2.8	Tâche d'évaluation d'une technique de sélection selon les normes ISO 9241-9. . .	36
2.9	Taxinomie des Interaction homme-machine (IHM), centrée sur l'utilisabilité tiré de Pincioli (2016).	38
2.10	Illustration d'un EV où le <i>steering</i> n'est pas autorisé via écran tactile (Okahashi et al., 2013).	45
3.1	Différentes mesures issues d'un Électrocardiogramme (ECG).	65
4.1	Algorithme pour le dépistage des troubles cognitifs légers en fonction du temps mis pour réaliser la tâche (Zygouris et al., 2015).	96
5.1	Illustration du VMT.	109
5.2	Cadre d'application du VMT.	110
5.3	Exemple de données enregistrées par le VMT : itinéraire du participant et photo de la table.	113
5.4	Scénario "facile" avec une charge mentale faible.	118
5.5	Modélisation du scénario avec une charge mentale élevée	119

5.6	Avatar dans le VMT.	122
5.7	Illustration du surlignage des items interactifs dans le VMT.	123
5.8	Illustration d'un objet 3D déposé dans la main de l'avatar pour la souris et le Razer Hydra.	123
5.9	Utilisation des techniques d'interaction proposées et développées dans le VMT.	124
5.12	Comparer les techniques d'interaction en fonction d'une situation immersive ou du non-immersive.	124
5.10	Comparer les techniques d'interaction en fonction d'une sélection en 2D et en 3D.	125
5.11	Comparer les techniques d'interaction en fonction d'une rotation de la caméra avec les mouvements de la main ou avec les mouvements de la tête.	126
5.13	Illustration d'une partie du rapport d'analyses.	128
6.1	Capture d'écran de la tâche de sélection.	132
6.2	Illustration de la récolte des données cardiaques.	133
6.3	Procédure de l'expérimentation. Les techniques d'interaction sont numérotées de 1 à 5. 1 : souris/clavier. 2 : gamepad. 3 : Razer Hydra. 4 : Razer Hydra et <i>tracking</i> de tête. 5 : Razer Hydra et HMD.	135
6.4	Répartition des participants par conditions et par techniques d'interaction.	136
7.1	Dendogramme des participants ayant utilisés la souris.	141
7.2	Dendogramme des participants ayant utilisés le gamepad.	142
7.3	Dendogramme des participants ayant utilisés le Razer Hydra.	143
7.4	Dendogramme des participants ayant utilisés le Razer Hydra et le <i>Head mounted display</i> (HMD).	144
7.5	Dendogramme des participants ayant utilisés le Razer Hydra et le <i>head tracking</i>	144
7.6	Score au <i>throughput</i> en fonction du degré d'expertise des participants.	146
7.7	Répartition des techniques d'interaction pour les participants <i>ChargeMentale_{Low}</i> et <i>ChargeMentale_{High}</i>	148
7.8	Utilisation de l' <i>input</i> de locomotion en fonction de la technique d'interaction.	149
7.9	Score au <i>throughput</i> en fonction de la technique d'interaction.	150
7.10	Score au SUS en fonction de la technique d'interaction.	151
7.11	Temps de réalisation de la phase de familiarisation en fonction de la technique d'interaction.	155
7.12	Temps de réalisation du VMT en fonction des conditions de charge mentale et des techniques d'interaction.	156

7.13	Score au VMT en fonction des conditions de charge mentale et des techniques d'interaction.	157
7.14	Score à la tâche du poulet (v1), de l'épicerie (v2) et de la table (v3) pour la condition charge mentale faible.	158
7.15	Score à la tâche du poulet (v1), de l'épicerie (v2) et de la table (v3) pour la condition charge mentale élevée.	159
7.16	Distance parcourue dans le VMT en fonction des techniques d'interaction.	160
7.17	Graphiques de validation du modèle de régression.	161
7.18	Graphiques de validation du modèle de régression.	161
7.19	Graphiques de validation du modèle de régression.	162
7.20	Graphiques de validation du modèle de régression.	163
A.1	Exemple de journal d'informations issu du VMT.	209
G.1	Recette de la tâche de poulet.	227

Liste des tableaux

2.1	Avantages et inconvénients de l'interaction par souris, écran tactile et interface tangible tiré de Besançon, Issartel, Ammi, et Isenberg (2016) et traduit en français.	42
2.2	Etudes comparant les techniques d'interaction.	48
2.2	Etudes comparant les techniques d'interaction.	49
2.2	Etudes comparant les techniques d'interaction.	50
2.2	Etudes comparant les techniques d'interaction.	51
2.2	Etudes comparant les techniques d'interaction.	52
2.2	Etudes comparant les techniques d'interaction.	53
2.2	Etudes comparant les techniques d'interaction.	54
3.1	Etudes évaluant la charge mentale.	78
3.1	Etudes évaluant la charge mentale.	79
3.1	Etudes évaluant la charge mentale.	80
3.1	Etudes évaluant la charge mentale.	81
3.1	Etudes évaluant la charge mentale.	82
4.1	Etudes évaluant les IADL en virtuel.	91
5.1	Actions possibles dans le VMT en fonction des objets.	112
5.2	Présentation des tâches du VMT.	113
5.3	Types de tâches dans le VMT.	116
5.4	Verbatims des coups de téléphone annonçant une tâche dans le VMT.	117
7.1	Constitution des groupes d'experts ou de novices en fonction des techniques d'interaction.	145
7.2	Variables différenciant un groupe de novices et d'experts en fonction des techniques d'interaction.	145

7.3	Performance au VMT et score au <i>NASA Task Load Index</i> (NASA-TLX) en fonction des groupes de profil de charge mentale.	147
7.4	Score au <i>throughput</i> en fonction des techniques d'interaction.	149

Glossaire

- ACP** Analyse en composante principale. 67
- ALCog** Aides Logicielles Cognitives. 31
- CLT** *Cognitive Load Theory*. 59, 60
- CRISP-DM** *Cross Industry Standard Process for Data Mining*. 68
- DDL** Degrés de liberté. 9, 27–29, 31, 42, 178, 185
- ECG** Électrocardiogramme. 9, 62, 64, 65, 67
- EEG** Électroencéphalogramme. 65, 66, 68, 175
- EV** Environnement virtuel. 18, 21, 23–33, 35, 38–41, 43, 55, 61, 69, 72–75, 77, 85–87, 90, 92, 96, 98–106, 109, 110, 112, 113, 127–131, 133, 135, 168, 170–173, 181–184, 189, 190
- HMD** *Head mounted display*. 10, 45, 46, 88, 89, 139, 144, 154, 189
- HR** *Heart rate*. 64
- HRV** *heart rate variability*. 64, 65, 140, 145, 162
- IADL** *Instrumental activities of daily living*. 18, 21, 86, 87, 90, 97, 171, 172, 184, 189, 190
- IHM** Interaction homme-machine. 9, 38, 41, 57, 73, 93
- IRM** Imagerie par résonance magnétique. 17
- LRS** *Landmark, Route, Survey*. 26, 33
- MCI** *Mild cognitive impairment*. 46, 54, 172
- MET** *Multiple Errands Test*. 18
- MMSE** *Mini Mental State Examination*. 54, 171
- MP** Mémoire prospective. 109, 110, 114, 170

- NASA-TLX** *NASA Task Load Index*. 14, 62, 63, 67, 68, 73, 74, 146, 147, 150, 151, 165–169, 178
- PASAT** *Paced Auditory Serial Addition Test*. 66
- RazerHMD** Razer Hydra et visiocasque. 136, 165–167, 172, 178, 179, 182, 188
- RazerHT** Razer Hydra et *tracking* de tête effectué avec la Kinect. 136, 137, 139, 148–151, 153–155, 157–159, 165–167, 178, 179, 182, 188
- RV** Réalité virtuelle. 18, 19, 23, 26, 30, 31, 41, 66, 72, 74, 77, 85, 90–93, 96, 99, 103, 104, 115, 182, 183, 189
- S-SWAT** *Simplified-Subjective Workload Assessment Technique*. 62
- SUS** *Systeme Usability Scale*. 131, 151, 172, 177, 182
- SVM** *support vector machine*. 68
- SWAT** *The Subjective Workload Assessment Technique*. 62–64
- VMT** Virtual Multitasking Test. 21, 109, 110, 130, 166, 167, 169, 173, 177, 181
- WP** *Workload profile*. 62–64, 68

Introduction générale

L'évaluation neuropsychologique a pour objectif de rendre compte des liens entre des comportements observables et le fonctionnement du cerveau (Lezak, 1995). T. D. Parsons (2015) expose l'évolution des mesures neuropsychologiques en reprenant les formulations de Bilder (2011). Les auteurs rapportent trois phases nommées respectivement :

- « Neuropsychologie 1.0 » (1950-1979). Cette période est centrée sur l'analyse clinique de patients cérébrolésés à l'aide de tests neuropsychologiques. Ceux-ci apportent une aide au diagnostic de la pathologie cérébrale ainsi qu'à la localisation d'une éventuelle lésion (Long & Kibby, 1995).
- « Neuropsychologie 2.0 » (1980- à nos jours). Cette période utilise les nouvelles technologies, comme les techniques d'Imagerie par résonance magnétique (IRM) pour faire la distinction entre les fonctions cognitives atteintes et préservées des patients cérébrolésés.
- « Neuropsychologie 3.0 » (de nos jours). Cette période représente le futur envisagé où les mesures neuropsychologiques sont validement écologiques. C'est-à-dire que les performances du patient cérébrolésé à un test neuropsychologique sont représentatives de ce qu'il aurait produit dans une situation similaire dans sa vie quotidienne.

Une tâche écologiquement valide est une tâche dans laquelle la performance du sujet est proche de celle qu'il peut produire dans la vie quotidienne (Chaytor & Schmitter-Edgecombe, 2003). Pour Franzen et Wilhelm (1996), un outil neuropsychologique est considéré écologiquement valide, s'il répond aux deux exigences suivantes : une exigence de verisimilitude et une exigence de véridicité. La verisimilitude désigne le fait que la tâche neuropsychologique, utilisée pour l'évaluation cognitive du patient, demande autant de ressources cognitives qu'une action de la vie quotidienne. Alors que la véridicité désigne le fait que la performance du patient cérébrolésé à la tâche neuropsychologique doit refléter celle qu'il aurait obtenu en condition réelle.

L'évaluation neuropsychologique doit donc être effectuée dans des situations proches de la vie réelle et éloignée des contextes artificiels de laboratoires. En effet, les tests papier-crayons et/ou leurs versions numériques ont un degré écologique modéré (Chaytor & Schmitter-Edgecombe, 2003) et des patients peuvent avoir une performance tout à fait satisfaisante aux tests papiers crayons, mais présenter des difficultés dans des tâches de la vie quotidienne (Lewis, Babbage, & Leathem, 2011).

Par exemple, des patients ne sont pas déficitaires à des tests neuropsychologiques qui évaluent les fonctions exécutives mais ont des difficultés pour réaliser le *Multiple Errands Test* (MET) (Shallice & Burgess, 1991) qui est un test écologique qui demande aux patients d'effectuer des achats dans un quartier qu'ils ne connaissent pas. Ces patients qui ne présentent donc pas de troubles apparents lors de l'évaluation neuropsychologique avec des tests classiques se montrent pourtant en difficulté dans des situations de la vie quotidienne comme celle de faire ses courses. Plus généralement, des auteurs observent que les tests neuropsychologiques classiques sont peu prédictifs des performances des patients lors des activités de la vie quotidienne (Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Burgess et al., 2006). Il est donc primordial d'évaluer les patients lors d'activité de la vie quotidienne comme les *Instrumental activities of daily living* (IADL) de sorte à ce que les résultats obtenus lors de l'évaluation puissent être représentatif de l'adaptabilité et de l'autonomie des patients dans leur vie quotidienne. Bien que des questionnaires existent pour évaluer les IADL, leur expertise n'est pas pleinement satisfaisante car des patients ont tendance à sous ou surestimer leurs capacités à performer des IADL. De plus, évaluer les patients dans des conditions réelles demande de fortes ressources humaines pour assurer le bon suivi de l'expérimentation et génère un manque de standardisation des tâches. En effet, il est difficile d'évaluer de façon équivalente des personnes dans un supermarché car les expérimentateurs ne peuvent pas contrôler le nombre de clients ou les événements imprévus qui pourraient s'y produire.

Les auteurs proposent alors d'utiliser d'autres techniques d'expérimentation et de mesure, comme la Réalité virtuelle (RV). En effet, les Environnements virtuels (EV)s peuvent avoir une plus grande validité écologique que les tests papiers crayons car ils peuvent simuler la réalité (Rizzo, Schultheis, Kerns, & Mateer, 2004).

La RV est une technologie qui simule la réalité à l'aide d'un logiciel recréant les sensations d'une situation qui n'est pas matérialisée (Burdea & Coiffet, 1993). L'utilisateur peut alors interagir avec les objets qui composent la scène virtuelle (Bowman & Hodges, 1997). La RV peut être définie comme une simulation informatique du monde réel dans le sens où elle fait

intervenir les mêmes sens et ressentis physiques et psychologiques que dans la réalité. En effet, une personne souffrant d'arachnophobie a les mêmes réactions de peur devant une araignée réelle qu'une araignée virtuelle. Ainsi elle est affectée par l'environnement et les stimuli virtuels qu'il contient même si elle a conscience que la situation n'est pas réelle (Gregg & Tarrier, 2007).

Le concept de RV est parfois confondu avec celui d'EV jusqu'à être considéré par certains comme un synonyme (Stanney, 2002). Un EV désigne un environnement basé sur le monde réel ou imaginaire dans lequel l'utilisateur peut naviguer et interagir par le biais d'une technique d'interaction (Sayers, 2004; Stanney, 2002). Il y a trois types d'EVs : les immersifs, les semi-immersifs et les non immersifs (Bowman, Kruijff, LaViola Jr, & Poupyrev, 2001). On parle d'un EV immersif lorsque les sens sont largement mobilisés, où l'utilisateur interagit en première personne. Les EVs semi-immersifs permettent à l'utilisateur d'interagir à la fois dans l'environnement réel et virtuel. Un EV non-immersif est affiché par le biais d'un matériel *desktop*, comme un écran d'ordinateur.

A notre sens, une distinction doit être faite entre les deux termes, car si les environnements de RV sont des EV, tous les EVs ne sont pas de la RV. Il est nécessaire de relever que, contrairement à certain EV, la RV a pour ambition de solliciter les sens de l'utilisateur de sorte à ce qu'il agisse de façon similaire en virtuel et en réel (Lee, Kim, Suh, & Koh, 2016). Pour ce faire, les concepteurs d'environnements de RV font appel à des caractéristiques immersives comme l'utilisation d'un affichage en stéréoscopie. Tandis que dans le cas de la psychologie, il n'y a pas forcément de distinction entre la RV et les EVs. Mais les deux termes renvoient à une simulation du monde dans laquelle un individu peut s'entraîner ou apprendre en réalisant un exercice autant de fois que nécessaire (Chittaro & Ranon, 2007). Ainsi, dans le cadre de nos travaux, pour apporter une meilleure précision, nous n'emploierons le terme RV que lorsque le système comprend des caractéristiques immersives.

Les EVs et les systèmes de RV permettent de repenser des situations d'évaluation des fonctions cognitives au cours de situations écologiques. En effet, l'évaluation comportementale des participants au cours de la réalisation d'IADL en réel est chronophage, difficile à mettre en place par la mobilisation de ressources humaines (e.g. différents expérimentateurs) et matérielles (e.g. utilisation d'aliments périssables) et limitée dans la précision des variables enregistrées. Par exemple le temps de complétion peut être enregistré à la main, avec l'usage d'un chronomètre, ou en extrayant l'information du film retraçant la performance du participant. L'analyse comportementale du participant est fastidieuse puisqu'elle demande à ce qu'un expérimentateur formé observe et annote les vidéos une par une. L'utilisation des EV permet de proposer des

situations issues de la vie quotidienne sans condition de ressources matérielles. Il devient donc possible de créer des scénarios de tests difficilement réalisables en réel comme ceux prenant position dans des supermarchés ou des appartements virtuels et de les doter d'objets 3D avec lesquels interagir sans avoir à penser à leur durée de vie. Les tests de RV permettent également de faciliter le travail de l'expérimentateur car les données recueillies comme des temps de réaction précis sont directement insérées dans une base de données informatique et sont prêtes à être traitées par l'expérimentateur (Lee, Ku, Kim, Kim et al., 2003 ; Cromby, Standen et Brown, 1996). En outre les tests de RV n'ont pas besoin d'être administrés par un expérimentateur formé car le participant, après avoir été entraîné dans un EV à naviguer et à sélectionner des objets, peut se débrouiller seul dans l'EV dont la passation des épreuves naturelles est standardisée par le système informatique. Ainsi le biais potentiel dû à l'administrateur que l'on peut trouver dans les tests classiques est fortement réduit. De plus, simuler en virtuel la vie réelle permet de proposer aux participants d'être évalués de façon sécuritaire lors de tâches pouvant être répétées et/ou diversifiées (Bellani, Fornasari, Chittaro et Brambilla, 2011).

L'éloge des outils de RV n'est donc plus à faire. Cependant de nombreuses questions demeurent dans le cas de la RV comme outil d'évaluation cognitive ou de dépistage cognitif. Ces préoccupations psychométriques apparaissent fondamentales sur le plan de la validité de la mesure. En effet, la performance obtenue en virtuelle n'est pas aisée à appréhender et il n'y a pas de mesures étalonnées pour savoir ce qui est ou pas une bonne performance. De plus, lorsque l'on compare les performances en réel et en virtuel à une même tâche, les résultats peuvent s'avérer contradictoires. Par exemple, des participants font plus d'erreurs et mettent plus de temps pour préparer une tasse de café en virtuel qu'en réel (Allain et al., 2014). Alors qu'à une autre tâche de cuisine où des participants préparent un plat de spaghetti, ils réalisent la recette plus rapidement en virtuel qu'en réel et le nombre d'erreurs n'est pas significativement différent entre les 2 conditions (Vallejo et al., 2017). Les méthodes d'interaction utilisées dans l'EV pourraient venir moduler la performance du participant en générant des erreurs d'utilisabilité et en modulant la charge mentale allouée. La charge mentale est définie par l'effort mental investi par le sujet pour mener à bien une tâche (Parasuraman, 2003). Autrement dit, la charge mentale fait référence aux ressources cognitives mobilisées pour réaliser une tâche à un moment donné (Oviatt, 2006).

Ce travail de thèse a pour objectif d'étudier l'impact du système d'interaction sur la performance de l'utilisateur lors de la réalisation de tâches dans un EV. Dans ce contexte, nous

proposons une approche permettant de qualifier, de façon la plus pertinente possible, la performance du participant en prenant en compte la charge mentale allouée à la réalisation d'une tâche. La prise en compte de la charge mentale comprend les caractéristiques propres à l'individu, la difficulté de la tâche et la manière d'interagir dans un EV.

Si les EVs peuvent mesurer automatiquement des variables telles que le temps de complé-
tion, il est difficile de pouvoir interpréter objectivement la performance du participant compte
tenu de la variabilité intra-individuelle. En effet, le résultat du participant au test est dépen-
dant aussi bien de ses habiletés à interagir avec un EV que de ses capacités cognitives. Dans
un premier temps il est nécessaire de constater si la technique d'interaction et/ou l'immersion
induit une augmentation de la charge mentale du participant. Pour ce faire nous comparerons
différentes techniques d'interaction. Dans un second temps il s'agit de comprendre dans quelles
circonstances la charge mentale liée au système de RV peut être croissante au point d'entraver
la bonne réalisation d'une tâche en RV. Pour répondre à notre objectif, nous proposons une
expérimentation comprenant deux conditions pour mesurer l'impact d'un système de RV sur
la charge mentale du participant : une avec une haute charge mentale et l'autre avec une faible
charge mentale.

Le manuscrit de thèse s'organise en deux parties. Au cours de la première partie, qui com-
prend quatre chapitres, nous proposons, dans un premier temps une analyse de la littérature
concernant les environnements virtuels et traitons en particulier la problématique de l'inter-
action et de l'évaluation de la performance humaine. Le chapitre 1 définit les techniques d'in-
teraction et leurs méthodes d'évaluation. Dans ce contexte nous proposons une méthode de
sélection des techniques d'interaction que nous comparerons au cours du protocole expérimen-
tal. Le chapitre 2 définit la charge mentale et ses méthodes d'évaluation. Dans ce chapitre, nous
présentons la charge mentale comme un concept multidimensionnel qui est influencé par le profil
du participant, la difficulté de la tâche et l'environnement d'évaluation. Le chapitre 3 présente
une revue de littérature qui concerne plus particulièrement la performance d'un participant au
cours de la réalisation d'IADL dans un EV. Le chapitre 4 traite de l'EV utilisé dans le cadre
du protocole expérimental : le Virtual Multitasking Test (VMT). C'est un appartement virtuel
dans lequel nous avons implémenté différentes techniques d'interaction pour les besoins de notre
étude. A savoir : la souris, le gamepad, le Razer Hydra, le Razer Hydra avec un système de
head tracking et le Razer Hydra avec un visiocasque. Ce chapitre fait le lien entre la théorie
utilisée dans les chapitres précédents et le contexte d'application.

La seconde partie de la thèse présente la partie expérimentale pour évaluer la charge mentale lors de la réalisation d'IADL. Au cours de ce protocole, pour comprendre comment la charge mentale vient moduler la performance du participant, nous comparons cinq techniques d'interaction au cours d'une condition de charge mentale élevée et d'une condition de charge mentale faible. Les résultats obtenus sont ensuite analysés et discutés afin de comprendre si les techniques d'interaction modulent la charge mentale mentale du participant et si oui, de quelles façons. Enfin, nous concluons et dégageons quelques perspectives de recherche.

Interaction en environnement virtuel

2.1 Introduction

Le défi des professionnels de la santé qui souhaitent exploiter les EVs à des fins cliniques, par exemple pour évaluer les capacités cognitives d'un individu, est de pouvoir utiliser des techniques d'interaction 3D peu coûteuses et faciles à manipuler qui vont permettre à l'utilisateur d'interagir dans l'EV, de manière efficace et en consommant un minimum de ressources cognitives. Si la technique d'interaction remplit ces différents critères, l'utilisateur pourra se concentrer exclusivement sur la réalisation des activités au sein de l'EV car la technique d'interaction ne mobilisera pas de charge mentale (le concept de la charge mentale est développé dans le chapitre 3).

En effet, pour interagir dans un EV, l'utilisateur doit passer par l'intermédiaire d'une technique d'interaction. Une technique d'interaction est la combinaison de logiciels (*software*) et de matériels informatiques (*hardware*) permettant la réalisation d'une tâche (Tucker, 2004). Le matériel permet à l'utilisateur de transmettre à l'ordinateur des informations d'entrée (*input*) qui seront analysées et retranscrites sous forme d'informations de sortie (*output*) par le logiciel. C'est l'ensemble des deux qui va permettre à l'utilisateur d'interagir avec l'EV. Par exemple quand l'utilisateur appuie sur un bouton du gamepad, une information est transmise au logiciel. Celui-ci l'analyse puis en transmet les conséquences telle que la sélection d'un item dans l'EV.

Le modèle $3I^2$ (Fuchs & Moreau, 2006) découpe en trois niveaux les processus d'immersion et d'interaction (I_2) pour proposer une analyse des systèmes de RV (Fig.2.1). Le premier niveau, I^2 sensorimotrices, présente le fait que l'utilisateur interagi physiquement avec l'application via une technique d'interaction et donc via les stimuli perçus par l'utilisateur lors de l'interaction. Le deuxième niveau, I^2 cognitives, permet à l'utilisateur de comprendre l'EV dont il perçoit les stimuli et donc de pouvoir envisager d'y agir. La partie logicielle de ce niveau gère les

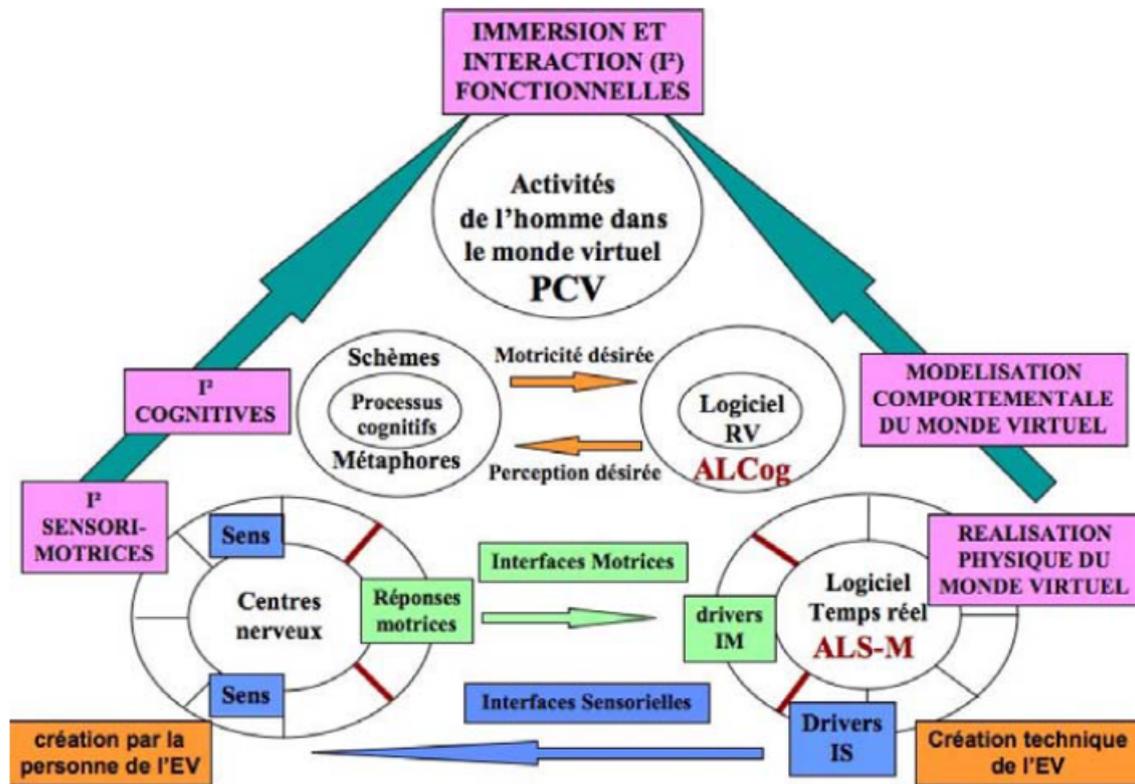


FIGURE 2.1 – Modèle des $3I^2$ de Fuchs et Moreau (2006).

comportements et les animations des objets et des événements présents dans l'EV. C'est à ce niveau que l'utilisateur prépare l'action qu'il souhaite accomplir en fonction de la compréhension qu'il a de l'EV. Ce niveau part du constat que l'utilisateur transfère des actions du monde réel au monde virtuel à l'aide de schèmes comportementaux intégrés. Ainsi, si l'EV présente des objets similaires en réel et en virtuel, l'utilisateur s'attendra à agir avec eux de la même façon dans les deux mondes. Au troisième niveau, I^2 fonctionnelles, l'utilisateur est immergé dans l'EV et peut réaliser une tâche. Lors de la conception de l'application, toutes les fonctionnalités ou actions possibles doivent être spécifiées au préalable.

Il existe une grande variété de techniques d'interaction et la plupart sont adaptées à un type de technologie. Par exemple, le stylet, qui a été mis en œuvre pour être utilisé avec des écrans tactiles, n'est pas forcément adaptable à des technologies non tactiles. Pour aider les concepteurs, les développeurs et les chercheurs à rassembler les techniques d'interaction, des taxinomies ont été créées. La multitude de taxinomies existantes pour classer les techniques d'interaction montre la réelle difficulté à regrouper efficacement les connaissances dans le domaine. Certaines vont être centrées sur la technique d'interaction dans son ensemble, alors que d'autres vont s'intéresser à un pan particulier de la technique d'interaction comme la navigation (Tan et al., 2001). Les techniques d'interaction 3D peuvent être caractérisées par trois termes :

la navigation, la sélection et la manipulation et, sur un autre registre, le contrôle (Jankowski & Hachet, 2013). Le contrôle renvoie au dialogue entre l'utilisateur et l'application, comme réaliser des modifications dans l'application via l'utilisation d'un menu (Bowman & Hodges, 1997). Dans le cas de nos travaux, nous n'allons pas aborder le contrôle de l'application puisque l'utilisateur n'est pas censé y être confronté lorsqu'il interagit dans l'EV lors d'une phase de familiarisation ou d'évaluation. Cependant, les autres concepts seront traités plus en détail.

Ce chapitre présente les techniques d'interaction utilisées dans un EV. Il commence par définir son usage (section 2.2) puis ses méthodes d'évaluation (section 2.3). Enfin, les avantages et les inconvénients liés aux techniques d'interaction seront présentés (section 2.4). Le chapitre se termine par une synthèse qui met en exergue l'importance de bien définir la technique d'interaction employée pour interagir dans un EV.

2.2 Définition et classification des techniques d'interaction

2.2.1 La navigation en environnement virtuel

La navigation est l'action qui permet au sujet de se déplacer dans un EV tout en cherchant son chemin (Stanney, 2002). Elle se définit par une composante motrice, dite de locomotion (*travel*), qui permet de se déplacer d'un point A à un point B, et par une composante cognitive (*wayfinding*) qui rend possible la création d'une carte cognitive de l'EV de sorte que l'utilisateur puisse se repérer dans l'espace (Bowman et al., 2001). La navigation dans un EV 3D nécessite que l'utilisateur navigue par le biais d'un avatar, qui le représente. Dans le cas d'une navigation à la première personne, la caméra est placée sur la tête de l'avatar et celui-ci peut être visible ou non. Dans le cas d'une navigation en troisième personne, la caméra est située derrière l'avatar de sorte que l'utilisateur voit le dos de son avatar (Denisova & Cairns, 2015). Il existe donc une vision subjective à la première personne qui permet à l'utilisateur de se sentir plus immergé dans l'EV (Denisova & Cairns, 2015) et une vision objective qui permet de voir globalement l'EV.

Composante cognitive de la navigation : *wayfinding*

Le terme de carte cognitive, introduit pour la première fois par Tolman (Tolman, 1948), désigne une représentation mentale des informations spatiales utilisées pour la navigation. Ainsi l'élaboration d'une carte cognitive d'un endroit nécessite d'acquérir et de retenir des informa-

tions spatiales à propos de ce lieu. Une fois ces informations acquises, l'individu peut alors les utiliser en tant que connaissances pour s'orienter dans l'espace. Le modèle *Landmark, Route, Survey* (LRS) (Goldin & Thorndyke, 1982 ; Siegel & White, 1975) est un modèle linéaire à trois étapes, proposé pour décrire la création d'une carte cognitive :

- *Familiarisation* : l'individu découvre l'environnement et prend en compte des points de repère différenciés les uns des autres. Ceux-ci permettent d'identifier des endroits spécifiques dans l'EV (Couclelis, Golledge, Gale, & Tobler, 1987).
- *Exposition* : après une période d'exposition dans l'environnement, l'individu peut agencer les différents points de repère entre eux dans le but de créer des trajets.
- *Configuration (survey representation, configurational knowledge)* : l'individu a une représentation de l'environnement flexible et adaptable au but. Ainsi il peut avoir une vision d'ensemble de l'environnement, localiser facilement des endroits et créer des routes alternatives entre ces endroits (Hintzman, O'Dell, & Arndt, 1981).

Composante motrice de la navigation : locomotion

La locomotion permet à l'utilisateur de se déplacer dans l'EV, tout en regardant autour de lui, sans prendre en compte les connaissances spatiales de celui-ci. Il existe plusieurs méthodes de locomotion dans l'EV. Dans la majorité des cas, elle s'effectue via une translation et une rotation de la caméra de façon continue. Mais d'autres techniques existent, comme la téléportation. Nous allons aborder dans cette sous-section différentes techniques de locomotion existantes tandis que nous évoquerons les aspects liés à la rotation de la caméra (*steering*) dans la sous-section suivante.

La navigation, dans les applications de RV, peut s'effectuer de façon réelle : l'utilisateur se déplace donc physiquement dans une zone d'interaction délimitée au préalable et ses mouvements physiques sont reproduits par l'avatar qui le représente. La marche réelle demande de prendre en compte des aspects techniques considérables. En effet, l'un des enjeux de la RV est de pouvoir retranscrire la navigation virtuelle en sachant que les navigations "réelle" et "virtuelle" n'utilisent pas les mêmes dimensions. Pour pallier cette contrainte avec la marche réelle, il est possible d'établir un EV à la même échelle que la zone de déplacement réelle dans laquelle l'utilisateur pourra se déplacer de sorte que les limites de la pièce réelle soient les mêmes que celles de l'EV (1), de proposer à l'utilisateur de marcher sur un tapis roulant omnidirectionnel

(2), ou d'utiliser des algorithmes de redirection de sorte à ce que l'utilisateur se déplace en cercle dans l'environnement réel (3).

La marche réelle est une technique de navigation intéressante, car elle est écologique, c'est-à-dire représentative d'une situation réelle, et ne demande aucun apprentissage mais qui n'est pas aisée à mettre en œuvre ; la navigation s'effectue donc plus généralement par le biais d'un contrôleur comme la manette de jeux associée au visiocasque. Un aperçu des techniques les plus courantes est présenté ci dessous :

- *Contrôleur/joystick* : la navigation est basée sur les entrées du joystick. Par exemple, pousser le manche du joystick de façon continue vers l'avant permet d'avancer. La navigation est intuitive puisqu'il s'agit de diriger le joystick vers la direction souhaitée. Il existe deux types de joystick : le joystick positionné sur un socle et les joystick positionnés sur une manette de jeux vidéo, ou gamepad (pour une description du gamepad, voir chapitre 5).
- *Simulation de la marche Walking-in-place* (Slater, Usoh, & Steed, 1995) : la navigation est basée sur la reconnaissance des mouvements du corps de l'utilisateur. L'utilisateur peut avancer dans l'EV lorsqu'il simule une marche sur place en levant ses jambes ou en effectuant un mouvement de balancier avec ses bras.
- *La téléportation* : la navigation est basée sur les entrées du contrôleur qui va indiquer la direction dans laquelle le déplacement est souhaité. La méthode consiste à autoriser l'utilisateur à se transporter automatiquement d'un endroit à un autre. Pour se faire, il pointe et sélectionne un endroit dans l'espace virtuel et y est téléporté (Fig. 2.2).
- *Les techniques hybrides* : plusieurs types de navigation sont mêlés et peuvent être utilisés de façon alternatives. Par exemple, les multiples *inputs* sur les manettes de jeux permettent de naviguer grâce à la téléportation puis avec de la translation.

L'orientation de la caméra (*steering*)

Les degrés de liberté (DDL) renvoient à la possibilité de mouvements dans l'espace en terme de rotations et de translations. Lorsque les mouvements d'un objet ne sont pas contraints, il possède 6 DDL : 3 translations et 3 rotations (Fig. 2.3). La gestion de l'orientation de la caméra n'est donc pas toujours aisée à rendre avec des techniques d'interaction 2D et le *steering*

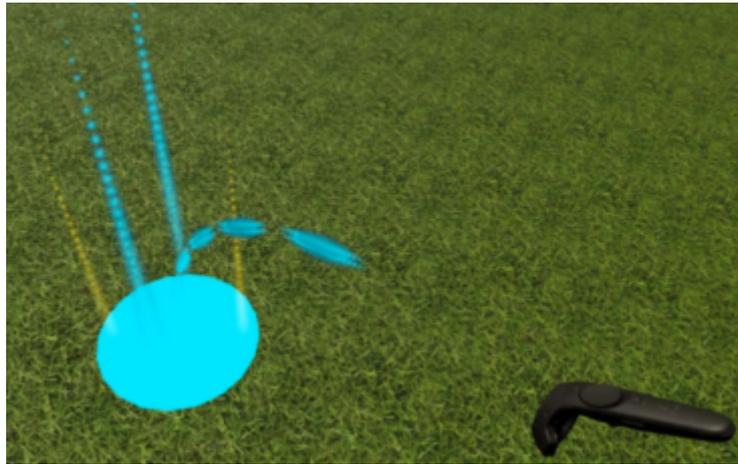


FIGURE 2.2 – Illustration d'une téléportation (Frommel et al., 2017).

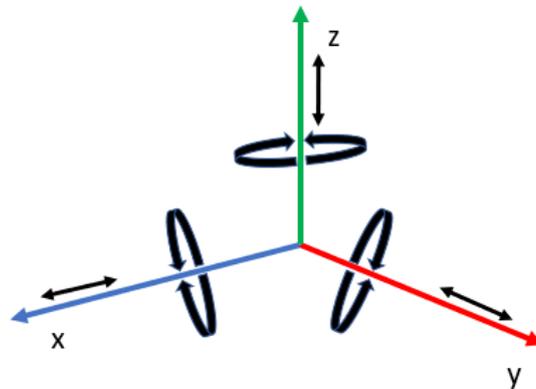


FIGURE 2.3 – Illustration des six DDL.

comporte aussi bien de l'exploration visuelle dans les grands EVs que de la précision lorsqu'il s'agit de regarder un objet qui compose la scène virtuelle (Jankowski & Hachet, 2013).

Les schèmes de contrôle les plus fréquents sont les schèmes couplés et découplés (Martel & Muldner, 2017). Les commandes transfèrent les données entrantes créées par les techniques d'interaction (*input*) en fonctionnalité de jeux (*output*) et autorisent le joueur à interagir avec l'EV en pouvant bouger l'orientation de la caméra pour faire un panorama des environs, spécifier la direction des mouvements de l'avatar (*steering*) pour naviguer et sélectionner des cibles (Martel & Muldner, 2017). Dans le schéma couplé, la caméra et les mouvements de l'avatar sont couplés de sorte que l'utilisateur bouge dans la direction que pointe la caméra. Il peut changer le point de vue de la caméra et l'orientation de l'avatar en même temps. Alors qu'avec le schéma découplé, l'utilisateur peut regarder où il veut dans l'environnement sans que cela affecte la direction des mouvements de l'avatar (Martel & Muldner, 2017). Par exemple, la commande couplée « mouselook » permet à l'utilisateur de contrôler simultanément la vue de la caméra, la sélection et l'orientation de son point de vue alors que le clavier gère les

déplacements (Martel & Muldner, 2017). La plupart des EVs destinés à l'évaluation cognitive utilisent un schéma couplé où l'orientation du point de vue et le curseur de sélection sont orientés dans la même direction que la caméra de sorte que l'utilisateur puisse se sentir présent dans l'EV.

2.2.2 Sélection et manipulation d'objets virtuels

La sélection dans un EV permet de sélectionner un item alors que la manipulation correspond aux déplacements, à la rotation et au changement d'échelle des objets 3D (Jankowski & Hachet, 2013).

Deux types de sélection existent dans les EVs : la sélection en 2D et la sélection en 3D (Fig. 2.4). La majorité des EVs non-immersifs utilise une sélection en 2D, avec 2 DDL, en permettant à l'utilisateur de mouvoir son curseur de sélection dans les directions horizontales et verticales, x et y. La sélection en 3D est majoritairement utilisée dans les EVs immersifs. Elle est plus complexe puisqu'elle requiert le contrôle de la technique d'interaction sur au moins 3 DLL, à savoir les axes x, y et z. L'axe z représentant la profondeur.



FIGURE 2.4 – Illustration d'une sélection en 2D et en 3D.

Sélectionner une cible avec une technique 2D

Deux techniques fondamentales existent pour sélectionner une cible en 2D : le pointage (*point and click*) et le glisser-déposer (*drag and drop*) (MacKenzie, Sellen, & Buxton, 1991).

- *Pointer* : l'utilisateur positionne le curseur de sélection sur la cible puis clique.
- *Glisser-déposer* : l'utilisateur sélectionne un item en pressant le bouton de la souris, déplace celui-ci jusqu'à une donnée puis relâche le bouton de la souris pour déposer l'item.

Si ces deux tâches sont utilisées de façon quotidienne par des utilisateurs d'ordinateur, elles ne font pas appel au même degré de difficulté. Les utilisateurs sont plus efficaces avec la technique de pointage que le glisser-déposer, qu'il s'agisse d'adultes (MacKenzie et al., 1991) ou d'enfants (Inkpen, 2001).

Sélectionner une cible avec une technique 3D

Les techniques de sélection utilisées dans les EVs en 2D peuvent être utilisées dans les EVs en 3D. Il faut toutefois noter que l'utilisation de visiocasques ne rend pas l'utilisation des techniques d'interaction traditionnelles, comme le clavier, confortable voir possible puisque le champ de vision de l'utilisateur ne peut pas dépasser ce qui est affiché par le casque de RV. Il n'est donc pas en mesure de voir le clavier. Pour y pallier et interagir de façon plus naturelle, les EVs, qui utilisent des technologies de RV, et ainsi reproduire le plus fidèlement possible les mêmes gestes dans l'environnement réel et en virtuel, d'autres techniques de sélection ont été développées. Deux principales métaphores de sélection sont utilisées en RV : le rayon (*ray casting*) et la main virtuelle (Argelaguet & Andujar, 2013). Ces deux techniques sont comparables aux gestes « pointer » ou « saisir » utilisés dans le monde réel. Le rayon virtuel s'étend de la main (Bowman & Hodges, 1997) (Fig. 2.5), des yeux de l'avatar, ou du curseur de sélection et autorise une sélection et/ou une manipulation lorsqu'il entre en collision avec la cible. Dans le cas de la main virtuelle, une représentation en 3D de la main du participant doit directement entrer en collision avec la cible pour autoriser une sélection et/ou une manipulation.

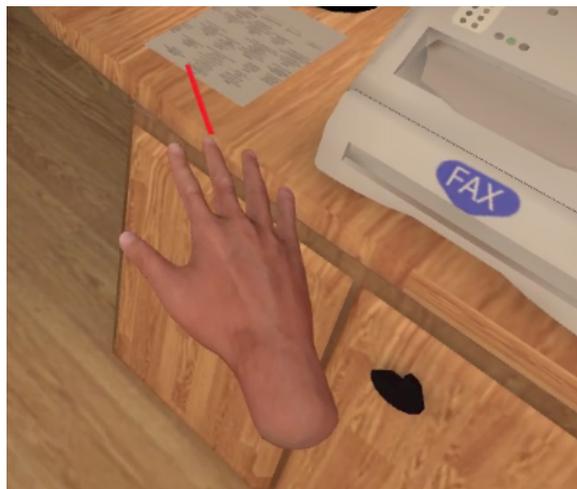


FIGURE 2.5 – Illustration de la technique du rayon.

Faciliter les interactions avec les aides logicielles cognitives.

Les aides logicielles cognitives Aides Logicielles Cognitives (ALCog) de motricité et de perception (Fuchs, 2016) permettent de faciliter les interactions entre l'utilisateur et l'application. Par exemple, comme expliquer lors de la différence entre une sélection en 2D et en 3D, la gestion des DDL va permettre de faciliter ou non les manipulations d'objets. De plus, l'ajout d'un dispositif haptique peut aussi permettre de faciliter les interactions. Ainsi, un retour haptique peut permettre à l'utilisateur de sentir lorsque deux objets entrent en collision. En effet, dans un EV, la profondeur n'est pas toujours bien perçue par les utilisateurs, il peut donc être plus difficile pour eux de détecter le point de collision entre un objet et une surface. L'ajout de dispositif haptique peut alors venir faciliter l'interaction en se combinant aux indices perceptifs. Un autre exemple d'ALC utilisé pour faciliter la navigation peut être proposé. Il s'agit de contraindre la navigation de l'utilisateur à un tracé prédéfini de sorte à ce qu'une fois que le sujet a indiqué la direction dans laquelle il souhaite se diriger, il parcourt automatiquement le chemin comme s'il naviguait sur des rails invisibles. Une illustration de cet exemple dans un magasin virtuel est proposée ultérieurement (section 2.4.2).

2.2.3 Classification des approches proposées

Interagir dans un EV nécessite d'utiliser au moins une technique d'interaction. Celles-ci ont été classées en différentes taxinomies pour faciliter le travail des concepteurs lors de la création de l'EV et permettre des évaluations comparatives. Ces taxinomies sont nombreuses. Elles peuvent être générales et prendre en compte la navigation, la sélection et la manipulation, ou bien plus restrictives en étant centrées, par exemple, seulement sur des techniques de locomotion en RV (Boletsis, 2017).

Certaines taxinomies décomposent les techniques d'interaction en sous-tâches (Bowman & Hodges, 1999). D'autres mettent en avant l'importance de la partie *hardware* et *software* que ce soit dans la gestion du ratio control/display, pour le nombre de DDL ou pour l'espace d'interaction disponible (Argelaguet & Andujar, 2013). D'autres ont classé les techniques d'interaction selon 2 grandes catégories : les techniques égocentriques et allocentriques (ou exocentriques) (Jung et al., 2014 ; Poupyrev, Ichikawa, Weghorst, & Billinghurst, 1998).

Par exemple, Tan et al. (2001) classent d'abord la navigation en fonction du type de tâche à effectuer (Fig. 2.6). La navigation peut être décomposée en 3 sous-tâches (Bowman et al., 2001 ; Tan et al., 2001) :

- *Explorer* : il s’agit de découvrir visuellement un EV, aucun autre but n’est requis.
- *Rechercher* : il s’agit de se déplacer vers une cible prédéterminée.
- *Manœuvrer* : il s’agit de réaliser une tâche spécifique en orientant sa trajectoire afin de permettre sa réalisation. Nous parlerons de manœuvre si l’objectif de la navigation est d’orienter sa trajectoire pour réaliser une tâche spécifique.

Le fractionnement de la navigation en trois tâches permet d’insister sur la nécessité de distinguer la navigation en tant que but ou en tant que moyen. Dans le premier cas elle implique une totale concentration de l’utilisateur sur la navigation alors que dans le seconde, la navigation doit être la plus aisée possible de sorte que le sujet puisse naviguer tout en mobilisant ses ressources cognitives vers la tâche à accomplir (Santos et al., 2009). Tan et al. (2001) proposent ensuite de s’intéresser au contrôle du déplacement. Le déplacement dans l’EV comprend les mouvements du point de vue, c’est-à-dire la position, l’orientation et la vitesse de la caméra virtuelle. Enfin, l’interface utilisateur est le média entre les informations d’entrée qui sont renseignées par l’utilisateur dans l’EV et les informations de sortie qui sont générées par l’EV suite aux actions de l’utilisateur. C’est grâce au contrôle du déplacement qu’il peut indiquer où et comment il souhaite se déplacer.

Les techniques egocentriques permettent à l’utilisateur d’être immergé dans l’EV et d’interagir directement dessus (Poupyrev et al., 1998). Ces sont les techniques d’interaction les plus fréquemment utilisées car elles sont intuitives et ne nécessitent pas de formation au préalable pour pouvoir les utiliser (Jung et al., 2014). Les techniques allocentriques, au contraire, comportent un point de vue extérieur à l’EV (Poupyrev et al., 1998). Cela suppose que le sujet ne soit pas totalement immergé dans l’EV (Jung et al., 2014).

La taxinomie des techniques d’interaction peut donc s’effectuer selon le point de vue de l’utilisateur, à savoir s’il est immergé ou non dans l’EV (technique egocentrique et allocentrique) ou en type de sous-tâches (Argelaguet & Andujar, 2013 ; Bowman & Hodges, 1999). Ces taxinomies imbriquent des caractéristiques liées aux outils d’interaction et au logiciel. D’autres catégorisations peuvent être employées en se basant principalement sur le *hardware* en lui-même. Par exemple, Skalski, Tamborini, Shelton, Buncher, et Lindmark (2011) ont proposé une taxinomie des techniques d’interaction tangibles. Ils catégorisent les techniques d’interaction en 4 catégories selon leur *mapping* (fonction de transfert entre les mouvements réel et virtuels), c’est-à-dire selon la correspondance entre les mouvements physiques et ceux obtenus en virtuel.

- *Correspondance naturelle directionnelle (directional natural mapping)* : c'est la forme de *mapping* la plus basique, il s'agit d'avoir une correspondance directionnelle entre les mouvements de la main et ceux effectués par le curseur de la souris. Ainsi si l'utilisateur pousse le joystick gauche du Razer Hydra (voir chapitre 5 pour une description du Razer Hydra) vers la gauche, le curseur de sélection sera aussi déplacé vers la gauche.
- *Correspondance naturelle kinésique (kinesic natural mapping)* : ce type de *mapping* intervient quand les mouvements physiques de l'utilisateur sont directement retranscrits dans l'EV, par exemple grâce à une caméra.
- *Correspondance naturelle tangible partielle (incomplete tangible natural mapping)* : l'utilisateur manipule un objet qui ressemble partiellement à l'objet réel de même fonction. Par exemple, la manette de wii qui simule une raquette de tennis.
- *Correspondance naturelle tangible complète (realistic tangible natural mapping)* : le dispositif d'interaction ressemble par la forme et la manipulation à un objet réel de même fonction. Par exemple, le stylet pour écrire sur un écran tactile endosse le rôle d'un crayon.

La multitude de taxinomies pouvant être utilisées pour caractériser les techniques d'interaction montre l'actuelle difficulté à comparer les techniques d'interaction entre elles. Cependant, cette même multitude autorise plusieurs angles de vue pour étudier les techniques d'interaction, en particulier en les fractionnant en méthode de translation et d'orientation de la caméra.

2.3 Evaluation des techniques d'interaction

2.3.1 Techniques de navigation

Composante cognitive

L'évaluation des connaissances spatiales liées à la navigation s'inspire majoritairement du modèle LRS (Goldin & Thorndyke, 1982 ; Siegel & White, 1975). Les expérimentateurs vont évaluer la qualité de la carte cognitive des participants en vérifiant l'élaboration des points de repère, des routes et de la vue d'ensemble de l'EV (Van der Ham, Faber, Venselaar, van Kreveld, & Löffler, 2015). La mise en avant et la relation entre ces informations sont très pertinentes dans le cas d'un grand EV où le sujet a la possibilité d'élaborer plusieurs itinéraires d'un point à l'autre en fonction des routes et points de repères établis au préalable. Cependant, dans le cas

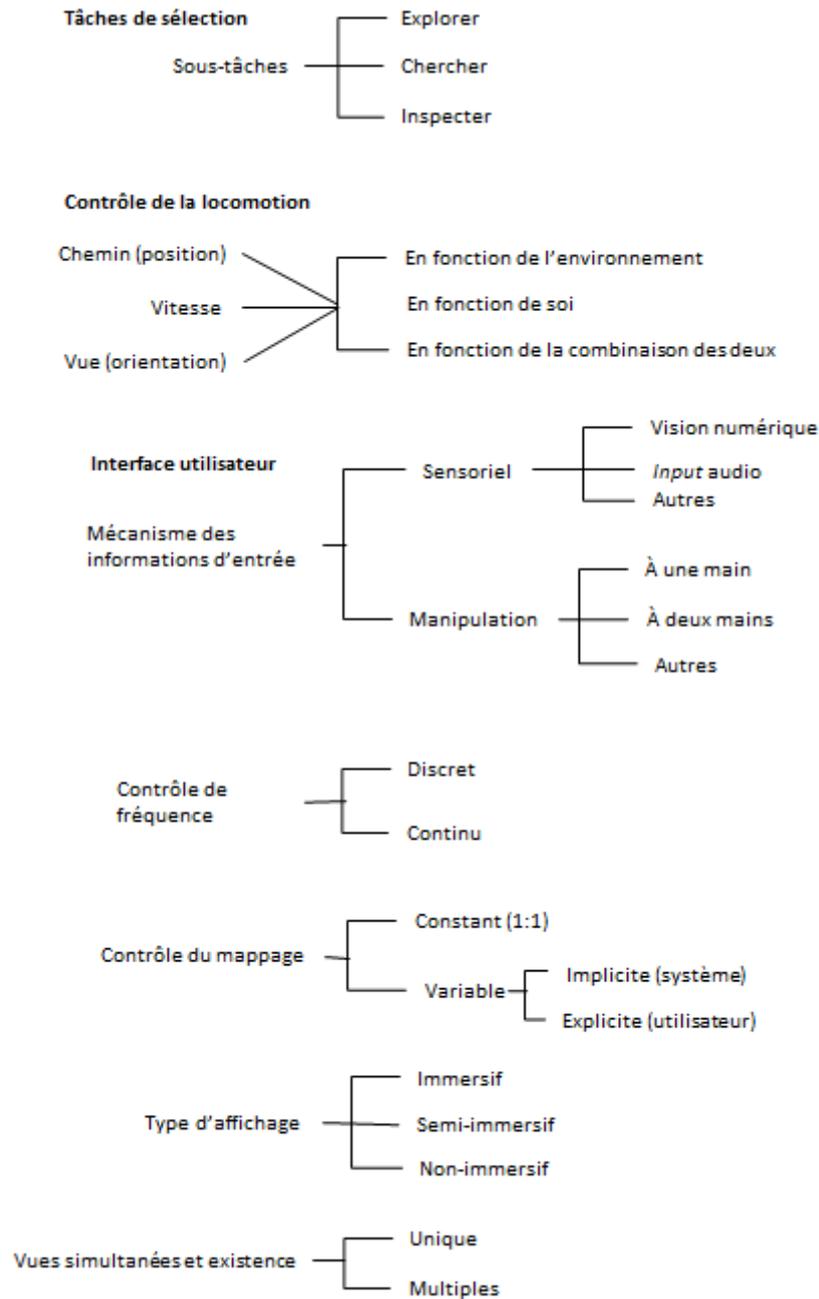


FIGURE 2.6 – Taxinomie des techniques de navigation tiré de Tan et al. (2001) et traduit en français.

d'un EV qui se limite, par exemple, à quelques pièces, il peut être préférable de demander au sujet de réaliser un plan, ou *sketch map*, de l'EV. En effet, dans ces petits EVs l'individu n'a pas nécessairement le besoin d'élaborer des relations entre les différents points de repère puisque l'orientation du champ visuel peut suffire à se repérer dans l'EV. Ces *sketch map* comprennent différents scores : le nombre d'objets représentés et leur position, ainsi que le résultat total impliquant la moyenne de ces 2 scores (Zanbaka, Lok, Babu, Ulinski, & Hodges, 2005).

Composante motrice

Dans le second cas où ce sont les aspects de locomotion de la navigation qui sont évalués, il est considéré que le temps mis pour réaliser la tâche équivaut à la performance du sujet (Patel & Vij, 2010). Mais d'autres variables peuvent être retenues comme le nombre de collisions et la distance parcourue. Les performances de navigation sont dans ce cas intimement liées à la technique d'interaction employée pour interagir dans l'EV.

Chittaro et al. (2006) ont développé un outil pour analyser la navigation et la locomotion du participant (Fig. 2.7). Analyser et visualiser les données d'un grand nombre de participants constitue un défi de taille pour les chercheurs actuellement. Surtout lorsque l'on étudie la navigation qui est une tâche qui requiert des habiletés cognitives et des habiletés motrices. Par exemple la méthode Chittaro et al. (2006) est efficace dans le cas d'un EV où il y a un trajet à suivre. On voit clairement que les participants partent du point A pour aller vers le B. Mais comment gère-t-on la visualisation des données lorsqu'il y a des aller-retours dans l'EV ? Pour pallier cette difficulté, une possibilité est de modéliser les déplacements sous forme de graphes (Verhulst et al., 2017). L'ordre des lieux visités et/ou des tâches réalisées par les utilisateurs sont ensuite facilement visualisés sur le graphe et des comparaisons entre groupes sont alors visuellement possibles.

2.3.2 Techniques de sélection

La norme ISO 9241-9 propose une tâche de sélection où le sujet doit sélectionner le plus rapidement possible les différentes cibles (Fig. 2.8). Au cours de la tâche de sélection, l'utilisateur doit sélectionner, le long de chaque essai, 13 cibles une à une, en suivant leur numérotation, comme indiqué sur la figure 2.8. Selon les essais, la taille et la position des cibles peut varier. Cette tâche a été développée en 2D et en 3D (Teather, Stuerzlinger, & Pavlovych, 2014) et s'inspire de la loi de Fitts (Fitts, 1954) qui permet de prédire la performance d'un individu pour sélectionner une cible. Elle se formule de la façon suivante :

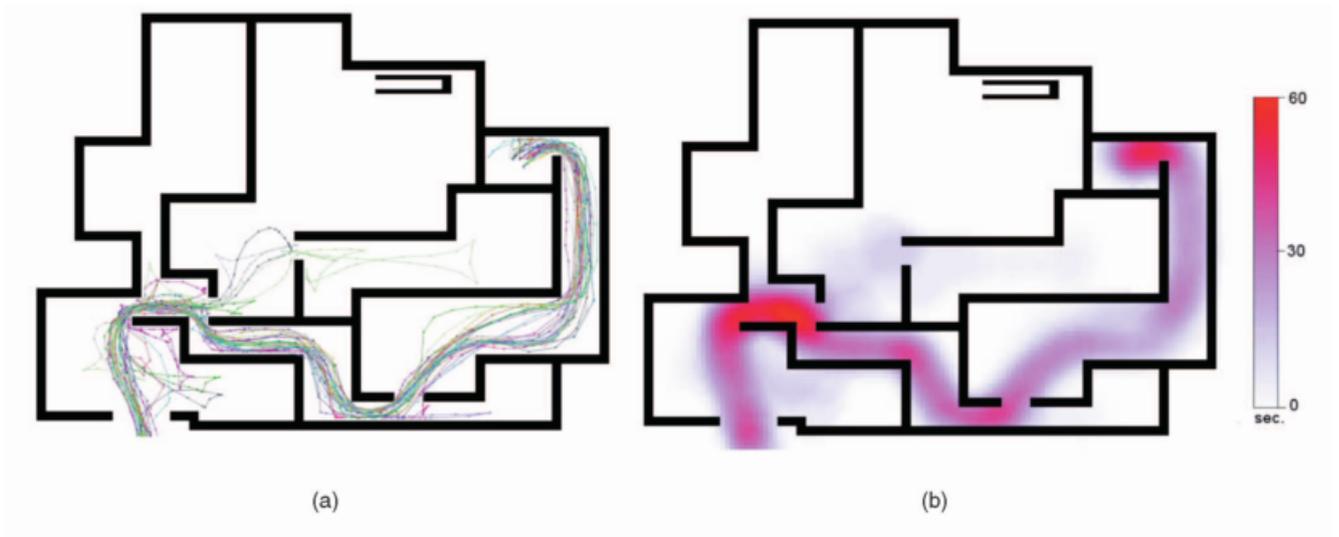


FIGURE 2.7 – Visualisation de la navigation (Chittaro et al., 2006) avec le tracé suivi par tous les utilisateurs (a) et visualisation de la durée totale passée par endroits (b).

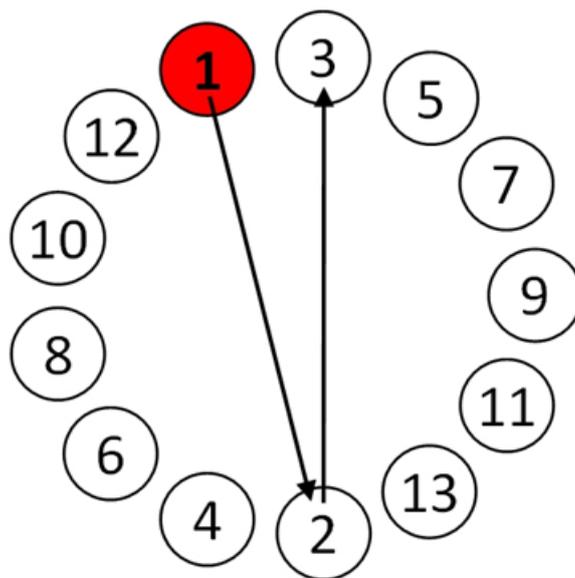


FIGURE 2.8 – Tâche d'évaluation d'une technique de sélection selon les normes ISO 9241-9.

$$MT = a + b \cdot \log_2 \left(\frac{D}{W} + 1 \right) \quad (2.1)$$

MT correspond au temps de réalisation moyen en seconde pour sélectionner une cible, D correspond à la distance pour atteindre la cible et W à la largeur de la cible. Le \log correspond à l'index de difficulté en bit. Les coefficients a et b sont déterminés de façon empirique grâce à une régression linéaire.

La loi de Fitts a ensuite été adaptée de différentes façons car prise isolément, elle ne permet pas de prédire la performance à une tâche de sélection en 3D (Murata & Iwase, 2001) ni celle d'un individu avec un déficit moteur (Almanji, Payne, Amor, & Davies, 2015). En effet, le modèle ne prend pas en compte un possible temps de pause entre les mouvements (Soukoreff & MacKenzie, 2004). La loi de Fitts a donc été réadaptée selon la formule de Shannon ci-dessous puis par le *throughput*¹.

$$IDE = \log_2 \left(\frac{De}{4.133 * SD_x} + 1 \right) \quad (2.2)$$

De correspond à la distance moyenne parcourue par le curseur de sélection pour atteindre une cible lors de la séquence d'essais. Pour sa part, SD_x est l'écart-type des distances entre le centre de la cible et celui où la sélection a été effectuée. Le *throughput*² s'obtient en divisant la justesse des sélections et le temps moyen pour sélectionner les cibles. La formule est indiquée ci-dessous :

$$Throughput = \frac{\log_2 \left(\frac{De}{4.133 * SD_x} + 1 \right)}{MT} \quad (2.3)$$

Le *throughput* est donc une méthode satisfaisante pour évaluer la performance du participant à une tâche de pointage. En effet, elle permet de combiner la vitesse et la précision de l'utilisateur avec une technique d'interaction lors de la sélection de la cible. De plus, le *throughput* permet de comparer l'utilisation des techniques d'interaction entre différentes études (MacKenzie, 1992), ce que la loi de Fitts n'autorise pas.

2.3.3 Utilisabilité et heuristique

L'utilisabilité peut se définir comme la qualité de l'interface utilisateur (Stanney, 2002), selon le modèle ISO 9241, c'est aussi un concept qui comprend des aspects quantifiables :

- *Efficiency* : ressources utilisées (temps, effort) pour compléter la tâche.

1. Le terme *throughput* peut être traduit par la capacité de traitement, la cadence. Par commodité nous continuerons d'employer le mot *throughput* dans notre exposé.

2. Une méthode pour calculer le *throughput* en Java est disponible dans la référence de MacKenzie, 2015. <http://www.yorku.ca/mack/hcii2015a.htm>.

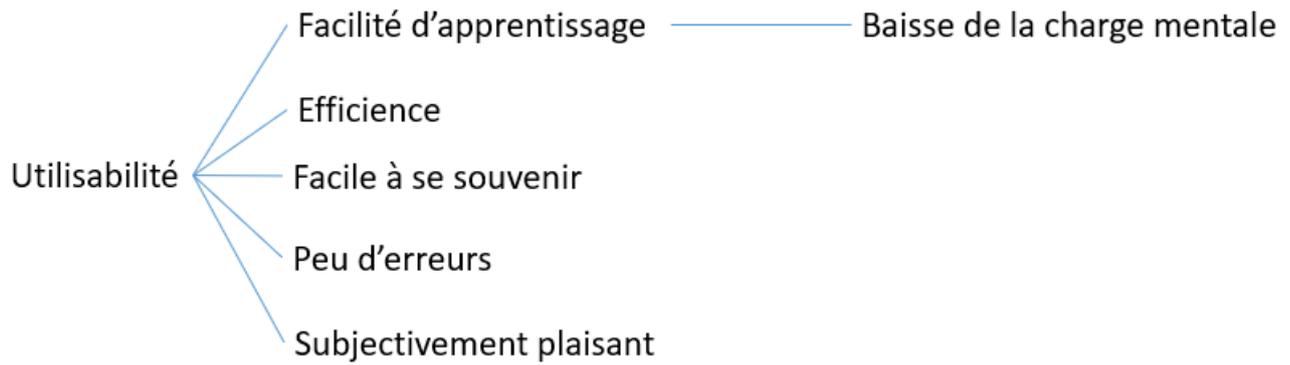


FIGURE 2.9 – Taxinomie des IHM, centrée sur l'utilisabilité tiré de Pincioli (2016).

- *Effectiveness* : justesse avec laquelle les utilisateurs complètent la tâche.
- *Satisfaction* : plaisir à interagir avec le système.

Diviser l'utilisabilité en catégories permet d'évaluer ses caractéristiques de façon séparée. Il devient primordial de comprendre les différentes caractéristiques qui composent les IHM car celles-ci entretiennent des relations entre elles de sorte que lorsqu'un aspect de l'IHM est amélioré, il peut restreindre la qualité d'une autre sous-caractéristique. Les relations entre les différentes caractéristiques ont été étudiée par Wieggers et Beatty (2013; cités par Pincioli (2016)). Par exemple il est possible d'améliorer la performance de l'utilisateur lors de l'utilisation d'un système, mais au détriment de la portabilité et de l'interopérabilité de ce même système.

L'amélioration d'une caractéristique peut donc affecter à la baisse la qualité d'une autre. Ainsi, lorsqu'une caractéristique nécessite une évolution, il faut d'abord mesurer son impact sur les autres caractéristiques du système afin d'évaluer si l'amélioration sera bénéfique ou au détriment des autres caractéristiques. La taxinomie liée à l'utilisabilité dans les IHM décompose aussi l'utilisabilité (Fig. 2.9).

Cette décomposition de l'utilisabilité est intéressante car elle fait intervenir la charge mentale. La sous-catégorie « facilité d'apprentissage » induit que l'utilisation de la technique d'interaction peut entraîner une charge mentale plus élevée, mais que celle-ci devrait baisser une fois l'utilisation de la technique d'interaction maîtrisée. Un EV doit donc permettre une interaction intuitive de sorte à mobiliser un minimum de ressources cognitives. L'intégration du concept de charge mentale dans les taxinomies de l'utilisabilité suppose de séparer les problèmes liés à

l'utilisabilité en 2 catégories : la première où l'interaction avec le système n'est pas intuitive et génère une hausse de la charge mentale chez l'utilisateur et la seconde où l'interaction n'est pas aisée et induit un effort physique de la part de l'utilisateur. Ces 2 catégories montrent que les techniques d'interaction peuvent engendrer une fatigue mentale et une fatigue physique qui ne module pas l'implication des ressources cognitives (Reinhardt & Hurtienne, 2017). Selon Reinhardt et Hurtienne (2017) il est nécessaire de séparer les problèmes d'intuitivité de l'interaction de ceux d'utilisabilité, plutôt que de les regrouper sous la même catégorie de problème d'utilisabilité, car une interaction non intuitive vient augmenter la charge mentale de l'utilisateur et l'empêche de compléter pleinement l'activité.

Le problème de la notion d'utilisabilité est qu'elle contient beaucoup de critères et que plus le nombre de critères augmente, plus son évaluation est compliquée. Pour y pallier, il est possible d'utiliser des heuristiques qui permettent d'effectuer une évaluation rapide de l'EV car elles contiennent un nombre limité de principes d'évaluation.

L'heuristique de Sutcliffe et Gault (2004) repose sur le fait que l'EV doit représenter le monde réel. Elle contient les échelles suivantes :

- *Engagement naturel* : les interactions doivent représenter des interactions du monde réel pour l'utilisateur. Idéalement les interactions virtuelles devraient se confondre avec celles du réel. L'interprétation de cette échelle dépend de l'engagement et du sentiment de présence de l'utilisateur.
- *Compatibilité avec la tâche de l'utilisateur* : les comportements des objets présents dans l'EV doivent correspondre à ceux du monde réel.
- *Expression naturelle de l'action* : la représentation de l'utilisateur dans l'EV doit lui permettre d'agir dans l'EV d'une manière naturelle et non restreinte. Si les dispositifs d'interaction ne comportent pas de retour haptique, l'expression naturelle de l'action en pâti d'office.
- *Coordination entre l'action et sa représentation* : le temps de latence entre l'action de l'utilisateur et la réponse dans l'EV doit être inférieure à 200 ms pour empêcher les symptômes de cybermalaise et permettre à l'utilisateur de se sentir présent.
- *Feedback réalistes* : les actions de l'utilisateur doivent être immédiatement visibles dans l'EV et répondre aux attentes de l'utilisateur de façon cohérente avec la réalité.

- *Fidélité du point de vue* : l'orientation du point de vue de l'utilisateur ne doit pas souffrir de latence.
- *Orientation et navigation* : l'utilisateur doit toujours savoir où il se trouve et pouvoir retrouver aisément son chemin.
- *Entrées et sorties* : les méthodes d'entrée et de sortie dans l'EV doivent être clairement définies.
- *Prévisions de départ* : quand il y a des compromis par rapport aux intentions initiales, ils doivent être consistants et clairement définis.
- *Support d'apprentissage* : les items actifs doivent comporter un indice visible et être expliqués si nécessaire pour permettre un apprentissage dans l'EV.
- *Instructions claires* : quand le système envoie des informations, celles-ci doivent être clairement signalées et permettent des échanges entre lui-même et l'utilisateur.
- *Présence* : l'utilisateur doit avoir l'impression de se sentir « là », dans l'EV.

Chaque branche de l'heuristique est évaluée par une échelle à 4 points : (1) sévère (la tâche est empêchée), (2) perturbant (l'utilisateur doit cesser la tâche le temps d'apprendre à surmonter le problème rencontré), (3) distrayant (l'utilisateur peut trouver la solution au problème après quelque temps et grâce à un indice) et (4) pratique (les utilisateurs peuvent régler le problème rencontré eux-même).

L'évaluation d'une technique d'interaction n'est pas aisée car elle peut comprendre de nombreux critères. C'est particulièrement le cas avec les concepts d'utilisabilité et d'heuristique qui sont décomposables en différentes notions.

2.4 Analyses des techniques d'interaction

2.4.1 Peut-on privilégier une technique d'interaction ?

Chaque technique d'interaction présente des avantages et des inconvénients. Il s'agit de pouvoir les repérer et les quantifier dans les différentes situations afin de pouvoir proposer la meilleure technique d'interaction selon les EVs et les tâches à y effectuer. Si cet objectif s'énonce clairement, il n'est pour autant pas facile à accomplir. En effet, les avantages et inconvénients

liés aux techniques d'interaction sont dépendants de plusieurs facteurs tels que les tâches et les caractéristiques de l'utilisateur. Ces enchevêtrements rendent difficile l'isolation des atouts et des désagréments des techniques d'interaction. Par exemple, différentes techniques d'interaction peuvent induire différents degrés de charge mentale, mais il est difficile de généraliser les résultats (cf. chap. 3). La même difficulté existe lorsque l'on compare la performance de l'utilisateur avec différentes techniques d'interaction. Si bien qu'il peut être non aisé de trouver dans la littérature des résultats applicables au protocole souhaité.

Il existe un grand nombre de techniques d'interaction et nous ne pouvons pas toutes les tester dans notre protocole expérimental. Nous avons donc dû faire un choix. Pour nous guider, nous avons pris en compte les techniques les plus utilisées dans le cadre des EVs avec des vocations d'évaluation cognitive. Valladares-Rodríguez, Pérez-Rodríguez, Anido-Rifón, et Fernández-Iglesias (2016) ont répertorié plusieurs EVs destinés à l'évaluation neuropsychologique et ont montré que 47.7% des études utilisent la souris, 20.75% l'écran tactile et peu le joystick, pour interagir dans l'EV. Différentes techniques d'interaction ont donc été proposées dans les EVs destinés à l'évaluation cognitive. Pour comprendre si et comment celles-ci peuvent venir impacter la performance du participant, il s'agit de bien saisir les avantages et inconvénients de celles-ci. Pour cela, les IHM ont comparé les techniques d'interaction à travers des tâches simples comme une activité de locomotion ou de sélection. Par exemple, Besançon et al. (2016) ont comparé plusieurs techniques d'interaction *desktop* : la souris, l'écran tactile et l'interface tangible. Ils ont répertorié leurs avantages et inconvénients (Tab. 2.1). Mais ces études ne prennent pas forcément en compte tous les concepts associés à la RV tel que le sentiment de présence ou les cybermalaises. En effet, une technique d'interaction peut être associée à des cybermalaises si elle est utilisée dans un EV immersif ou si cet EV est mal configuré.

2.4.2 Choix des techniques d'interaction pour le VMT

Parmi le choix existant, les techniques d'interaction les plus utilisées dans les EVs pour les évaluations neuropsychologiques sont la souris et l'écran tactile (Valladares-Rodríguez et al., 2016). Celles les plus fréquemment utilisées en RV, tous types d'EVs confondus, sont le joystick et le *Walking-in-place* (Boletsis, 2017). L'ensemble des études qui suivent sont décrites dans le Tab. 2.2.

L'écran tactile présente l'intérêt majeur d'être une technique d'interaction intuitive. En ef-

TABLE 2.1 – Avantages et inconvénients de l’interaction par souris, écran tactile et interface tangible tiré de Besançon et al. (2016) et traduit en français.

	Avantages	Inconvénients
Souris	<ul style="list-style-type: none"> • accessibilité • précision apparante • familière • séparation des DDL • peu de fatigue physique • modifications si tâche complexe 	<ul style="list-style-type: none"> • <i>mapping</i> difficile • interaction la plus lente • modifications nécessaires
Écran tactile	<ul style="list-style-type: none"> • accessibilité • précision apparante • hausse des mouvements rapides • plus rapide que la souris • <i>mapping</i> plus facile • différents <i>mapping</i> possibles pour les tâches complexes 	<ul style="list-style-type: none"> • manque de clarté quant à l’adéquation du <i>mapping</i> • plus lente que l’interaction tangible • fatigue physique • erreur de sortie
Interaction tangible	<ul style="list-style-type: none"> • interaction la plus rapide • <i>mapping</i> intuitif • impression de contrôle • caractère innovant 	<ul style="list-style-type: none"> • tâches complexes non réalisables • s’appuie sur un <i>tracking</i> 3D • nécessite un objet séparé (tangible) • <i>mapping</i> des interactions non flexible • toujours en marche, des modifications doivent être apportées pour mettre le dispositif en arrêt • erreur de sortie

fet, lors d'une tâche de sélection avec l'écran tactile, les utilisateurs ont un meilleur temps de réalisation qu'avec la souris (Besançon et al., 2016 ; Forlines, Wigdor, Shen, & Balakrishnan, 2007). De part sa rapidité et son intuitivité, l'écran tactile est toutefois une technique appréciée par les concepteurs d'EVs. Ce principalement pour les EVs 2D. En effet, les EVs évaluant les IADL ne requièrent pas nécessairement de la navigation. Ainsi, dans le cas d'une tâche consistant à faire du café dans une cuisine virtuelle, l'utilisateur n'a pas besoin de naviguer. Par exemple, dans les études de Martínez-Pernía et al. (2017), tous les items sont présentés devant lui et il lui suffit de les sélectionner en les pressant grâce à un écran tactile. Dans les cas où la réalisation de la tâche ne nécessite pas de navigation l'usage de l'écran tactile semble particulièrement approprié. Cependant dans les EVs où la navigation est indispensable, comme dans les supermarchés virtuels, l'utilisation de l'écran tactile semble autoriser moins de liberté que d'autres techniques d'interaction. La navigation via écran tactile s'effectue en appuyant sur un bouton, par exemple de la forme d'une empreinte de pas (Zygouris et al., 2015) et n'autorise donc pas une navigation en dehors des marques proposées. De plus, l'orientation de la caméra n'est pas toujours facilement exploitable avec l'utilisation de l'écran tactile. Par exemple dans l'étude de Okahashi et al. (2013), le *steering* n'est pas autorisé dans l'EV, de sorte que l'utilisateur ne peut qu'avancer ou faire demi-tour (Fig. 2.10).

L'EV que nous développons a pour vocation d'évaluer les fonctions cognitives. Ainsi, la première idée était d'utiliser des techniques naturelles (*Natural Gesture Interface*) pour interagir à l'intérieur de l'environnement choisi. Ce choix semblait le meilleur, d'une part car les techniques naturelles autorisent un haut sentiment de présence et, d'autre part, car elles permettent à l'utilisateur de naviguer et de sélectionner des items sans l'utilisation d'un dispositif d'interaction. C'est donc le corps de celui-ci qui renseigne les informations d'entrées à l'EV. Par exemple, le *Walking-in-place* a l'avantage de reproduire en partie ou totalement les mêmes façons de naviguer ou de sélectionner qu'en réel puisqu'il suffit de simuler une marche pour avancer. Cette technique est donc perçue comme fortement naturelle (Ferracani, Pezzatini, Bianchini, Biscini, & Del Bimbo, 2016). Dans le cas des grands EVs, les techniques naturelles permettent aussi à l'utilisateur d'acquérir de meilleures connaissances spatiales qu'avec des techniques non-naturelles comme le joystick (Williams, Bailey, Narasimham, Li, & Bodenheimer, 2011). L'acquisition de connaissances spatiales est d'ailleurs de même qualité que lors de la marche réelle (Williams et al., 2011). Les techniques d'interactions tentent de reproduire les interactions naturelles de sorte à pouvoir proposer des interactions intuitives et efficaces. Cependant, une technique d'interaction naturelle n'est pas forcément liée à une meilleure per-

formance qu'une technique d'interaction non-naturelle. Par exemple, les utilisateurs mettent moins de temps à compléter une tâche qui comprend une partie importante de navigation, avec une technique d'interaction non-naturelle (la manette de jeux vidéo) qu'avec une technique d'interaction semi-naturelle (Nabiyouni, Saktheeswaran, Bowman, & Karanth, 2015 ; Verhulst, Richard, Richard, Allain, & Nolin, 2016). Les techniques d'interaction, pour être efficaces, n'ont donc pas forcément besoin de copier la réalité et d'être naturelles. En effet, une autre étude a montré que les dispositifs tangibles n'ont pas besoin d'être ultra-réalistes (Reinhardt & Hurtienne, 2018). Les auteurs ont demandé à des jeunes adultes de jouer à un jeu vidéo de tennis. Les utilisateurs sont répartis en trois groupes : le premier qui joue avec un contrôleur en forme de raquette de tennis, le second avec un contrôleur qui représente seulement le manche de la raquette et le troisième groupe où les utilisateurs n'utilisent pas de raquette. Ceux-ci effectuent des gestes avec la main qui sont reconnus par une Kinect. Les utilisateurs ont peu apprécié l'interaction avec la main, ils préfèrent interagir avec un contrôleur que ce soit la raquette complète ou le manche de raquette. Il n'y a pas d'effet bénéfique à utiliser un contrôleur avec une forme complète de raquette par rapport à l'utilisation du contrôleur en forme de manche de raquette. En effet, il n'y a pas de différence significative au niveau de la performance, de la présence ou encore de la charge mentale entre ces 2 groupes (Reinhardt & Hurtienne, 2018).

Cependant, les techniques naturelles peuvent fatiguer physiquement l'utilisateur (Santos et al., 2009 ; Verhulst et al., 2016). En effet, au cours de ces 2 études, les utilisateurs devaient produire des mouvements continus, avec les bras et/ou les jambes, pour naviguer. Les utilisateurs ont ensuite renseigné à un questionnaire de charge mentale, une fatigue physique importante. Du fait de la forte demande physique liée aux techniques naturelles, nous avons choisi de les écarter et de nous diriger vers d'autres techniques. De plus, selon Kulik (2009), les techniques d'interaction n'ont pas à mimer la réalité mais elles doivent être efficaces ou amusantes, et les deux dans le meilleur des cas. Cette remarque est primordiale dans le sens où nous sommes dans une période où nous souhaitons mimer la réalité dans un souci d'écologie de la tâche. Il est important de réussir à concilier ces deux mondes que sont le monde réel et le monde virtuel, pour parvenir à tirer profit de ce que chacun peut proposer. Par exemple en virtuel il est possible de parcourir une grande distance sans être physiquement fatigué grâce à des techniques comme la téléportation. Il s'agit donc de pouvoir proposer une technique d'interaction avec laquelle l'utilisateur sera le plus efficace possible. Pour cela nous avons choisi de nous centrer sur l'utilisation de la souris, du gamepad et d'une interaction tangible : le Razer Hydra (voir chap. 5 pour une description du Razer Hydra). Nous avons sélectionné deux techniques

d'interaction *desktop* (la souris et le gamepad) qui sont présentés comme faciles d'utilisation et efficaces. Ainsi qu'une technique d'interaction tangible qui présente aussi l'avantage d'être rapide lors de son utilisation mais aussi inhabituelle. Elle présente donc un aspect de motivation pour le participant qui va le découvrir (Besançon et al., 2016).



FIGURE 2.10 – Illustration d'un EV où le *steering* n'est pas autorisé via écran tactile (Okahashi et al., 2013).

2.4.3 Comparaison des techniques

Le gamepad est une technique d'interaction non-naturelle, ou artificielle, avec laquelle l'utilisateur a un temps de réalisation plus rapide qu'avec une technique semi-naturelle telle que dans le Virtusphere³, qui permet à l'utilisateur de marcher sur place (Nabiyouni et al., 2015). Le gamepad ne présente toutefois pas d'avantage sur le temps de réalisation par rapport à la marche réelle (Nabiyouni et al., 2015). Ces résultats sont issus d'une tâche de navigation où les utilisateurs devaient suivre une ligne droite. Lorsque l'on compare le gamepad à la souris, nous pouvons constater que les temps sont comparables dans le cas d'une tâche de sélection (Zaraneck, Ramoul, Yu, Yao, & Teather, 2014). D'autres auteurs ont trouvé que le temps de réalisation était plus long en utilisant le gamepad plutôt que la souris pour le même type de tâche (Natapov, Castellucci, & MacKenzie, 2009; Ramcharitar & Teather, 2017). En version immersive, avec le port d'un HMD, le gamepad est aussi plus lent que la souris (Farmani & Teather, 2017). Quant à la comparaison entre le gamepad et le Razer Hydra, nous pouvons constater que ce dernier est plus lent que le gamepad et la souris (Farmani & Teather, 2017). Cependant, si l'on s'intéresse à la *Wii mote*TM, qui est un dispositif proche, d'un point de vue

3. <http://www.virtusphere.com/>

technique, du Razer Hydra, nous pouvons constater des résultats contradictoires, le gamepad a un temps de réalisation plus long que la *Wiimote*TM (Natapov et al., 2009). En effet, le Razer Hydra et la *Wiimote*TM consistent en une technique semi-naturelle qui autorise l'utilisateur à pointer son bras et la manette vers la cible à sélectionner. Lorsque le rayon virtuel du Razer Hydra ou de la *Wiimote*TM rencontre une cible, une sélection de celle-ci est autorisée si l'utilisateur confirme la sélection en appuyant sur un bouton. Les deux techniques fonctionnent donc de façon équivalente. De plus, le gamepad a un moins bon *throughput* qu'avec la *wiimote*TM et la souris à une tâche de sélection (Natapov et al., 2009). Les utilisateurs font aussi plus d'erreurs avec le gamepad qu'avec la souris (Natapov et al., 2009; Zaranek et al., 2014).

En terme d'effort d'apprentissage, des questionnaires subjectifs ont permis de renseigner que le gamepad demande davantage d'effort d'apprentissage que le *Walking-in-place* (Verhulst et al., 2016) alors que d'autres auteurs ont trouvé que les efforts entre la marche réelle et l'utilisation du gamepad sont équivalents (Nabiyouni et al., 2015). La souris demande peu voire pas d'effort d'apprentissage (Besançon et al., 2016). C'est de plus une technique d'interaction utilisée de façon quotidienne dans le monde du travail ou privé avec les ordinateurs. Les utilisateurs préfèrent la *wiimote*TM au gamepad (Natapov et al., 2009).

En effet, toute personne familiarisée avec un ordinateur l'est aussi avec la souris qui est la technique d'interaction de base. Cependant si la souris relève d'une utilisation habituelle et aisée pour beaucoup, elle peut aussi être difficile à utiliser de façon efficiente pour certaines populations. Par exemple, le patient avec des troubles cognitifs légers (*Mild cognitive impairment* (MCI)) est anxieux à l'idée d'utiliser un ordinateur (Wild et al., 2012). De plus, Seelye et al. (2015) ont analysé les schémas des mouvements de la souris chez des patients MCI et des sujets âgés sains et ont montré que les sujets MCI utilisaient la souris de façon moins efficiente que les sujets sains. Ils effectuent aussi plus de pauses entre les mouvements successifs de la souris. Dans d'autre cas, la souris ne semble pas être la méthode la plus adéquate comparée à d'autres techniques d'interaction comme celles très immersives utilisant un HMD. En effet, la souris est une technique d'interaction qui demande de la pratique pour être utilisée correctement. En effet, pour ceux qui ne sont pas familier avec les ordinateurs, l'utilisation de la souris peut interférer avec la réalisation de la tâche (Zucchella et al., 2014). Ainsi plusieurs techniques d'interactions ont été comparées entre elles afin de proposer la plus appropriée selon l'environnement, la tâche ou encore la population ciblée.

Les résultats ci-dessus nous indiquent que la souris semble être la technique d'interaction

qui apporte les meilleurs résultats mais que cela peut dépendre de la tâche à réaliser et du profil des participants. En effet, les personnes âgées ou présentant des difficultés motrices et/ou cognitives peuvent avoir des difficultés à utiliser une technique d'interaction qui semble simple d'utilisation pour d'autres.

TABLE 2.2 – Etudes comparant les techniques d’interaction.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Besançon et al. (2016)	Évaluer la performance et l'utilisabilité de techniques d'interaction	36 adultes (âge : 30.2 (8.7))	Tâche de manipulation avec la souris, une interface tangible et un écran tactile puis complétion NASA-TLX.	Les participants sont plus rapides avec la technique tangible qu'avec l'écran tactile. Ils sont plus rapides avec l'écran tactile qu'avec la souris. La technique tangible demande moins de charge mentale que l'écran tactile qui en demande moins que la souris.	Chaque technique d'interaction a ses avantages et inconvénient en terme de performance et de charge mentale générée.
Farmani et Teather (2017)	Comparaison de techniques d'interaction à une tâche de sélection	9 participants	Tâche de sélection avec la souris, le gamepad et le Razer Hydra alors que les participants portent un visiocasque.	Les participants sont significativement plus rapide avec la souris et plus lent avec le Razer Hydra. Dans cet exemple c'était un schéma découpé donc les rotation se font via la technique d'interaction et non par les mouvements de la têtes.	Chaque technique d'interaction a ses avantages et inconvénient.

TABLE 2.2 – Etudes comparant les techniques d’interaction.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Ferracani et al. (2016)	Evaluer les techniques d’interaction naturelles en RV.	19 participants (âge : 26.4 (5.8)).	<p>Les participants réalisent une tâche de navigation et de sélection avec 4 technique d’interaction :</p> <ul style="list-style-type: none"> — Walking-in-Place — Arm Swinging (avancer avec un mouvement de balancier des bras) — TAP (un Leap Motion est posé sur le visiocasque et les participants pointe du doigt la direction dans laquelle ils souhaitent se diriger) — Push (un Leap Motion est posé sur le visiocasque et les participant ouvrent ou ferment la main pour naviguer, comme s’ils abaissaient un levier) 	Le Walking-in-Place et le TAP sont les 2 techniques les plus rapide et avec lesquelles il y a le moins de collisions. Le Walking-in-Place est perçu comme étant la technique la plus naturelle et le TAP est préféré par les utilisateurs.	Chaque technique d’interaction a ses avantages et inconvénients.

TABLE 2.2 – Etudes comparant les techniques d'interaction.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Forlines et al. (2007)	Evaluer les différences de performance entre deux types de technique d'interaction	12 participants (âge : de 22 à 47 ans)	Mesure de la performance à une tâche de tâche et une tâche de glisser-déposer avec la souris pour un groupe et l'écran tactile pour le second.	La souris et l'écran tactile ont des performances similaires. Si les participants sont un peu plus rapide avec l'écran tactile ils font aussi plus d'erreurs.	Chaque technique d'interaction a ses avantages et inconvénients.
Martínez-Pernía et al. (2017)	Appréhender la possibilité d'utiliser les nouvelles technologies avec une population de patients cérébrolésés.	7 dyades de patients cérébrolésés/neuropsychologues	Les patients doivent réaliser une tasse de thé dans un EV : le SBS-COT. Les variables mesurées sont : les erreurs d'omissions et de commissions, les erreurs dans la séquence d'action, les erreurs de persévération et le temps de réalisation. Les participants interagissent avec un écran tactile.	les patients réussissent très bien la tâche de thé, ils ne font pas de persévération. Les patients ne relèvent pas de difficulté dans l'utilisation de l'outil et ont trouvé la tâche trop facile.	Les EVs et les écrans tactiles peuvent être utilisés auprès de patients cérébrolésés.

TABLE 2.2 – Etudes comparant les techniques d’interaction.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Nabiyouni et al. (2015)	Comparaison du joystick, marche réelle et virtusphère pour avoir une comparaison entre une technique naturelle, semi-naturelle et une technique non naturelle	12 participants de 18 à 35 ans.	Les patients doivent réaliser une tâche de locomotion en suivant une ligne droite ou bien une ligne avec des angles droits.	En terme de précision (mesurée par la déviation par rapport à une ligne droite), il n’y a pas de différence entre joystick et marche réelle mais ces deux là sont meilleurs que la technique semi-naturelle. Pour le temps de réalisation, il n’y a pas de différence entre marche et joystick mais ces deux là mettent moins de temps que le virtusphere. Le virtusphere plus dure à apprendre que le gamepad et marche réelle et aussi plus fatigant.	Chaque technique d’interaction a ses avantages et inconvénients.
Natapov et al. (2009)	Comparaison de techniques d’interaction à une tâche de sélection	15 participants de 18 à 28 ans (âge moyen : 22.2).	Les patients doivent réaliser une tâche de sélection avec la <i>Wii moteTM</i> , le gamepad et la souris.	La souris a le meilleur <i>Throughput</i> suivi de la <i>Wii moteTM</i> et du gamepad. Les participants préfèrent l’utilisation de la <i>Wii moteTM</i> à celle du gamepad.	Chaque technique d’interaction a ses avantages et inconvénients.

TABLE 2.2 – Etudes comparant les techniques d’interaction.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Okahashi et al. (2013)	Appréhender la possibilité d’utiliser un supermarché virtuel avec des patients cérébrolésés.	10 patients cérébrolésés et 10 contrôles. 10 jeunes adultes et 10 personnes âgées saines.	Les personnes doivent faire des achats dans un supermarché virtuel. Ils interagissent avec l’écran tactile. Ils effectuent au préalable des tests neuropsychologiques.	Les patients regardent plus souvent la liste, ont plus d’indice et plus de mouvements vers l’avant. Ils mettent aussi plus de temps au total, plus de temps sur la route et dans les magasins. Les résultats montrent que les âgés mettent seulement plus de temps total, plus de temps sur la route et dans le magasin. Le temps total est corrélé avec des mesure de l’attention et de mémoire.	Les EVs peuvent être utilisés avec des patients cérébrolésés.
Ramcharitar et Teather (2017)	Comparaison de techniques d’interaction à une tâche de sélection.	5 participants de 21 à 30 ans.	Les participants doivent réaliser une tâche de sélection avec la souris et la manette de jeux de Steam. Cette manette se compose d’un joystick, d’un pavé tactile et d’un capteur gyroscopique.	La souris a le meilleur <i>Throughput</i> et le moins d’erreurs alors que quand les utilisateurs interagissent avec le capteur gyroscopique ou le pavé tactile, ils ont un score au <i>Throughput</i> équivalent. C’est avec le joystick que les participants ont le moins bon score.	Chaque technique d’interaction a ses avantages et inconvénients.

TABLE 2.2 – Etudes comparant les techniques d’interaction.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Reinhardt et Hurtienne (2018)	Comparer une technique naturelle, semi-naturelle et non-naturelle.	93 étudiants (âge : 21.33 (2.10)).	Les participants jouent à un jeu de tennis avec : la main (kinect, <i>kinesic</i>), la manette de wii (<i>incomplete tangible</i>) ou une manette qui ressemble à une racquette (<i>realistic tangible</i>)	Il n’y a pas de différence en terme de charge mentale entre les techniques d’interaction mais il y a une corrélation négative entre charge mentale et le fait d’être un homme et une corrélation entre l’expérience de jeux et la charge mentale. De plus, la raquette complète ou incomplète est la technique préférée et la main celle qui est le moins apprécié.	Les utilisateurs préfèrent tenir un contrôleur que de ne pas en tenir. Il n’est pas forcément nécessaire d’avoir une raquette complète ou réaliste car ceci n’améliore pas forcément l’expérience utilisateur.
Verhulst et al. (2016)	Évaluer la performance et l’utilisabilité de techniques d’interaction.	80 jeunes adultes (âge : 23 3.45)	Tâche dans un supermarché virtuel puis passation du NASA-TLX.	Les participants sont plus rapides avec le gamepad et le <i>head tracking</i> . Il y a moins de frustration avec le gamepad que la kinect et moins d’effort avec le gamepad que la kinect.	Chaque technique d’interaction a ses avantages et inconvénient en terme de performance et de charge mentale générée
Williams et al. (2011)	Comparer la marche réelle, la navigation avec un gamepad et le Walking-in-Place sur la Wii Balance Board.	12 participants.	Les participants réalisent des tâches de navigation avec le gamepad, le Walking-in-Place sur la Wii Balance Board ou en marche réelle.	Les participants ont de meilleurs résultats en orientation spatiale en marche réelle et en Walking-In-Place qu’avec le gamepad.	Chaque technique d’interaction a ses avantages et inconvénients.

TABLE 2.2 – Etudes comparant les techniques d’interaction.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Zaraneck et al. (2014)	Comparer des techniques d’interaction à une tâche de sélection.	12 étudiants (âge : 20.9 (2.2)).	Les participants les sujets doivent sélectionner des cibles de tailles variées avec la souris, le gamepad, la Kinect et le Move de Playstation.	Les participants sont plus rapides avec la souris et ils sont les plus lents avec la Kinect. La souris a le meilleur <i>Throughput</i> suivi du gamepad, de la Kinect puis du Move.	Chaque technique d’interaction a ses avantages et inconvénients.
Zygouris et al. (2015)	Appréhender la possibilité d’utiliser un supermarché virtuel comme outil de dépistage des MCI.	21 personnes âgées et 34 patients MCI.	Les personnes doivent faire des achats dans un supermarché virtuel. Ils interagissent avec l’écran tactile. Ils effectuent au préalable une batterie de tests neuropsychologiques.	Le supermarché virtuel peut être utilisé pour discriminer les patients MCI des personnes âgées saines. Le temps de réalisation du supermarché virtuel est corrélé avec des tests neuropsychologiques dont le MMSE.	Les EVs peuvent être utilisés comme outils de dépistage des MCI.

2.5 Synthèse

Il n'est pas toujours aisé de choisir la bonne technique d'interaction dans un EV. Si la demande semble unique (accomplir telle ou telle tâche), les actions pour la mener à bien peuvent être complexes, et les techniques d'interaction peuvent elles aussi se décomposer en types d'action bien distincts. Par exemple la navigation peut être séparée en activité de translation ou de rotation de la caméra. Ces actions peuvent être issues d'informations d'entrée différentes. Il est donc nécessaire de bien renseigner la façon dont la technique d'interaction est implémentée dans l'EV pour permettre de comprendre à la fois son fonctionnement, mais aussi les comparaisons entre différentes méthodes d'interaction. Dans le cas de nos travaux de thèse nous avons choisi de nous centrer sur l'utilisation de la souris, du gamepad et du Razer Hydra pour évaluer la charge mentale associée à l'utilisation d'une technique d'interaction dans un EV, car ce sont des techniques fréquemment utilisées, qui demandent un apprentissage modéré et qui sont rapides et aisés à mettre en œuvre. Ces techniques sont décrites plus en détail dans le chapitre 5. Nous allons développer le concept de charge mentale et la façon dont il doit être lié aux techniques d'interaction dans le chapitre suivant.

La charge mentale en environnement virtuel

3.1 Introduction

La charge mentale est un concept qui a été largement étudié dans différents domaines, que ce soit celui des IHM, de la psychologie ou de l'éducation. Ce concept est de première importance car il permet de prédire la performance du sujet lors de la réalisation d'une tâche, et donc d'adapter les situations pour garantir une performance optimale. Cependant, la charge mentale possède de multiples modalités et est donc un sujet d'étude complexe. En particulier lorsqu'il s'agit de l'évaluer. Différents critères existent pour expertiser un outil de mesure. La qualité de l'outil sera d'autant meilleure qu'il comporte des critères variés. Selon Moustafa, Luz, et Longo (2017), ceux-ci comprennent :

- *La fiabilité* : la méthode doit pouvoir détecter les modulations de charge mentale lors du test ;
- *La validité* : l'outil mesure bien ce qu'il doit mesurer, ici la charge mentale ;
- *La sensibilité* : l'outil de mesure est sensible aux variantes de difficultés de la tâche et aux autres facteurs influençant la charge mentale. La méthode ne doit pas être affectée par des facteurs qui ne sont pas reliés à la charge mentale ;
- *Le diagnostic* : la méthode doit pouvoir identifier les causes des modulations de la charge mentale ;
- *L'intrusivité* : la méthode ne doit pas être intrusive ou venir interférer avec la performance du participant ;
- *L'équipement* : la méthode d'évaluation doit demander un minimum de matériel pour

éviter d'interférer avec la performance du participant ;

— *L'acceptabilité* : la méthode doit être acceptée par les participants.

La majorité des outils proposés ne permettent pas de remplir tous ces critères. Il s'agit donc de bien comprendre les spécificités de chaque contexte d'expérimentation et des outils d'évaluation pour proposer la meilleure mesure de la charge mentale.

Dans ce chapitre, les approches théoriques et la définition de la charge mentale seront exposées (section 3.2). Puis, les différentes méthodes d'évaluation de la charge mentale seront listées (section 3.3). Enfin, nous passerons en revue les apports de la littérature à propos des possibilités de prédire la charge mentale (section 3.4). Pour terminer, une synthèse mettant en avant l'intérêt d'évaluer la charge mentale grâce à différentes méthodes sera exposée.

3.2 Définition et approches théoriques

3.2.1 Définition

La charge mentale comprend plusieurs définitions et approches théoriques transposables d'un domaine à l'autre sont disponibles dans la littérature. Pour certains auteurs, la charge mentale comprend des synonymes tels que les "ressources mentales" et les "capacité cognitive". Pour d'autres, ces notions comportent des subtilités (Barrouillet, 1996) qui les différencient. Selon Barrouillet (1996), les ressources cognitives désignent l'énergie cognitive disponible pour l'individu, à un moment donné et pour une tâche particulière. La capacité cognitive est la quantité maximale de ressources cognitives, c'est-à-dire un ensemble de fonctions cognitives de haut et de bas niveau que peut mobiliser l'individu en vue de réaliser une tâche. La charge mentale est la charge cognitive générée par une tâche. D'autres auteurs ne font pas la distinction entre ces termes et prennent la charge mentale dans sa dimension globalisante, c'est-à-dire en considérant que la charge mentale mesure la quantité de ressources mentales mobilisées par un sujet lors de la réalisation d'une tâche (Tricot & Chanquoy, 1996). La charge mentale se définit facilement mais ses multiples dimensions en font un concept complexe (Chanquoy, Tricot, & Sweller, 2007).

3.2.2 Théorie des ressources cognitives limitées (Kahneman, 1973)

Selon la théorie des ressources cognitives limitées (Kahneman, 1973), les ressources allouées à la réalisation de la tâche dépendent de la difficulté de celle-ci. Ainsi, une activité complexe demande plus de ressources cognitives même si celle-ci est automatique de sorte que le sujet ne puisse pas réussir plusieurs tâches si elles consomment l'intégrité des ressources cognitives pouvant être allouées à leur réalisation. Dans ce cas, l'individu ne pourra pas recevoir toutes les informations nécessaires et risque de commettre des erreurs. L'idée d'une capacité limitée des ressources disponibles permet d'expliquer les erreurs liées à une différence de charge mentale selon les tâches. A la suite de la théorie des ressources cognitives, nous pouvons introduire les travaux de Eysenck et Calvo (1992) qui font une distinction entre qualité de la performance et qualité des processus cognitifs. La qualité de la performance représente l'accomplissement du participant en termes de réussite : a-t-il réussi à réaliser la tâche ou non ? La qualité des processus cognitifs est basée sur la relation entre la qualité de la performance et la quantité de ressources cognitives utilisées par le participant pour atteindre un certain niveau de performance. Ainsi lorsque le participant complète la tâche, si l'utilisation du système a été demandeuse de ressources cognitives, nous pouvons supposer qu'il mettra plus de temps pour réaliser la tâche qu'un participant très à l'aise avec l'utilisation du système. S'il n'arrive pas à terminer la tâche ou commet des erreurs, cela peut être dû à une surcharge cognitive, le participant n'ayant pas assez de ressources cognitives disponibles pour traiter toutes les informations présentes.

3.2.3 Théorie de la charge cognitive (Paas & Van Merriënboer, 1994)

La théorie de la charge cognitive (*Cognitive Load Theory* (CLT), Paas et Van Merriënboer (1994)), qui est utilisée dans le domaine des apprentissages et de l'éducation, est une bonne illustration des composantes multiples de la charge mentale. Cette théorie a été développée pour fournir des méthodes d'apprentissage qui prennent en compte la limitation des ressources cognitives disponibles par les apprenants. La CLT est basée sur le fait que les individus ont des capacités limitées en mémoire de travail qui interagissent avec une mémoire à long terme aux capacités illimitées. La CLT définit la charge cognitive comme une composante multimodale qui représente la charge générée par la réalisation d'une tâche (Paas & Van Merriënboer, 1994). Elle comprend les caractéristiques suivantes :

- *La charge mentale*. C'est un aspect de la charge cognitive qui est générée par les caractéristiques de la tâche et du sujet. Elle prend en compte les connaissances du sujet en lien

avec celles demandées pour compléter la tâche. Prendre en compte la charge mentale permet d'avoir une estimation de la charge cognitive demandée pour réaliser la tâche.

- *L'effort mental*. C'est un aspect de la charge cognitive qui correspond aux efforts nécessaires pour réaliser la tâche. Il est mesuré lorsque le sujet effectue la tâche.
- *La performance*. C'est un aspect de la charge cognitive qui correspond à la réalisation de l'apprenant. Elle se mesure en nombre de bonnes réponses, nombre d'erreurs, temps de réalisation. La performance peut être déterminée en temps réel ou après la réalisation de la tâche.

De plus, la CLT distingue 3 types de charge cognitive : charge intrinsèque, charge extrinsèque et charge effective. La charge intrinsèque est déterminée par l'interaction entre la tâche et les connaissances du sujet. La charge extrinsèque est la charge due à une mauvaise compréhension de la tâche. La charge effective est la charge qui contribue à la construction et à l'automatisation des schémas en mémoire à long terme. La charge mentale est donc modulée par différents facteurs dont ceux reliés à l'individu lui-même (charge intrinsèque) et ceux à la tâche (charge extrinsèque).

3.2.4 Cadre théorique proposé

Dans le cas de cette thèse, la définition retenue pour la charge mentale demeure générale. Nous envisageons la charge mentale comme la quantité de ressources cognitive dépensée à un moment donné pour une tâche ou un ensemble de tâche. L'individu possède donc une réserve de ressources cognitive dans laquelle il vient puiser des ressources de façon non homogène en fonction des situations et de leurs exigences. La charge mentale peut donc être modulée en fonction de l'individu et de ses connaissances (1), de la complexité de la tâche (2) et des facteurs environnementaux (3) associés à celle-ci. Par exemple, une personne familière avec les ordinateurs allouera moins de charge mentale à effectuer une recherche sur le web qu'une autre non familière, pour les raisons suivantes :

1. La personne habituée saura aisément réaliser la tâche, surtout si elle est quotidienne, en faisant appel à des schémas d'actions mémorisés (Chanquoy et al., 2007).
2. La recherche sur le web peut être plus ou moins complexe en fonction des requêtes.

3. L'utilisation de la souris ne viendra pas gêner la réalisation de la tâche car l'individu familier avec les ordinateurs sait la manipuler facilement.

La charge mentale sera donc différente, à tâche équivalente, pour une personne familière et une non accoutumée. En effet, la charge mentale liée à la tâche sera augmentée pour cette dernière car elle devra à la fois gérer la charge mentale liée à l'utilisation de la souris, celle liée au degré de complexité de la tâche sans pouvoir bénéficier d'aide en mémoire permettant de faire appel à des schémas.

Une charge mentale élevée peut donc avoir des conséquences négatives sur la performance du participant. En effet, les ressources cognitives sont limitées (Kahneman, 1973) et si l'ensemble des ressources disponibles sont consommées lors de la réalisation d'une ou plusieurs tâches, l'individu pourra ne pas recevoir toutes les informations nécessaires et risquera de commettre des erreurs. L'augmentation de la charge mentale est un indice de la répartition des ressources cognitives, de sorte que plus elle augmente, plus de ressources cognitives sont impliquées dans la réalisation d'une tâche. Les activités de la vie quotidienne, même si elles sont routinières et automatiques pour la plupart, consomment des ressources cognitives et des erreurs peuvent survenir quand les ressources cognitives sont limitées à cause d'une complication neurologique, par exemple suite à une lésion cérébrale, ou tout simplement de l'inattention (Kessler, Giovannetti, & MacMullen, 2007a).

Pour comprendre comment la charge mentale peut venir moduler la performance du participant dans des EVs destinés à l'évaluation cognitive, et en particulier la charge mentale induite par l'utilisation de la technique d'interaction. Nous avons choisi de simplifier au maximum les interactions dans l'EV afin de permettre aux participants d'être focaliser une la réalisation de la tâche. Il s'agit de constater si différentes techniques d'interaction génèrent des modulations dans la charge mentale et si celles-ci peuvent venir impacter la performance du participant.

3.3 Evaluation de la charge mentale

La charge mentale est un processus facile à définir mais qui est complexe à cause de ses différentes dimensions. Evaluer la charge mentale est donc une préoccupation majeure (Gevins & Smith, 2003). Il s'agit de comprendre quels paramètres, c'est-à-dire les liaisons entre l'individu, la tâche et les facteurs environnementaux, viennent moduler la charge mentale et de quelles façons. Différentes méthodes existent pour mesurer la charge mentale :

1. *Les mesures subjectives* : par exemple l'utilisation de questionnaires.
2. *Les mesures objectives*.
 - Les mesures physiologiques : par exemple les données recueillies par un ECG.
 - Les mesures comportementales : par exemple les erreurs effectuées lors de la tâche.
 - Les indices de performance correspondent à un assemblage de plusieurs variables généralement comportementales et physiologiques.
 - Les analyses basées sur les données : elles sont effectuées à posteriori avec des méthodes d'apprentissages supervisés ou non.

3.3.1 Les mesures subjectives

L'utilisation des questionnaires revient à demander aux sujets de quantifier, par eux-mêmes, l'effort mental fourni pour la réalisation de la tâche une fois que celle-ci est terminée. Différents questionnaires existent dont le NASA-TLX (Hart & Staveland, 1988), le *The Subjective Workload Assessment Technique* (SWAT) (Reid & Nygren, 1988), et sa version simplifiée le *Simplified-Subjective Workload Assessment Technique* (S-SWAT) (Luximon & Goonetilleke, 2001), ou le *Workload profile* (WP) (Tsang & Velazquez, 1996).

Le NASA-TLX est le questionnaire le plus utilisé pour mesurer la charge mentale (de Winter, 2014). Il comprend six sous-échelles : la demande mentale, la demande physique, la demande temporelle, la performance, l'effort et la frustration. Chaque sous-échelle est caractérisée par deux bornes : faible et élevée. La récolte des données issues du NASA-TLX nécessite trois étapes. Dans un premier temps il est nécessaire de donner un poids à chacune des sous-échelles pour les prioriser. En effet, chaque tâche ne fait pas appel aux mêmes composantes. Ainsi, certaines peuvent demander plus d'implication physique que d'autres. L'utilisateur remplit ensuite chaque sous-échelle, le score allant de 0 à 100. Le calcul de la charge mentale est indiqué par l'équation ci-dessous :

$$TLX = \left(\sum_{i=1}^6 di \times wi \right) \frac{1}{15} \quad (3.1)$$

Avec di qui représente le score donné à l'échelle par l'utilisateur après la tâche. Et avec wi qui est le poids associé à la dite échelle.

Une modification du NASA-TLX peut être proposée via le NASA-RTLX. Dans cette version les différentes sous-échelles ne sont pas pondérées. L'utilisation du questionnaire est donc simplifiée et plus rapide.

Le SWAT comprend 3 dimensions : le temps, l'effort mental et le stress (*psychological stress load*). Le participant se positionne sur chacune de ces échelles selon trois gradients (faible à élevé).

Le questionnaire WP, quant à lui, comprend 8 dimensions que le participant évalue entre 0 et 1 en fonction des ressources attentionnelles consommées lors de la tâche. Les 8 échelles sont les suivantes :

- *Processus central*. Il contient les ressources attentionnelles destinées à percevoir l'environnement et destinées à la résolution de problèmes et à la prise de décision.
- *Processus de réponse*. Il contient les ressources attentionnelles destinées à la sélection d'une réponse et à son exécution.
- *Processus spatial*. Il contient les ressources attentionnelles destinées au traitement de l'information spatiale.
- *Processus verbal*. Il contient les ressources attentionnelles destinées au traitement l'information verbale.
- *Processus auditif*. Il contient les ressources attentionnelles destinées au traitement de l'information auditive.
- *Réponses manuelles*. Cette dimension comprend les ressources attentionnelles consommées par la réalisation d'une activité motrice.
- *Réponses orales*. Cette dimension comprend les ressources attentionnelles consommées par la prise de parole.

Le score de ces échelles est ensuite additionné.

Les différents questionnaires ont chacun leurs spécificités. Ainsi Rubio, Díaz, Martín, et Puente (2004) recommandent de choisir le type de questionnaire sur la charge mentale selon les cas suivants :

1. Pour une comparaison de la charge mentale entre plusieurs tâches de niveaux de diffi-

- cultés différents, il est préférable d'utiliser le WP.
2. Pour prédire la performance d'un individu à une tâche, il est préférable d'utiliser le NASA-TLX.
 3. Pour analyser la demande cognitive ou les ressources attentionnelles allouées pour une tâche, il est possible d'utiliser le WP ou le SWAT.

Les mesures subjectives présentent l'avantage d'être non intrusives, facile à utiliser et peu coûteuses. De plus, les différentes échelles présentes dans les questionnaires sont une aide précieuse pour cibler l'origine de la charge mentale. Cependant l'évaluation survient en fin de tâche et est donc dépendante de la perception finale du participant (Wilson & Eggemeier, 1991). Elle n'est donc pas sensible aux changements d'état de charge mentale au cours de la tâche.

3.3.2 Les mesures physiologiques

Le système nerveux autonome est un régulateur des fonctions physiologiques. Il se compose du système nerveux sympathique qui module les ressources du corps dans des conditions stressantes et du système nerveux parasympathique qui permet de ramener les fonctions physiologiques à la normale suite à un événement stressant (Malik et al., 1996). Le rythme cardiaque (*Heart rate* (HR)) ainsi que la variabilité de la fréquence cardiaque (*heart rate variability* (HRV)) sont sensibles aux changements d'états du système nerveux autonome et peuvent donc être utilisés pour étudier la charge mentale. Le rythme cardiaque augmente lors des situations à forte charge mentale. Il correspond aux nombres de battements du cœur par minutes. Il augmente lorsque la tâche est plus complexe et requiert donc plus de charge mentale (Banville et al., 2017; Fallahi, Motamedzade, Heidarimoghadam, Soltanian, & Miyake, 2016; Jimenez-Molina & Lira, 2017; Knaepen et al., 2015). La moyenne du rythme cardiaque est de 76.90 ± 3.7 battements par minute et de 82.70 ± 4.8 battements par minute pour une tâche à forte charge mentale (Jimenez-Molina & Lira, 2017). La moyenne du rythme cardiaque est plus élevée lorsque la tâche demande une forte charge mentale qu'une faible charge mentale (Jimenez-Molina & Lira, 2017). Les données HRV se calculent par l'intervalle de temps entre 2 contractions du cœur est aussi un indicateur de la charge mentale (Jasper, Sibley, & Coyne, 2016). En effet, lorsque la charge mentale augmente, la variabilité de la fréquence cardiaque diminue (Byrne & Parasuraman, 1996; Cinaz, Arnrich, La Marca, & Tröster, 2013). Les données HRV sont aussi corrélées négativement avec le nombre d'erreurs, de sorte que plus il y a d'erreurs, plus la fréquence

cardiaque diminue (Muthukrishnan, Gurja, & Sharma, 2017; Teo et al., 2017). Si bien que les données HRV peuvent être utilisées comme un indicateur du fonctionnement cognitif (Muthukrishnan et al., 2017). Les différentes mesures pouvant être issues d'un ECG sont présentées dans le tableau 3.1.

Mesure	Calcul
HR	Nombre de battements par minute
HRV	Intervalle moyen entre les pics R (ms)
SDNN	Ecart-type des intervalles entre les pics R (ms)
MaxRR	Plus grand intervalle entre les pics R (ms)
MinRR	Plus petit intervalle entre les pics R (ms)
MinMaxRR	Différence entre le MaxRR et le MinRR (ms)
SDHR	Ecart-type du HR (bpm)
RMSSD	Racine carrée moyenne des différences entre les pics R
NN50	Nombre d'intervalles entre les pics R supérieurs à 50 ms
PNN50	NN50 divisé par le nombre d'intervalles entre les pics R
LF	Puissance spectrale basse fréquence des pics R entre 0.04 et 0.15 Hz
HF	Puissance spectrale haute fréquence des pics R entre 0.15 et 0.4 Hz
LF/HF	Ratio LF/HF

FIGURE 3.1 – Différentes mesures issues d'un ECG.

D'autres mesures physiologiques se sont montrées sensibles dans l'étude de la charge mentale, comme le diamètre des pupilles, ou les données issues d'un Électroencéphalogramme (EEG). Par exemple, des mesures EEG ont été effectuées sur des travailleurs qui contrôlent le trafic. Les résultats montrent que lors des situations complexes, générant beaucoup de charge mentale, l'activité de la bande thêta diminue (Fallahi, Motamedzade, Heidarimoghadam, Soltanian, Farhadian, & Miyake, 2016). Ce qui montre que les travailleurs fatiguent. Ces outils de mesures ne seront pas abordés en détail dans le cadre de cette thèse car leurs études nécessitent l'utilisation

d'outils de mesure qui ne sont pas forcément compatibles selon les protocoles expérimentaux. Par exemple, nous ne possédons pas de visiocasque avec *tracking* des yeux.

Les indices physiologiques de la charge mentale sont sensibles et permettent de voir la modulation de la charge mentale en temps réel. Toutefois cela demande du matériel dédié à la prise de mesures. Ce matériel est souvent intrusif, nécessite des compétences pour l'analyse des résultats et peut être incompatible avec un système de RV. Par exemple le port du visiocasque peut empêcher la bonne mise en place des électrodes pour l'EEG.

3.3.3 Les mesures comportementales et les index de performance

La charge mentale peut être mesurée par la performance à la tâche, que ce soit dans le cas de tâches simples ou dans d'un paradigme de double tâche. Dans le premier cas la charge mentale est mesurée en fonction de la réussite à la tâche. Tandis que dans le second cas, c'est la performance à la tâche concurrente qui permet d'établir le degré de charge mentale. En effet, la charge mentale augmente lorsque la tâche devient plus complexe et lorsque la charge mentale impliquée dans la réalisation d'une tâche devient trop importante, la performance à celle-ci diminue. Différents degrés de complexité d'une même tâche peuvent être proposés pour comprendre à quel moment la charge mentale vient moduler négativement la performance du participant. Par exemple une même situation peut être déclinée en 3 versions : une avec une charge mentale dite neutre, une seconde avec une charge mentale peu élevée et une troisième avec des tâches de plus en plus complexes. Les variables mesurées pour qualifier la performance sont généralement le temps de réalisation et les erreurs. Les situations de charge mentale élevée comportent une hausse du temps de réalisation et davantage d'erreurs que lors des conditions générant peu de charge mentale.

Yang et al. (2017) ont utilisé deux conditions expérimentales : une où le participant devait tracer un cercle avec un stylo avec un retour d'effort intégré (*stylus haptic phantom*) et une autre où il était en double-tâche et devait tracer le cercle tout en réalisant le *Paced Auditory Serial Addition Test* (PASAT) (Gronwall & Wrightson, 1974)). Le PASAT est une tâche qui génère de la charge mentale car elle consomme un grand nombre de ressources en mémoire de travail. Lors du PASAT, il est demandé au participant d'additionner le chiffre entendu avec celui qui a été annoncé précédemment. Les résultats montrent qu'en double-tâche la performance du participant baisse durant la tâche motrice. Les participants mettent plus de temps et forment

un cercle peu précis. La hausse de la charge mentale vient donc influencer négativement la performance du participant.

Les mesures comportementales permettent de mesurer la charge mentale de façon objective et d'isoler les moments où la condition est trop complexe. Ce qui est primordial, par exemple, avec des EVs de rééducation cognitive. Cependant, elles peuvent être lourdes méthodologiquement et non écologiques avec des protocoles expérimentaux qui comportent plusieurs conditions dont certaines qui ne répondent pas à des situations issues de la vie quotidienne des participants. Pour y pallier, il est possible de combiner plusieurs variables en un index de performance en prenant en compte des mesures complémentaires de la charge mentale. De plus, la multimodalité de la charge mentale en fait un concept que les auteurs préfèrent mesurer avec un ensemble de variables plutôt qu'avec une variable unique (Longo & Barrett, 2010).

Miyake (2001) a utilisé un index de charge mentale. Pour ce faire, il a utilisé trois mesures physiologiques, à savoir les données cardiaques, la pression sanguine et la conductance cutanée, et un score subjectif issu des résultats au NASA-TLX en prenant en compte les sous-échelles demande mentale, demande temporelle et effort. Suite à une Analyse en composante principale (ACP), un indice de charge mentale peut être proposé pour j tâche et k observations. Où W_{ik} est le coefficient de l'axe principal, P_{1jk} le ratio basse et haute fréquence de l'ECG, P_{2jk} l'amplitude de la pression sanguine, P_{3jk} la sudation de la peau et S_{1jk} le score subjectif.

$$\text{ChargeMentale} = W_{1k}P_{1jk} + W_{2k}P_{2jk} + W_{3k}P_{3jk} + W_{4k}S_{1jk} \quad (3.2)$$

Les index de performance permettent d'avoir une bonne idée de la charge mentale générée par l'utilisateur lors de la tâche, mais il peut être compliqué pour l'expérimentateur d'assembler les différentes données en une équation.

3.3.4 Les mesures basées sur les données

L'évaluation de la charge mentale basée sur les données permet de faire une classification de la charge mentale a posteriori, une fois les données récoltées avec l'aide des apprentissages supervisés et non-supervisés pour faire émerger des groupes dans les données.

Les apprentissages non-supervisés vont créer des groupes en fonction de caractéristiques communes non connues au préalable. Ces méthodes peuvent être utilisées pour faire émerger des patterns dans la structure des données. Les résultats obtenus par les apprentissages non-supervisés peuvent permettre de débiter un apprentissage supervisé. Les méthodes d'apprentissages supervisés comprennent des techniques de classification et de régression. La régression est utilisée pour prédire les variables quantitatives en fonction d'informations connues au préalable et les techniques de classification permettent de catégoriser les variables qualitatives, par exemple catégoriser un mail en spam ou non spam.

Pour construire des modèles de charge mentale à partir des données, l'approche *Cross Industry Standard Process for Data Mining* (CRISP-DM) (Chapman, Clinton, Khabaza, Reinartz, & Wirth, 1999) peut être utilisée. Six étapes sont suivies pour créer un modèle : l'élaboration du but, la compréhension des données, la préparation des données, la création du modèle, l'évaluation du modèle et son amélioration. L'utilisation des apprentissages supervisés pour l'évaluation de la charge mentale donne des résultats prometteurs, en particulier pour l'analyse des données physiologiques. Par exemple les données EEG comprennent de nombreuses variables et l'utilisation des algorithmes peut être une aide précieuse pour mieux repérer des patterns dans les données (Baldwin & Penaranda, 2012 ; Caywood, Roberts, Colombe, Greenwald, & Weiland, 2017). D'autres mesures physiologiques ont aussi été prises en compte comme le rythme cardiaque et la saturation en oxygène dans le sang (Jimenez-Molina & Lira, 2017). Ces deux mesures permettent de discriminer différents degrés de difficulté de la tâche et de les classer en deux niveaux de charge mentale grâce à l'algorithme *support vector machine* (SVM).

Si la majorité des évaluations basées sur les données utilise des variables issues de mesures physiologiques, d'autres études montrent que d'autres types de données peuvent être utilisées. Par exemple, Moustafa et al. (2017) ont utilisé une base de données comprenant les scores au NASA-TLX, au WP et de performance à des tâches de navigation sur le web et ont conduit des analyses via des algorithmes d'apprentissages supervisés. Les résultats indiquent d'une part que la méthode random forest donne des résultats plus sensibles que les questionnaires subjectifs seuls et d'autre part que les résultats obtenus par l'algorithme sont plus proches de ceux du NASA-TLX que de ceux du WP.

Les mesures basées sur les données sont sensibles et permettent d'isoler des patterns liés à différents degrés de charge mentale. Ces techniques sont très utiles lorsque beaucoup de variables

sont enregistrées, par exemple plus de 50 variables. Cependant elles demandent de pouvoir baser leurs apprentissages sur un nombre d'observations important puisqu'il faut séparer la base de données en deux parties : une plus importante dédiée à l'entraînement de l'algorithme et une autre moins importante pour le tester.

3.4 Estimation de la charge mentale

3.4.1 La charge mentale associée à la tâche

Lorsque la difficulté de la tâche augmente, la charge mentale augmente aussi. La complexité d'une tâche peut donc augmenter la charge mentale (Chanquoy et al., 2007) de sorte qu'une tâche simple demande peu de ressources cognitives tandis qu'une tâche complexe en mobilise davantage (Kramer & Parasuraman, 2007). Dans le cas des EVs destinés à l'évaluation cognitive il est nécessaire de maîtriser la charge mentale associée à la tâche, donc sa complexité. En effet, si la complexité de la tâche est trop élevée, elle peut générer une charge mentale trop importante. Cette surcharge mentale peut empêcher le participant de réussir correctement la tâche, qu'il ait des difficultés cognitives ou non. Par exemple, dans les métiers de la santé il est important de maîtriser la charge mentale des infirmier(e)s pour éviter des erreurs dont la conséquence pourrait être dramatique. Le ou la professionnel(le) soumis(e) à une surcharge mentale pourrait se tromper dans la distribution ou le dosage des médicaments. C'est en effet dans les situations de surcharge mentale, lorsque la réalisation de l'activité requiert plus de ressources cognitives que l'individu n'en a de disponible, qu'il risque de commettre des erreurs comme un mauvais séquençage des actions à compléter ou des oublis. L'ensemble des études présentées dans ces sections sont répertoriées dans le tableau (Tab. 3.1)

La charge mentale a été étudiée de nombreuses fois dans le monde professionnel que ce soit auprès des pilotes d'avions, dans les métiers de l'armée ou auprès d'infirmières. Le but était de démontrer que, lors de la réalisation de tâches complexes, la charge mentale est liée à une baisse de la performance et à une augmentation des erreurs (Yousef Zade, Mazloumi, Abbasi, & Akbar Zade, 2016). De plus, une forte charge mentale est associée à une fatigue mentale chez le sujet (Fallahi, Motamedzade, Heidarimoghadam, Soltanian, Farhadian, & Miyake, 2016).

Si les tâches complexes conduisent à une charge mentale élevée, il est nécessaire d'être conscient que la réalisation de chaque tâche, peu importe son degré de difficulté, implique des

ressources cognitives et induisent de la charge mentale. Par exemple, choisir la main avec laquelle sélectionner une cible implique une modulation de la charge mentale (Liang, Wilkinson, & Sainburg, 2017). Cependant, toute hausse de la charge mentale n'impacte pas forcément négativement la performance du participant (Ayaz et al., 2012 ; Liang et al., 2017). En effet, lorsque le participant réalise une tâche et que sa charge mentale est élevée, il ne commettra pas forcément d'erreurs. C'est en particulier le cas pour les personnes expérimentées dans un domaine. Elles peuvent réaliser une tâche complexe et générer beaucoup de charge mentale sans pour autant que cela vienne impacter leur performance. Les conséquences d'une charge mentale élevée ou trop élevée sont toutefois admises et se constatent essentiellement avec des comportements de type erreur. Cependant, ces comportements d'erreurs ne se produisent que lorsque la majorité ou la totalité des ressources cognitives sont déjà mobilisées. Les conséquences d'une charge mentale moyenne ou faible sont moins visibles, et donc plus difficiles à repérer. L'augmentation de la charge mentale viendrait moduler la performance du participant lorsqu'elle est très élevée et lorsque la tâche est complexe et/ou lorsque la personne est en situation de multitâches.

Dans le cas de la psychologie, le multitâche (*multitasking*) renvoie à une situation dans laquelle le sujet doit réaliser au moins deux tâches en même temps, donc en double-tâche, soit en alternance (Stoet, O'Connor, Conner, & Laws, 2013). Par exemple, en positionnant le curseur de souris à l'endroit désiré, puis en la lâchant, pour pouvoir saisir du texte à l'aide du clavier. La première tâche est l'utilisation de la souris tandis que la seconde est l'ajout d'informations textuelles. Selon d'autres études, il n'y a pas de *multitasking* à proprement parler chez l'humain, mais de l'alternance rapide entre la réalisation de plusieurs tâches (*task switching*) (Kirschner & De Bruyckere, 2017). En effet, si l'on compare le cerveau humain à un processeur, celui-ci n'aurait qu'un cœur. Ainsi, l'homme ne serait pas capable de réaliser plusieurs tâches en même temps à proprement parler mais pourrait en effectuer plusieurs de façon successive, la succession pouvant être rapide et donner l'impression de simultanéité. En effet, mener une activité implique un contrôle moteur, perceptif et cognitif. Par exemple, un individu qui travaille en répondant à ses SMS ne fait pas les deux tâches en même temps. Dans un premier temps il doit laisser ce qu'il faisait pour diriger un contrôle moteur vers le téléphone, mobiliser ses sens et son attention pour lire le SMS et y répondre pour enfin retourner vers sa tâche initiale qui a durant ce temps été délaissée.

Lors de la réalisation de tâches, plusieurs buts peuvent être activés en même temps mais dans une majorité de cas, l'activation de ces différents buts entraîne une hausse des ressources

cognitives impliquées dans la réalisation des tâches, et donc une augmentation du temps de réalisation, puisque toutes les informations ne peuvent pas être traitées en même temps de façon optimale (Salvucci & Taatgen, 2010). En effet, les ressources cognitives sont limitées et la réalisation de plusieurs tâches demande à l'individu de pouvoir prioriser ses actions en fonction de ses ressources cognitives disponibles. Lorsque les tâches consomment trop de ressources cognitives, l'individu risque de manquer des informations, de commettre des erreurs et/ou de ne pas aboutir à la finalité recherchée. La superposition des tâches à réaliser vient consommer de plus en plus de ressources mentales jusqu'à surcharger la cognition et ainsi moduler négativement la performance du participant. Ainsi lorsque l'on compare le même type de tâche mais avec des niveaux différents de difficulté on constate que la charge mentale augmente avec la difficulté de la tâche alors que la performance diminue (Jaquess et al., 2017). Dans cette étude, 63 jeunes adultes ont réalisé une situation multitâche dans un simulateur de vol. Les utilisateurs étaient confrontés à une des trois conditions de difficulté :

- *Une condition facile.* Il s'agissait de maintenir une altitude, un cap et une vitesse constant(e) par un temps clément.
- *Une condition médium.* Il s'agissait de maintenir un cap et une vitesse constant(e) tout en effectuant des changements d'altitude régulier. Le temps est nuageux.
- *Une condition difficile.* Il s'agissait de maintenir la vitesse de vol en ajustant l'altitude et le cap. Le temps est nuageux avec des vents modérés.

Il est important de comprendre les processus mis en œuvre lors des situations de multitâches. En effet, les difficultés en multitâche ont des répercussions négatives sur la vie quotidienne (Burgess, Veitch, de Lacy Costello, & Shallice, 2000) ainsi que dans le champ des apprentissages académiques, car la réussite aux tâches est vite compromise lorsque plusieurs tâches sont réalisées de façon simultanée (Dux et al., 2009 ; Law, Logie, & Pearson, 2006 ; Monsell, 2003). En effet, les sujets mettent plus de temps à réaliser la situation en multitâches qu'en simple tâche (Courage, Bakhtiar, Fitzpatrick, Kenny, & Brandeau, 2015 ; Rubinstein, Meyer, & Evans, 2001 ; Salvucci & Taatgen, 2010). De plus, à travers une méta-analyse, les auteurs (Jeong & Hwang, 2016) suggèrent que la performance lors de situations en *multitasking* est impactée de façon hétérogène. Ces mêmes auteurs se sont donc intéressés à savoir comment la présentation de l'information (tâche simple ou multitâches) venait impacter la performance des participants (Hwang & Jeong, 2018). Pour ce faire, ils ont étudié l'impact de la hiérarchie de la tâche, le degré d'interférence sensorielle et la réponse comportementale (une réponse est-elle attendue ou

non). La hiérarchie de la tâche correspond à la tâche que l'on priorise. Elle sera alors appelée tâche primaire, tandis que l'autre ou les autres seront qualifiées de tâches secondaires (Wang et al., 2015). Il est difficile de réguler ce paramètre car même si on dit aux sujets quelle est la tâche primaire, il n'est pas possible d'être assuré qu'ils aient considérés cette tâche comme l'activité primaire (Hwang & Jeong, 2018).

Dans le cas des EVs destinés à l'évaluation cognitive dans des situations de multitâche, il est donc primordial de proposer des tâches suffisamment complexes pour permettre au sujet de s'impliquer. Mais le degré de complexité doit être contrôlé, de sorte à ne pas venir détériorer la performance du participant en épuisant ses ressources cognitives disponibles.

3.4.2 La charge mentale associée aux facteurs environnementaux

La relation entre la charge mentale et les différents concepts liés aux EV et la RV est peu claire et les études ne permettent pas réellement d'en rendre compte. Il est cependant indispensable de comprendre comment l'EV est construit pour comprendre comment son utilisation peut venir moduler la charge mentale de l'utilisateur. Différents types d'EV peuvent impliquer différents degrés de charge mentale. C'est en effet, ce qu'ont montré différents auteurs en mettant à jour une corrélation entre le sentiment de présence et la charge mentale ressentie (George, Demmler, & Hussmann, 2018 ; Ma & Kaber, 2006). Ces résultats indiquent que les deux concepts sont liés et suggèrent qu'un participant se sentant plus présent dans un EV, va s'y impliquer davantage et donc que ses ressources cognitives seront davantage sollicitées.

L'utilisation de la technique d'interaction nécessite un contrôle gestuel (Argelaguet & Andujar, 2013 ; Bowman et al., 2001) effectué principalement par les mains. Mais chaque partie du corps peut être potentiellement impliquée pour interagir avec un objet 3D via l'utilisation d'une technique d'interaction. Par exemple, dans le cas des casques de RV, les mouvements de la tête, autrement dit les informations d'entrée, sont renseignés pour permettre l'orientation de la caméra et donc du point de vue dans l'EV. En fonction de l'interface avec laquelle l'utilisateur interagit, la performance à une même tâche peut être plus ou moins bonne (Oviatt, 2006). Différentes interfaces et techniques d'interaction peuvent donc faire intervenir différents degrés de charge mentale et venir impacter de manière positive ou négative la performance surtout lorsque leur utilisation n'est pas suffisamment intuitive. L'intuitivité de l'interaction avec l'EV peut être mesurée en comparant un test d'utilisabilité simple avec un autre en para-

digme de double-tâche. Lors de la double-tâche, la tâche secondaire ne doit pas venir corrompre la réussite à l'activité primaire en impliquant l'usage des mains et les ressources visuelles qui sont déjà allouées pour la réussite de la tâche 1. La tâche secondaire demande donc à l'utilisateur de produire un schéma régulier de gestes en tapant avec ses pieds (par exemple : tap-tap-pause-pause/tap-tap-pause-pause) (Park & Brünken, 2015). Cette méthode, basée sur le rythme, indique que la performance à la tâche secondaire reflète le niveau de charge mentale consommée par la tâche principale de sorte qu'une activité principale non intuitive générera un rythme irrégulier avec des temps de réaction plus long à la tâche concurrente. Comme les ressources cognitives sont limitées, une grande sollicitation de celles-ci peut entraîner des erreurs. Ce phénomène de surcharge cognitive est connu comme venant impacter négativement la performance du participant dans le domaine des IHM (Olson & Olson, 1990). Une technique d'interaction complexe pourrait donc augmenter la charge mentale du participant et ainsi le fatiguer et lui faire commettre plus d'erreurs, comme des erreurs d'utilisation de type faux clic.

Pour comprendre le rôle des différentes techniques d'interaction dans un EV, des études comparatives ont été publiées, mais peu comparent le degré de charge mentale généré par la technique d'interaction en elle-même. Parmi celles-ci, Besançon et al. (2016) ont montré des différences de charge mentale entre une technique d'interaction tangible, un écran tactile et la souris. Les participants devaient déplacer un objet 3D et le repositionner selon une orientation spécifique. Pour ce faire ils devaient utiliser une interface tangible qui se présentait sous la forme d'un item réel dont les mouvements sont mappés sur ceux de l'objet 3D, la souris ou l'écran tactile, mais dont les commandes ne sont pas usuelles. Par exemple pour effectuer une rotation de l'objet 3D avec la souris, l'utilisateur devait appuyer sur le bouton gauche et sur le droit pour effectuer une translation. Les résultats en charge mentale, obtenus avec le NASA-TLX, montrent que la charge mentale est la plus faible dans les cas où l'interaction s'effectue avec la technique d'interaction tangible. C'est en particulier le cas de la sous-échelle « demande mentale » qui est plus élevée pour la souris que pour la technique tangible. Cependant, l'utilisation de la souris consomme moins de ressources physiques que la technique tangible et l'écran tactile. Dans une autre étude comparant une activité dans un supermarché virtuel, des techniques d'interaction avec une manette de jeu vidéo et la Kinect (les utilisateurs effectuent une action motrice, reconnue par la Kinect puis mappée dans l'EV) ont été analysées via les données issues du NASA-TLX (Verhulst et al., 2016). Les participants produisaient plus d'effort et ressentaient plus de frustration dans les conditions utilisant la Kinect que dans celles avec la manette de jeux vidéo. Une troisième étude montre que la charge mentale, mesurée encore avec

le NASA-TLX, est plus faible avec le casque de RV et le Leap Motion, et plus élevée lors de la même tâche, mais cette fois effectuée devant un écran et où les mouvements sont reconnus par la kinect (Speicher, Daiber, Kiefer, & Krüger, 2017). Ainsi l'immersion ne semble pas être associée à une hausse de la charge mentale, mais c'est la difficulté de la tâche ou le manque d'intuitivité qui entraîne une hausse.

Dans une autre recherche, qui compare l'utilisation du Leap Motion¹ et du bras haptique Novint Falcon², on constate qu'il n'y a pas de différence de charge mentale sensible entre les deux dispositifs chez la personne âgée mais que la différence se creuse significativement chez la personne âgée avec un trouble moteur (Ramírez-Fernández, Morán, & García-Canseco, 2015). Une technique d'interaction intuitive, qui se base sur des activités habituelles de l'individu, pourrait réduire la charge mentale associée à l'interaction. Par exemple, la marche est une activité naturelle pour l'homme. Grâce à l'avancée de la technologie la marche peut être utilisée pour naviguer dans un EV de RV. En plus de permettre d'augmenter le sentiment de présence, la marche réelle semble consommer moins de ressources cognitive qu'une navigation avec le joystick avec ou non le port du visiocasque (Zanbaka et al., 2005). Dans cette étude, les auteurs ont demandé aux participants de visiter une pièce virtuelle pendant cinq minutes et de réaliser ensuite un plan (*sketch map*) de l'EV. Les participants pouvaient naviguer en immersif, c'est-à-dire avec un casque de RV, en se déplaçant physiquement ou avec un joystick, ou en non-immersif avec un joystick. Zankaba et al. (2005) ont montré que les participants avaient dessiné des plans plus précis lorsqu'ils naviguaient en marche réelle qu'avec le joystick (avec ou sans casque de RV). De plus, la phase de familiarisation durait moins longtemps en marche réelle que dans les conditions avec le joystick. Ces résultats laissent supposer que la marche réelle requiert moins de charge mentale et que la méthode d'interaction est intégrée plus facilement que l'utilisation du joystick. Ainsi les participants avaient plus de ressources cognitives disponibles pour mémoriser les éléments composant l'EV et ensuite les rappeler. Une autre étude a comparé la charge mentale perçue par des soldats lors d'un entraînement virtuel, présenté sur un ordinateur portable avec une interaction via le clavier et la souris, et un entraînement en réel (Stevens, Mondesire, Maraj, Badillo-Urquiola, & Maxwell, 2016). Les 98 participants devaient réaliser quatre tâches en collectif. Ces tâches ont été choisies car elles étaient réalisables aussi bien en réel qu'en virtuel. Par exemple, ils devaient réagir à une embuscade. Puis, ils complétaient le NASA-TLX. Les analyses ont montré que la charge mentale ressentie était plus

1. <https://www.leapmotion.com/en/>

2. <https://haptichouse.com/pages/novints-falcon-haptic-device>

élevée en virtuel qu'en réel, sauf pour les sous-échelles "demande physique" et "effort" (Stevens et al., 2016). Ainsi une tâche présentée sur différents supports, réel ou virtuel, peut entraîner une charge mentale différente. Cependant la façon dont la technologie vient moduler la charge mentale reste opaque. En effet, chaque tâche et technique d'interaction est susceptible de générer une charge mentale particulière spécifique. La performance à une tâche en virtuel est donc fortement influencée par l'utilisation de la technique d'interaction et celle-ci doit être facilement utilisable pour permettre à l'individu de compléter la tâche demandée (Ramírez-Fernández et al., 2015).

Les résultats issus des études précédentes sont intéressants, mais restent difficiles à mettre en pratique car ils sont difficilement transposables d'une étude à l'autre. Par exemple, la charge mentale est plus élevée avec l'utilisation de la manette de jeux vidéo car les novices ont besoin de fournir des efforts pour apprendre à manier correctement la manette de jeux vidéo (Verhulst et al., 2016). La même limite apparaît dans l'étude de Besançon et al. (2016), qui montre que si la charge mentale est plus élevée avec la souris, cela est dû en partie au fait que l'utilisation de la souris comportait des commandes usuelles, si bien qu'il fallait inhiber celles-ci pour apprendre de façon adéquate les nouvelles commandes.

3.4.3 La charge mentale associée à l'utilisateur

La charge mentale invoquée par l'utilisateur pour compléter une tâche dépend de la difficulté de la tâche, des facteurs environnementaux mais aussi les paramètres liés à l'individu, de sorte qu'en fonction de l'expertise de l'individu, le degré de charge mentale sera plus ou moins élevé. Par exemple, un individu expert dans un domaine pourra réaliser une tâche complexe sans que cela ne génère d'effort particulier car le sujet sait parfaitement bien réaliser la tâche et aura intégré les différentes étapes de réalisation de la tâche en mémoire à long terme. Tandis, qu'un individu débutant fournira beaucoup de ressources cognitives pour réaliser la tâche complexe. Lorsque la charge mentale est étudiée lors de la réalisation d'une tâche, le degré de familiarité de l'individu avec la tâche, mais aussi avec les mécanismes d'interaction dans l'EV, devraient être pris en compte.

L'utilisation d'une technique d'interaction est sous-tendue par des fonctions sensori-motrices et cognitives (Wong, Chan, Li-Tsang, & Lam, 2009) dont des habiletés spatiales (Santos et al., 2009). En effet, l'utilisation de la technique d'interaction nécessite un mouvement moteur de

l'utilisateur pour contrôler le curseur de sélection affiché à l'écran. La trajectoire du curseur de sélection comprend un point de départ, des sous-mouvements séparés par des pauses et une arrivée sur la cible (Almanji, Davies, & Stott, 2014). Un sous-mouvement comprend une longueur et une direction (Hourcade, 2006) et plusieurs sont effectués avant d'atteindre la cible. De façon générale, un premier sous-mouvement avec une grande vélocité est effectué pour rapprocher le curseur de la cible, puis, d'autres sous-mouvements, plus lents, sont réalisés pour corriger la trajectoire du curseur jusqu'à la cible (Balakrishnan, 2004 ; Thompson, McConnell, Slocum, & Bohan, 2007). De plus, Hourcade, Nguyen, Perry, et Denburg (2010) ont montré que, chez la personne âgée, les performances à une tâche informatique de sélection de cibles sont corrélées avec des scores aux tests cognitifs. Vingt personnes âgées ont participé à cette étude. Il leur était demandé de sélectionner des cibles le plus rapidement possible avec une souris d'ordinateur. Lors de la tâche, une aide à la sélection pouvait se déclencher lorsqu'un algorithme détectait que l'utilisateur avait besoin d'une assistance. Une corrélation négative existe entre un score d'intelligence non-verbale et le nombre de fois où l'assistance est déclenchée (Hourcade et al., 2010). De mauvaises capacités sensori-motrices (Wong et al., 2009) et des fonctions cognitives altérées peuvent donc entraîner une utilisation non efficiente de la technique d'interaction.

Nous pouvons suggérer qu'une personne familière avec une technique d'interaction ressentira moins de charge mentale lors de son utilisation. En effet, même si l'apprentissage d'une technique d'interaction basée sur la capture des mouvements comme la kinect est aisée et s'apprend rapidement, on observe que son utilisation demande des efforts (D'Souza, Pathirana, McMeel, & Amor, 2011 ; Verhulst et al., 2016). Le degré de familiarité de l'individu intervient aussi avec d'autres types de dispositif électronique, comme l'utilisation d'un MP3 (N. E. Kang & Yoon, 2008). Dans cette étude, les participants ont été observés en train de réaliser des tâches, entre autre, sur un MP3. Par exemple les participants devaient trouver une chanson dans un dossier. N. E. Kang et Yoon (2008) ont montré que les sous-échelles "effort" et "frustration" du NASA-RTLX sont influencées par les connaissances antérieures des individus avec la technologie.

La charge mentale est aussi impactée par l'âge de l'individu selon les situations. Une personne âgée ressentira plus de charge mentale qu'un adulte plus jeune (Gonçalves & Cameirão, 2016). Ainsi, dans une situation de trafics complexes (intersections, partage de la route), les personnes âgées expérimentent plus de charge mentale que les adultes lorsqu'ils se déplacent sur un vélo électrique (Vlakveld et al., 2015). Au cours d'une autre étude, la charge mentale a été comparée lors d'une tâche sur un MP3 entre des jeunes adultes (21-29 ans) et des adultes d'âge

moyen (46-59 ans). Les adultes d'âge moyen ont rapporté, via le NASA-RTLX, avoir ressenti plus de charge mentale que les jeunes adultes. C'est en particulier le cas pour les sous-échelles « demande temporelle » et « performance » qui sont influencée par les différences d'âge. Au regard des résultats, les auteurs indiquent que les adultes d'âge moyen ont sous-estimé leur performance alors que leur temps de réalisation n'était pas significativement différent de celui des jeunes adultes (N. E. Kang & Yoon, 2008). De plus, des différences de charge mentale lors d'une tâche d'achats en ligne en RV ont été mise en avant selon le sexe des participants (Zhao, Shi, You, & Zong, 2017). Les hommes fournissent plus d'effort mental pour acheter un produit en ligne, surtout si celui-ci a un coût élevé, que les femmes (Zhao et al., 2017).

La charge mentale allouée à une tâche informatisée, comme celle de RV semble être modulée par la technique en elle-même ainsi que par le profil de l'utilisateur : son expérience antérieure avec les interactions dans les EVs, son âge et son sexe. Ainsi selon les techniques d'interaction utilisées et les prédispositions des utilisateurs, la charge mentale pourrait être plus ou moins élevée. Le profil de l'utilisateur devrait donc faire l'objet d'une attention particulière dans les études sur la charge mentale, par exemple en effectuant une *baseline* du degré de familiarité en "novice" ou "expert".

TABLE 3.1 – Etudes évaluant la charge mentale.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Besançon et al. (2016)	Évaluer la performance et l'utilisabilité de techniques d'interaction	36 adultes (âge : 30.2 (8.7))	Tâche de manipulation avec la souris, une interface tangible et un écran tactile puis complétion NASA-TLX.	Les participants sont plus rapides avec la technique tangible qu'avec l'écran tactile. Ils sont plus rapides avec l'écran tactile qu'avec la souris. La technique tangible demande moins de charge mentale que l'écran tactile qui en demande moins que la souris.	Chaque technique d'interaction a ses avantages et inconvénients en terme de performance et de charge mentale générée.
Fallahi, Motamedzade, Heidarimoghadam, Soltanian, Farhadian, et Miyake (2016)	Évaluer la charge mentale des opérateurs lors d'une activité de surveillance en centre de contrôle de la circulation urbaine	16 opérateurs (âge : 29,4 (2,61))	Mesure d'un électroencéphalogramme au cours de 4 conditions : repos, trafic léger, trafic dense et récupération. Mesure de la charge mentale avec le S-SWAT.	La charge mentale est plus élevée en situation de trafic dense que léger. Les modulations de charge mentale sont notables par le S-SWAT et une modification dans les bandes thêta et alpha.	Les situations plus complexes génèrent plus de charge mentale que les situations moins complexes.
George et al. (2018)	Appréhender la relation entre la présence, la charge mentale et le niveau attentionnel	non renseigné	Deux EVs sont utilisés : l'EV1 comprend une charge mentale élevée où le sujet doit repérer des cibles tandis que l'EV2 est une visite dans un musée. Le temps de réaction face à des stimuli a été enregistré. Le NASA-TLX et l'IPQ ont aussi été utilisés.	Il y a une différence de charge mentale entre l'EV 1 (tirer sur cible) et l'EV 2 (visite de musée). Il y a aussi une corrélation entre le sentiment de présence et la charge mentale dans l'EV1.	Il y a une relation entre la charge mentale et la présence.

TABLE 3.1 – Etudes évaluant la charge mentale.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Gonçalves et Cameirão (2016)	Évaluer l'utilisabilité d'une technique d'interaction.	90 âgés (âge : 70.2 (5.3))	Réalisation d'une tâche de sélection avec les pieds ou les mains puis passation du NASA-TLX	Il y a une meilleure utilisabilité pour les participants qui interagissent avec les pieds mais on ne trouve pas de différence au niveau de la charge mentale. Lorsqu'il s'agit de faire du glisser-déposer, les participants ont une meilleure utilisabilité et moins de charge mentale avec les pieds qu'avec les mains.	Les participants âgés interagissent de façon plus efficace avec les pieds qu'avec les mains.
Jaquess et al. (2017)	Évaluer la relation entre la charge mentale et les ressources attentionnelles	63 participants entre 19 et 26 ans	Tâche dans un simulateur de vol selon 3 degrés de difficultés (facile, neutre et difficile). Des données physiologiques sont enregistrées (ECG et EEG) pendant le simulateur puis le NASA-TLX est complété.	Les résultats au NASA-TLX augmentent entre les 3 conditions de difficulté de la tâche. Il n'y a pas de différence dans le HRV entre les 3 conditions de difficulté de la tâche. La performance au simulateur de vol diminue entre les 3 conditions de difficultés de la tâche. On a une corrélation négative entre la charge mentale et la réserve attentionnelle (mesurée par EEG).	Lorsque la charge mentale augmente, les ressources attentionnelles diminuent.

TABLE 3.1 – Etudes évaluant la charge mentale.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
N. E. Kang et Yoon (2008)	Appréhender la performance des participants en fonction de leur âge.	Jeunes adultes (20–29 ans) et adultes d'âges moyen (46–59 ans).	Réalisation de tâches sur un MP3 puis passation du NASA-RTLX.	Les âgés ressentent plus de charge mentale que les jeunes (demande temporelle et performance). La frustration est liée à la familiarité avec l'usage des ordinateurs. Il n'y pas de différence significative entre les âges pour le temps de complétion.	La charge mentale est liée à l'âge des participants.
Ma et Kaber (2006)	Appréhender la présence des éléments sensoriels sur le sentiment de présence et ses relations avec la performance à une tâche dans l'EV.	32 étudiants (âge : 23.7) qui sont habitués aux jeux video mais non à jouer au basket.	Les participants font un jeu de basket virtuel puis remplissent les questionnaires de présence et de charge mentale.	La vue en première personne génère plus de charge mentale que la vue en troisième. La présence et la charge mentale sont corrélées. La charge mentale négativement corrélée à la performance.	Il y a une relation entre la charge mentale et la présence.
Ramírez-Fernández et al. (2015)	Évaluer la performance et l'utilisabilité de techniques d'interaction dans le cadre de la rééducation motrice.	15 adultes d'âge moyen sains (âge : 55.73 (7.56)) et 15 adultes âgés avec des déficits moteurs (âge : 78.80 (11.30)).	Les sujets réalisent une tâche motrice avec un dispositif haptique et le Leap Motion. Le NASA-TLX est ensuite complété.	Les résultats montrent que si les sujets âgés sains trouvent que la charge mentale générée par le dispositif haptique Novint Falcon et le Leap Motion est sensiblement la même. Les sujets avec un déficit moteur trouvent que le Leap Motion génère plus de charge mentale que le dispositif haptique. Il n'y a pas de différence pour le temps de complétion.	Chaque technique d'interaction a ses avantages et inconvénients en terme de performance et de charge mentale générée.

TABLE 3.1 – Etudes évaluant la charge mentale.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Speicher et al. (2017)	Évaluer la performance dans un dispositif immersif et non-immersif.	12 adultes (âge = 25.8 (8.6))	Réalisation d'une tâche de manipulation avec un dispositif immersif (HMD + leap motion) et un non-immersif (kinect). Le NASA-TLX est ensuite complété.	Les participants sont meilleurs avec la kinect mais préfèrent le dispositif immersif. En terme de charge mentale, elle est moins élevée en immersif.	Chaque technique d'interaction a ses avantages et inconvénients en terme de performance et de charge mentale générée.
Stevens et al. (2016)	Appréhender la possibilité de réaliser un entraînement militaire en virtuel.	98 soldats (spécialistes ou sergents) américains (âge : 27.7 (4.7))	Réalisation d'un entraînement militaire en réel et en virtuel puis passation du NASA-TLX.	A une tâche d'entraînement, les soldats ont ressenti plus de charge mentale en virtuel (clavier/souris) qu'en réel sauf pour les sous échelle effort et physique du NASA.	La technique d'interaction peut générer de la charge mentale
Verhulst et al. (2016)	Évaluer la performance et l'utilisabilité de techniques d'interaction.	80 jeunes adultes (âge : 23 (3.45))	Tâche dans un supermarché virtuel puis passation du NASA-TLX.	Les participants sont plus rapides avec le gamepad et le <i>head tracking</i> . Il y a moins de frustration avec le gamepad que la kinect et moins d'effort avec le gamepad que la kinect.	Chaque technique d'interaction a ses avantages et inconvénients en terme de performance et de charge mentale générée
Vlakveld et al. (2015)	Appréhender la charge mentale en fonction de l'âge des participants et de la complexité des situations.	29 adultes âgés (âge : 69,6 (4,2)) et 29 adultes d'âge moyen (âge : 37,7 (4,3))	Promenade à vélo en situation de trafic léger ou dense puis passation du NASA-TLX.	La charge mentale est plus élevée en trafic dense que léger. La charge mentale est plus élevée pour les participants âgés.	La charge mentale est plus élevée en situation complexe et chez les personnes âgées.
Yousef Zade et al. (2016)	Appréhender la relation entre les erreurs professionnelles des infirmiers et la charge mentale perçue.	150 infirmiers	Passation du NASA-TLX et du Cognitive Failure Questionnaire (CFQ).	Il y a une corrélation entre la demande physique et le CFQ et une corrélation entre la frustration et le CFQ.	La hausse de la charge mentale chez les infirmiers peut conduire à des erreurs lors de la réalisation de leur tâches professionnelles

TABLE 3.1 – Etudes évaluant la charge mentale.

Auteur(s)	Objectif	Population	Méthode	Résultats	Discussion
Zhao et al. (2017)	Appréhender la charge mentale en fonction du sexe et de la tâche.	36 jeunes adultes (âge : 29.81 (2.61))	Les participants font une tâche de shopping en ligne, ils doivent acheter un sofa (produit cher) et un miroir (produit peu cher). Ils complètent ensuite le NASA-TLX.	La charge mentale, mesurée avec le NASA-TLX, des hommes est plus élevée que celle des femmes en VR lors d'une tâche d'achats en ligne. La charge mentale des hommes est plus élevée principalement lors des achats chers car les hommes apparaissent plus concentré sur les aspects pratiques de l'achat et les femmes plus sur les aspects émotionnel, donc la charge des hommes est plus élevée.	La charge mentale dépend du sexe des participants et de la tâche à effectuer.

3.5 Synthèse

La charge mentale est un concept multimodal que l'on peut définir comme la quantité de ressources cognitives nécessaires pour compléter une tâche. Le degré de charge mentale requis varie selon les caractéristiques (1) de la tâche, (2) de l'environnement et (3) de l'individu. Il existe plusieurs façons d'évaluer la charge mentale, cela inclue des mesures : comportementales, physiologiques, subjectives et basées sur l'analyse des données. Chaque méthode d'évaluation apporte des avantages et des inconvénients. Ainsi, combiner plusieurs types de mesures apporte des résultats plus sensibles. Par exemple, les mesures cardiaques sont certes objectives mais ne sont pas assez spécifiques car elles varient avec d'autres phénomènes comme l'état émotionnel (Chanquoy et al., 2007) et également d'une personne à l'autre (Miyake, 2001). Pour une mesure plus sensible de la charge mentale, les mesures physiologiques peuvent donc être analysées avec d'autres variables, comme celles de performance et de charge mentale subjective (Cegarra & Chevalier, 2008). Dans le cas de notre procédure expérimentale, nous nous centrerons sur une mesure de la charge mentale objective avec les signaux ECG et subjective avec l'utilisation du questionnaire le NASA-TLX.

Evaluation de la performance lors d'IADL

4.1 Introduction

L'utilisation de la RV et des EVs est de plus en plus fréquente pour apporter une aide à l'évaluation cognitive et plus spécifiquement neuropsychologique du participant. En effet, les EVs permettent d'évaluer celui-ci dans des situations issues de la vie quotidienne qui sont standardisées et contrôlées. Les tests virtuels sont donc utilisés pour renseigner la performance du sujet durant l'évaluation, pour ensuite la transposer dans l'environnement réel.

Si l'utilisation des EVs pour étudier la performance d'un individu soulève de nombreuses questions, elle présente plusieurs avantages dont la possibilité d'enregistrer un grand nombre de données. Cependant, la multitude de données et de variables associées rend la performance du participant plus difficile à analyser et soulève des interrogations. Par exemple, quelles sont les variables les plus sensibles pour prédire la performance du participant et son appartenance à un groupe présentant des difficultés d'ordre cognitives. Ainsi Bowman et Hodges (1999) posaient la question : *"but what is performance ?"*¹. Il est en effet primordial de se questionner sur la finalité de l'EV utilisé pour mesurer la performance, car c'est ce qui va guider le développement de l'EV d'évaluation mais aussi la collecte des données qui serviront à qualifier la performance de l'individu. Dans les EVs d'évaluation cognitive, la performance est majoritairement considérée comme le temps de réalisation mais d'autres variables devraient être aussi prises en compte tel le choix de la technique d'interaction. En effet, la performance peut être influencée par différents facteurs tels que (Bowman & Hodges, 1999) :

- *Les caractéristiques de la tâche* : par exemple, la distance parcourue ou la taille des objets manipulés.
- *Les caractéristiques de l'environnement* : par exemple, la difficulté.

1. Mais qu'est-ce que la performance ?

- *Les caractéristiques de l'utilisateur* : par exemple, ses capacités cognitives et physiques.
- *Les caractéristiques du système*. Par exemple, la fréquence de rafraîchissement.

La performance du participant relève donc de plusieurs facteurs qu'il est nécessaire de prendre en compte pour l'appréhender correctement.

Dans ce chapitre nous proposons une stratégie de sélection des variables pour qualifier la performance du participant (4.2). Ensuite les différentes variables utilisées dans des EVs écologiques sont décrites de façon détaillée (4.3) et, enfin, elles seront mises en relation avec l'utilisabilité de la technique d'interaction (4.4).

4.2 Sélection des variables pour la qualification de la performance

4.2.1 Sélection des objets d'étude

L'usage d'EVs destinés à l'évaluation neuropsychologique est de plus en plus développé. Une revue de ces tests a été proposée par Zygouris et Tsolaki (2015). Deux types de tâches sont généralement proposés pour les mesures neuropsychologiques : les écologiques évaluant les IADL et les non-écologiques. Parmi ces derniers, une majorité sont des EVs 2D qui sont des numérisations de tests neuropsychologiques classiques ou de tests qui en sont proches. Au cours de ces évaluations cognitives, les participants doivent répondre à différentes consignes et réaliser plusieurs tâches. Leur production est ensuite quantifiée et analysée par les expérimentateurs. Par exemple, le *Trail Making Test* peut être numérisé. Ainsi, le temps est enregistré par l'ordinateur et les erreurs peuvent aussi être automatiquement additionnées puis notifiées par le programme. L'expérimentateur peut alors directement se focaliser sur l'appréhension des résultats et ne perd pas de temps à effectuer les corrections des tests. Lors de ces évaluations, la performance du participant correspond généralement, d'une part, au temps mis pour donner une réponse à la consigne et d'autre part à l'exactitude de cette réponse. Les modalités d'évaluation de la performance du participant sont donc décidées par les concepteurs du test et ne sont pas forcément représentatives d'une production qui aurait été obtenue en réel. En effet, les tests peuvent être éloignés, en termes de demande cognitive et de sens, des activités de la vie quotidienne.

Dans le cas où il est souhaité évaluer la performance du participant en condition écologique, lors d'IADL, il est nécessaire de définir la façon d'évaluer l'utilisateur. En effet, les participants sont familiarisés avec ces tâches qu'ils réalisent dans leur vie quotidienne. Ils ont donc déjà des scripts d'activités de la réalisation de la tâche en mémoire.

Il est donc possible de proposer aux participants différents types de consignes : des consignes fermées où le participant doit réaliser la tâche de la manière dont elle est demandée par l'expérimentateur (par exemple : veuillez préparer une tasse de café avec du lait et du sucre) ou des consignes ouvertes où le participant peut réaliser la tâche selon ses habitudes propres (par exemple : veuillez préparer une tasse de café). Dans le cas où les consignes données sont ouvertes, l'expérimentateur doit avoir bien défini auparavant les attendus et les actions pouvant ou non être considérées comme des erreurs. En effet, la production du participant peut différer de celle que l'expérimentateur attend car elle ne correspond pas aux habitudes de ce dernier. La production du participant n'est pour autant pas une erreur. Par exemple, un participant peut dresser la table en ajoutant des petites cuillères et un autre ne pas en mettre. Si la consigne n'est pas explicite à ce sujet, aucun participant ne commet d'erreur. L'expérimentation lors d'IADL n'est donc pas aisée et suppose que les concepteurs du test réfléchissent longuement en amont pour définir les consignes et la façon dont les tâches vont être quotées.

Pour mieux comprendre la façon dont la performance du participant est mesurée lors d'IADL en virtuel, nous avons effectué une revue de littérature de ces EVs. Les EVs simulant la conduite automobile ont été exclus de la compilation d'articles pour des raisons de caractéristiques du système. En effet dans ces études, la tâche est principalement réalisée dans un simulateur de conduite. Les comparaisons de résultats entre études ne sont donc pas possible du fait de la spécificité des techniques d'interaction. Il existe peu d'EVs évaluant le comportement des participants lors de la réalisation d'IADL ; bien que ces études soient primordiales afin de soulever de possibles difficultés dans leur vie quotidienne. Actuellement trois types d'EVs semblent destinés à cette forme d'évaluation : les supermarchés virtuels, les cuisines virtuelles et les bâtiments virtuels comme une bibliothèque virtuelle (Tab. 4.1).

Auteurs	Interaction	EV	Variables enregistrées
Allain et al. (2014)	souris/clavier	Cuisine	- Temps de réalisation - Nombre et type d'erreurs
Aubin et al. (2015)	souris/clavier	Supermarché	- Temps de réalisation - Distance parcourue - Nombre d'achats - Nombre d'actions correctes - Nombre d'erreurs - Nombre de pauses - Durée des pauses - Temps pris pour payer
Besnard et al. (2016)	souris/clavier	Cuisine	- Temps de réalisation - Score d'accomplissement - Nombre et type d'erreurs
Dehn et al. (2018)	Joystick et écran tactile (octavis)	Supermarché	- Temps de réalisation - Distance parcourue - Nombre d'achats
Gamito et al. (2015)	<i>inconnu</i>	Cuisine	- Temps de réalisation - Nombre d'erreurs
Greenwood et al. (2016)	Joystick	Supermarché	- Temps de réalisation - Nombre d'erreurs - Nombre d'allées empruntées
Josman et al. (2008)	souris/clavier	Supermarché	- Temps de réalisation - Distance parcourue - Nombre d'achats - Nombre d'actions correctes - Nombre d'erreurs - Nombre de pauses - Durée des pauses - Temps pris pour payer
N. E. Kang et Yoon (2008)	joystick et HMD	Supermarché	- Temps de réalisation

Auteurs	Interaction	EV	Variables enregistrées
			<ul style="list-style-type: none"> - Distance parcourue - Erreurs d'interaction - Nombre de collisions - Score d'accomplissement
Klinger et al. (2006)	Souris/clavier	Supermarché	<ul style="list-style-type: none"> - Temps de réalisation - Distance parcourue - Nombre d'achats - Nombre d'actions correctes - Nombre d'erreurs - Nombre de pauses - Durée des pauses - Temps pris pour payer
Kurtz et al. (2006)	Joystick	Appartement	<ul style="list-style-type: none"> - Temps de réalisation - Nombre d'erreurs - Distance parcourue - Utilisation d'indices nombre de consultations de la liste et durée
Larøi et al. (2010)	Gamepad (flèches directionnelles)	Supermarché	<ul style="list-style-type: none"> - Temps de réalisation - Nombre d'erreurs - Nombre d'achats - Nombre d'erreurs corrigées - Nombre d'aller-retour dans une allée - Nombre d'entrée dans une allée non pertinente - Nombre de fois où une allée non pertinente est visualisée - Nombre de fois où la liste de course est visualisée - Utilisation d'indices nombre de consultations

Auteurs	Interaction	EV	Variables enregistrées
			de liste et sac et durée
Martínez-Pernía et al. (2017)	Ecran tactile	Cuisine	- Temps de réalisation - Nombre et type d'erreurs
McGeorge et al. (2001)	souris/clavier	Université	- Nombre de tâches réalisées - Score de planification
Martono et al. (2016)	Ecran tactile	Cuisine	- Temps de réalisation - Nombre et type d'erreurs - Vitesse des mouvements des doigts
Nir-Hadad et al. (2017)	Ecran tactile	Supermarché	- Temps de réalisation - Distance parcourue - Nombre d'erreurs - Temps avant le premier achat
Okahashi et al. (2013)	Ecran tactile	Supermarché	- Temps de réalisation - Utilisation d'indices nombre de consultations liste et sac
			- Nombre de mouvements vers l'avant -demi-tour -Nombre d'achats corrects -Temps total -Temps dans les magasins -Temps sur la route -Temps moyen par magasin
Ouellet et al. (2018)	Marche réelle et HMD	Supermarché	- Temps de réalisation -Temps d'initiation au jeu - Nombre d'achats - Nombre d'erreurs
T. Parsons et McMahan (2017)	HMD + manettes	Supermarché	- Utilisation d'indice nombre de consultation de liste et carte - Nombre d'achats - Nombre d'erreurs

Auteurs	Interaction	EV	Variables enregistrées
			- Total dépensé
Renison et al. (2012)	Gamepad	Bibliothèque	- Score d'accomplissement
Vallejo et al. (2017)	Ecran tactile	Cuisine	- Temps de réalisation - Nombre d'erreurs
Werner et al. (2009)	Souris/clavier	Supermarché	- Temps de réalisation - Nombre d'achats - Nombre d'erreurs - Nombre de pauses - Durée des pauses
Zygouris et al. (2015)	Ecran tactile	Supermarché	- Temps de réalisation - Nombre d'erreurs - Nombre d'achats

TABLE 4.1 – Etudes évaluant les IADL en virtuel.

Nous avons répertoriés les EVs présentés dans le tableau (4.1) pour tenter de mieux comprendre le fonctionnement de l'évaluation de la performance du participant au cours d'IADL. En effet, de nombreuses questions sont soulevées dans le cas de la RV comme outil d'évaluation cognitive. Ces préoccupations psychométriques apparaissent fondamentales sur le plan de la validité de la mesure. La validité réfère à la capacité d'un instrument à mesurer ce qu'il prétend mesurer. Elle doit, au terme du processus, être représentative du construit théorique auquel l'instrument réfère. En effet peu d'études comparent les performances en réel et en virtuel pour une même activité complexe de la vie quotidienne et les résultats s'avèrent contradictoires. Il est donc difficile de savoir si les EVs mesurant les IADL sont des outils valides et s'ils représentent bien la performance qu'aurait obtenu le participant dans une situation équivalente en réel.

4.2.2 Comparaison de la performance à une même tâche en réel et en virtuel

Pour comprendre la performance du participant lors de la réalisation d'IADLs dans un EV et vérifier si cette mesure est valide, les chercheurs ont voulu comparer la production de l'utilisateur à une même tâche en virtuel et en réel. Nous avons répertorié, à ce jour, six études qui comparent virtuel et réel à une même tâche : la première étude est un test d'errances multiples réalisé dans des bureaux d'un département de psychologie d'une université anglaise (McGeorge et al., 2001), la seconde se déroule dans une bibliothèque (Renison et al., 2012), puis trois évaluent la performance des participants lors de la réalisation de tâches dans une cuisine virtuelle (Allain et al., 2014 ; Besnard et al., 2016 ; Vallejo et al., 2017) et une dernière lors d'une tâche d'achats dans un magasin (Greenwood et al., 2016).

Les études de McGeorge et al. (2001), de Renison et al. (2012) et de Vallejo et al. (2017) ne relèvent pas de différence de performances que la tâche soit réalisée en réel ou en virtuel. La performance obtenue en virtuel est donc représentative de celle en réel. Cependant, les autres études présentent des résultats contradictoires entre les deux conditions. En effet, lors de la préparation du café, les participants font plus d'erreurs et mettent plus de temps pour réaliser l'action en virtuel qu'en réel (Allain et al., 2014 ; Besnard et al., 2016). Malgré ces différences, nous constatons que les scores obtenus dans les deux conditions sont significativement corrélés, ce qui suggère un lien entre les deux conditions. Enfin, lors de la tâche dans le magasin, les participants mettent plus de temps en réel qu'en virtuel à accomplir les actions demandées (Greenwood et al., 2016).

Ces différences de résultats peuvent s'expliquer de plusieurs façons. D'une part l'activité d'achats dans un supermarché est une activité plus complexe que celle de préparer un café qui est une tâche en 2D, sans navigation dans l'espace et sans gestion de la profondeur. En effet, lorsqu'un individu réalise ses courses en réel, il est sollicité par de nombreuses stimulations sensorielles comme le volume sonore, le grand nombre de produits et doit adapter son comportement en fonction des nombreux autres clients. Tandis que dans la tâche de préparation du café dans une cuisine virtuelle, le participant perçoit moins d'informations et n'a donc pas à gérer d'informations non-pertinentes. La grande complexité du monde réel est difficilement reproductible en termes d'interaction. Ainsi, nous concevons que lors d'une activité complexe en termes d'interaction (navigation, sociale, etc.), le participant présente plus de difficultés en réel qu'en virtuel car il aura plus de stimuli distrayants. Dans le cas d'une activité moins complexe comme préparer une tasse de café, le participant est soumis à moins de stimulations sensorielles puisqu'il peut agir seul et que la présence des expérimentateurs ne cause généralement pas de bruits ni de distractions, et ne nécessite pas d'y être attentif.

D'autre part l'utilisation d'une technique d'interaction, comme la souris, en RV, peut entraîner des erreurs d'utilisabilité. En effet, dans les études de Allain et al. (2014) sur des patients avec la maladie d'Alzheimer et des contrôles, et de Besnard et al. (2016) sur des patients cérébrolésés et des contrôles, les participants font plus d'erreurs en virtuel qu'en réel et les auteurs suggèrent que cela pourrait être dû à l'utilisation de la souris. Ainsi, l'utilisation d'une technique d'interaction pour réaliser une tâche sur un ordinateur peut venir gêner le participant au point de lui faire commettre plus d'erreurs.

Pour mieux comprendre les différences de résultats entre les études et approfondir le fait que certaines études trouvent des différences entre réel et virtuel et d'autres pas, il est nécessaire de s'intéresser à la façon dont la performance est mesurée. Pour rappel, les études de McGeorge et al. (2001), de Renison et al. (2012) et de Vallejo et al. (2017) ne relèvent pas de différence de performances en réel et en virtuel. Les deux premières études étaient réalisées auprès d'une population de patients cérébrolésés et des contrôles, tandis que la dernière étude était faite des personnes âgées. Ces études mesurent la performance par un score total. Ainsi le nombre de bonnes actions effectuées est évalué en réel et en virtuel et la comparaison entre ces scores ne présentent pas de différence significative (McGeorge et al., 2001 ; Renison et al., 2012). La dernière étude compare le nombre d'erreurs en réel et en virtuel et le résultat est équivalent dans les deux conditions (Vallejo et al., 2017). Le temps de réalisation est aussi mesuré et les

auteurs indiquent un temps moyen de 19 minutes pour réaliser le plat de spaghetti à la sauce tomate en réel contre 3 minutes en virtuel. Le temps peut être plus long en réel puisque les EVs ont la possibilité de simplifier les interactions, ainsi les différents paramètres de la simulation ne sont pas forcément réalistes. Par exemple, il n'est pas forcément nécessaire d'attendre plusieurs minutes que l'eau entre en ébullition en virtuel.

Appréhender la performance du participant semble dépendre de la façon dont celle-ci est mesurée et du nombre de variables utilisées. Si la performance de l'utilisateur est mesurée globalement, par exemple, par un score d'accomplissement ou le nombre de bonnes actions. On ne trouve pas de différence entre réel et virtuel et les résultats obtenus dans un des deux environnements sont donc transposables à l'autre. En revanche si la performance est mesurée par de nombreuses variables, comme le temps de réalisation et la distance parcourue, des différences entre réel et virtuel vont apparaître. Ainsi, lorsque l'on s'intéresse au succès de la tâche, elle semble aussi bien réussie dans les deux environnements réel et virtuel. Alors que d'autres variables vont montrer, par exemple, que les utilisateurs sont plus lents en virtuel. Il est donc nécessaire de comprendre ce que mesure les différentes variables pour appréhender celles qui sont le plus représentatives de la performance de l'utilisateur.

4.2.3 Limites de l'approche

La comparaison d'une même tâche effectuée en réel et en virtuel indique que les performances peuvent être différentes entre les deux conditions en fonction de la technique d'interaction employée, du type d'environnements (activités en présence d'autrui ou non) et des variables utilisées pour mesurer la performance (peu ou de nombreuses variables). Les résultats issus de la section précédente nous indiquent que la performance peut être meilleure en virtuel qu'en réel car les EVs permettent de simplifier les interactions. Par exemple il y a moins de stimuli dans un supermarché virtuel. A contrario, la performance du participant peut aussi être moins bonne en virtuel du fait de l'utilisation d'une technique d'interaction. L'utilisation de la RV lors d'une tâche d'évaluation cognitive peut donc induire des erreurs liées à l'utilisabilité de la technique d'interaction et à la charge mentale générée par les IHM ou au contraire les minimiser grâce à un contrôle des stimuli par rapport à la réalité.

Pour alimenter cette réflexion, il est intéressant de mettre en relation l'étude de Besnard et al. (2016) où des patients cérébrolésés réalisaient une tâche en EV consistant à faire une

tasse de café et celle de Martínez-Pernía et al. (2017) qui ont développé une tâche similaire où des patients cérébrolésés devaient préparer une tasse de thé, en interagissant avec un écran tactile. Ceux-ci n'ont présenté aucune difficulté pour compléter la tâche et ont pu facilement interagir avec l'écran tactile. Mises en relation, ces deux études laissent supposer qu'il est plus facile d'interagir avec un écran tactile qu'avec la souris pour des sujets présentant des troubles cognitifs. Cette hypothèse peut être alimentée par la littérature en IHM sur les travaux portant sur la comparaison d'interfaces et de techniques d'interaction. En effet, lorsque l'on compare l'écran tactile et la souris comme technique d'interaction pour une même tâche, il apparaît que les participants sont plus rapides avec l'écran tactile qu'avec la souris (Besançon et al., 2016 ; Forlines et al., 2007), et ce, sûrement car les utilisateurs sont familiers avec les écrans tactiles qu'ils utilisent au quotidien avec les téléphones portables. Ces résultats sont transférables lors des tâches de la vie quotidienne comme la conduite automobile. En effet, si l'on compare l'action de garer sa propre voiture en réel, ou une voiture virtuelle sur un écran d'ordinateur ou en RV, il est possible de constater que les individus mettent plus de temps à compléter la tâche sur un écran d'ordinateur qu'en réel, alors qu'on ne trouve pas de résultats significativement différents entre la condition réelle et de RV (Read & Saleem, 2017). L'utilisation d'un dispositif de simulation non familier peut compliquer la tâche par rapport à une tâche en réel pour laquelle les utilisateurs manipulaient leur véhicule personnel, un outil donc familier.

Pour appréhender au mieux la performance d'un participant, il devient primordial de pouvoir différencier les erreurs liées à une perturbation cognitive de celles dues à l'utilisation d'une technique d'interaction ainsi que de prendre en compte les différences de charge mentale induites lors de la réalisation d'une tâche en réalité virtuelle. Pour mieux comprendre les erreurs commises en RV, il est nécessaire de prendre en compte l'utilisabilité d'une technique d'interaction ainsi que ses liens avec les habiletés cognitives et motrices. Pour ce faire, une attention a été portée aux variables utilisées pour renseigner la performance d'un sujet dans une tâche de RV.

Les variables utilisées pour caractériser la performance du participant commise lors de la réalisation de la tâche sont majoritairement le temps de réalisation et les erreurs. Une analyse plus approfondie de ces deux variables est proposée ci-dessous.

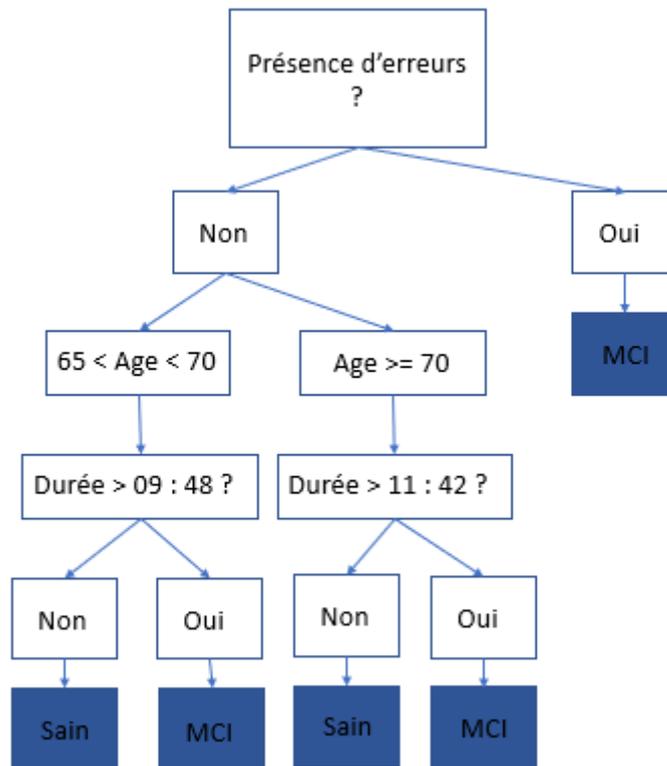


FIGURE 4.1 – Algorithme pour le dépistage des troubles cognitifs légers en fonction du temps mis pour réaliser la tâche (Zygouris et al., 2015).

4.3 Variables utilisées dans les EVs lors d'IADL

4.3.1 Temps de réalisation

Le temps de réalisation de la tâche est une variable essentielle à considérer car le temps mis pour réaliser une tâche de navigation équivaut à la performance du sujet (Patel & Vij, 2010). En effet, plusieurs études ont montré que le temps de réalisation était corrélé à des mesures cognitives comme le MMSE (Allain et al., 2014 ; Tong, Chignell, Tierney, & Lee, 2016 ; Zygouris et al., 2015) et la WAIS (Besnard et al., 2016) et qu'il permettait de différencier des groupes d'individus jeunes ou âgés (Banville et al., 2017 ; Sangani, Fung, Kizony, Koenig, & Weiss, 2013 ; Verhulst et al., 2017 ; Zygouris et al., 2015) et sains ou pathologiques (Allain et al., 2014 ; Tong et al., 2016 ; Werner et al., 2009 ; Zygouris et al., 2015). Par exemple Zygouris et al. (2015) proposent un algorithme (Fig. 4.1) qui permet de dépister les troubles cognitifs légers chez les sujets âgés à l'aide de la variable temps de réalisation, de sorte qu'un temps de réalisation plus long est lié à des fonctions cognitives plus ou moins perturbées.

Selon la méthode de dépistage de Zygouris et al. (2015), il est intéressant de voir que l'algo-

rithme présente le temps de réalisation comme une variable pouvant discriminer un participant au vieillissement normal d'un autre au vieillissement cognitif perturbé en se basant sur le temps mis pour compléter la tâche et l'âge du participant. En effet, plus le sujet est âgé plus il mettra de temps pour réaliser la tâche. D'une part, car il doit s'adapter à une interface qu'il ne connaissait pas avant et, d'autre part, car le vieillissement est caractérisé par un ralentissement global du traitement de l'information. Avec l'avancée en âge, le participant a besoin de plus de temps pour traiter les informations. Cet algorithme permet donc de présenter un seuil à partir duquel la performance du participant est modulée par un déclin cognitif de celle qui ne l'est pas. La variable temps de réalisation s'avère, dans cet exemple, être primordiale pour caractériser la performance d'un participant. Cependant, il faut noter que le temps de réalisation sera modulé par la familiarité d'un participant avec une technique d'interaction. En effet, il est supposé qu'un participant familier avec une technique d'interaction sera plus à l'aise et mettra moins de temps pour compléter la tâche qu'un participant novice.

4.3.2 Types d'erreurs

Un système de cotation circonspect permet d'appréhender correctement la performance de l'utilisateur et d'apporter une aide, pour mieux comprendre le comportement, cibler le diagnostic et orienter la rééducation (Seligman, Giovannetti, Sestito, & Libon, 2014). Dans cette optique, des erreurs à des tâches de la vie de tous les jours ont été modélisées et analysées (Giovannetti, Bettcher, Brennan, Libon, et al., 2008 ; Giovannetti, Bettcher, Brennan, Libron, et al., 2008 ; Giovannetti et al., 2012). Ainsi le modèle omission-commission proposé par Giovannetti et son équipe, repose sur le principe que des déficits en mémoire épisodique et des fonctions exécutives entraînent des difficultés dans la réalisation des actions de la vie quotidienne (Bailey, Kurby, Giovannetti, & Zacks, 2013 ; Giovannetti, Bettcher, Brennan, Libron, et al., 2008 ; Giovannetti et al., 2012 ; Kessler, Giovannetti, & MacMullen, 2007b). Une erreur d'omission se produit quand le participant omet la réalisation d'une étape nécessaire à la réalisation de la tâche tandis qu'une erreur de commission est produite lorsqu'une ou plusieurs des sous-étapes ne sont pas réalisées correctement mais n'empêche pas pour autant de mener la tâche à terme. Par exemple, dans une tâche préparation d'un café, mettre en marche la cafetière avant d'y ajouter l'eau est une erreur de commission : les étapes ne sont pas réalisées dans le bon ordre mais tant que l'eau est mise dans la cafetière, le café pourra être préparé. Alors que le fait de ne pas mettre d'eau du tout dans la machine ne permet pas d'avoir de café, c'est donc une erreur d'omission.

Les erreurs commissives peuvent se décliner en différentes sous-catégories :

- *Les persévérations* où les participants réalisent plus de fois que nécessaire une même étape (fermer puis rouvrir le réservoir à filtre).
- *Les omissions-anticipations* où les participants réalisent une étape dans un ordre non conventionnel (mettre du café avant le filtre à café).
- *Les substitutions d'objets* où les participants utilisent un objet alternatif à celui utilisé de façon conventionnelle (utiliser un couteau pour mélanger son café).

D'autres erreurs commissives existent mais sont anodines, particulièrement en RV, car elles s'intéressent aux praxies du participant qui ne sont pas sollicitées dans un EV dans lequel le sujet interagit avec une technique d'interaction telle que la souris. Par exemple, les substitution de gestes où le participant utilise un objet avec des gestes inappropriés comme prendre une cuillère à l'envers. De plus, la production de ces erreurs demande une liberté d'action qui n'est pas forcément possible avec la façon dont l'EV a été codé. En effet, les EVs peuvent être développés de sortes à faciliter les interactions, par exemple en ne permettant pas de déplacer des items que sur les axes vertical et horizontal (x et y). L'utilisateur ne doit donc pas gérer en plus la profondeur (axe z) et des erreurs peuvent être évitées, comme une erreur de mauvaise orientation spatiale où le participant met le filtre dans le mauvais sens dans le réceptacle à filtres.

Dans la suite de leurs travaux, les auteurs ont proposé une nouvelle analyse des erreurs, basée cette fois sur les micro-erreurs qui sont plus sensibles pour repérer le déclin cognitif (Seligman et al., 2014). Ces micro-erreurs sont des actions en trop dans la réalisation de la tâche, ou un mauvais séquençage des étapes mais qui ne rentrent pas dans les critères des erreurs de commission, ou encore l'initiation d'une erreur mais dont l'action n'est pas aboutie. En RV ces erreurs peuvent être très proches d'une mauvaise utilisabilité de la technique d'interaction. En effet, le curseur de la souris peut se rapprocher nettement d'une cible sans pour autant la sélectionner car le participant maîtrise mal la technique d'interaction ou car il pensait sélectionner la cible mais s'est rétracté. Dans le premier cas c'est une erreur d'utilisabilité et dans le second une micro-erreur.

Une autre analyse des erreurs a été réalisée par Hendry et al. (2016) avec le *Home-Based Cooking Task* (HBCT) qui est une version de la tâche de la cuisine réalisable au domicile du

participant et non spécifiquement en institution comme à l'hôpital (Chevignard et al., 2008). Lors de cette dernière, les participants se situent dans une cuisine dans un hôpital et doivent réaliser un gâteau au chocolat et une omelette. Le HBCT se base donc sur les mêmes instructions et matériels que la tâche de Chevignard et al. (2008) mais se déroule au domicile du participant. Le HBCT est donc un test écologique des fonctions exécutives qui prend en compte les variables suivantes :

- *Erreur d'omission.*
- *Addition* : une étape est réalisée alors qu'elle n'est pas nécessaire à la réalisation de la tâche (mettre des ingrédients en trop).
- *Commentaire/question* : poser des questions ou avoir des commentaires adressés à l'expérimentateur tandis que les participants doivent agir comme s'ils étaient seuls (où est le sucre?).
- *Substitution* : erreur dans la séquence d'étapes à suivre, utiliser le mauvais objet.
- *Estimation* : mésestimer les quantités demandées.
- *But* : la tâche est-elle complétée ou incomplète?
- *Avoir un comportement dangereux envers soi ou l'entourage.*
- *Temps de réalisation.*

Les variables utilisées par Chevignard et al. (2008), puis Hendry et al. (2016) sont proches de celles utilisées par Giovannetti, Bettcher, Brennan, Libron, et al. (2008) ; Giovannetti et al. (2012), mais différent dans le fondement théorique. En effet, selon le modèle de Giovannetti, Bettcher, Brennan, Libron, et al. (2008) ; Giovannetti et al. (2012), une difficulté à réaliser une IADL est due à des déficits en mémoire épisodique et des fonctions exécutives, tandis que selon ces auteurs, ces difficultés sont dues non pas à un domaine cognitif spécifique mais à des perturbations dans un ensemble de domaines cognitifs. Pour d'autres auteurs, la réalisation d'une erreur d'omission est liée à un déficit en mémoire sémantique (Hendry et al., 2016). En effet, la réalisation d'une tâche de la vie quotidienne demande que le participant ait des connaissances sur la façon de mener à bien la tâche mais aussi les capacités de retenir en mémoire de travail les consignes liées à l'activité ainsi que les étapes nécessaires pour la compléter (Hartmann, Goldenberg, Daumüller, & Hermsdörfer, 2005). Ainsi les omissions dans la réalisation de la tâche

peuvent être dues à un déficit en mémoire sémantique du participant (Bier & Macoir, 2010). Par exemple, celui-ci n'arrive pas à se souvenir qu'il est nécessaire de mettre, entre autres, des assiettes sur la table pour pouvoir manger. Cependant, ne pas mettre de petite cuillère en prévision d'un dessert n'est pas en soi une erreur d'omission reliée à la mémoire sémantique puisque les personnes mettent ou ne mettent pas de petite cuillère sur la table en fonction de leurs habitudes. Notons tout de même que, si la consigne avait stipulé de disposer l'ensemble des couverts sur la table, petites cuillères comprises, ne pas les avoir mises serait bien une erreur d'omission. Mais dans ce cas, non plus relié à la mémoire sémantique mais à la mémoire de travail. Le participant aurait oublié une partie de la consigne. Cela laisse supposer que lorsqu'on guide le participant par une consigne, cela n'induit plus chez lui les mêmes processus cognitifs. En effet, les éléments d'une consigne font appel à de la mémoire de travail alors que la réalisation d'une activité avec une consigne ouverte fait appel à des scripts des activités de la vie quotidienne en mémoire à long terme. Ainsi, demander au participant de faire un café avec du lait et du sucre ne fait pas forcément appel aux habitudes de vie de celui-ci et le maintient de la consigne consomme des ressources en mémoire de travail.

Lorsque l'on souhaite évaluer l'adaptabilité de celui-ci dans sa vie quotidienne, il est important de mesurer, non pas si celui-ci sait suivre une consigne, mais s'il arrive à mener à bien une activité, comme il le ferait chez lui, en faisant appel à ses habitudes et scripts de réalisation des activités. Pour cela, il est intéressant de donner des consignes générales aux participants afin de les laisser agir au maximum comme dans la vie réelle, mais aussi des consignes spécifiques pour surcharger la mémoire de travail du participant et donc la complexité du test.

L'analyse des erreurs en virtuel est liée aux types de consignes. Ainsi, dans le cas d'une consigne spécifique, les erreurs peuvent être définies à l'avance par l'examineur puisqu'elles correspondent à un non-respect de la consigne ou à une déviation de celle-ci. Cependant, dans le cas où la consigne est ouverte, par exemple « Mettez la table pour deux personnes », et ne donne pas de précisions spécifiques, le participant réalisera a priori la tâche comme il a l'habitude de l'effectuer à son domicile. De plus, l'utilisation des EVs peut faciliter les interactions avec les objets 3D de sorte qu'il n'est pas possible d'observer toutes les erreurs possibles en réel, en virtuel. Par exemple, pour simplifier les interactions dans un EV, les concepteurs ont pu faire le choix de ne pas implémenter de manipulation d'items. Ainsi lorsque le participant sélectionne un item puis le repose, ce dernier prend automatiquement la position et la rotation attendue. Le sujet ne pourra donc pas déposer une fourchette à l'envers, sur la table. Il y a d'autre part moins d'objets, particulièrement d'objets inutiles ou non rangés, en RV que dans la réalité d'une cuisine.

Les erreurs et le temps de réalisation sont deux variables majeures pour qualifier la performance du participant et détecter si l'individu présente des troubles cognitifs ou non. Cependant, ces 2 variables sont aussi dépendantes de l'utilisation de la technique d'interaction. Ces 2 seules variables pourraient cependant être insuffisantes pour caractériser correctement la performance du participant et demandent à être complétée par d'autres critères comme le profil de celui-ci. Par exemple, il est indispensable de comprendre, dans un premier temps le niveau d'expertise de l'utilisateur avec une technique d'interaction et avec l'utilisation des environnements virtuels 2D et 3D.

4.3.3 Autres variables dépendantes

Distance parcourue

La distance parcourue peut être mesurée (Aubin et al., 2015; Dehn et al., 2018; Josman et al., 2008; Y. J. Kang et al., 2008; Klinger et al., 2006; Kurtz et al., 2006; Nir-Hadad et al., 2017). La distance parcourue par l'utilisateur montre, dans le cas du supermarché virtuel, la stratégie utilisée pour aller chercher les produits dans les rayons. C'est donc une mesure de planification de l'action d'achats. Un individu parcourant plus de distance, aura fait plus d'aller-retours dans les rayons ou n'aura pas emprunté le chemin le plus court (Nir-Hadad et al., 2017). C'est aussi une variable qui permet de discriminer un groupe de patients avec la maladie de Parkinson qui parcourent plus de distance que les sujets contrôles (Klinger et al., 2006).

Temps de pause

Le nombre de pauses et leur durée sont des données enregistrées, à notre connaissance seulement dans le *Virtual Action Planning-Supermarket* (VAP-S) (Aubin et al., 2015; Josman et al., 2008; Klinger et al., 2006; Werner et al., 2009), et cet exemple illustre bien le besoin de décrire précisément la déclaration de chaque variable. En effet, le temps de pause peut renvoyer plusieurs conceptions. Par exemple, un temps de pause peut être défini lorsque le participant ne navigue pas dans l'EV ou lorsqu'il n'utilise pas la technique d'interaction du tout, etc. La durée du temps de pause doit aussi être indiquée pour ne pas créer de faux positifs, c'est-à-dire enregistrer un temps de pause alors que ce n'en est pas un. Par exemple, lors d'une manipulation, il est présumé normal que l'individu fasse un grand nombre de pauses, mais

que le temps de pause ne dépasse pas les millièmes de secondes selon la façon dont les pauses sont enregistrées. Il apparaît donc nécessaire de bien définir la façon dont sont enregistrées les différentes variables dans l'EV pour permettre d'identifier leur rôle dans la caractérisation de la performance du participant.

Les études portant sur le VAP-S ont montré que le nombre de pauses pouvait être utilisé pour différencier des groupes de populations dont des patients avec des troubles cognitifs légers par rapport à ceux qui présentent un trouble du spectre de la schizophrénie (TSS) ou une lésion cérébrale (Josman et al., 2008). Les 3 groupes de patients devaient effectuer des achats dans le supermarché virtuel. Les résultats indiquent que les individus avec des troubles cognitifs légers sont ceux qui ont le moins bien réussi l'activité : ils ont un temps de réalisation plus long, effectuent plus de pauses et font plus d'erreurs. Des patients avec des troubles cognitifs légers ont ensuite été comparés, dans le même EV, à des participants âgés. Les patients avec des troubles ont effectué des pauses plus longues et ont mis plus de temps pour faire les achats dans le supermarché virtuel (Werner et al., 2009).

Consultation des indices lors de la réalisation de la tâche

L'accès aux aides pour réaliser la tâche peut aussi être comptabilisé. Il comporte différentes variables qui viennent supporter le participant lors de son expérience en virtuel. Par exemple, lors d'une situation de multitâches parmi lesquelles le participant doit monitorer sa prise de médicaments, ces aides peuvent être caractérisées par le nombre de fois où l'heure et la liste des tâches sont regardées (Kurtz et al., 2006). Ces variables permettent de savoir si le participant est en mesure de compléter les différentes tâches en faisant appel seulement à ses ressources cognitives ou s'il a besoin d'une assistance car la situation vécue requiert trop de charge mentale pour pouvoir se souvenir de l'intégralité des éléments à traiter. Lors de leur étude, Kurtz et al. (2006) ont montré que les individus avec un trouble du spectre schizophrénique consultaient plus souvent la liste des tâches et l'heure que les participants contrôles. Ces résultats sont corroborés avec d'autres études qui ont montré que les patients présentant un trouble du spectre schizophrénique (Larøi et al., 2010) et ceux avec une lésion cérébrale (Okahashi et al., 2013) consultaient plus de fois et pendant plus longtemps la liste des achats à effectuer dans un supermarché virtuel que les participants contrôles.

4.4 Interaction entre processus cognitif et utilisabilité de la technique

4.4.1 Problématique

Les EVs permettent de proposer des paradigmes d'évaluations cognitives valides et écologiques. Cependant, ces outils sont relativement nouveaux et soulèvent des questions sur le développement et l'interprétation des données recueillies. En effet, lorsque le comportement d'un individu est mesuré à l'aide d'EVs, de nombreuses données peuvent être enregistrées ce qui rend la performance de l'utilisateur difficile à interpréter (Levy, 2013). D'une part car l'importance des variables n'est pas hiérarchisée et d'autre part car leur signification n'est pas toujours bien définie. Par exemple, dans la section précédente, nous avons montré que le temps de réalisation était représentatif des caractéristiques cognitives du participant, mais aussi de l'utilisabilité de la technique d'interaction.

Les données enregistrées peuvent être de l'ordre de la performance ou du chemin réalisé pour atteindre cette performance. Ce chemin est représenté dans les logs qui contiennent une multitude d'informations comme les *inputs* issus de la technique d'interaction, le temps, etc. (Rupp et al., 2012). Des évaluations et analyses doivent donc être conduites afin d'isoler les variables pertinentes permettant de qualifier la performance et le profil cognitif et comportemental d'un individu. Cependant, la réflexion qui doit être conduite autour de la pertinence des variables à enregistrer pour qualifier la performance d'un participant est complexe. Il est nécessaire d'avoir aussi bien une maîtrise des processus cognitifs mis en œuvre dans la réalisation des tâches que d'une connaissance de l'EV et de la façon dont celui-ci est développé. Par exemple, pour bien appréhender la performance de l'utilisateur, l'expérimentateur doit comprendre comment l'EV est construit et comment les variables sont enregistrées. Par exemple, pour assurer une reproductibilité des résultats, il est important de définir la façon dont les erreurs ont été enregistrées. Est-ce grâce à de l'observation ? Est-ce une détection automatique via le système ? Dans ce cas, une bonne pratique serait de fournir les fonctions principales des scripts en pseudo-code pour s'assurer qu'il n'y ait pas d'erreur dans l'enregistrement des mesures.

Les variables précédemment citées (temps, erreur, distance parcourue, etc.) sont principalement utilisées pour décrire les processus cognitifs de l'individu évalué. Ainsi un participant qui réussit le test virtuel, donc le complète en peu de temps et sans erreur, peut être considéré

comme ayant un bon fonctionnement cognitif, alors que ce n'est pas le cas pour un individu ayant mis beaucoup de temps ou réalisé des erreurs. Toutefois, il est nécessaire de prendre en compte que l'utilisateur interagit dans l'EV via un médium : la technique d'interaction. Ainsi, selon le degré de familiarité de l'utilisateur avec la technique d'interaction, celui-ci pourrait mettre plus ou moins de temps pour compléter les tâches virtuelles, indépendamment de ces capacités cognitives. En effet, l'individu non familier avec un joystick devra allouer des ressources cognitives aussi bien à l'utilisation de celui-ci qu'à la réalisation du test. Sa charge mentale sera donc plus élevée qu'un individu familier et il risque de mettre plus de temps pour compléter l'EV, voire même de commettre des erreurs. Pour qualifier correctement le profil de l'utilisateur et déceler s'il présente des perturbations cognitives, il est indispensable de prendre en compte la façon dont il a pu interagir avec l'EV et le fait de savoir s'il a présenté des difficultés d'utilisabilité. Cette question est encore trop rarement prise en compte dans les études mesurant les fonctions cognitives à l'aide d'EVs. En effet, la façon d'interagir dans l'EV est parfois insuffisamment décrite (Verhulst et al., 2017) et ne permet donc pas de connaître le degré de familiarité de l'individu avec l'utilisation de la technique d'interaction. Pour cela, il est possible de tirer parti de l'utilisation de la technique d'interaction en enregistrant des variables directement liées à leur manipulation.

L'appréhension de la performance du participant demande donc une bonne compréhension de l'EV d'évaluation, des variables mesurées, de la technique d'interaction utilisée et du profil du participant en relation avec l'utilisation de cette dernière.

4.4.2 Profil du participant

Il est nécessaire de comparer différentes techniques d'interaction de façon à proposer la plus appropriée selon les tâches à réaliser et la population ciblée.

Peu d'études font le lien avec le niveau d'aptitude à utiliser une technique d'interaction que ce soit au travail avec le plus souvent l'utilisation de la souris et du clavier, ou pour le plaisir avec des périphériques plus variés comme le gamepad, la Kinect, le visiocasque etc. Au fur et à mesure que le participant utilise une technique d'interaction, son utilisation devient automatique et ne consomme plus autant de ressources attentionnelles. Ainsi, nous pouvons supposer qu'un utilisateur expert avec la souris, aura de meilleurs résultats à une tâche dans un EV car il fera moins d'erreurs liées à l'utilisabilité de la technique d'interaction qu'un utilisateur

novice, c'est-à-dire peu familier avec la technique d'interaction. Nous pouvons donc aisément supposer que les techniques d'interaction les plus fréquemment utilisées dans la vie quotidienne vont être celles qui sont associées à une meilleure performance de la part des utilisateurs. Parmi celles-ci il y a la souris et l'écran tactile. En effet, au cours du chapitre 1, nous avons montré qu'un utilisateur familier avec un système impliquera moins de charge mentale lors de la réalisation d'une tâche qu'un non-familier avec les techniques d'interaction.

Pour appuyer ce propos, Ramcharitar et Teather (2017) ont montré que les participants ont une performance plus élevée avec le touchpad du steamcontroller² qu'avec le joystick. Cela serait dû au fait que les participants utilisent fréquemment les surfaces tactiles, que ce soit avec le pad de l'ordinateur ou bien avec leur téléphone portable. De plus, pour les individus qui ne sont pas familiers avec les ordinateurs, l'utilisation de la souris peut interférer avec la réalisation de la tâche Zucchella et al. (2014). Quant aux techniques d'interaction moins communes, comme les visiocasques ou la Kinect, on peut observer que les novices apprennent très vite à utiliser la Kinect (Roupé, Bosch-Sijtsema, & Johansson, 2014; Verhulst et al., 2016) et le visiocasque (Lee et al., 2016). Un individu familier avec l'utilisation des techniques d'interaction sera moins gêné par son utilisation qu'un utilisateur non-familier. En effet, il existe une corrélation positive entre le degré de familiarité et la performance du participant à la tâche (Y. J. Kang et al., 2008). A contrario, les individus familiers avec l'utilisation de la souris réussissent mieux les tâches de sélection (Hourcade et al., 2010) et/ou mettent moins de temps à compléter une tâche avec le gamepad (Verhulst et al., 2016).

Le temps de réalisation d'une tâche pourrait donc être modulé par la familiarité du sujet avec l'utilisation de la technique d'interaction. Il est donc nécessaire d'être vigilant quant à la possibilité de qualifier la performance du participant en regardant principalement la variable temps de réalisation sans la mettre en relation avec le profil du participant (expert ou novice). Pour pallier l'hétérogénéité des participants avec l'usage des EVs lors d'une expérimentation, les chercheurs proposent, avant chaque expérimentation dans un EV, que le participant réalise une phase de familiarisation d'au plus 15 minutes durant laquelle il apprend à se servir de la technique d'interaction ainsi que les possibilités d'interaction avec l'EV. Mais les recherches présentées précédemment indiquent que cette phase de familiarisation n'est pas forcément suffisante. Il devient donc primordial de comprendre comment la familiarité et le niveau d'expertise de l'utilisateur avec une technique d'interaction vient impacter son utilisation et comment il

2. Le touchpad du steamcontroller correspond au pavé tactile disposé au centre des manettes fournies avec le casque de RV HTC Vive.

est possible de caractériser le profil d'un individu.

Une première approche de caractérisation du profil des utilisateurs peut être proposée par l'étude de Rosa, Morais, Gamito, Oliveira, et Saraiva (2016). Les auteurs ont demandé aux participants de jouer à un jeu en RV, avec un visiocasque, et ont récolté de nombreuses données démographiques. Les résultats ont été analysés à l'aide d'analyses en composante principale et de cluster. Ils relèvent trois groupes d'utilisateurs :

- *Les joueurs sur PC*. Ce groupe est majoritairement composé d'hommes dont les techniques d'interaction de prédisposition sont la souris et le clavier. Ils sont plus prédisposés à être immergés dans un EV et sont moins sujets aux symptômes de cybermalaise. Ils jouent fréquemment seul aux jeux vidéo.
- *Les joueurs sur consoles*. Ce groupe est majoritairement composé de jeunes adultes dont la technique d'interaction de prédisposition est le gamepad. Ils jouent fréquemment à plusieurs aux jeux vidéo.
- *Les non-joueurs*. Ce groupe est majoritairement composé de femmes et est plus soumis aux symptômes de cybermalaise.

La caractérisation des utilisateurs selon leur profil, par exemple expert ou novice avec l'utilisation d'une technique d'interaction est encore un fait rare qui demande d'être approfondi. En effet, les résultats qui accompagnent les profils d'utilisateurs sont une aide à l'interprétation de la performance de l'individu dans un EV.

4.4.3 Intégration de variables liées à la technique d'interaction

L'interprétation de la performance du participant dans un EV destiné à l'évaluation cognitive nécessite de prendre en compte aussi bien le degré de réussite aux tâches évaluées que l'utilisation de la technique d'interaction par l'utilisateur. Il est en effet primordial de définir si une mauvaise utilisation de la technique d'interaction est venue impacter ou non la réussite aux tâches.

L'utilisation d'une technique d'interaction implique des processus cognitifs et donc de la charge mentale pour pouvoir être utilisée correctement. Par exemple, Hourcade et al. (2010) ont montré que les performances aux tâches pointage étaient corrélées avec des tests cognitifs chez la personne âgée. Ainsi, plus les aptitudes cognitives sont élevées, plus l'utilisateur est

performant en tâche de sélection et moins il a besoin d'assistance à la sélection par le biais d'un algorithme. De plus, les individus avec des troubles cognitifs légers utilisent de façon moins efficace la souris que des personnes âgées sans troubles. Ils font, par exemple, plus de pauses entre les mouvements de la souris (Seelye et al., 2015). Il semble donc que les personnes sans troubles cognitifs réussissent mieux à interagir avec un EV et que la façon d'utiliser la technique d'interaction peut donc être un critère assez sensible pour détecter des problèmes chez un individu. Ainsi, dans les cas où le participant réalise la tâche sans erreur, nous pouvons supposer qu'un temps de réalisation long est le signe d'une mauvaise utilisation de la technique d'interaction suite à un ralentissement cognitif et/ou moteur. En effet, comme énoncé précédemment, le temps de réalisation est une variable suffisamment sensible pour classer les personnes âgées avec des troubles cognitifs légers de celles qui n'en ont pas (Zygouris et al., 2015).

Une majorité des EVs utilisés pour mesurer la performance du participant en psychologie ne prennent pas encore appui sur l'utilisation de la technique d'interaction. Toutefois, certaines études ont su tirer profit des nouvelles technologies pour effectuer des enregistrements de meilleure qualité et prendre en compte des variables plus nombreuses. Par exemple, Martono et al. (2016) ont enregistré les mouvements des doigts de participants lorsqu'ils réalisaient une tâche de café sur un écran tactile. Le participant portait un marqueur au bout de son index dont la position est reconnue par le capteur infra-rouge Leap Motion positionné au dessus de la tablette. Ainsi, il est possible de savoir lorsque le participant s'apprête à sélectionner un item mais se rétracte avant de toucher la surface tactile. Les enregistrements comportant les mouvements du doigt des participants ont ensuite été classés en trois groupes par une classification non-hiérarchique. Lorsque l'on compare le nombre d'erreurs effectuées durant la tâche de café en fonction des trois groupes, nous pouvons constater que les groupes réalisent la tâche de café en différents laps de temps et que leurs types d'erreurs sont différents (Martono et al., 2016).

D'autres auteurs proposent de mesurer un index de performance en fonction de l'utilisabilité de la technique d'interaction (Y. J. Kang et al., 2008). L'index de performance se calcule selon l'équation suivante où les erreurs d'interaction est le ratio entre le nombre de clics effectués dans l'EV et le nombre de clics correctes (qui rencontrent une cible valide).

$$IndexPerformance = temps + distance\ parcourue + collision - nombre\ d'achats + erreur\ d'interaction \quad (4.1)$$

Les participants devaient collecter différents items dans un supermarché virtuel. Ainsi l'index de performance prend en compte le nombre d'achats mais cette variable peut être aisément transposée à d'autres applications en la modifiant par exemple par le nombre de bonnes actions. L'étude a comparé la performance d'individu avec une lésion cérébrale et de sujets contrôles. Les résultats montrent que les patients ont un index de performance moins élevé que les sujets contrôles mais aussi qu'ils réalisent plus d'erreurs d'interaction.

L'interaction dans un EV nécessite l'usage d'une technique d'interaction. Si celle-ci peut être difficile à utiliser pour le participant, elle peut aussi représenter une nouvelle façon de récolter des mesures et d'interpréter la performance du participant. En effet, la façon d'utiliser la technique d'interaction a permis d'apporter un nouvel angle d'approche dans la caractérisation des participants. Par exemple, la vitesse des mouvements est un bon critère pour différencier des groupes (Martono et al., 2016).

4.5 Synthèse

L'utilisation des EVs dans l'évaluation cognitive demande aux expérimentateurs de développer de nouvelles pratiques. Il est nécessaire, dans un premier temps, de comprendre si la performance du participant obtenue en virtuel est représentative de celle qu'il obtient en réel à tâche équivalente. En effet, il peut y avoir des différences de performances entre réel et virtuel en fonction de la technique d'interaction utilisée, mais aussi de la façon dont la performance est mesurée. Par exemple, est-ce qu'il y a beaucoup de variables ou pas. L'évaluation de la performance du participant demande donc de comprendre les variables utilisées pour caractériser la production de l'utilisateur, mais aussi les liens qu'elles entretiennent entre elles. Il s'agit de pouvoir discriminer un participant qui a échoué à la réalisation de la tâche pour des raisons cognitives de celui qui n'a pas réussi pour des raisons de difficulté avec l'utilisation de la technique d'interaction. Pour aider à l'interprétation de la performance du participant, une caractérisation du profil de celui-ci et l'ajout de variables issues de la technique d'interaction en elle-même sont donc nécessaires. De plus, une bonne description des techniques et de la façon dont les variables sont enregistrées est attendue. C'est donc à partir de ces constats que nous avons développé notre EV et choisi d'y intégrer des variables permettant a posteriori de qualifier le profil de l'utilisateur.

Cas d'application : le *Virtual Multitasking Test*

Test

5.1 Introduction

L'EV utilisé au cours de cette thèse est un appartement virtuel nommé le VMT. C'est un EV destiné à l'évaluation cognitive de participants. Il contient plusieurs pièces : une entrée, deux chambres, un bureau, une salle à manger avec une cuisine ouverte, un salon, un couloir et une salle de bain (Fig. 5.1).



FIGURE 5.1 – Illustration du VMT.

Le VMT a pour vocation d'évaluer les participants dans un contexte écologique, proche de la vie de tous les jours, et en particulier durant des tâches de Mémoire prospective (MP). La

MP peut être définie comme la mémoire des actions futures, c'est-à-dire la mémoire qui permet de se souvenir des activités à effectuer plus tard ou à un moment donné, comme aller faire ses courses en rentrant du travail. Il est primordial d'évaluer la MP car c'est cette mémoire qui est la plus déficitaire lorsque les performances d'un individu sont mesurées dans un contexte de vie quotidienne (Kliegel & Martin, 2003).

Le projet du VMT s'inscrit dans une collaboration entre l'Université d'Angers avec le Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS) et le Laboratoire de Psychologie des Pays de la Loire (LPPL), l'Université du Québec à Rimouski (UQAR) et le Centre de développement et de recherche en imagerie numérique (CDRIN). L'outil de développement actuel est Unity 3D version 5.6.1. Le cadre d'application du VMT comprend un scénario d'évaluation du participant durant lequel les actions du participant sont enregistrées dans des logs. A la fin du test, les analyses des productions de l'utilisateur sont automatisées par des scripts en Python et sont transmises à l'expérimentateur sous forme de documents de format .html (Fig. 5.2). Au cours de la thèse, j'ai développé deux scénarii dans le VMT, implémenté différentes techniques d'interaction ainsi que des scripts pour établir un recueil de données et une analyse qualitative de celles-ci.

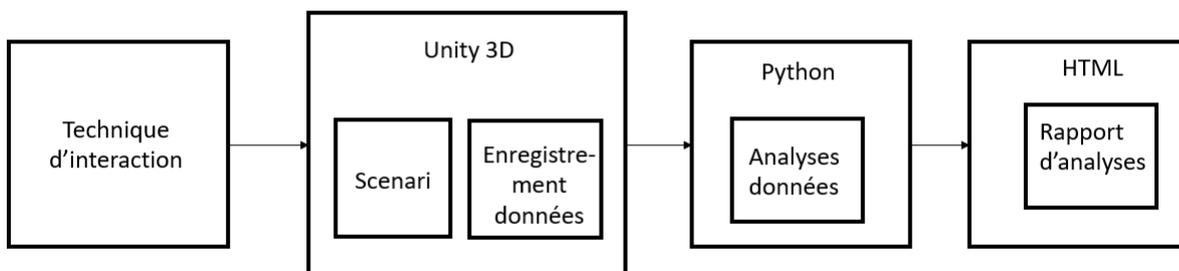


FIGURE 5.2 – Cadre d'application du VMT.

5.2 Spécificités du VMT

Le VMT contient plusieurs fonctionnalités :

- **La navigation** : cette fonctionnalité permet à l'utilisateur de se déplacer dans l'EV.

La navigation suppose que la simulation soit lancée au préalable par l'expérimentateur et que l'utilisateur renseigne un *input* via la technique d'interaction. Par exemple, appuyer sur une touche fléchée du clavier. Les déplacements demandés par l'utilisateur sont effectués par un avatar à la première personne.

- **L'interaction avec un objet (sélection)** : cette fonctionnalité permet à l'utilisateur de sélectionner des items dans l'EV. L'interaction avec un objet suppose que la simulation soit lancée au préalable par l'expérimentateur, que l'utilisateur soit suffisamment proche de l'item, que ce dernier soit interactif et enfin que l'utilisateur renseigne un *input* via la technique d'interaction. Par exemple, appuyer sur le bouton gauche de la souris. Les objets présents dans le VMT sont soumis à plusieurs actions possibles (Tab. 5.1).
- **Les scenari** : cette fonctionnalité renseigne plusieurs tâches qui peuvent être activées ou désactivées par l'expérimentateur dans l'interface de Unity 3D (Tab. 5.2).
- **L'enregistrement des données** : cette fonctionnalité permet d'enregistrer les actions de l'utilisateur lors de son interaction dans le VMT. Deux types d'informations sont enregistrées : les données liées au comportement de l'utilisateur et les données liées à l'utilisation de la technique d'interaction. Les données liées au comportement comprennent l'action effectuée (ramasser ou déposer), le nom de l'item et l'endroit où est déposé l'item (si l'action est déposer). Le temps où l'action est réalisée est également notifié. Les données liées à l'utilisation de la technique d'interaction comprennent : le temps, l'action et le détail de l'action.
- **L'analyse des données** : cette fonctionnalité permet de fournir à l'expérimentateur des fichiers directement exploitables. Ainsi l'expérimentateur peut obtenir le temps de réalisation, le score aux tâches, le tracé des mouvements de l'utilisateur dans le VMT et la photo de la table (Fig. 5.3). Pour faciliter la compréhension des données recueillies, une analyse plus approfondie est fournie a posteriori via des scripts en Python.

TABLE 5.1 – Actions possibles dans le VMT en fonction des objets.

Objets	Actions
Portes (portes/placards/frigo)	Ouvrir/fermer
Fenêtres chambre d'ami et cuisine	Ouvrir/fermer
Sécheuse	Ouvrir/fermer/démarrer
Nourriture poisson	Verser dans l'aquarium
Chemise	Prendre/déposer
Produits consommables	Prendre/déposer
Couverts	Prendre/déposer
Billets	Prendre/déposer
Téléphones	Décrocher/raccrocher

EV	Tâches
	Le fax : au lancement de la tâche, les objets 3D nécessaires à la réalisation de la tâche sont instanciés dans l'EV. Une feuille interactive est instanciée à côté du fax, dans le bureau. L'utilisateur doit faxer trois feuilles. Pour ce faire il sélectionne la feuille puis le fax. Une fois la feuille faxée, l'utilisateur la sélectionne pour que celle-ci soit déposée à nouveau sur le bureau.
	La sècheuse : au lancement de la tâche, les objets 3D nécessaires à la réalisation de la tâche sont instanciés dans l'EV. Une chemise est instanciée sur le lit de la chambre principale. L'utilisateur doit prendre la chemise puis la déposer dans la sècheuse qui est située dans le placard du couloir. Une fois la sècheuse ouverte puis refermée pour y mettre la chemise, l'utilisateur programme la machine pour 5 minutes de séchage. A la fin des 5 minutes, la sècheuse lance un signal sonore. L'utilisateur peut alors récupérer la chemise pour la redéposer sur le lit de la chambre principale.
	Les billets : au lancement de la tâche, les objets 3D nécessaires à la réalisation de la tâche sont instanciés dans l'EV. Trois lettres et un billet sont instanciés sur la table à côté de l'entrée. L'utilisateur doit soulever les lettres pour trouver le billet.
	Les poissons : les objets de cette tâche sont instanciés par défaut. L'utilisateur doit nourrir les poissons. L'aquarium est positionné dans la chambre principale et les granulés sur l'aquarium. L'utilisateur doit sélectionner 4 fois la boîte pour nourrir les poissons.

EV	Tâches
	<p>L'orage : au lancement de la tâche, les fenêtres de la cuisine et de la chambre d'amis sont ouvertes. Au début de l'orage, l'utilisateur l'entend gronder. L'utilisateur doit fermer les fenêtres.</p>
	<p>Le poulet rôti : au lancement de la tâche, les objets 3D nécessaires à la réalisation de la tâche sont instanciés dans l'EV. Neuf produits consommables sont instanciés sur le comptoir de la cuisine. Le poulet à rôtir est instancié à côté du four de la cuisine, sur une assiette. A côté de cette assiette est instancié un plat de cuisson vide. L'utilisateur doit déposer le poulet dans le plat. Ajouter du sel, du poivre et du sirop d'érable. Laisser reposer le poulet 5 minutes. Il doit ensuite mettre le four à préchauffer puis ajouter les légumes, l'huile d'olive et la sauce sur le poulet. Il peut enfin déposer le plat dans le four et faire cuire le poulet pendant 1h30.</p>
	<p>La table : au lancement de la tâche, les objets 3D nécessaires à la réalisation de la tâche sont instanciés dans l'EV. Douze verres, 8 fourchettes, 8 couteaux et 8 petites cuillères, 8 assiettes et 4 serviettes de table sont instanciés dans les placards de la cuisine, à côté du four. L'utilisateur doit prendre les ustensiles un à un puis les déposer sur la table de la cuisine. Il doit mettre la table pour deux personnes.</p>
	<p>L'épicerie : au lancement de la tâche, les objets 3D nécessaires à la réalisation de la tâche sont instanciés dans l'EV. Seize produits consommables sont instanciés sur le comptoir de la cuisine. L'utilisateur doit sélectionner, un à un, les items consommables sur le comptoir de la cuisine puis les ranger à l'endroit souhaité.</p>

TABLE 5.2 – Présentation des tâches du VMT.



FIGURE 5.3 – Exemple de données enregistrées par le VMT : itinéraire du participant et photo de la table.

5.3 Implémentation des tâches dans le VMT

5.3.1 Contexte théorique lié aux tâches

Pour avoir un aperçu des capacités préservées de la personne et de son autonomie dans le quotidien, des tâches de MP ont été mises en place dans un appartement virtuel (Kliegel & Martin, 2003).

Plusieurs auteurs se sont intéressés à la façon d'évaluer la MP chez les patients cérébrolésés en sachant que la MP fait appel aux fonctions exécutives telles que la planification, l'inhibition, l'organisation, etc. (Burgess, Gonen-Yaacovi, & Volle, 2011). Selon le modèle du *multiprocess Framework* (McDaniel & Einstein, 2007), la capacité à se souvenir de réaliser des actions dans le futur dépend de deux actions : le monitoring qui est le suivi de l'activité et l'exécution spontanée. On parle d'exécution spontanée quand la tâche est réalisée naturellement lorsqu'un indice prospectif est repéré dans l'environnement. Par exemple, s'arrêter faire les courses quand on voit un magasin. Tandis que le monitoring demande à l'individu de garder une attention sur les cibles qui déclenchent les traces mnésiques liées à l'action à entreprendre. Par exemple, appeler une personne à un moment précis de la journée.

Les actions de monitoring sont en lien avec le *Système Attentionnel de Supervision* (SAS, Norman et Shallice (1986)). Le SAS permet de contrôler la réalisation d'une action non routinière en activant ou inhibant des schémas pertinents pour la réalisation de cette action. Ainsi, le SAS solliciterait les fonctions exécutives nécessaires pour repérer l'indice prospectif, stopper l'action en cours et entreprendre l'action (McDaniel & Einstein, 2007). Ainsi, certaines tâches peuvent être réalisées automatiquement, comme prendre un médicament que l'on prend quotidiennement et d'autres demandent une attention latente jusqu'à la réalisation de l'activité. Dans un second modèle théorique, le monitoring est toujours nécessaire (*Preparatory Attention and Memory Model* (PAM, Smith (2003))), ce qui sous-entend que le rappel prospectif n'est pas automatique. En effet, les processus attentionnels sont engagés dans une tâche concurrente jusqu'à ce que l'action nécessite d'être réalisée ou jusqu'à ce qu'une occasion de réaliser l'action se présente.

Dans une tâche de MP, l'individu doit organiser différentes actions afin de réaliser un but et se rappeler de la tâche à réaliser lorsqu'il est possible de l'effectuer. Il y a donc un délai d'attente avant que le participant puisse réaliser la tâche, en plus de la tâche concurrente (Burgess et al., 2011 ; McDaniel & Einstein, 2007). L'individu doit aussi se souvenir lui-même qu'il a une action à réaliser en surveillant les indices prospectifs présents dans l'environnement, comme l'heure

(Burgess et al., 2011). La consigne est donc donnée au début de l'évaluation et il n'y a pas de rappel de la tâche demandée. Kvavilashvili et Ellis (1996) ont décrit plusieurs types de tâche pour mesurer les capacités de MP :

- *La tâche basée sur le temps (time-based)*. C'est une tâche à réaliser à un moment précis ou après un laps de temps. Par exemple, allumer la radio à 17h10 pour écouter une émission. Ce type de tâche demande de surveiller l'écoulement du temps et d'être attentif à ce que l'environnement contient des horloges.
- *La tâche basée sur l'évènement (event-based)*. On se souvient de la tâche grâce à un indice prospectif dans l'environnement. Par exemple, en passant devant la machine à laver, je me souviens que je dois absolument laver cette chemise car j'en aurai besoin en fin de semaine.
- *La tâche basée sur l'activité (activity-based)*. Le participant doit réaliser une action avant ou après la réalisation de la tâche. Par exemple, fermer le robinet d'eau après avoir lavé la vaisselle.

Pour mêler ces différentes tâches dans une seule session de test, le multitâche apparaît être tout à fait approprié. La MP est en effet une composante du multitâche (Shallice & Burgess, 1991) et se révèle être utilisable en RV pour mesurer la MP de façon écologique (Banville et al., 2010). De façon expérimentale, une situation de multitâche consiste à proposer aux sujets une situation dans laquelle ils doivent planifier, résoudre des problèmes et/ou choisir entre leurs performances à différentes tâches (Burgess et al., 2000). Par exemple, lorsqu'un individu cuisine, il doit être attentif à couper les aliments tout en surveillant leur cuisson et l'ordre dans lesquels il va les faire revenir.

Les études proposant des situations de multitâche en RV doivent comporter un nombre accru de distracteurs pour être prédictives des performances dans la réalisation des tâches de la vie quotidienne en réel (Burgess et al., 2006). Le VMT propose la possibilité d'intégrer plusieurs scénari avec une situation de multitâche. Pour que le sujet ne porte pas l'attention essentiellement sur les tâches de MP, on lui propose une tâche concurrente qui est imbriquée dans les tâches de MP. La tâche concurrente doit être suffisamment longue et demandeuse en charge cognitive pour être efficace. Des événements interviendront en temps voulu comme par exemple un coup de téléphone demandant au sujet de réaliser une action. L'ajout de ces événements a pour but d'augmenter la charge mentale du sujet.

Trois types de tâches sont donc implémentés dans le VMT : des tâches de MP, des tâches concurrentes et des tâches interférentes (Tab. 5.3). Les tâches concurrentes sont indiquées à l'utilisateur dès le début du VMT de manière à ce qu'il pense que ce sont les seules tâches à réaliser. Les tâches de MP et les tâches interférentes sont inattendues et ajoutées dans le VMT par des appels téléphoniques (Tab. 5.4). Chacune de ces tâches peut être divisée en sous-tâches et c'est la décomposition des tâches en sous-tâches qui permet d'imbriquer une tâche de MP dans la tâche concurrente. L'activité du poulet est la tâche de MP imbriquée dans la tâche concurrente. Dans ce cas, l'ajout du poulet au four fait partie de l'activité concurrente alors que l'ajout des ingrédients, le temps de la marinade et le préchauffage et la programmation du four font partie d'une tâche prospective.

TABLE 5.3 – Types de tâches dans le VMT.

Tâche	Type de tâche
Table	Concurrente
Epicerie	Concurrente
Poulet - Mettre le poulet au four	Concurrente
Poulet - Marinade	MP - basée sur le temps
Poulet - Ajout d'ingrédients	MP - basée sur l'évènement
Poulet - Préchauffage	MP - basée sur l'activité
Fax	MP - basée sur le temps
Sécheuse	MP - basée sur l'évènement
Poissons	Interférente
Orage	Interférente
Billet	Interférente

5.3.2 Présentation des scénari

Le VMT comporte deux scénari : le premier est construit pour générer une charge mentale faible chez le participant alors que le second est mis en œuvre pour en générer une forte. Ainsi, le degré de difficulté du scénario varie au cours des deux conditions.

TABLE 5.4 – Verbatims des coups de téléphone annonçant une tâche dans le VMT.

Tâches	Verbatims
Fax	« <i>Allo, j'ai oublié un document important sur mon bureau ce matin. Est-ce que tu pourrais me le faxer SVP ? Comme c'est un document volumineux, il faudra que tu me l'envoies par petites parties et toutes les 5 minutes. Merci et à ce soir</i> »
Sécheuse	« <i>Allo, c'est encore moi ! Afin d'être fin prêt(e) pour ce soir, j'aurais besoin que tu passes ma chemise 5 minutes à la sécheuse pour la défroisser. J'aimerais que tu la remettes sur mon lit par la suite. Merci encore une fois !</i> »
Poissons	« <i>Salut, je viens de réaliser que mon poisson n'a pas été nourri depuis 3 jours. Pauvre petit ! Peux-tu le nourrir pour moi Svp ? Il faut lui donner 4 doses</i> »
Billets	« <i>Bon, c'est la dernière fois que je t'appelle avant ce soir : Peux-tu me dire si les billets sont sur la table à l'entrée ? Vas-voir je t'attends...</i> [au retour de l'utilisateur] <i>Ils y sont ! Super merci et à ce soir</i> »

Scénario avec une charge mentale faible

Au cours de la condition de charge mentale faible, il est demandé au participant de réaliser trois tâches : l'épicerie, le poulet et la table. Ces trois tâches forment ce qu'on appelle l'activité concurrente. Elles peuvent être réalisées sans ordre pré-défini (Fig. 5.4).

Scénario avec une charge mentale élevée

Au cours de la condition charge mentale élevée, il est demandé au participant de réaliser les trois tâches concurrentes. Cependant au fur et à mesure des actions réalisées par l'utilisateur, d'autres tâches apparaissent. L'apparition des tâches sont basées sur les actions de l'utilisateur (Fig. 5.5). C'est-à-dire que certaines actions déclenchent des appels téléphoniques et donc de nouvelles tâches à effectuer. C'est l'addition de ces tâches non prévues qui va augmenter la charge mentale du participant. Sur la figure 5.5, un système avec des jetons a été utilisé pour modéliser le scénario. Ainsi le début du VMT génère trois jetons (la génération des jetons est représenté par un symbole de couleur noir) qui sont distribué aux trois premières tâches (poulet, table et épicerie). Tant que le jeton n'est pas détruit (une tâche dans un cercle symbolise une destruction de jeton), la tâche est toujours en cours et l'utilisateur peut y revenir. Ainsi le participant réalise de façon alternative autant de tâche qu'il a de jeton. Les jetons sont déplacés au fur et à mesure que les sous-tâches sont complétées et certaines activités peuvent produire

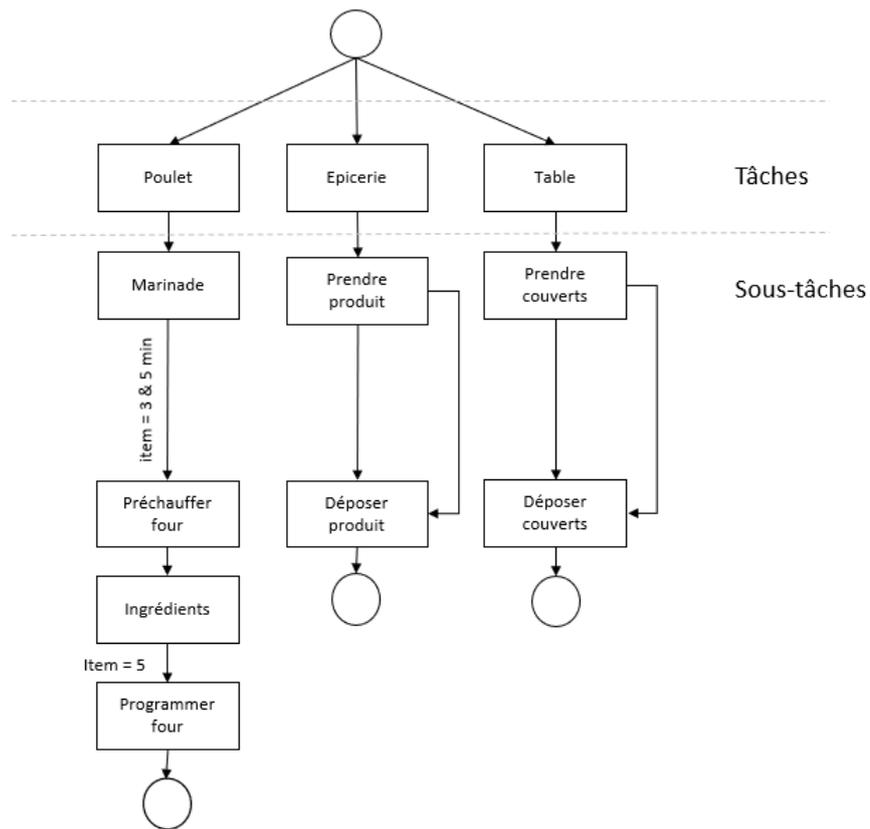


FIGURE 5.4 – Scénario "facile" avec une charge mentale faible.

des jetons. Ce type de modélisation par diagramme permet d'illustrer un scénario multitâche.

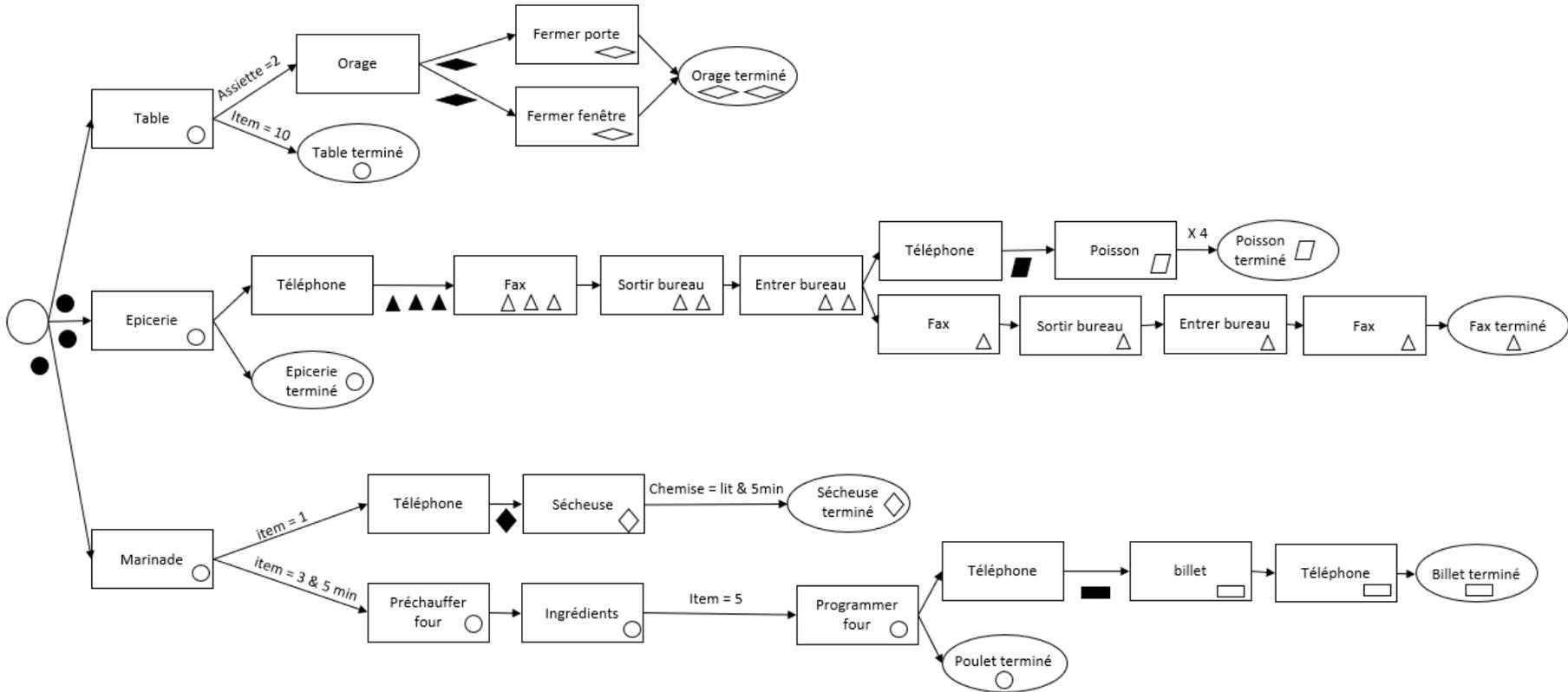


FIGURE 5.5 – Modélisation du scénario avec une charge mentale élevée

5.3.3 Enregistrement des variables

Dans le VMT, le temps de réalisation, la distance parcourue, l'ensemble des actions effectuées par l'utilisateur et le score aux tâches sont enregistrés. Chaque tâche comprend son propre système de notation :

- Tâche du fax : 4 points maximum par fax envoyé. Un point est enlevé pour chaque dizaine de secondes supplémentaire de différence entre le temps où le fax est réalisé et le temps idéal (un fax toute les 5 minutes)(Alg. 1). Total de 12 points.
- Tâche de la sècheuse : 1 point si la chemise a été séchée pendant 5 minutes et si elle a été reposée sur le lit de la chambre principale. Total d'1 point.
- Tâche du billet : 1 point si le billet a été sélectionné. Total d'1 point.
- Tâche des poissons : 1 point par sélection du paquet de granulés pour poissons. Total de 4 points.
- Tâche du poulet : cette tâche contient plusieurs sous-tâches qui contiennent chacune un score. Total de 8 points.
 - Marinade : 4 points si le poulet a mariné 5 minutes. Un point est enlevé par dizaine de secondes supplémentaire de différence entre le temps où le dernier produit de la marinade a été ajouté et l'ajout d'un ingrédient. Le principe algorithmique est le même que celui pour la tâche du fax (Algo. 1).
 - Ingrédients : 1 point par ingrédient ajouté dans le plat du poulet. Total de 10 points.
 - Four : le four contient 2 scores : 1 score de préchauffage et 1 score de programmation.
 - * Quatre points si le four est mis à préchauffer moins de 10 secondes après l'ajout du poulet dans le plat. Un point est enlevé pour chaque dizaine de secondes supplémentaires de différence entre le temps où le four est mis à préchauffer et le dépôt du poulet dans le plat. Total de 4 points.
 - * Quatre points si le four est programmé moins de 10 secondes après le dépôt du poulet dans le four. Un point est enlevé pour chaque dizaine de secondes supplémentaires de différence entre le temps où le four commence à être programmé et le moment où le poulet est déposé dans le four. Total de 4 points.
- Tâche de la table : 1 point par item correct placé sur la table. Les items corrects sont : les couverts (cuillère, couteau, fourchette), les deux sortes de verres, les serviettes, les

assiettes, le sel, le poivre et le pain.

- Tâche de l'épicerie : 1 point par item bien rangé. Huit produits vont au frigo et 8 dans les placards. Chaque produit est labellisé avec une location attendue. Si le produit est déposé sur la location attendue, 1 point est comptabilisé. Sinon aucun point n'est comptabilisé.

```

temps ideal = 300sec (5min);
temps = temps fax - temps ideal;
while nombre fax <= 3 do
  if temps <= 10 then
    | Score = 4;
  end
  else if temps <= 20 then
    | Score = 3;
  end
  else if temps <= 30 then
    | Score = 2;
  end
  else if temps <= 60 then
    | Score = 1
  end
  else if tt then
    | Score = 0
  end
end
end

```

Algorithm 1: Algorithme d'évaluation de la tâche du fax.

5.4 Implémentation des interactions dans le VMT

5.4.1 Approche proposée

Les interactions dans le VMT permettent à l'utilisateur de réaliser des actions et des tâches. Ces interactions comprennent de la sélection et de la navigation (translation et rotation de la caméra). La navigation s'effectue en première personne avec un avatar non-visible et la technique de sélection est le rayon (Fig. 5.6). L'origine du rayon est le curseur de sélection pour la souris et le gamepad et la main virtuelle est le curseur de sélection pour les techniques avec le Razer Hydra. Afin de faciliter les interactions dans le VMT, celles-ci ont été simplifiées par rapport à ces mêmes interactions en réel. Ainsi, dans le VMT, l'utilisateur n'effectue pas de manipulation d'objets 3D. Par exemple, lorsque l'utilisateur dépose un couteau sur le set de table, il n'a pas à gérer l'orientation du couteau. Le couvert vient se positionner avec l'orientation attendue.

De plus, des actions n'ont pas lieu dans le VMT, alors qu'elles sont indispensables en réel. Ainsi, l'utilisateur n'a pas à découper les légumes avant de les ajouter dans le plat de cuisson. Ceux-ci se découpent et se positionnent automatiquement dans le plat lorsque l'utilisateur les

y dépose.

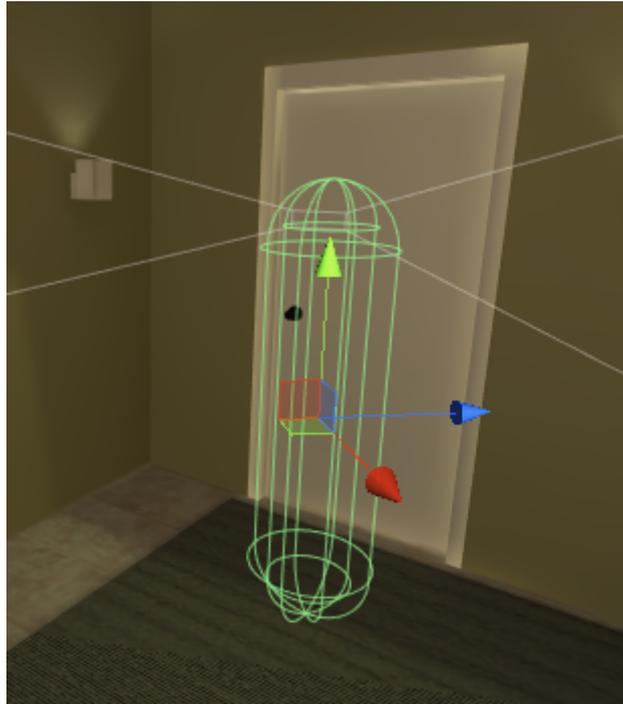


FIGURE 5.6 – Avatar dans le VMT.

Pour guider l'utilisateur dans les interactions, des informations visuelles apparaissent lorsqu'un objet 3D est interactif. En effet, si le curseur de sélection rencontre un objet qui porte le tag 'interactif', et que cet objet est suffisamment proche de l'utilisateur, alors l'objet se surligne en jaune (Fig. 5.7). L'utilisateur peut alors le sélectionner en cliquant dessus. L'objet vient alors se déposer dans la main du participant (Fig. 5.8) en attendant d'être déposé sur une surface. L'utilisateur ne peut sélectionner qu'un seul objet à la fois.

5.4.2 Présentation des techniques d'interaction

Les techniques d'interaction ont été choisies pour permettre de les comparer dans le VMT en fonction des différentes informations d'entrée qui peuvent être renseignées par la technique d'interaction (Tab. 5.9).

Les techniques d'interaction pourront être comparées selon la méthode de :

- Sélection : sélection en 2D versus sélection en 3D (Fig. 5.10).
- Rotation de la caméra : mouvements de la main versus mouvements de la tête (Fig. 5.11).



FIGURE 5.7 – Illustration du surlignage des items interactifs dans le VMT.



FIGURE 5.8 – Illustration d'un objet 3D déposé dans la main de l'avatar pour la souris et le Razer Hydra.

— Immersion : non-immersif versus immersif (Fig. 5.12).

Techniques	Translation	Rotation	Sélection
Clavier/souris	Pression des touches fléchées	Mouvements de la souris	Clic gauche de souris
Gamepad	Pression du joystick gauche	Pression du joystick droit	Pression du bouton A
Razer Hydra	Pression du joystick gauche	Pression du joystick droit	Gâchette de la manette droite
Razer Hydra et <i>tracking</i> de tête avec Kinect	Pression du joystick gauche	Mouvements de la tête	Gâchette de la manette
Razer Hydra et HTC Vive	Pression du joystick gauche	Mouvements de la tête	Gâchette de la manette

FIGURE 5.9 – Utilisation des techniques d’interaction proposées et développées dans le VMT.



FIGURE 5.12 – Comparer les techniques d’interaction en fonction d’une situation immersive ou du non-immersive.

5.4.3 Enregistrement des variables

Le VMT enregistre des informations liées à l’utilisation de la technique d’interaction. Toutes ces informations sont enregistrées dans un fichier .csv pour les cinq techniques d’interaction.

Les différentes actions sont notées :

— `[TechniqueInteraction]_travelling_pressed` : un *input* destiné à la translation de la

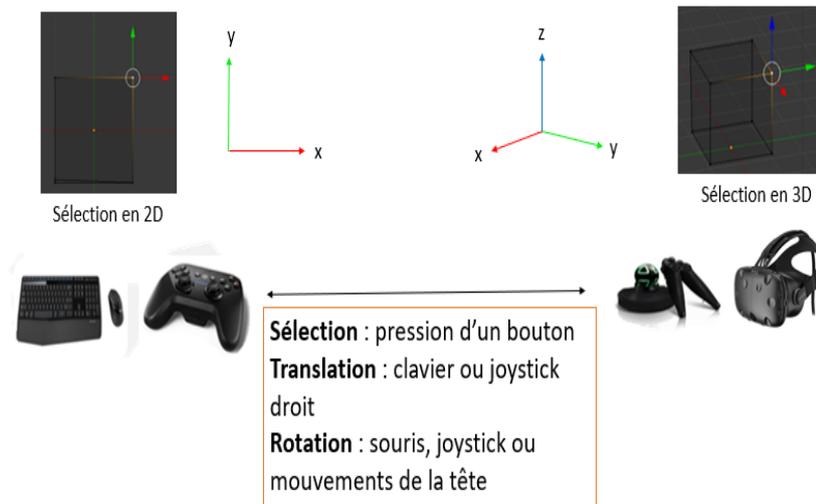


FIGURE 5.10 – Comparer les techniques d'interaction en fonction d'une sélection en 2D et en 3D.

caméra est enclenché. Par exemple, pour la souris, cette action est notifiée lorsqu'une flèche du clavier est pressée. La position de l'avatar est indiquée en détail.

- `[TechniqueInteraction]_travelling_released` : aucun *input* destiné à la translation de la caméra est enclenché. Par exemple, pour la souris, cette action est notifiée lorsqu'aucune flèche du clavier n'est pressée. La position de l'avatar est indiquée dans la colonne intitulée détail.
- `[TechniqueInteraction]_steering_pressed` : un *input* destiné à la rotation de la caméra est enclenché. Par exemple, pour le gamepad, cette action est notifiée lorsque le joystick droit est poussé. La rotation de l'avatar est indiquée dans la colonne intitulée en détail.
- `[TechniqueInteraction]_steering_released` : aucun *input* destiné à la rotation de la caméra n'est enclenché. Par exemple, pour le gamepad, cette action est notifiée lorsque le joystick droit n'est pas poussé. La rotation de l'avatar est indiquée dans la colonne intitulée en détail.
- `[TechniqueInteraction]_click` : un *input* destiné à la sélection est enclenché. Par exemple, cette action est notifiée lorsqu'un clic gauche de la souris est effectué. Le nom de l'item ou de la location sélectionné est indiqué en détail.
- `[TechniqueInteraction]_MissClick` : un clic gauche est effectué sur une cible non valide. Par exemple, l'utilisateur clique dans le vide.



FIGURE 5.11 – Comparer les techniques d’interaction en fonction d’une rotation de la caméra avec les mouvements de la main ou avec les mouvements de la tête.

5.5 Implémentation du rapport d’analyses

Les actions de l’utilisateur dans le VMT sont enregistrées sous forme de données textuelles dans un journal d’informations (annexe A). Un premier résumé de ces informations est extrait et indique le temps de réalisation du VMT, la distance parcourue et le score aux tâches.

Les données inscrites dans le journal d’informations du VMT sont trop denses pour pouvoir être lues par l’expérimentateur. Pour faciliter le travail de ce dernier, un traitement du journal d’informations est effectué a posteriori. Ce traitement est réalisé en Python et génère ensuite un rapport d’analyses en .html (Fig. 5.13). Le rapport d’analyses comprend les informations suivantes :

1. **Sommaire du test** : il indique la référence du participant, le temps de réalisation ainsi que la distance parcourue lors du VMT. Le nombre d’actions totales et le nombre de fois où les portes et les placards sont ouverts sont aussi indiqués.
2. **Sommaire des tâches** : il indique le nom de chaque tâche effectuée, dans leur ordre d’apparition, le début, la fin et la durée de ces tâches.
3. **Evaluation des tâches** :
 - (a) **Epicerie** : il indique le score à la tâche avec le décompte des points et l’endroit où est déposé l’item, le nombre d’actions effectuées pour réaliser la tâche et le nombre de fois où le frigo a été ouvert.

- (b) Poulet : il indique les différents scores à la tâche : le score d'ingrédients, le score macération, le score de préchauffage et le score de programmation. Le nombre d'actions réalisées pour compléter la tâche est aussi inscrit. Les différents scores sont ensuite approfondis :
- i. Ingrédients : le détail du score indique si des ingrédients ont été rangés avant le début de la recette et s'il manque des ingrédients. Le temps écoulé entre l'ajout du premier et du dernier ingrédient est aussi renseigné. L'ordre d'ajout des ingrédients dans le plat de cuisson est ensuite présenté.
 - ii. Macération : le détail du score indique si tous les ingrédients de la marinade ont été ajoutés et s'il n'y a pas d'ajout de mauvais ingrédients. Le début, la fin et la durée de la macération est ensuite ajouté.
 - iii. Préchauffage : le détail du score indique le temps de latence entre l'ajout du poulet dans le plat et le moment où le four est mis à préchauffer.
 - iv. Programmation : le détail du score indique le temps de latence entre le moment où le four est déposé dans le four et celui où le four est terminé d'être programmé. Toutes les étapes liées à la programmation du four sont ensuite affichées.
- (c) Table : il indique le nombre d'actions effectuées pour réaliser la tâche. Le détail des items déposés sur la table est ensuite indiqué.
- (d) Fax : il indique le nombre de fax réalisés et le score pour chaque fax avec le temps de latence entre le moment où le fax est effectué et les 5 minutes attendues.
- (e) Chemise : il indique le nombre d'actions effectuées pour réaliser la tâche. Le temps entre l'appel et la programmation de la sècheuse et le temps entre l'appel et le dépôt de la chemise sur le lit.
4. **Alternance des tâches** : il indique le nombre d'alternance des tâches et le détail de ces alternances.

5.6 Synthèse

Le VMT est un EV destiné à l'évaluation cognitive de participants lors d'IADLs. Pour ce faire, différentes tâches doivent être réalisées dans l'appartement virtuel. Différents scénari peuvent être implémentés. Dans notre cas deux scenari ont été créés : une condition de charge mentale faible et une condition de charge mentale élevée. Dans le premier cas, la difficulté de

Evaluation Poulet

Sommaire

Score ingrédients: 10 / 10

Score macération (err= 4 sec): 4 / 4

Score préchauffage: 0 / 4

Score programmation: 0 / 4

Nb actions réalisées pour compléter la tâche du poulet : 49

Détails score ingrédients

Ingrédients rangés avant leur utilisation dans la recette (0): aucun

Ingrédients manquant (0): aucun

Temps écoulé entre l'ajout du premier et du dernier ingrédient: 06:39

Ordre d'ajout des ingrédients dans la rôtissoire

	log
42	3:07 : Déposer Poulet -> Rotissoire
45	3:15 : Déposer Sirop d'erable -> Versage
51	3:33 : Déposer Sel -> Versage
56	3:47 : Déposer Poivre -> Versage
121	8:51 : Déposer Carottes -> Rotissoire
124	9:01 : Déposer Oignons -> Rotissoire

FIGURE 5.13 – Illustration d'une partie du rapport d'analyses.

la tâche est faible, l'utilisateur doit réaliser trois tâches (tâche du poulet, de la table et de l'épicerie). Tandis que dans le second cas, l'utilisateur doit réaliser des activités inattendues en plus des trois tâches initiales.

Toutes les actions de l'utilisateur dans l'EV sont enregistrées. Celles-ci contiennent des informations liées à la réalisation des tâches, comme le score, et des informations reliées à l'utilisation de la technique d'interaction. Par exemple le temps passé à effectuer des rotations de la caméra.

Les informations liées à la réalisation de la tâche sont ensuite retranscrites dans un rapport d'analyses de manière à faciliter le travail de l'expérimentateur.

Méthode

6.1 Hypothèses

Pour répondre à l'objectif de l'expérimentation et comprendre l'impact d'un système de RV sur la performance cognitive du participant, il est envisagé de comparer le niveau de charge mentale impliqué à travers des techniques d'interactions et des conditions d'immersion différentes.

— **Hypothèse 1 : la charge mentale est modulée par la technique d'interaction.**

Différentes techniques d'interaction peuvent induire différents degrés de charge mentale, que ce soit au niveau de l'immersion, de la sélection ou de la méthode pour effectuer des rotations de la caméra. Ainsi, nous supposons que :

1. Une condition immersive génère plus de charge mentale qu'une condition non-immersive. Pour répondre à cette hypothèse nous comparerons la performance de l'utilisateur avec le Razer Hydra avec *head tracking* et le Razer Hydra avec le HMD.
2. Une technique d'interaction avec une sélection en 2D génère moins de charge mentale qu'une sélection en 3D. Une majorité des techniques de sélection permettent une sélection sur un plan 2D. Pour faciliter l'interaction avec le système, la profondeur n'est pas prise en compte. Cependant les casques de RV plus récents sont fournis avec des manettes permettant une sélection en 3D, celles-ci conduiraient à se sentir plus immergé dans l'EV, mais aussi à créer une situation plus écologique en tentant de reproduire les différentes dimensions du monde réel. Il a donc été choisi de comparer la charge mentale en fonction d'une sélection en 2D et en 3D. La souris et le gamepad seront comparés au Razer Hydra.
3. Une technique d'interaction où les rotations de la caméra sont effectuées par les mou-

vements de la main génère moins de charge mentale qu'une où elles sont effectuées par les mouvements de la tête. Les techniques d'interaction utilisées le plus fréquemment dans le quotidien des personnes que ce soit au travail ou pour jouer effectuent les rotations de la caméra avec les mouvements de la main. Cependant pour améliorer l'immersion et reproduire plus fidèlement la réalité, il est préférable d'orienter sa vue grâce aux mouvements de la tête. Pour comparer ces deux modes d'orientation du champ visuel, la souris et le gamepad seront comparé au Razer Hydra avec *tracking* de tête (Kinect 2 et HTV Vive).

- **Hypothèse 2 : la charge mentale est modulée par le profil de l'utilisateur.**
Nous supposons qu'un utilisateur expert avec la souris, aura de meilleurs résultats à une tâche en réalité virtuelle car il fera moins d'erreurs liées à l'utilisabilité de la technique d'interaction qu'un utilisateur novice. Nous pouvons penser que les utilisateurs experts d'une technique d'interaction feront moins d'erreurs liées à l'utilisabilité de la technique d'interaction et pourront se concentrer sur la tâche à accomplir seulement.
- **Hypothèse 3 : la charge mentale est modulée par la difficulté de la tâche.**
Nous supposons que la charge mentale sera plus élevée lorsque la difficulté de la tâche augmente. Ainsi, les utilisateurs auront plus de charge mentale dans la condition charge mentale élevée du VMT, avec l'ajout des tâches inattendues, que dans la condition charge mentale faible.
- **Hypothèse 4 : une charge mentale trop élevée impacte négativement la performance du participant dans l'EV.** La hausse de la charge mentale induit :
 1. Un temps de réalisation plus long. Les individus avec une charge mentale élevée auront un temps de réalisation plus long que les individus avec une charge mentale faible.
 2. Des erreurs d'interaction plus nombreuses. Les individus avec une charge mentale élevée effectueront plus de clics sur des cibles non valides qu'un individu avec une charge mentale faible.

6.2 Outils pour le recueil de données

Les outils utilisés dans le cadre de l'expérimentation sont les suivants :

- **Un questionnaire général** (annexe B). Ce questionnaire a été construit pour avoir un aperçu du participant (âge, sexe, etc.) et de son utilisation des techniques d'interaction qu'elles soient utilisées pour le travail ou pour le divertissement.
- **Une tâche de sélection**. Cette tâche a pour but de tester les habiletés des participants avec l'utilisation d'une technique d'interaction. Elle est basée sur les normes ISO 9241-9 où le sujet doit sélectionner le plus rapidement possible les 13 cibles en forme de cylindre les unes à la suite des autres (Fig. 6.1). Au moment de la sélection, un retour auditif survient avec deux sons différents : un pour une sélection réussie et un second pour une mauvaise sélection. A l'issue de la tâche, le *throughput* est calculé.
- **Le VMT**. Cet EV est un appartement virtuel qui comprend différentes tâches pour évaluer la performance d'un utilisateur au cours d'IADL. L'intégralité de l'EV est décrit dans le chapitre précédent.
- **Le *Systeme Usability Scale* (SUS)** (Brooke et al., 1996) (annexe C). Ce questionnaire évalue l'utilisabilité d'un système de RV et de la technique d'interaction à travers une série de 10 affirmations sur lesquelles les participants se positionnent via une échelle de Likert à cinq points. Le score au SUS est compris entre 0 et 100. Le score moyen, tout type d'interface confondu, est de 69.5 (Bangor, Kortum, & Miller, 2009). Pour les besoins de l'étude, j'ai traduit le questionnaire en français.
- **Le *Presence Questionnaire* (PQ, Witmer et Singer (1998))** (annexe D), traduit en français par le laboratoire de cyberpsychologie de Université du Québec en Outaouais. Ce questionnaire mesure le sentiment de présence à l'aide d'un score global et avec cinq sous-échelles : le réalisme, la possibilité d'agir et la possibilité d'examiner liée à la technique d'interaction, la qualité de l'interface et l'auto-évaluation de la performance.
- **Le *NASA Task Load Index* (NASA-TLX, Hart et Staveland (1988))** (annexe E). Ce questionnaire est le plus utilisé pour évaluer la charge mentale (de Winter, 2014). Il a été décrit en intégralité dans le chapitre 3.
- **Le *Simulator Sickness Questionnaire* (SSQ, (Kennedy, Lane, Berbaum, & Lilienthal, 1993))** (annexe F) traduit en français par l'Université du Québec en Outaouais. Ce

questionnaire permet de quantifier la présence de symptôme de cybermalaise lors d'une expérience virtuelle. Il se compose de 16 questions divisées en trois échelles : oculomoteur, désorientation spatiale et nausée. La somme de chaque échelle est multiplié par un poids. Le résultat est ensuite additionné pour chaque échelle et est multiplié par un nouveau poids, ce qui donne le score de cybermalaise. Un score en dessous de 10 indique des symptômes minimaux alors qu'un score supérieur à 20 indique des problèmes de cybermalaise. Dans sa version française, il comprend 2 sous-échelles (nausée et oculomoteur) permettant d'évaluer les sensations de mal-être suite à l'immersion dans un système de RV. Le questionnaire se réalise en pré puis en post-immersion. En effet, la première réalisation sert de baseline pour comparer l'apparition d'éventuels symptômes avant et après l'immersion. Malgré le fait que ce soit un ancien questionnaire et que les systèmes de RV ont eu une importante évolution technologique, il est toujours pertinent de nos jours (Balk, Bertola, & Inman, 2013).

- **Le Kit BiTalino**¹. Il permet l'enregistrement de données physiologiques à l'aide d'électrodes à usage unique. Les données cardiaques sont relevées par le système BiTalino puis envoyées par bluetooth au logiciel OpenSignal qui se charge de les retranscrire dans un fichier .txt. Les données cardiaques sont ensuite analysées dans un logiciel externe (Fig. 6.2).



FIGURE 6.1 – Capture d'écran de la tâche de sélection.

6.3 Procédure expérimentale

La procédure de l'expérimentation est représentée sur la figure suivante (Fig. 6.3). Après avoir rempli l'accord de participation, les participants commencent l'expérimentation en répon-

1. <http://www.bitalino.com>

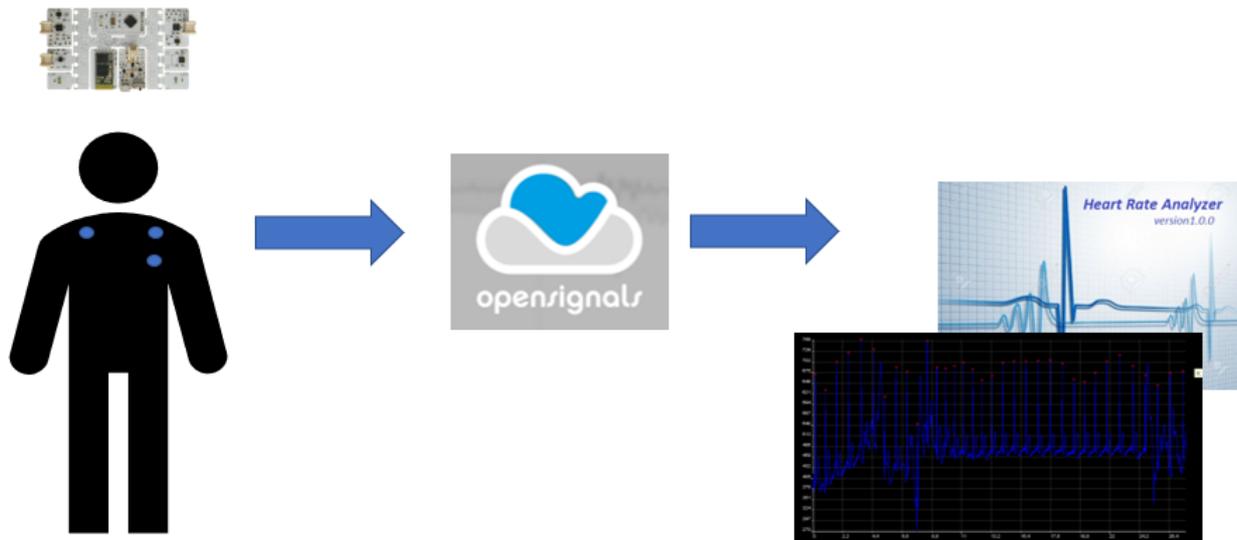


FIGURE 6.2 – Illustration de la récolte des données cardiaques.

dant au questionnaire général puis au SSQ. Ils effectuent ensuite la tâche de sélection avec l'une des cinq techniques d'interaction (le choix de la technique est généré aléatoirement). Il est placé, sur les participants, les électrodes à usage unique. Une fois l'enregistrement du rythme cardiaque au repos terminé, les participants réalisent une phase de familiarisation pour apprendre à interagir dans l'EV.

Au cours de la phase de familiarisation, les participants apprennent à interagir dans le VMT, c'est-à-dire à naviguer et à sélectionner des objets géométriques (des cubes). Cet entraînement permet aux participants de visiter l'appartement virtuel et de s'entraîner à naviguer d'un point A à un point B en créant une carte mentale de l'EV. La visite de l'appartement, avant l'expérimentation, est un point nécessaire pour garantir une navigation efficace (Sayers, 2004). Lors de la visite de l'EV, les participants doivent suivre la consigne suivante :

« Vous allez maintenant pouvoir visiter l'appartement de votre ami. Dans un premier temps, je vais vous demander de réaliser les actions que je vous indique. Dans un second temps, vous pourrez vous déplacer où vous le souhaitez. Cependant, il vous est demandé de ramasser les cubes que vous verrez. Il y a trois cubes. Comme vous ne pouvez porter qu'un cube à la fois vous devrez déposer le premier sur la table de la cuisine avant de prendre le second. Pour ramasser un cube, il vous suffit de cliquer dessus. »

« Maintenant avancez et allez ouvrir la porte qui est devant vous, en cliquant dessus. Vous voici dans l'entrée et voici le salon. Allez dans le bureau, sur votre gauche. Faxer un document. Maintenant revenez sur vos pas et ouvrez la porte. Voici la salle de bain. Continuer

dans le couloir et ouvrez cette porte. Voici la chambre d'ami. Ouvrez maintenant le placard sur le côté. Ouvrez la machine à laver et programmez la en appuyant sur le bouton ici. Maintenant continuez le couloir et allez dans la cuisine. Allez préchauffer le four, l'ouvrir et le programmer. Maintenant vous pouvez vous rendre dans la chambre principale. Nous avons terminé la visite. Vous pouvez vous rendre où vous le souhaitez, et déplacer les trois cubes.

Dites-moi quand vous avez terminé. »

Les participants disposent de cinq minutes pour apprendre la recette de cuisine permettant de réaliser la tâche du poulet (annexe G.1). Suite à cela ils sont prêts à réaliser le VMT dans la condition charge mentale faible ou dans la condition charge mentale élevée avec une des techniques d'interaction. Le choix de la condition de charge mentale et de la technique d'interaction est généré aléatoirement.

Dans la condition de charge mentale faible il est annoncé au participant :

« Vous allez vous rendre dans le même appartement virtuel que celui que l'on a visité ensemble. C'est l'appartement de votre amie. Ce soir vous irez à un spectacle ensemble. Comme le spectacle début tôt et que votre amie termine tard son travail, elle vous a demandé de l'aide à la maison. Elle vous a demandé de ranger les courses, préparer le repas, c'est-à-dire faire la recette du poulet rôti, et de mettre la table pour deux personnes. »

Pour la condition de charge mentale élevée, le même discours est tenu au participant. Une phrase est toutefois ajoutée :

« De plus, si le téléphone sonne, elle vous demande d'y répondre. »

6.4 Dispositifs d'interaction

Le test est affiché sur un écran de 117 cm pour les conditions non-immersives et sur le casque HTC Vive pour la condition immersive. Les outils d'interaction sont les suivant :

- *Souris*. Une souris optique de la marque Dell est utilisée.
- *Gamepad*. Un gamepad Xbox 360 a été utilisé. La manette de jeux vidéo Xbox consiste en deux joysticks analogiques destinés à être utilisé par les pouces de la main droite et gauche. Le joystick droit est utilisé pour effectuer les translations, donc le *travelling* dans

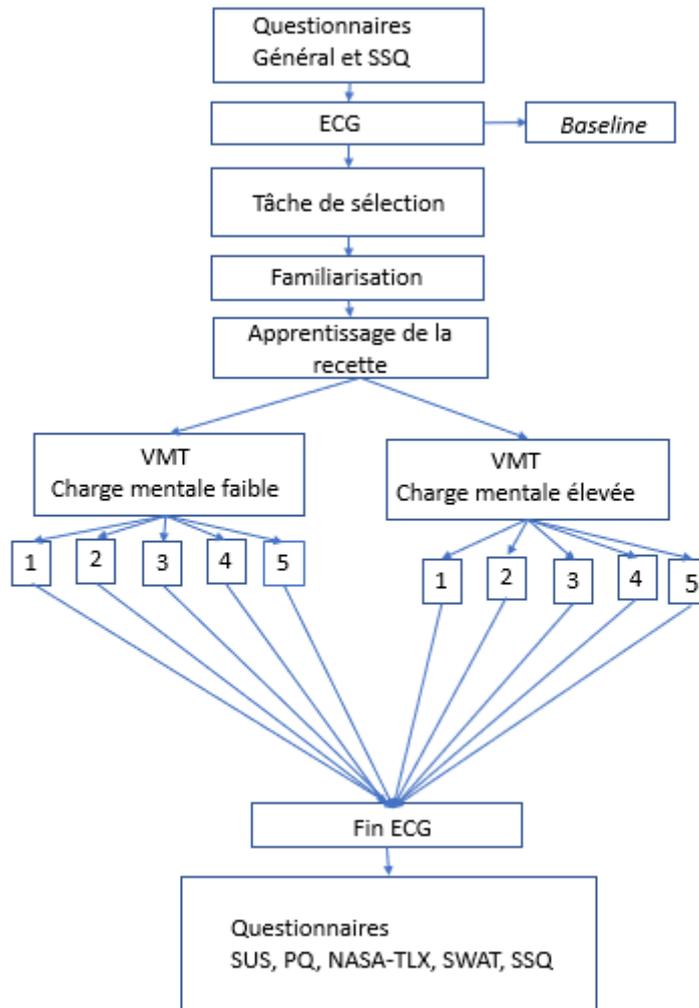


FIGURE 6.3 – Procédure de l’expérimentation. Les techniques d’interaction sont numérotées de 1 à 5. 1 : souris/clavier. 2 : gamepad. 3 : Razer Hydra. 4 : Razer Hydra et *tracking* de tête. 5 : Razer Hydra et HMD.

l’EV, tandis que le gauche permet d’effectuer des rotations, c’est-à-dire le *steering* du point de vue de l’utilisateur.

- *Razer Hydra*. Le Razer Hydra comprend deux manettes de jeu vidéo avec chacune un joystick, une gâchette et 5 boutons. Chacun des joysticks sont attachés à une station de base qui suit leur position et leur orientation grâce à un champ magnétique. La présence des joysticks et des boutons sur les manettes donne un point de référence aux personnes habituées à utiliser des gamepads. Lors de l’expérimentation, une table en bois a été utilisée pour éviter les conflits électro-magnétique avec le Razer Hydra. De plus, les objets métalliques étaient éloignés et il était demandé aux participants de ne pas garder leur téléphone portable sur eux ou les autres objets pouvant provoquer des interférences.

- *Kinect 2*. La Kinect 2 est un outil de capture de mouvements qui fonctionne grâce à une caméra et à un émetteur infrarouge. L'outil permet de capter la position et l'orientation du corps de l'utilisateur ainsi que celles de son visage.
- *HTC Vive*. Le casque de RV HTC Vive est un dispositif d'affichage immersif. L'orientation et la position du casque sont reconnus par deux émetteurs infrarouges. La zone couverte par les émetteurs va jusqu'à 4M x 3M.

6.5 Participants à l'étude

Cinquante-six sujets ont participé à l'étude. Les participants sont des étudiants de l'Université d'Angers bénévoles et volontaires. Je ne les connaissais pas au préalable. Ceux-ci ont été répartis de façon aléatoire en 2 conditions de charge mentale : charge mentale faible (noté LowWorkload) et charge mentale élevée (noté HighWorkload). Vingt-deux participants ont réalisé la condition charge mentale élevée et 34 la condition charge mentale faible. La répartition des participants par condition et technique d'interaction est générée aléatoirement par un programme. Il était attendu de une vingtaine de participants par groupe mais la campagne de recrutement n'a bien marché ce qui a entraîné une répartition non homogène des participants au sein des groupes. Au cours d'une des deux conditions, les participants ont pu interagir avec l'une des techniques d'interaction suivante : souris, gamepad, Razer Hydra, Razer Hydra et Razer Hydra et *tracking* de tête effectué avec la Kinect (RazerHT) ou Razer Hydra et visiocasque (RazerHMD) (Fig. 6.4).

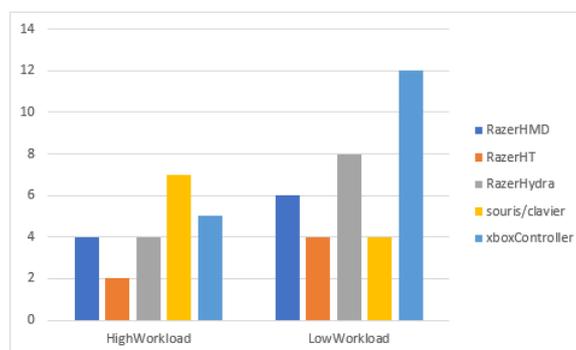


FIGURE 6.4 – Répartition des participants par conditions et par techniques d'interaction.

Les participants sont divisés en plusieurs groupes :

- *Condition de charge mentale faible* :
 - Souris (noté *Souris_{Low}*) : 4 participants (âge moyen : 25.4 ; écart-type : 5.7) dont 1 femme et 3 hommes.
 - Gamepad (noté *Gamepad_{Low}*) : 12 participants (âge moyen : 23.2 ; écart-type : 3.9) dont 4 femmes et 8 hommes.
 - Razer Hydra (noté *Razer_{Low}*) : 8 participants (âge moyen : 24 ; écart-type : 5.1) dont 4 femmes et 4 hommes.
 - Razer Hydra avec le HMD (noté *RazerHMD_{Low}*) : 6 participants (âge moyen : 23 ; écart-type : 4.3) dont 3 femmes et 3 hommes.
 - Razer Hydra avec le *head tracking* (noté *RazerHT_{Low}*) : 4 participants (âge moyen : 23.9 ; écart-type : 4.5) dont 2 femmes et 2 hommes.
- *Condition de charge mentale élevée* :
 - Souris (noté *Souris_{High}*) : 7 participants (âge moyen : 22.4 ; écart-type : 4.3) dont 4 femmes et 3 hommes.
 - Gamepad (noté *Gamepad_{High}*) : 5 participants (âge moyen : 22 ; écart-type : 1.5) dont 2 femmes et 3 hommes.
 - Razer Hydra (noté *Razer_{High}*) : 4 participants (âge moyen : 23.3 ; écart-type : 2.5) dont 3 femmes et 1 homme.
 - Razer Hydra avec HMD (noté *RazerHMD_{High}*) : 4 participants (âge moyen : 22.7 ; écart-type : 3.3) dont 3 femmes et 1 homme.
 - Razer Hydra avec le *head tracking* (noté *RazerHT_{High}*) : 2 participants (âge moyen : 20.1 ; écart-type : 2.8) dont 0 femme et 2 hommes.

Résultats

7.1 Présentation de la base de données

La base de données concernant notre étude expérimentale rassemble les résultats de 53 observations. Trois participants ont été exclus des analyses car ils n'ont pas pu terminer le test à cause de symptômes de cybermalaise. Ces participants interagissaient dans le VMT avec le gamepad.

Deux façons permettent de faire référence aux techniques d'interaction : leur nom et un regroupement selon leurs caractéristiques. Ainsi, les techniques d'interaction ont aussi été regroupées selon leurs méthodes d'immersion, d'orientation de la caméra et de sélection.

Les données ont été séparées en deux groupes selon le type d'immersion utilisé :

- Non-immersif : Razer Hydra avec *head tracking* (noté RazerHT)
- Immersif : Razer Hydra avec HMD (noté RazerHMD)

Les données ont été séparées en deux groupes selon le type de *steering* employé :

- *Steering* avec des mouvements de la main : données issues de la souris, gamepad et du Razer Hydra
- *Steering* avec des mouvements de la tête : données issues du *RazerHT* et du *RazerHMD*

Les données ont été séparées en deux groupes selon le type de sélection employé :

- Sélection 2D : données issues de la souris et du gamepad
- Sélection 3D : données issues du Razer Hydra, *RazerHT* et du *RazerHMD*

Les analyses de profil ont été conduites par des analyses de clustering hiérarchiques et les

groupes ont été comparés entre eux par des méthodes non-paramétriques du fait d'une distribution ne suivant pas une loi normale. Les tests de Kruskal-Wallis et U de Mann-Whitney ont donc été utilisés pour les comparaisons entre groupes.

Dans un premier temps, la discrimination des observations en groupe selon leur degré d'expertise avec la technique d'interaction (novice ou expert) et selon leur profil de charge mentale ressenti au cours de l'expérimentation (faible ou fort) est effectué (chap. 7.2). Ensuite les résultats issus de la technique d'interaction sont présentés (chap. 7.3). Puis les analyses portent sur les questionnaires (chap. 7.4) et la performance au cours du VMT sera analysée (chap. 7.5). Enfin, les données cardiaques sont analysées (chap. 7.6).

7.2 Elaboration des groupes d'individus

7.2.1 Expertise du participant

Pour appréhender la performance du participant de la meilleure façon possible, il est nécessaire de prendre en compte le degré d'expertise de participant avec la technique d'interaction. En effet, la charge mentale est moins élevée lorsque l'individu a des connaissances préalables dans la réalisation de la tâche (Chanquoy et al., 2007). Pour comprendre le niveau d'expertise du participant avec l'utilisation d'une technique d'interaction nous nous sommes intéressés aux variables issues de la tâche de sélection, de la tâche de familiarisation et des variables subjectives sur la durée et la fréquence d'utilisation d'un ordinateur et des jeux vidéo. Les variables identifiées sont les suivantes :

- HRV Fitts : variabilité de la fréquence cardiaque lors de la tâche de sélection
- HRV Fami : variabilité de la fréquence cardiaque lors de la tâche de familiarisation
- Habituation : somme de la durée et de la fréquence d'utilisation des ordinateurs et des jeux vidéo
- *Throughput*
- Fitts temps : temps de réalisation de la tâche de sélection
- Fami temps : temps de réalisation de la tâche de familiarisation
- Fitts miss : nombre d'erreurs de clic à la tâche de sélection

- Fami miss : nombre d'erreurs de clic à la tâche de familiarisation
- Fami distance : distance parcourue lors de la tâche de familiarisation

Une fois les variables sélectionnées, une analyse de clustering hiérarchique (HCA) a été utilisée pour tenter d'identifier des structures de données et ainsi isoler au moins deux groupes de participants. Pour ce faire, chaque technique d'interaction a été étudiée isolément les unes des autres. Les groupes ont ensuite été comparés avec un test de Kruskal Wallis puis par un test de Mann-Withney pour comprendre les relations entre les groupes.

Discrimination des groupes avec l'utilisation de la souris

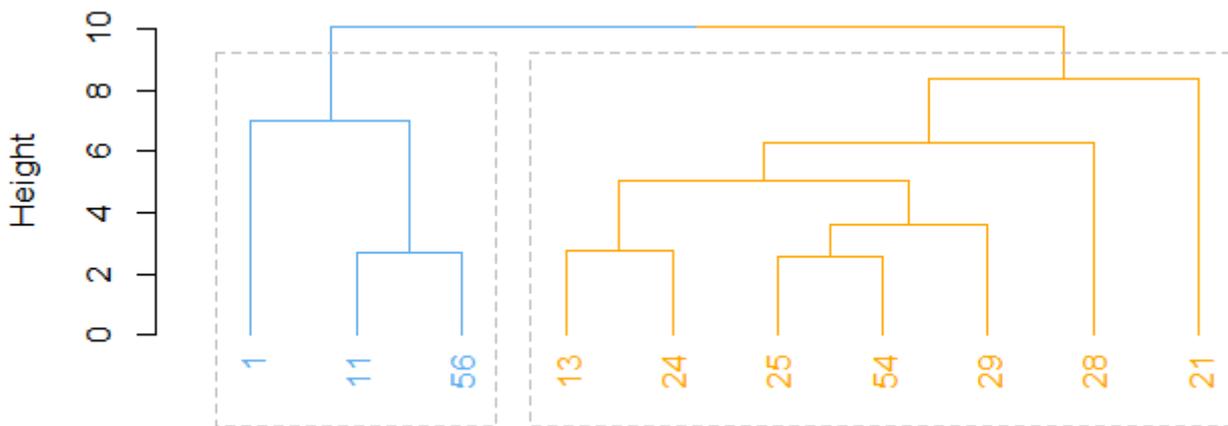


FIGURE 7.1 – Dendrogramme des participants ayant utilisés la souris.

La HCA sépare deux groupes (Fig. 7.1) : un avec trois participants et un autre avec sept participants. Ceux-ci sont nommés respectivement groupe « A » en bleu et « B » en orange.

Une différence significative pour la variable « habitude » est présente pour les différents groupes ($H(2) = 7.05; p = 0.02$). Le groupe « B » a un score d'habitude avec les ordinateurs et les jeux vidéo plus élevé que le groupe « A ». Le groupe « A » a un rythme cardiaque plus élevé que le groupe « B » ($U = 14; p = 0.05$). Le groupe « B » a un *throughput* plus élevé que le groupe « A » lors de la tâche de sélection ($U = 0; p = 0.01$).

Au regard des résultats, le groupe « B » peut être qualifié d'expert en comparaison au groupe « A » qui peut être qualifié de novice.

Discrimination des groupes avec l'utilisation du gamepad

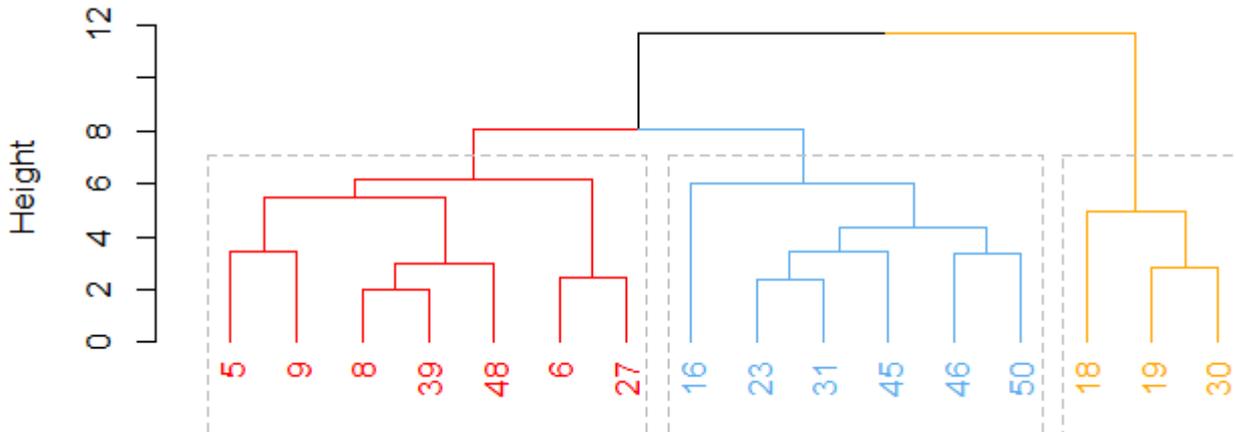


FIGURE 7.2 – Dendrogramme des participants ayant utilisés le gamepad.

La HCA sépare trois groupes (Fig. 7.2) : un avec sept participants, un avec six participants et un autre avec trois participants. Ceux-ci sont nommés respectivement groupe « A » en rouge, « B » en bleu et « C » en orange.

Une différence significative pour le rythme cardiaque est présente pour les différents groupes ($H(2) = 7.05; p = 0.02$). Le groupe « A » a un rythme cardiaque moins élevé que le groupe « B » ($U = 0; p = 0.002$) et que le groupe « C » ($U = 0; p = 0.02$). Pour la variable temps de réalisation à la tâche de sélection, une différence significative est présente pour les différents groupes ($H(3) = 8.63; p = 0.03$). le groupe « C » met moins de temps que le groupe « A » ($U = 21; p = 0.01$) et que le groupe « B » ($U = 18; p = 0.02$) pour sélectionner les cibles. Il y a une différence entre les trois groupes pour le *throughput* à la tâche de sélection ($H(3) = 8.6; p = 0.03$). Le groupe « C » a un meilleur *throughput* que le groupe « A » ($U = 0; p = 0.01$) et que le groupe « B » ($U = 0; p = 0.02$). Une différence significative pour les erreurs de clics lors de la phase de familiarisation est présente pour les différents groupes ($H(3) = 7.7; p = 0.05$). Nous pouvons aussi constater que le groupe « C » a plus d'erreur de clics dans la phase de familiarisation que le groupe « A » ($U = 0; p = 0.02$) et que le groupe « B » ($U = 0; p = 0.02$).

Les résultats indiquent que le groupe « C » peut être qualifié comme expert et que les groupes « A » et « B » peuvent être regroupés ensemble et être qualifié comme novice avec

l'utilisation du gamepad.

Discrimination des groupes avec l'utilisation du Razer Hydra

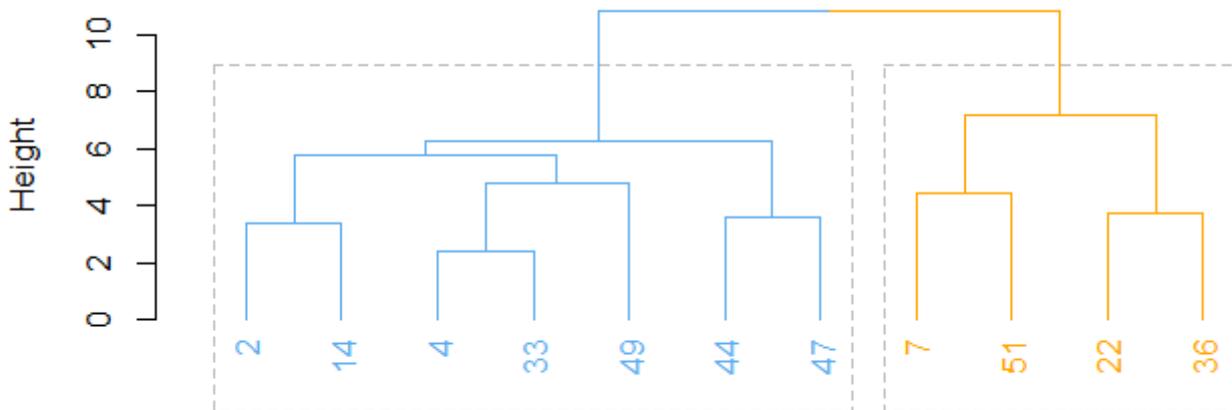


FIGURE 7.3 – Dendrogramme des participants ayant utilisés le Razer Hydra.

La HCA sépare deux groupes (Fig. 7.3) : un avec sept participants, un avec quatre participants. Ceux-ci sont nommés respectivement groupe « A » en bleu et « B » en orange.

La comparaison entre groupe indique que le groupe « A » met plus de temps pour compléter la tâche de sélection que le groupe « B » ($U = 27; p = 0.01$). Le *throughput* du groupe « B » est meilleur que celui du groupe « A » ($U = 0; p = 0.006$). Le groupe « B » parcourt plus de distance que le groupe « A » lors de la phase de familiarisation ($U = 0; p = 0.02$).

Les résultats indiquent que le groupe « B » peut être qualifié d'expert tandis que le groupe « A » peut être qualifié de novice. En effet, le groupe expert a un meilleurs temps de réalisation et un meilleur *throughput* à la tâche de sélection.

Discrimination des groupes avec l'utilisation du Razer Hydra avec le port du HMD

La HCA sépare deux groupes (Fig. 7.4) : un avec cinq participants, un avec quatre participants. Ceux-ci sont nommés respectivement groupe « A » en bleu et « B » en orange.

La seule variable discriminant les deux groupes est le *throughput* à la tâche de sélection.

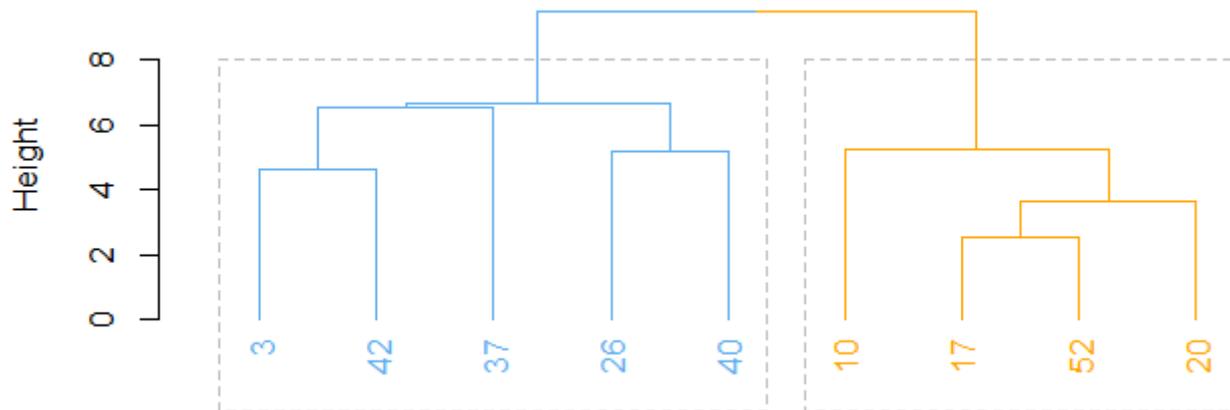


FIGURE 7.4 – Dendrogramme des participants ayant utilisés le Razer Hydra et le HMD.

Les résultats montrent que le groupe « B » a un meilleur *throughput* que le groupe « A » ($U = 0; p = 0.01$).

Le groupe « B » peut être qualifié d'expert tandis que le groupe « A » peut être qualifié de novice.

Discrimination des groupes avec l'utilisation du Razer Hydra avec le *head tracking*

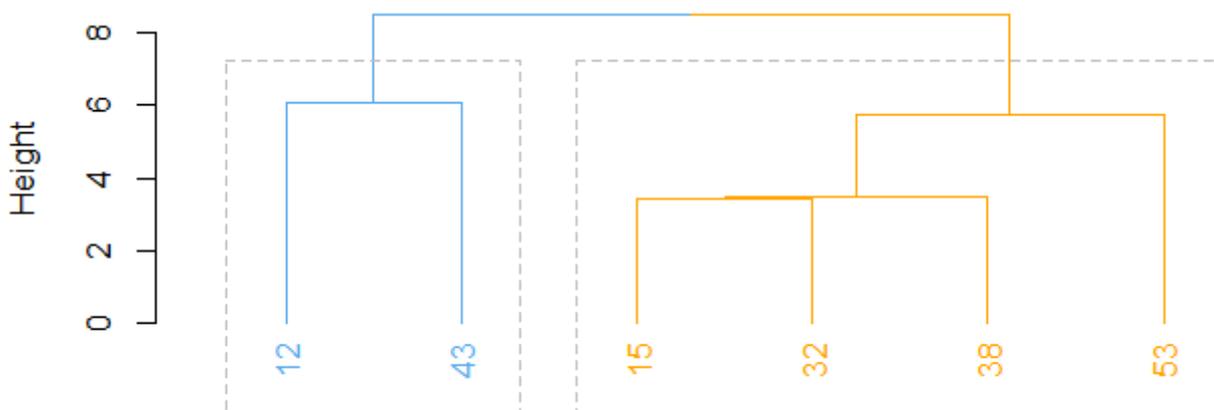


FIGURE 7.5 – Dendrogramme des participants ayant utilisés le Razer Hydra et le *head tracking*.

La HCA sépare deux groupes (Fig. 7.5) : un avec deux participants, un avec quatre participants. Ceux-ci sont nommés respectivement groupe « A » en bleu et « B » en orange.

TABLE 7.1 – Constitution des groupes d’experts ou de novices en fonction des techniques d’interaction.

	Expert	Novice
RazerHMD	4	5
Razer Hydra	4	7
Souris	7	3
Gamepad	3	13
Σ	18	28

TABLE 7.2 – Variables différenciant un groupe de novices et d’experts en fonction des techniques d’interaction.

	Souris	Gamepad	Razer Hydra	RazerHMD
HRV Fitts	-	-	-	-
HRV Fami	✓	✓	-	-
Habituation	✓	-	-	-
<i>Throughput</i>	✓	✓	✓	✓
Fitts temps	-	✓	✓	-
Fami temps	-	-	-	-
Fitts miss	-	-	-	-
Fami miss	-	✓	-	-
Fami distance	-	-	✓	-

Aucune variable ne présente de différence significative entre les deux groupes.

L’analyse indique que, ni le groupe « A » ni le groupe « B » ne peut être qualifié comme expert ou comme novice.

Synthèse des résultats sur l’expertise du participant

Il y a beaucoup d’experts avec la souris (Tab. 7.1). Les participants étant en grande majorité des étudiants, c’est une observation puisqu’ils utilisent les ordinateurs au quotidien.

La variable commune, pour discriminer les experts des novices, à l’ensemble des techniques d’interaction est le *throughput*. Comme le montre la figure suivante (Fig. 7.6), les experts ont

un meilleur *throughput* que les novices. Cette variable apparaît donc être très intéressante pour séparer un groupe de l'autre. Cette seule variable pourrait être utilisée pour catégoriser le degré d'expertise du participant. Cela permettrait de qualifier le participant plus rapidement qu'en utilisant un algorithme d'apprentissage non-supervisé comme le CAH qui demande une analyse plus longue.

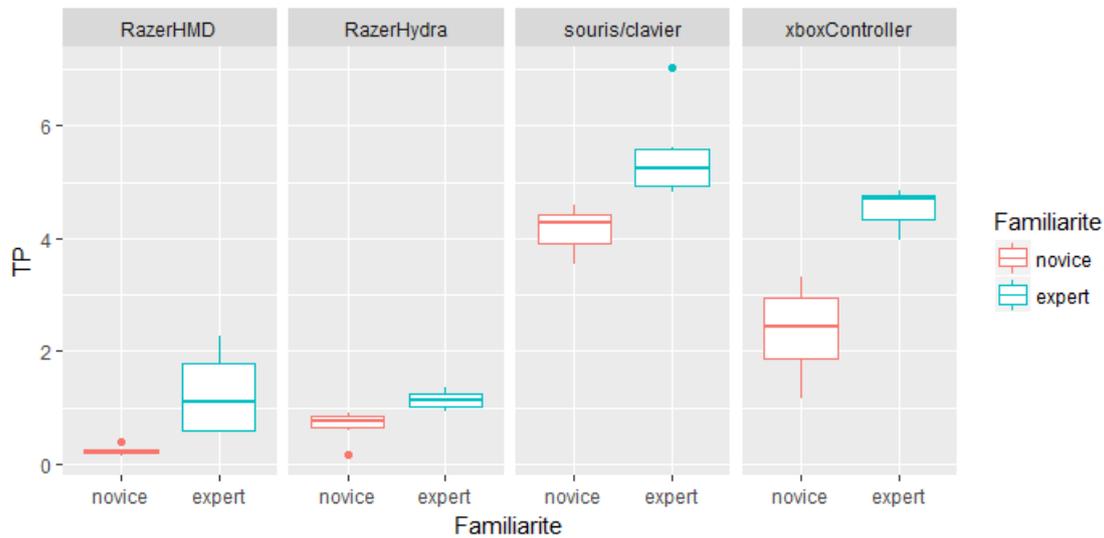


FIGURE 7.6 – Score au *throughput* en fonction du degré d'expertise des participants.

7.2.2 Elaboration du profil de charge mentale

Les variables associées à la charge mentale sont : le score au NASA-TLX qui contient la somme de l'ensemble des sous-échelles du questionnaire, le temps de réalisation du VMT et le score au VMT qui est l'addition du score aux trois tâches du VMT (la tâche du poulet, la tâche de la table et la tâche de l'épicerie). Pour catégoriser nos utilisateurs en charge mentale faible ou élevée nous avons commencé par standardiser les variables entre -1 et 1. Il est nécessaire de suivre cette procédure car les différentes variables reliées à la charge mentale n'ont pas les mêmes unités et ne sont donc pas a priori comparable entre elles. Ensuite la variable temps de réalisation et score au NASA-TLX ont été additionnées entre elles. Plus le score à ces variables est élevé plus nous pouvons supposer que la charge mentale est importante. C'est le contraire pour la variable score au VMT. Ainsi, le score au NASA-TLX et le temps ont été soustrait au score au VMT. De sorte que plus le score obtenu est important moins les utilisateurs présenteront de charge mentale et vice versa. Autrement dit, l'élaboration du profil de charge mentale des participants est faite à partir des étapes suivantes :

	Score au VMT	Temps au VMT	NASA-TLX
<i>ChargeMentale_{Low}</i>	40.71	14.05	25.68
<i>ChargeMentale_{High}</i>	35.29	27.96	57.08

TABLE 7.3 – Performance au VMT et score au NASA-TLX en fonction des groupes de profil de charge mentale.

1. Standardisation des variables A, B et C
2. $D = A + B$
3. Charge mentale = $C - D$

Le score de charge mentale a ensuite été observé en fonction des écart-types. Les données appartenant au troisième quartile ont le meilleur score et donc peu de charge mentale alors que les données du premier quartile ont le moins bon score et donc beaucoup de charge mentale.

Cette méthode permet d'isoler 14 sujets avec une charge mentale faible, 14 avec une charge mentale élevée et le reste des participants est classé en neutre. Si on compare les résultats obtenus avec ceux obtenus par une HCA, nous pouvons constater que 64% des réponses sont équivalentes entre les 2 groupes. C'est-à-dire que 64% sont classées dans les mêmes groupes avec les 2 méthodes employées. Les 36% restant sont majoritairement dû à ce que la première méthode ordonne les données et sélectionne les deux extrémités. Ainsi, toute une partie des données est autour de la moyenne et n'est pas labélisée comme les extrêmes, tandis que les analyses de clusters catégorisent l'ensemble des données. Comme nous connaissons les variables employées pour constituer les groupes et leur importance, nous pouvons rester avec la méthode utilisée en premier lieu.

7.3 Utilisation de la technique d'interaction

7.3.1 Répartition de l'utilisation des techniques d'interaction : Navigation

Selon l'angle théorique qui a été choisi pour décrire une technique d'interaction, celle-ci se définit par de la navigation et de la sélection. Les variables utilisées pour caractériser l'utilisation de la technique d'interaction lors de la navigation sont :

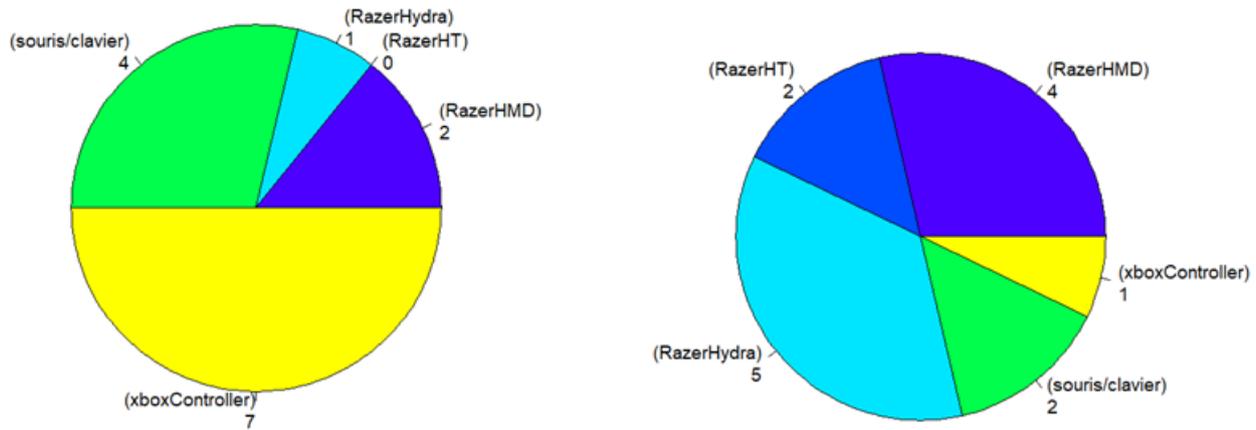


FIGURE 7.7 – Répartition des techniques d’interaction pour les participants *ChargeMentaleLow* et *ChargeMentaleHigh*.

- *Travelling* : correspond au temps où le participant navigue dans le VMT. Est exclu du temps de *travelling*, le temps où le participant effectue en même temps une rotation de la caméra (i.e., *steering*).
- *Steering* : correspond au temps où le participant effectue une rotation de la caméra dans le VMT. Est exclu du temps de *steering*, le temps où le participant effectue en parallèle une translation dans l’environnement (i.e., *travelling*).
- *Travelling + steering* : comprend le temps où le participant effectue de la locomotion et du *steering* en même temps.

Le *steering* et le *travelling* sont effectués en même temps plus souvent avec le RazerHMD que la souris ($U = 3.5; p = 0.005$) et le gamepad ($U = 14.5; p = 0.006$).

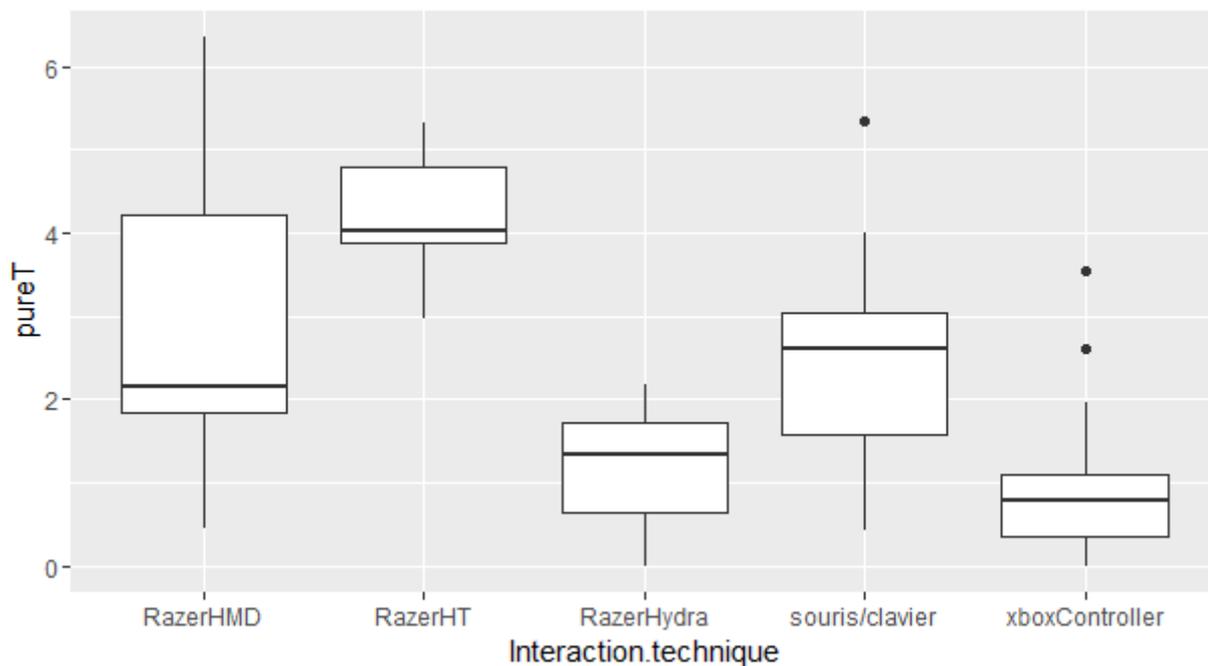
Les participants effectuent plus de *steering* avec le RazerHT que le Razer Hydra ($U = 52; p = 0.003$) et la souris ($U = 49; p = 0.01$). De plus ils en effectuent plus avec le gamepad qu’avec le RazerHMD ($U = 49; p = 0.02$), le Razer Hydra ($U = 171; p = 0.0001$) et la souris ($U = 170; p = 0.0001$). Les hommes réalisent plus de rotation de la caméra que les femmes ($U = 223; p = 0.03$) et ce, particulièrement dans la condition de charge mentale élevée ($U = 30; p = 0.05$).

Les participants utilisent moins de locomotion avec le Razer Hydra que le RazerHMD ($U = 15; p = 0.009$), le RazerHT ($U = 11; p = 0.03$) et la souris ($U = 21.5; p = 0.02$). Ils utilisent plus

TABLE 7.4 – Score au *throughput* en fonction des techniques d'interaction.

Technique d'interaction	moyenne du <i>throughput</i> (écart-type)
Souris	5 (0.88)
Gamepad	2.6 (1.13)
Razer Hydra	0.83 (0.3)
RazerHMD	0.7 (0.7)
RazerHT	0.2 (0.1)

de locomotion avec le RazerHT que la souris ($U = 42; p = 0.05$) et le gamepad ($U = 79; p = 0.01$). Il y a plus de locomotion avec le RazerHMD que le gamepad ($p = 0.01$). Ils utilisent aussi plus de locomotion avec la souris qu'avec le gamepad ($U = 131; p = 0.01$). Les résultats sont illustrés par le graphique (Fig. 7.8).

FIGURE 7.8 – Utilisation de l'*input* de locomotion en fonction de la technique d'interaction.

7.3.2 Score au *Throughput*

Le *throughput* a été comparé pour les différentes techniques d'interaction (Fig. 7.9).

Les résultats indiquent que le Razer Hydra a un meilleur *throughput* que le RazerHT ($U =$

62; $p = 0.003$). Le gamepad a un meilleur *throughput* que le Razer Hydra ($U = 7$; $p > 0.001$), le RazerHMD ($U = 10$; $p = 0.0004$) et le RazerHT ($U = 0$; $p = 0.0006$). La souris a un meilleur *throughput* que le Razer Hydra ($U = 0$; $p > 0.001$), que le RazerHMD ($U = 0$; $p = 0.0003$), que le RazerHT ($U = 0$; $p = 0.001$) et que le gamepad ($U = 9$; $p > 0.001$).

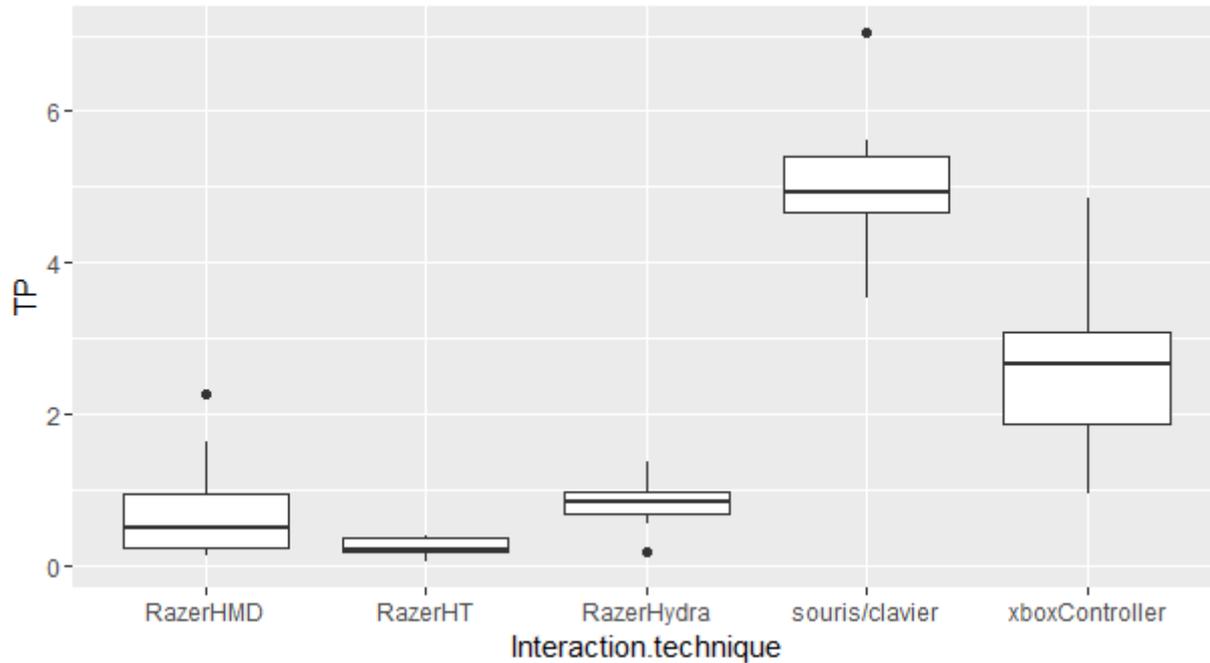


FIGURE 7.9 – Score au *throughput* en fonction de la technique d'interaction.

7.4 Analyses des questionnaires

7.4.1 Score au NASA-TLX

Les résultats au NASA-TLX indiquent que les techniques d'interaction qui utilisent une sélection en 3D génèrent un score de charge mentale plus élevé que les techniques avec une sélection en 2D ($U = 237$; $p = 0.01$). Cette différence est significative dans la condition de charge mentale faible ($U = 83$; $p = 0.03$) mais pas dans les conditions de charge mentale élevée. Les techniques d'interaction avec une sélection en 3D demandent plus de charge physique qu'avec une sélection en 2D ($U = 209.5$; $p = 0.002$). C'est particulièrement le cas pour les conditions avec une charge mentale faible ($U = 64.5$; $p = 0.005$). De plus, une sélection en 3D demande plus d'effort qu'une sélection en 2D ($U = 234$; $p = 0.009$).

Les résultats au NASA-TLX indiquent que les techniques d'interaction qui effectuent la rotation de la caméra via les mouvements de la tête génèrent un score de charge mentale plus

élevé que celles qui font les rotations de la caméra avec un mouvement de la main pour les conditions de charge mentale faible ($U = 68; p = 0.05$). Le *steering* avec la tête génère plus d'implication physique que celui avec la main ($U = 217.5; p = 0.05$). Il génère aussi plus d'efforts ($U = 192; p = 0.02$), cela est particulièrement le cas pour les conditions de charge mentale faible ($U = 59.5; p = 0.02$). Le *steering* avec la tête en condition de charge mentale faible génère plus de frustration qu'en charge mentale élevée ($U = 51; p = 0.02$).

7.4.2 Score au System Usability Scale

Le gamepad a un meilleur score au SUS que le RazerHMD ($U = 110; p = 0.04$) et le RazerHT ($U = 83.5; p = 0.03$). La souris a un score plus important que le RazerHMD ($U = 110; p = 0.01$) et le RazerHT ($p = 0.01$) (Fig. 7.10). De plus, le score au SUS est négativement corrélé avec le score au NASA-TLX ($r = -0.58; p > 0.001$).

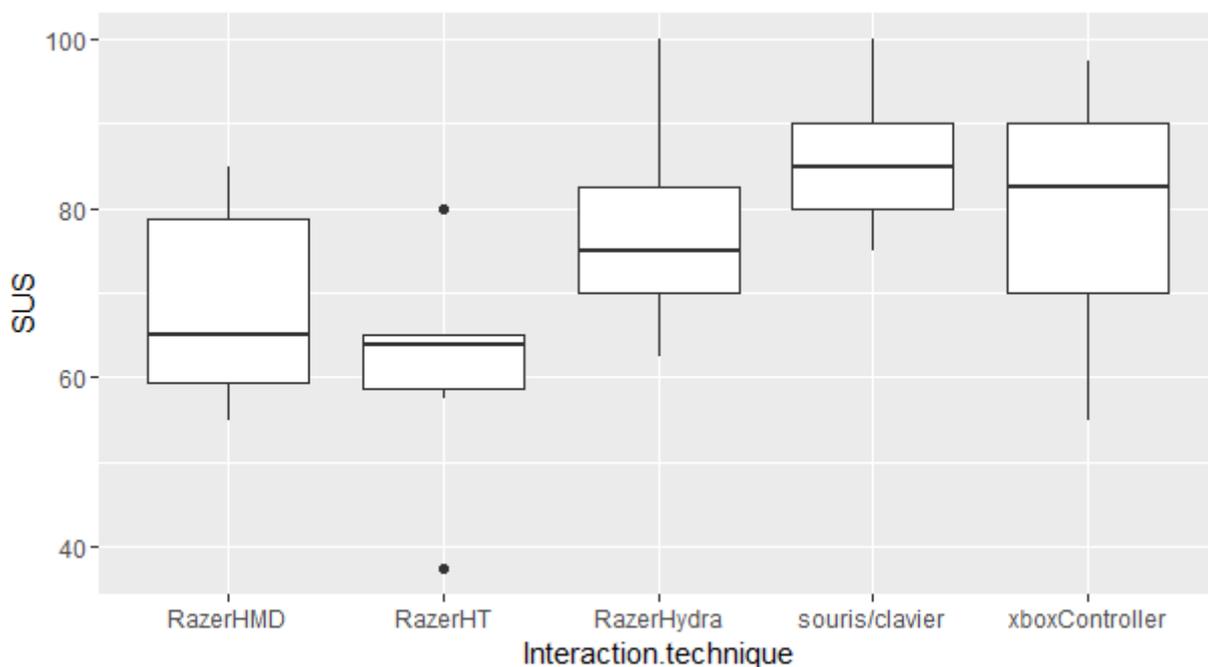


FIGURE 7.10 – Score au SUS en fonction de la technique d'interaction.

Le score au SUS est plus élevé avec une sélection 2D que 3D ($U = 537; p = 0.0009$). C'est particulièrement le cas pour les conditions de charge mentale élevée ($U = 106; p = 0.002$). Le SUS est plus élevé lorsque la rotation de la caméra est effectuée par les mouvements de la main que de la tête ($U = 467; p = 0.003$). Le SUS est plus élevé en immersif pour les conditions de charge mentale faible qu'en non-immersif ($U = 2; p = 0.05$). Les participants avec un score de charge mentale fort ont un score au SUS moins élevé que les participants avec un score de charge mentale faible ($t = 4; p = 0.0003$).

7.4.3 Score de présence

Les experts ont un meilleur score à la sous-échelle possibilité d'agir que les novices ($U = 315.5; p = 0.01$). La souris renseigne un meilleur score à la sous-échelle performance que le Razer Hydra ($U = 102; p = 0.05$).

Il y a plus de possibilité d'agir avec une sélection en 2D qu'en 3D ($U = 505.5; p = 0.005$). Cette différence est significative pour la condition charge mentale élevée ($U = 93.5; p = 0.02$), mais pas pour la condition de charge mentale faible. Les experts avec la technique d'interaction ont plus de possibilité d'agir que les novices ($U = 134.5; p = 0.02$). Les participants $ChargeMentale_{High}$ ont moins de possibilité d'agir que ceux $ChargeMentale_{Low}$ ($U = 2.67; p = 0.01$).

La qualité de l'interface est ressentie comme meilleure en condition de charge mentale élevée comparée à la condition de charge mentale faible pour la sélection en 2D ($U = 141.5; p = 0.01$). Elle est aussi meilleure avec une sélection en 2D qu'en 3D dans la condition de charge mentale élevée ($U = 104.5; p = 0.003$). Les participants $ChargeMentale_{Low}$ ont une meilleure qualité d'interface que ceux avec $ChargeMentale_{High}$ ($U = 34.5; p = 0.003$).

La performance des participants est ressentie comme meilleure avec une sélection en 2D qu'une sélection en 3D, ce dans les deux conditions de charge mentale ($U = 545.5; p = 0.0004$). Les participants avec $ChargeMentale_{Low}$ ont aussi une meilleure performance ressentie que ceux dans le profil $ChargeMentale_{High}$ ($U = 24; p = 0.0006$).

Le sentiment de présence total est meilleur avec une sélection en 2D qu'en 3D ($U = 466.5; p = 0.04$). En particulier pour les conditions de charge mentale élevée ($U = 94.5; p = 0.02$). Le sentiment de présence est plus élevé avec une rotation de la caméra effectuée par les mouvements de la main que de la tête ($U = 393; 0.03$). Ce particulièrement pour les conditions de charge mentale élevée ($U = 76.5; p = 0.03$). Les participants avec une charge mentale faible ont un sentiment de présence plus élevé que ceux avec une charge mentale forte ($U = 38; p = 0.006$).

7.4.4 Score de cybermalaise

Le sentiment de cybermalaise augmente entre le pré et le post-questionnaire ($p > 0.001$). Entre le pré et le post-test, on passe d'une moyenne de 3.6 à une moyenne de 8.7.

Le RazerHMD génère plus de symptômes de cybermalaise que la souris ($U = 1; p = 0.01$) et le gamepad ($U = 0.5; p = 0.02$). Et ce, principalement pour la condition de charge mentale élevée. Il y a donc plus de symptômes de cybermalaise en sélection en 3D qu'en 2D, et ce, principalement pour la condition de charge mentale élevée ($U = 18; p = 0.005$). Il y a aussi, plus de cybermalaise avec une rotation de la caméra effectuée par les mouvements de la tête que de la main. On trouve une différence significative pour la condition de charge mentale élevée ($U = 18; p = 0.02$), mais pas pour la condition de charge mentale faible.

Pour une meilleure compréhension du score de cybermalaise nous pouvons nous intéresser aux échelles « nausée » et « occulo-moteur ». Les symptômes « nausée » augmentent entre le pré et le post-questionnaire ($t = 4.43; p > 0.001$). Le RazerHMD a un score plus élevé à l'échelle « nausée » que le RazerHT ($U = 47; p = 0.04$), la souris ($U = 93.5; p = 0.005$) et le gamepad ($U = 122; p = 0.04$). Ces symptômes sont particulièrement développés lors des conditions de charge mentale élevée où le RazerHMD génère plus de nausées que la souris ($U = 1; p = 0.01$) et le gamepad ($U = 0; p = 0.01$). La sélection en 3D génère donc plus de symptômes, et ce particulièrement lors de la condition de charge mentale élevée ($U = 12.5; p = 0.001$). Il y a des résultats équivalents pour la rotation de la caméra avec les mouvements de la main qui génèrent moins de symptômes que lorsque c'est effectué avec les mouvements de la tête et ce, particulièrement pour la condition de charge mentale élevée ($U = 14; p = 0.01$).

Les symptômes occulo-moteur augmentent entre le pré et le post-questionnaire ($t = 2.89; p = 0.0002$). Le RazerHMD apporte un score occulo-moteur plus élevé que la souris ($U = 86; p = 0.03$). Il y a donc plus de symptômes occulo pour la sélection en 3D qu'en 2D. On trouve une différence significative pour la condition de charge mentale élevée ($U = 28; p = 0.03$) mais pas pour la condition de charge mentale faible. Il y a aussi plus de symptômes avec la rotation de la caméra effectuée par la tête que par les mains ($U = 179.5; 0.03$).

Les participants soumis à une *ChargeMentale_{High}* ont plus de symptômes de cybermalaise que ceux soumis à une *ChargeMentale_{Low}* ($U = 148.5; p = 0.02$). Ils ont plus de symptômes liés aux nausées ($U = 142; p = 0.04$) et aux troubles occulo-moteurs ($U = 153; p = 0.01$).

On ne trouve pas de différence significative en fonction des conditions et du degré de familiarité. Un score en dessous de 10 indique des symptômes minimaux alors qu'un score supérieur

à 20 indique des problèmes de cybermalaise (Kennedy et al., 1993). Il est donc possible de considérer que l'EV génère des symptômes minimaux de cybermalaise et que ceux-ci sont liés au port du HMD et au degré de charge mentale ressenti lors de la tâche.

7.5 Analyse de la performance au VMT

7.5.1 Phase de familiarisation

La durée de la phase de familiarisation n'est pas fixée. Chaque participant suit les mêmes consignes, mais on lui laisse ensuite la liberté de se balader dans l'appartement et de stopper lorsqu'ils se sentent prêt.

Temps de réalisation

Les participants ont une phase de familiarisation plus longue avec le RazerHT qu'avec le RazerHMD ($U = 4; p = 0.007$), la souris ($U = 60; p = 0.008$) et le gamepad ($U = 89; p = 0.0005$). Ils ont un temps de réalisation plus long avec le RazerHMD que pour la souris ($U = 79; p = 0.02$) et le gamepad ($U = 110.5; p = 0.008$). Avec le Razer Hydra, le temps est plus long que pour la souris ($U = 96; p = 0.007$) et que le gamepad ($U = 141; p = 0.001$). La figure 7.11 illustre les résultats.

Pour les variables issues de la performance au VMT, nous constatons que les participants mettent plus de temps pour compléter la tâche avec une sélection en 3D qu'une sélection en 2D ($U = 211; p = 0.01$). C'est particulièrement le cas pour les conditions de charge mentale faible ($U = 68; p = 0.04$).

Les participants *ChargeMentale_{High}* complètent la phase de familiarisation en plus de temps que les participants *ChargeMentale_{Low}* ($U = 146; p = 0.001$).

Nombre d'erreurs de clic

Les participants font plus d'erreurs de clics avec le Razer Hydra que le RazerHMD ($p = 0.01$), que la souris ($p = 0.0001$) et que le gamepad ($p > 0.001$). Il y a plus d'erreurs avec le RazerHMD qu'avec la souris ($p = 0.0002$) et le gamepad ($p = 0.0001$). Il y a plus d'erreurs avec le RazerHT qu'avec la souris ($p = 0.001$) et le gamepad ($p = 0.0007$).

Les participants *ChargeMentale_{High}* font plus d'erreurs de clics que les participants *ChargeMentale_{Low}* ($U = 122.5; p = 0.05$).

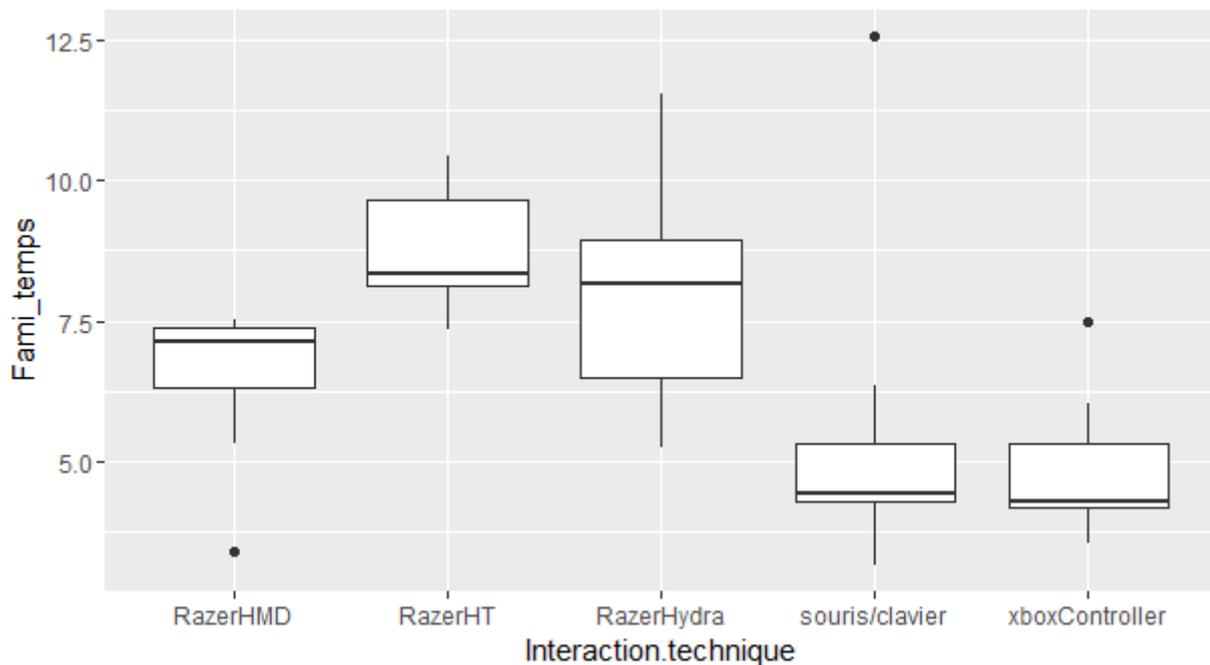


FIGURE 7.11 – Temps de réalisation de la phase de familiarisation en fonction de la technique d'interaction.

7.5.2 Temps de réalisation du VMT

Le VMT est réalisé plus rapidement, lors de la condition de faible charge mentale, avec le gamepad qu'avec le Razer Hydra ($U = 16; p = 0.04$) et le RazerHT ($U = 7; p = 0.05$) (Fig. 7.12). Le VMT est complété plus rapidement avec une sélection en 2D qu'en 3D ($U = 211; p = 0.01$). Cela est particulièrement vrai lors de la condition charge mentale faible ($U = 68; p = 0.04$).

7.5.3 Score au VMT

Le score au VMT est un score composite qui comprend le score aux trois sous-tâches : mettre la table, ranger l'épicerie et préparer le poulet rôti (Fig. 7.13).

Lors de la condition charge mentale faible, les participants qui interagissent avec la souris ont un meilleur score que ceux qui utilisent le RazerHT ($U = 15; p = 0.05$). Ceux qui utilisent le gamepad ont un meilleur score que ceux qui interagissent avec le Razer Hydra ($U = 72; p = 0.03$) et le RazerHT ($U = 43; p = 0.02$).

Dans la condition de charge mentale élevée, les participants qui utilisent la souris ont un meilleur score que ceux qui interagissent avec le RazerHMD ($U = 25.5; p = 0.03$). Les participants qui utilisent le gamepad ont aussi un meilleur score que ceux qui utilisent le RazerHMD

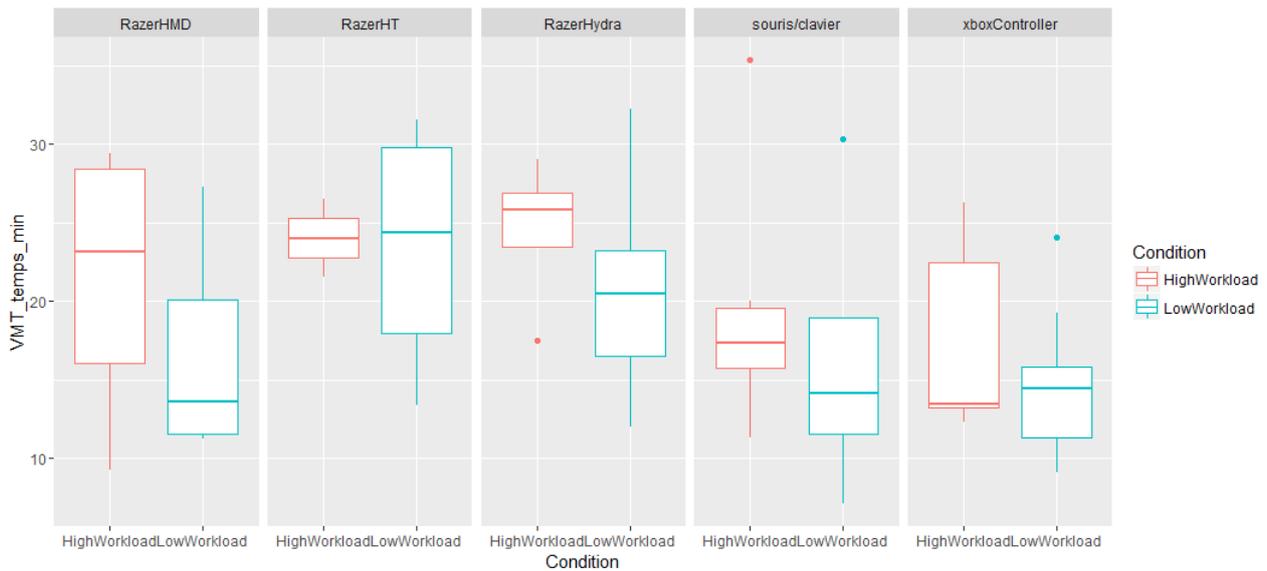


FIGURE 7.12 – Temps de réalisation du VMT en fonction des conditions de charge mentale et des techniques d'interaction.

($U = 18.5; p = 0.04$).

Les participants ont un meilleur score au VMT lorsqu'ils utilisent une technique avec une sélection en 2D qu'en 3D ($U = 584.5; p = 0.001$). Cela est le cas pour les conditions avec une charge mentale faible ($U = 21.5; p = 0.01$) et pour celles avec une charge mentale élevée ($U = 90.5; p = 0.04$). Si l'on décompose le score au VMT en sous-tâche, nous pouvons voir que les participants réussissent mieux la tâche du poulet ($U = 456; p = 0.05$) et la tâche de la table ($U = 474.5; p = 0.02$) avec une sélection 2D.

Pour les variables de performance, nous pouvons constater que les participants qui utilisent un *steering* avec les mouvements de la main ont un meilleur score que ceux qui l'effectuent avec les mouvements de la tête ($U = 457.5; p = 0.01$). C'est en particulier le cas pour les conditions où la charge mentale générée par les tâches est élevée ($U = 78; p = 0.02$). Si l'on décompose le score au VMT, nous pouvons voir que les résultats sont meilleurs avec une rotation de la caméra avec les mouvements de la main pour la tâche de la table ($U = 407; p = 0.01$), en particulier pour les conditions de charge mentale élevée ($U = 76.5; p = 0.03$).

Les conditions avec le Razer Hydra semblent générer plus de charge mentale que lorsque les participants utilisent la souris ou le gamepad. Pour mieux comprendre la performance du participants, le score au VMT peut être décomposé en sous-tâche (Fig. 7.14 et Fig. 7.15).

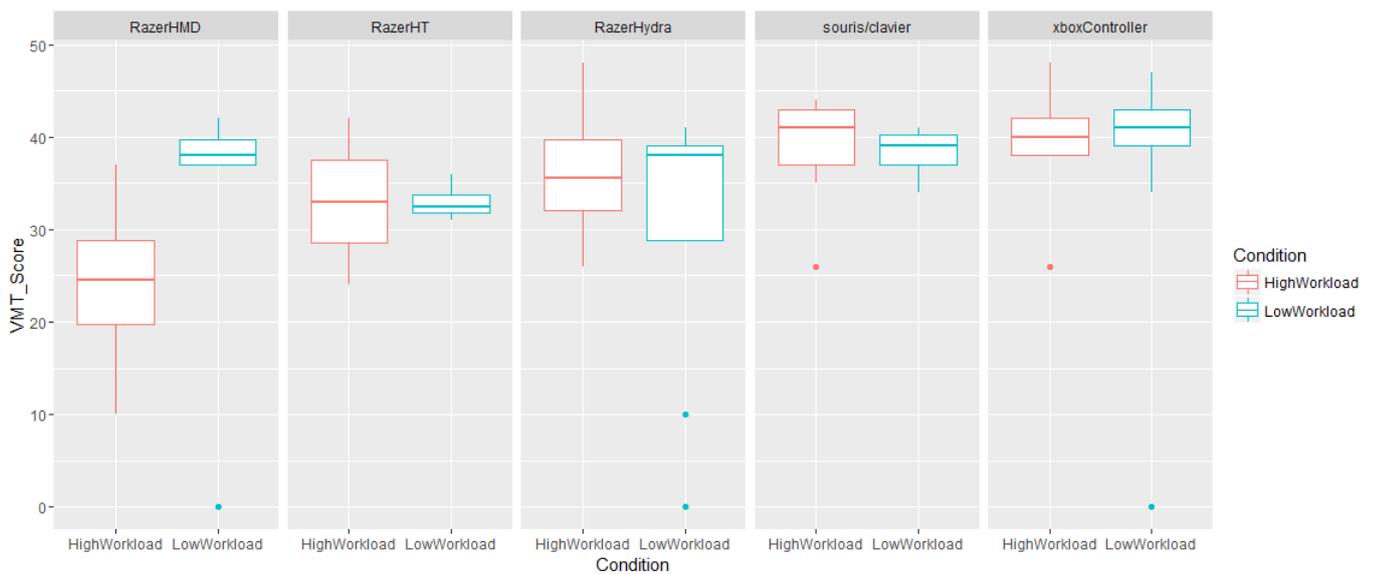


FIGURE 7.13 – Score au VMT en fonction des conditions de charge mentale et des techniques d’interaction.

Pour la tâche du poulet, lors de la condition charge mentale faible, le gamepad a un meilleur score que la souris ($U = 4; p = 0.01$) et que le Razer Hydra ($U = 59.5; p = 0.05$). Pour la tâche de la table, le RazerHT a de moins bons résultats que la souris ($U = 15.5; p = 0.03$) et que le gamepad ($U = 38.5; p = 0.03$). De plus, le RazerHMD a un meilleur score en condition charge mentale faible que charge mentale élevée ($U = 18.5; p = 0.04$).

Dans la condition charge mentale faible, c’est le gamepad qui présente les meilleurs résultats et les situations avec le RazerHT et le Razer Hydra les moins bons. Au cours des situations où la charge mentale est élevée, les participants avec le RazerHMD ont eu de moins bons résultats ; alors que ce n’était pas le cas dans la condition préalable. Lorsque la charge mentale augmente, les participants avec le RazerHMD semblent en pâtir le plus au niveau de la performance. Autrement dit, ils semblent avoir expérimenté le plus de charge mentale.

7.5.4 Autres variables

Nombre d’actions réalisées par tâches

Tâche de l’épicerie

La tâche de l’épicerie est réalisée en plus d’actions, dans la condition de charge mentale élevée, avec le gamepad que le Razer Hydra ($U = 19; p = 0.03$) et le RazerHMD ($U = 20; p = 0.01$). Elle est aussi réalisée en plus d’actions avec la souris que le Razer Hydra ($U = 24.5; p =$

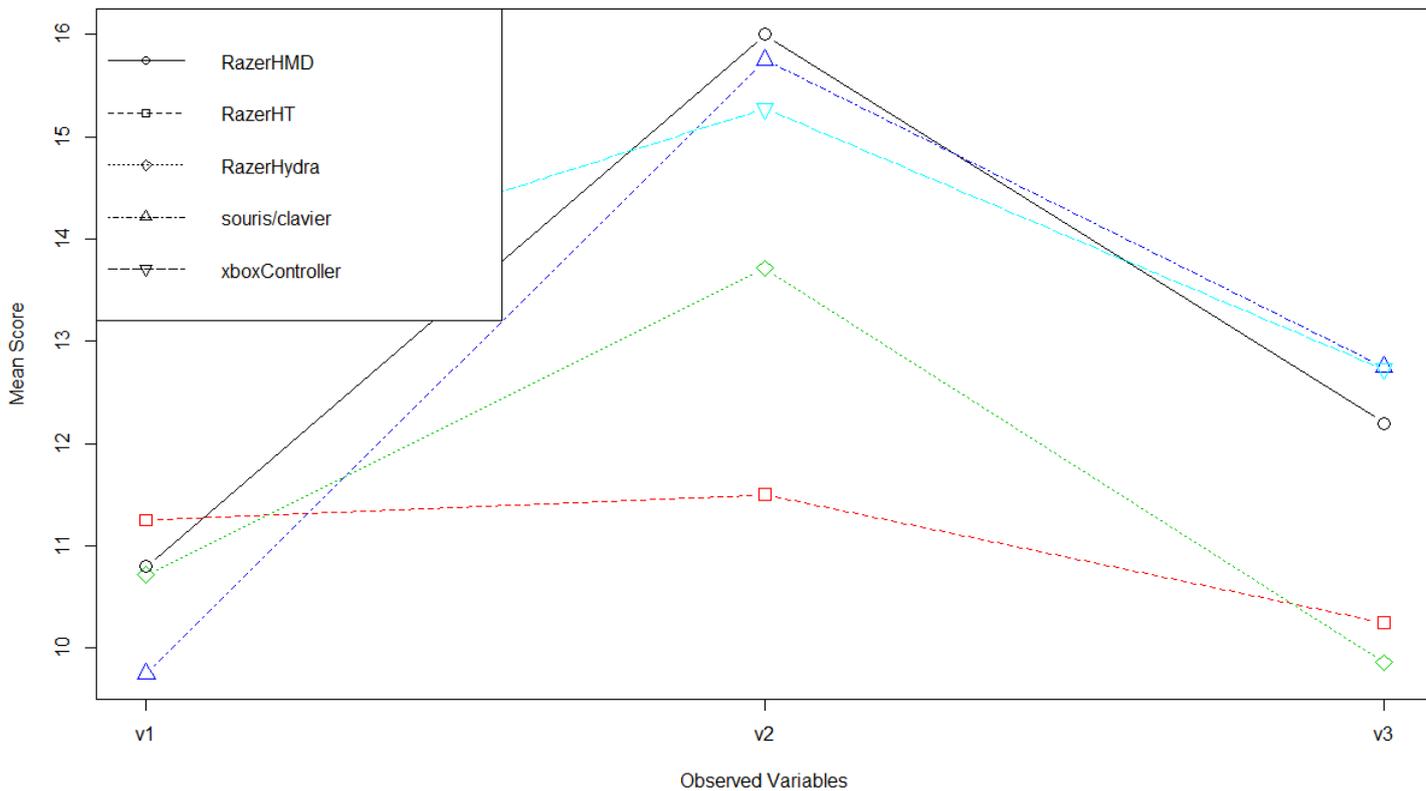


FIGURE 7.14 – Score à la tâche du poulet (v1), de l'épicerie (v2) et de la table (v3) pour la condition charge mentale faible.

0.05) et le RazerHMD ($U = 25; p = 0.03$). La tâche est donc réalisée en davantage d'actions avec une sélection 2D qu'une sélection 3D dans la condition charge mentale élevée ($U = 111; p > 0.001$) et avec un *steering* effectué par les mouvements de la main que par les mouvements de la tête ($U = 77.5; p = 0.03$). Les hommes réalisent la tâche de l'épicerie en plus d'action que les femmes ($U = 227.5; p = 0.02$). Les hommes réalisent la tâche avec une moyenne de 40 actions tandis que les femmes la réalisent avec une moyenne de 30 actions.

Tâche de la table

La tâche de la table est réalisée en plus d'actions avec un *steering* effectué par les mouvements de la tête dans la condition charge mentale faible que dans la condition charge mentale élevée ($U = 45; p = 0.03$). Les experts réalisent aussi plus d'actions que les novices ($U = 119; p = 0.009$). Pour le *steering* les résultats ne peuvent pas être comparés car lors de la condition de charge mentale élevée avec le RazerHT, 4 participants sur 6 n'ont pas complété la tâche de la table.

Tâche du poulet

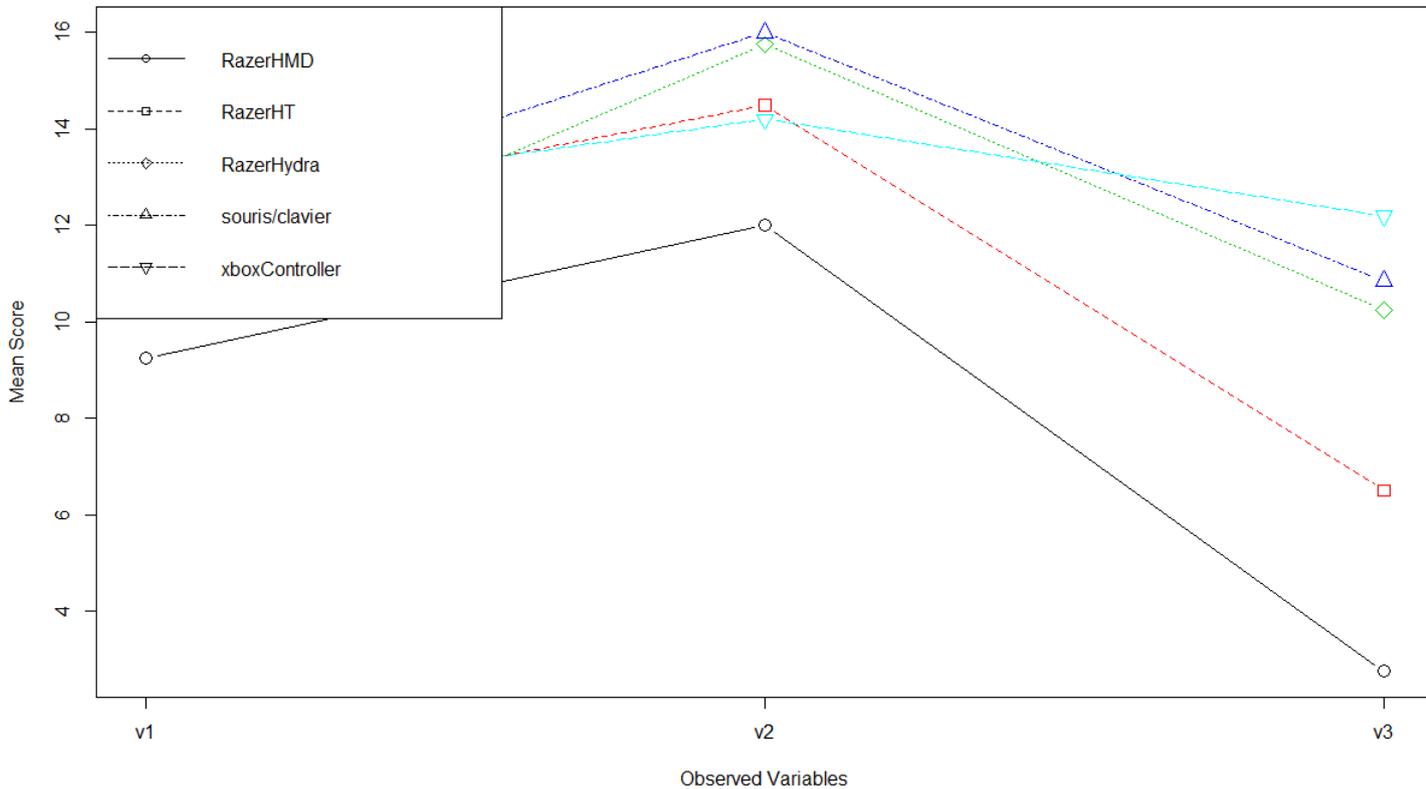


FIGURE 7.15 – Score à la tâche du poulet (v1), de l'épicerie (v2) et de la table (v3) pour la condition charge mentale élevée.

Les experts réalisent la tâche du poulet en plus d'actions que les novices ($U = 137.5; p = 0.03$). Les sujets *ChargeMentale_{High}* réalisent la tâche en plus d'actions que les participants *ChargeMentale_{Low}* ($t = 2.15; p = 0.04$).

Distance parcourue

Les utilisateurs, en condition de charge mentale faible, parcourent plus de distance avec le RazerHT que le gamepad ($U = 43; p = 0.02$) (Fig. 7.16). Les utilisateurs interagissant avec une technique d'interaction qui autorise une sélection en 3D parcourent plus de distance que ceux avec une sélection en 2D ($U = 40; p = 0.01$), c'est aussi le cas pour ceux qui effectuent les rotations de la caméra avec les mouvements de la tête ($U = 51; p = 0.03$). Les participants parcourent plus de distance en non-immersif qu'en immersif ($U = 0; p = 0.01$). Ces 3 derniers résultats ne sont valables que pour la condition de charge mentale faible.

7.5.5 Modèle de régression

Une régression multiple a été calculée pour prédire le score à la tâche du poulet du VMT en fonction de l'expertise du participant avec une technique d'interaction et la charge mentale

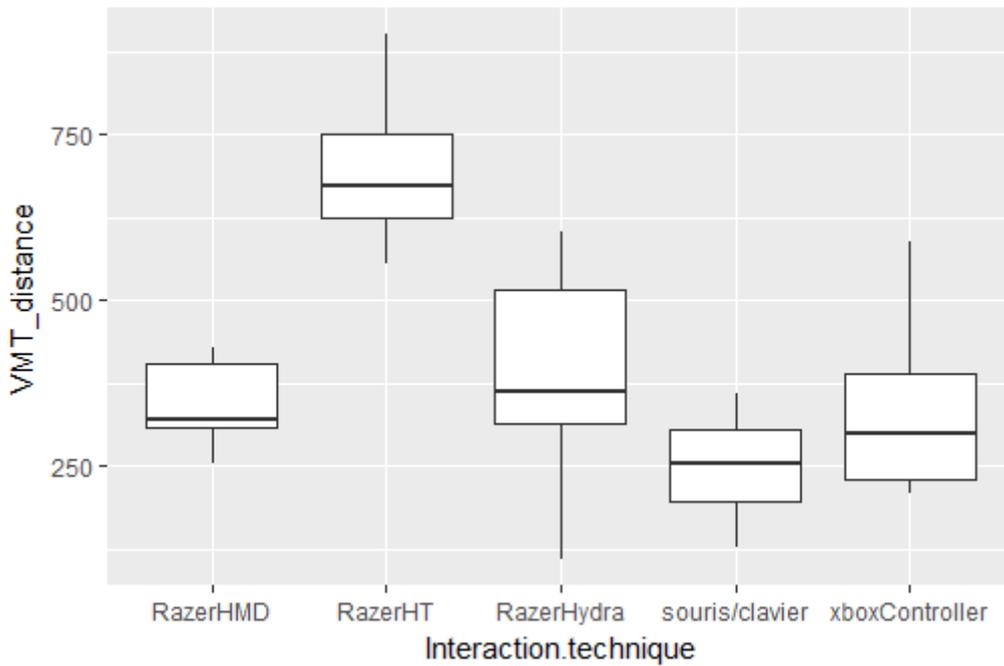


FIGURE 7.16 – Distance parcourue dans le VMT en fonction des techniques d’interaction.

lors du test. Une équation de régression significative a été trouvée ($F(5, 47) = 2.44, p < 0.05$) avec un R^2 de .20. On constate qu’un participant *ChargeMentale_{Low}* a 2 points de plus qu’un participant *ChargeMentale_{High}* lors de la tâche du poulet du VMT (Fig. 7.17).

Une régression multiple a été calculée pour prédire le nombre d’actions nécessaire pour la tâche de la table du VMT en fonction de l’expertise du participant avec une technique d’interaction et la charge mentale lors du test. Une équation de régression significative a été trouvée ($F(5, 41) = 4.01, p < 0.005$) avec un R^2 de .20. On constate qu’un participant avec un profil de charge mentale faible réalise 18 actions de moins qu’un participant avec un profil de charge mentale fort lors de la tâche de la table du VMT. De plus, un expert avec la technique d’interaction réalise 23 actions de plus qu’un novice (Fig. 7.18).

Une régression multiple a été calculée pour prédire la distance parcourue dans le VMT en fonction de l’expertise du participant avec une technique d’interaction et la charge mentale lors du test. Une équation de régression significative a été trouvée ($F(5, 47) = 5.65, p < 0.0005$) avec un R^2 de .37. On constate qu’un participant avec un profil de charge mentale faible parcourt 176 mètres de moins qu’un participant avec un profil de charge mentale fort lors de la tâche de la table du VMT. De plus, un expert avec la technique d’interaction parcourt 105 mètres de plus qu’un novice (Fig. 7.19).

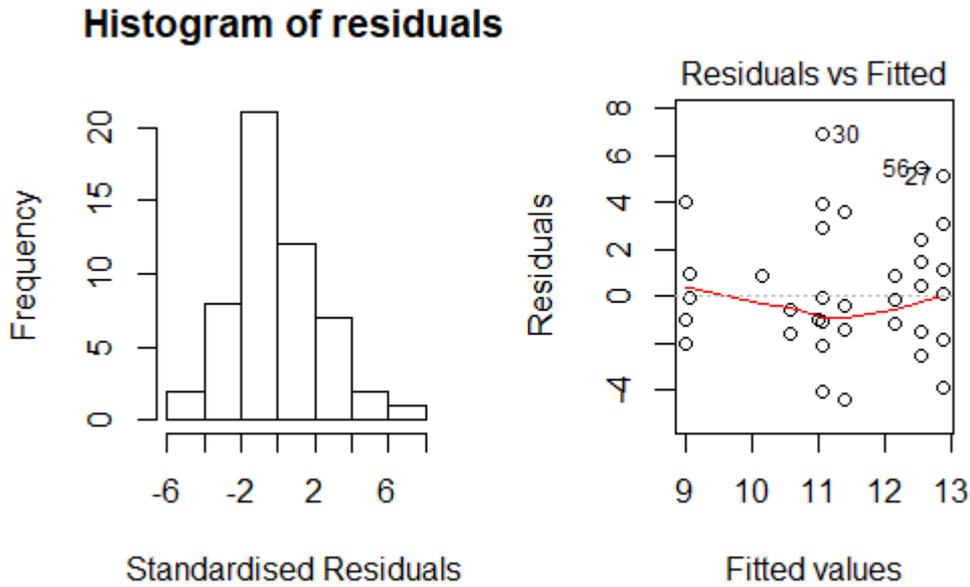


FIGURE 7.17 – Graphiques de validation du modèle de régression.

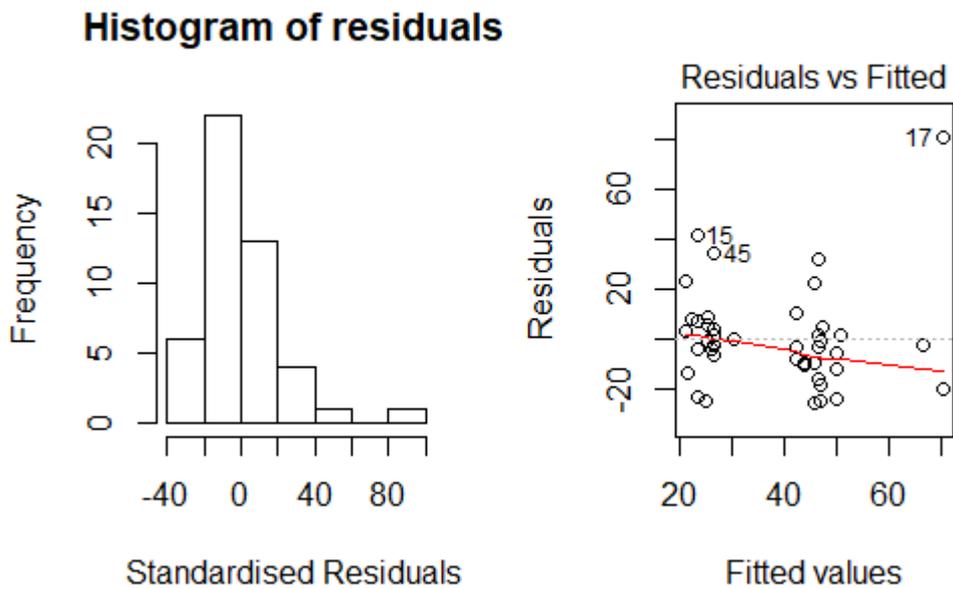


FIGURE 7.18 – Graphiques de validation du modèle de régression.

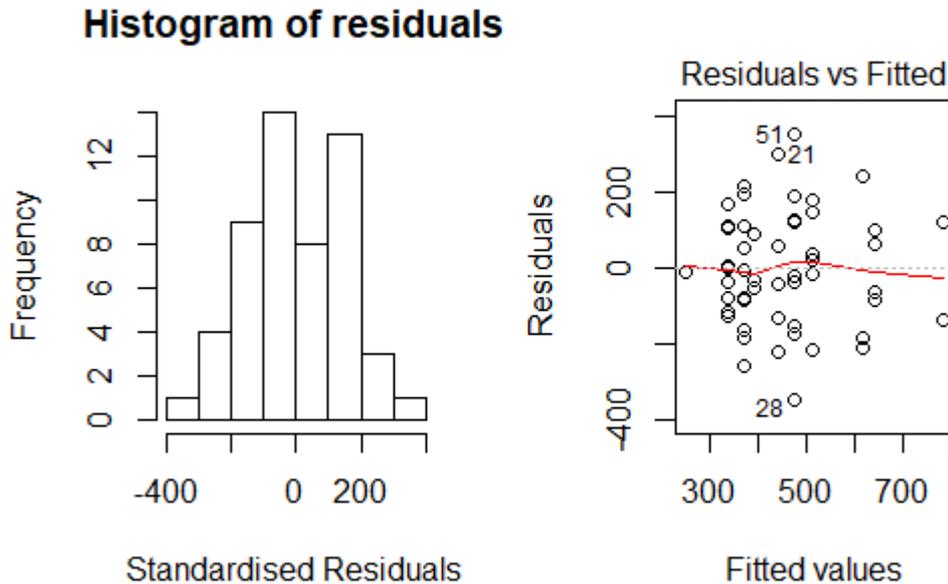


FIGURE 7.19 – Graphiques de validation du modèle de régression.

Une régression multiple a été calculée pour prédire le temps de réalisation du VMT en fonction de l'expertise du participant avec une technique d'interaction et la charge mentale lors du test. Une équation de régression significative a été trouvée ($F(5,47) = 18.06, p < 0.01$) avec un R^2 de .657. On constate qu'un participant avec un profil de charge mentale faible met 12 minute de moins qu'un participant avec un profil de charge mentale fort pour terminer le VMT (Fig. 7.20).

7.6 Analyse des données cardiaques

Nous ne trouvons pas de différence significative pour les données physiologiques entre les conditions ($U = 172; p = 0.08$) et les techniques d'interaction ($H(4) = 1.68; p = 0.7$). Nous ne trouvons pas de différence entre la baseline et le test pour le HRV ($U = 1080; p = 0.6$) et le rythme cardiaque moyen ($U = 1117; p = 0.9$). Il est donc difficile de supposer que les données physiologiques soient un indicateur de la charge mentale.

Si nous nous intéressons à la charge mentale évaluée par les mesures physiologiques, nous ne trouvons pas de différences significatives.

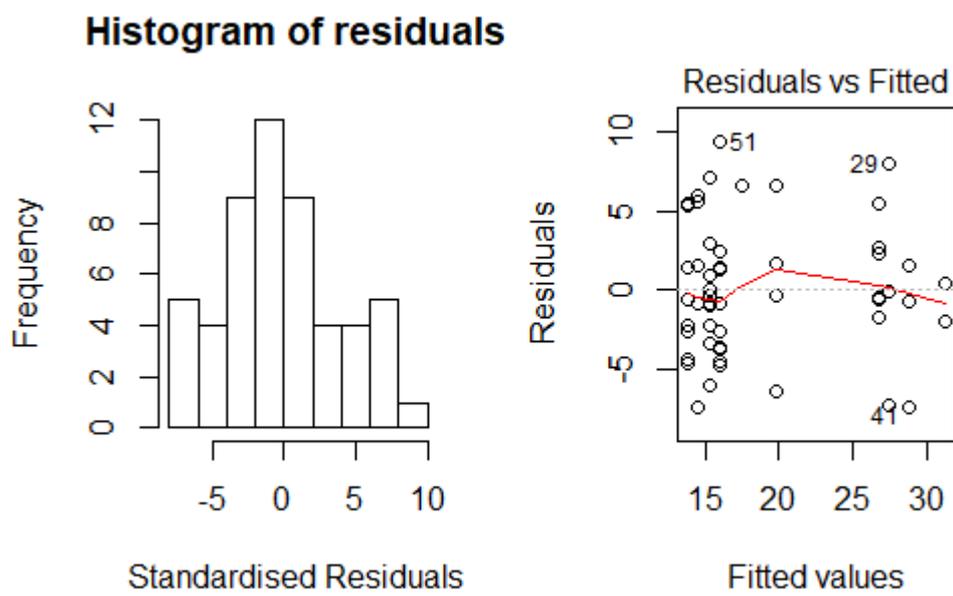


FIGURE 7.20 – Graphiques de validation du modèle de régression.

Discussion

8.1 Retour sur les résultats

8.1.1 Modulation de la charge mentale en fonction de l'environnement virtuel

La première hypothèse stipulait que la charge mentale pouvait être modulée par la technique d'interaction et les caractéristiques immersives de l'application. Elle supposait en particulier, que les conditions immersives, les techniques avec une sélection en 3D et celles avec un *steering* effectué par les mouvements de la tête, engendreraient plus de charge mentale.

Il y a trois façons de mesurer la charge mentale : les questionnaires, les données physiologiques et la performance obtenue.

Lorsque la charge mentale est mesurée par les signaux physiologiques, il n'y a pas de différence significative entre les techniques d'interaction. Lorsqu'elle est mesurée par le NASA-TLX, il n'y a pas de différence de charge mentale entre les différentes techniques d'interaction. Il y a toutefois une différence significative lors de la condition de charge mentale faible. Dans cette dernière condition, l'utilisation de la souris demande moins d'implication physique que le Razer Hydra et le gamepad en demande moins que le Razer Hydra et que le RazerHT. Il demande aussi moins d'effort au participant que le RazerHMD. Quand les techniques d'interaction sont regroupées selon leurs caractéristiques, nous constatons que les techniques d'interaction avec une sélection en 3D génèrent plus de charge mentale que les techniques avec une sélection en 2D, elles demandent en particulier plus d'implication physique. Les techniques avec une rotation de la caméra via les mouvements de la tête impliquent aussi davantage de charge mentale que les techniques qui autorisent les rotations de la caméra via les mouvements de la main.

Lorsque la charge mentale est mesurée par la performance au VMT, les utilisateurs avec le gamepad mettent moins de temps pour compléter la tâche et ont un meilleur score que ceux avec le Razer Hydra, le RazerHT et le RazerHMD. Le score avec le gamepad est surtout plus élevé à la tâche du poulet, tandis que le score est moins élevé avec le RazerHT à la tâche de l'épicerie. Quand les techniques d'interaction sont regroupées selon leurs caractéristiques, nous constatons que les techniques d'interaction en 2D ont un meilleur score que les techniques en 3D, en particulier, pour la tâche du poulet et de la table. Les techniques d'interaction avec un *steering* via les mouvements de la main ont aussi un meilleur score et une performance plus élevée à la tâche de la table que les techniques d'interaction qui utilisent les mouvements de la tête. Au niveau du temps de complétion, les utilisateurs des techniques avec une sélection en 3D mettent plus de temps pour compléter la tâche que ceux utilisant des techniques avec une sélection en 2D. Quant à la distance parcourue dans le VMT, les techniques avec une sélection en 3D et avec un *steering* via les mouvements de la tête entraînent une distance supérieure que les techniques 2D et celles avec une rotation par les mouvements de la main. De plus, les participants parcourent plus de distance en non-immersif qu'en immersif.

Les résultats présentés sont partiellement cohérents avec ceux de la littérature. En effet, les résultats de Besançon et al. (2016) indiquent que la souris consomme moins de ressources physiques qu'une technique d'interaction tangible mais plus de ressources mentales. Nous avons montré que l'utilisation de la souris demande moins d'implication physique que le Razer Hydra, et, de façon générale, une technique avec une sélection en 2D en consomme moins qu'une technique avec une sélection en 3D. En effet, les utilisateurs ont besoin de plus de temps pour sélectionner des cibles supérieures, en haut de l'écran, que des cibles inférieures avec une sélection en 3D (Murata & Iwase, 2001). Ainsi, l'utilisation de la sélection en 3D demande plus d'implication physique car les utilisateurs doivent lever ou baisser le bras qui tient la technique d'interaction selon les positions de la cible et cela se traduit par un temps de complétion plus long et avec un score important à la sous-échelle demande physique du NASA-TLX. Cependant, contrairement à Besançon et al. (2016), les résultats indiquent qu'une sélection en 2D consomme moins de ressources mentales qu'une sélection en 3D. Cette différence dans les résultats peut être due au fait que les méthodes d'interaction utilisées avec la souris dans l'étude de Besançon et al. (2016) ne sont pas usuelles. Ainsi, le participant doit désapprendre ses habitudes pour en acquérir de nouvelles avec le système proposé. En effet, dans l'étude de Besançon et al. (2016), les rotations et les translations des items étaient effectuées lorsque l'utilisateur appuyait respectivement sur le bouton droit et gauche de la souris. Alors qu'avec l'usage courant, l'utilisateur

va déplacer un item en bougeant physiquement la souris de droite à gauche sur la surface où elle est positionnée.

Contrairement à Speicher et al. (2017), nous ne trouvons pas de différence de charge mentale subjective entre immersif et non-immersif. La condition immersive ne semble pas ajouter de charge mentale par rapport à une condition non-immersive, bien que l'on trouve que les participants parcourent plus de distance en non-immersif. Nous pouvons supposer que la différence entre immersif et non-immersif est générée par les cybermalaises. En effet, on trouve plus de cybermalaise avec le RazerHMD qu'avec le RazerHT. Aussi pour éviter de ressentir ces symptômes les participants ont pu faire appel à des stratégies d'évitement pour limiter la locomotion dans le VMT.

Les résultats indiquent que les différentes techniques d'interaction semblent induire le même degré de charge mentale, du moins dans la condition où la tâche génère une charge mentale élevée. Lorsque les techniques sont regroupées selon leurs caractéristiques, quelques différences peuvent être mises en avant. En effet, il y a plus de charge mentale avec une technique en 3D qu'en 2D et cela se traduit de façon subjective avec une hausse du score au NASA-TLX et objective avec une baisse du score au VMT. Il en est de même pour les techniques qui impliquent les mouvements de la tête, elles requièrent plus de charge mentale que les techniques qui effectuent une rotation via les mouvements de la main. Le regroupement des techniques d'interaction selon leurs caractéristiques aident à mieux appréhender la charge mentale liée aux techniques d'interaction. En effet, il est possible d'implémenter plusieurs méthodes d'interaction avec un même outil. Décrire les *inputs* associés à une à une même technique permet d'être plus précis lors de la comparaison des techniques d'interaction.

8.1.2 Modulation de la charge mentale selon le profil de l'utilisateur

La seconde hypothèse stipulait que selon le degré d'expertise du participant avec une technique d'interaction, la charge mentale générée serait plus ou moins importante. Ainsi, un utilisateur expert avec l'utilisation d'une technique aurait moins de charge mentale qu'un participant novice.

On ne trouve pas de différence significative entre experts et novices avec l'utilisation d'une technique d'interaction au niveau du NASA-TLX et des données physiologiques. La seule va-

riable, liée à la performance, qui comprend une différence entre les deux groupes est le nombre d'actions nécessaires pour réaliser la tâche du poulet et la tâche de mettre la table.

On ne peut pas en déduire que les experts ont un degré de charge mentale moins élevé que les novices. Cependant, nous trouvons des différences dans l'utilisation de la technique en fonction du degré de familiarité du participant avec celle-ci. Ainsi, les participants experts utilisent plus le *steering* et la locomotion en même temps avec le Razer Hydra qu'avec la souris, tandis que les novices réalisent plus d'orientations de la caméra avec le gamepad qu'avec le Razer Hydra et la souris. Ces résultats suggèrent qu'un participant expert sera plus à même d'utiliser les deux *inputs* (locomotion et orientation de la caméra) en même temps qu'un novice. Le participant novice aura tendance à effectuer plus de mouvements de la caméra. De plus, si l'on regarde le score de présence à l'échelle « possibilité d'agir », nous constatons que les experts ont un score plus élevé que les novices. Ces derniers ont donc plus de difficultés pour interagir dans l'EV que les experts.

Le degré de charge mentale ressentie ne semble pas être influencé par le profil du participant. Ainsi, nous ne trouvons pas de différence au NASA-TLX en fonction du degré d'expertise du participant avec la technique d'interaction et de son sexe. Cependant, nous observons que les participants experts explorent plus l'EV car ils se sentent plus présents dedans que les participants novices. A ce stade, nous pouvons donc suggérer que peu importe le degré d'expertise de l'individu avec la technique, celui-ci peut interagir dans l'EV car au fur et à mesure, les différences de niveau entre les groupes s'estompent. Ainsi, les EVs peuvent être utilisés avec tout le monde.

8.1.3 Modulation de la charge mentale en fonction de la difficulté de la tâche

La troisième hypothèse stipulait que la réalisation d'une tâche difficile impliquait plus de charge mentale qu'une tâche facile.

Lorsque l'on regarde les différentes mesures de la charge mentale (performance, signaux physiologiques et questionnaires) nous pouvons constater que dans notre cas nous ne trouvons pas de différence de charge mentale entre la condition de charge mentale faible et de charge mentale élevée. Cependant, des différences existent au niveau de l'utilisation de la technique

d'interaction entre les deux conditions. En effet, dans la condition de charge mentale élevée on trouve des différences au niveau du score de présence : les techniques d'interaction avec une sélection en 2D ont un score de présence plus élevé (sous-échelles possibilité d'agir, qualité de l'interface et total) que les techniques avec une sélection en 3D. Tandis que lors de la condition de charge mentale faible, on trouve des différences au niveau des techniques d'interaction pour le score au NASA-TLX et les sous-échelles demande physique et effort. Les techniques avec une sélection en 2D ont un score moins élevé que les techniques avec une sélection en 3D, mais seulement dans la condition de charge mentale faible.

Les résultats précédents suggèrent que selon la difficulté de la tâche, la performance du participant va s'exprimer de différentes façons. Dans le cas où la difficulté est faible, l'utilisateur va être plus à même de renseigner la charge mentale ressentie. Ainsi, ils indiquent plus de charge mentale avec l'utilisation du Razer Hydra qu'avec la souris ou le gamepad. La charge mentale est d'ailleurs ressentie principalement par des aspects moteurs. En effet, les sous-échelles demande physique et effort présentent des différences significatives entre les techniques d'interaction mais pas la sous-échelle demande mentale.

Il est admis, par la littérature, que la charge mentale augmente avec la difficulté de la tâche (Chanquoy et al., 2007 ; Jaquess et al., 2017) mais que toute hausse de la charge mentale n'indique pas forcément une baisse de la performance (Ayaz et al., 2012 ; Liang et al., 2017). Dans le cas de notre étude, nous ne trouvons pas de différence de charge mentale subjective ou de performance entre les deux conditions d'expérimentation. Nous pouvons supposer que nous ne trouvons pas de différence de résultats entre les deux conditions car les participants n'étaient pas suffisamment impliqués dans la réalisation de la tâche simple.

8.1.4 Impact de la charge mentale sur la performance

La quatrième hypothèse stipulait que la hausse de la charge mentale aurait un impact négatif sur la performance du participant.

Lorsque l'on compare la performance des participants, toutes techniques d'interaction confondues, entre la condition de charge mentale faible et la condition de charge mentale élevée, nous ne trouvons pas de différence significative. Cependant, les participants *ChargeMentale_{High}* ont une performance moins élevée que les participants *ChargeMentale_{Low}*. Ce résultat est sans surprise puisque les variables temps de complétion et score au VMT ont été utilisées pour construire

le profil de charge mentale, il est donc logique que les participants *ChargeMentale_{High}* aient un temps de complétion plus long et un score moins élevé. Cependant d'autres variables permettent de faire la distinction entre les deux groupes. En effet, les modèles de régression indiquent qu'un participant *ChargeMentale_{High}* a un score moins élevé à la tâche du poulet, effectue la tâche de la table en plus d'actions, parcourt plus de distance. De plus, il effectue plus d'erreurs de clic dans l'EV.

La hausse de la charge mentale vient donc impacter négativement la performance du participant et en particulier pour la tâche du poulet qui fait appel à des fonctions cognitives variées. En effet, la tâche du poulet est la tâche de MP imbriquée dans la tâche concurrente. Sa mise en œuvre demande au sujet de retenir les ingrédients d'une recette de cuisine tout en contrôlant le temps entre les différentes sous-tâches. Ces résultats sont cohérents avec ceux obtenus dans la littérature qui indiquent que lorsque la charge mentale est élevée, la performance à la tâche en est impactée négativement (Yousef Zade et al., 2016). Les résultats obtenus sont enchevêtrés les uns avec les autres et nous admettons qu'il y a des évolutions de la charge mentale au cours des conditions et des techniques d'interaction, mais que celles-ci semblent faibles. Il n'est pas aisé de comprendre à quoi sont dues ces modulations de la charge mentale et nous pouvons nous demander si la charge mentale ressentie par le participant est causée par la tâche, la technique d'interaction ou bien par des caractéristiques intrinsèques de l'utilisateur.

8.2 Discussion générale : à quoi est due la charge mentale ?

Au cours de ce travail, 3 mesures de la charge mentale ont été utilisées : les signaux physiologiques, un questionnaire et la différence de performance au cours de deux conditions : une devant générer une faible charge mentale et une devant générer une forte charge mentale. Contrairement à la littérature, nous n'avons pas trouvé que le rythme cardiaque était plus élevé en fonction des conditions de la tâche et donc qu'il augmentait avec la hausse de la charge mentale (Jimenez-Molina & Lira, 2017). Nous n'avons pas non plus trouvé de différence entre la mesure au repos et la mesure lors de la tâche. Ainsi les résultats ne permettent pas d'indiquer que les données ECG sont sensibles aux modulations de la charge mentale. Les différences de résultats peuvent être dues au système d'enregistrement des mesures, au système d'analyse des

données ECG. Une autre hypothèse peut être proposée : les signaux physiologiques sont dépendants de plusieurs facteurs. Ainsi, la différence dans les signaux, si différence il y a, est difficile à mettre en évidence. De plus, les signaux physiologiques varient avec l'état émotionnel de l'individu (Chanquoy et al., 2007), mais sont aussi variables d'une personne à l'autre (Miyake, 2001).

La différence entre les techniques d'interaction et les conditions n'est pas très visible lorsque la charge mentale est mesurée par le NASA-TLX. Elle est plus visible quand elle est mesurée par la performance. Cependant, il est toujours intéressant de pouvoir comparer le ressenti subjectif et la performance objective du participant. Les résultats ont ainsi pu mettre en avant d'une part que la principale différence entre les techniques d'interaction était l'implication physique qu'elles demandaient et d'autre part que les différences entre techniques d'interaction sont plus visibles quand elles sont regroupées par leurs caractéristiques.

Quant à la différence de performance entre les deux conditions, les résultats ne permettent pas d'indiquer une charge mentale entre les deux conditions. Il semble donc nécessaire de proposer diverses mesures de la charge mentale pour avoir une idée plus précise de la façon dont celle-ci vient moduler l'expérience du sujet dans un EV. En effet, les résultats nous indiquent que les participants *ChargeMentale_{High}* ont une performance moins bonne que les utilisateurs *ChargeMentale_{Low}*. Il est donc important de pouvoir comprendre ce qui vient moduler la charge mentale chez le sujet, puisque celle-ci peut modifier négativement sa performance. Nos hypothèses en lien avec la littérature stipulaient que la charge mentale pouvait être modulée par le profil de l'utilisateur, la difficulté de la tâche et l'environnement soit la technique d'interaction dans le cas de l'expérimentation.

La distinction entre experts et novices avec l'utilisation d'une technique d'interaction n'a pas de lien avec la charge mentale, ni avec le score à la tâche, dans le cas de l'évaluation d'IADL. Cependant, nous pouvons supposer que dans le cas d'une application différente, cette variable peut avoir son importance, en particulier pour les EVs qui mesurent le temps de réaction des utilisateurs. Par exemple, le jeu sérieux *whack-a-mole* a été utilisé comme mesure de l'inhibition. Au cours de l'application, les sujets doivent sélectionner, le plus rapidement possible, les taupes cibles qui sortent de leur trou et inhiber les stimuli non-valides. Cette tâche est corrélée avec la tâche du Stroop chez les jeunes adultes sains (Tong & Chignell, 2014) et avec des mesures des fonctions exécutives, comme le MMSE, chez les utilisateurs adultes âgés (Tong et al., 2016). Dans le cas de ce type de tâche, qui fait intervenir de la vitesse de traitement, le degré d'expertise avec l'utilisation de la technique d'interaction pourrait avoir son importance car une personne

qui utilise mieux la technique d'interaction pourrait effectuer la tâche en moins de temps qu'un novice. De plus, la distinction entre novice et expert reste un critère intéressant à prendre en compte dans les études, car elle pourrait permettre de voir le temps d'apprentissage dans un EV en comparant le score de tous les utilisateurs puis de ceux qui sont novices et experts. En effet, les personnes âgées et, spécialement celles avec un MCI, peuvent se sentir anxieuses à l'idée d'utiliser une technique d'interaction (Wild et al., 2012). Il est donc nécessaire de leur proposer un temps d'apprentissage plus long avec une technique d'interaction afin de diminuer leurs craintes face à la nouvelle situation.

Ainsi, si le degré d'expertise avec l'utilisation de la technique d'interaction n'apparaît pas avoir de lien avec la charge mentale ressentie par le participant, cette donnée semble être importante dans le cas où le temps de réaction du participant est pris en compte et pour adapter la durée d'une phase de familiarisation des utilisateurs.

Les résultats de l'étude ont montré que la technique d'interaction n'engendrait pas forcément de charge mentale chez le participant et que celle-ci s'exprimait plutôt dans la condition de charge mentale faible que dans la condition de charge mentale élevée. Dans le cas où la difficulté de la tâche est élevée, les utilisateurs vont plutôt renseigner des facilités d'interaction avec certaines techniques d'interaction. Ainsi, il y a un score de présence et un SUS plus élevé avec les techniques d'interaction avec une sélection en 2D. De plus, un plus grand nombre de cybermalaises a été relevé avec le RazerHMD, seulement dans les conditions où la difficulté de la tâche est élevée.

Ces résultats suggèrent qu'il y a bien une différence de ressenti entre les deux conditions de difficulté de la tâche, mais que celle-ci ne s'exprime pas de la même façon. Dans le cas où la tâche doit générer une faible charge mentale, les utilisateurs renseignent une hausse subjective de la charge mentale avec des techniques d'interaction qui ont une sélection en 3D. Tandis que, dans le cas où la tâche doit générer une charge mentale élevée, ils vont renseigner des difficultés dans l'utilisabilité de la technique d'interaction.

Ainsi, lorsque le participant réalise une tâche facile, il va être plus à même de porter son attention sur l'outil d'interaction et les aspects qui peuvent venir le gêner. Tandis que, lorsque le participant réalise une tâche difficile, il va porter son attention sur la tâche et non plus sur la technique d'interaction. Ainsi, nous pouvons supposer que la différence de performance que l'on peut retrouver dans les études mesurant les IADL n'est pas due à la technique d'interaction

mais à l'attention et à l'intérêt que le participant porte à la réalisation de la tâche. De façon simplifiée, la motivation permet d'expliquer l'engagement d'un comportement (Fenouillet, Kaplan, & Yennek, 2009). Bien qu'il existe de nombreuses théories de la motivation, de façon générale, elles divisent la motivation en deux composantes : la motivation extrinsèque lorsque les patients sont motivés par des facteurs extérieurs à la tâche comme la reconnaissance de ses pairs (Wood, Griffiths, Chappell, & Davies, 2004) et la motivation intrinsèque lorsque l'individu produit des comportements libres et maintenus pour satisfaire un besoin d'autonomie (Deci & Ryan, 2000). L'une des théories de la motivation que nous pouvons mettre en avant pour expliquer la performance dans un EV est la théorie du *flow* (Csikszentmihalyi, 1990, 1997). Le *flow* s'expérimente lorsque l'implication du sujet dans la réalisation d'une tâche est telle qu'il oublie ce qu'il y a autour pour ne se concentrer plus que sur l'activité. Suite à l'absorption dans la tâche, l'individu perd alors la notion du temps et de lui-même (Csikszentmihalyi, 1990). L'auteur résume 7 composants du *flow* en deux catégories : les conditions et les caractéristiques. Les conditions sont les prérequis du *flow* et nécessitent d'avoir compris la tâche à effectuer, de recevoir des feedbacks sur son activité, de pouvoir être concentré dessus et que le but de l'action soit atteignable tout en demandant du challenge. Les caractéristiques désignent ce qui se passe quand on expérimente le *flow* : le contrôle de l'action et l'altération de soi et du temps. Si la présence décrit une immersion dans l'EV, le *flow* lui renvoie plutôt à une immersion dans l'activité (Weibel, Wissmath, Habegger, Steiner, & Groner, 2008). Pour d'autres auteurs, *flow* et présence sont des concepts similaires (Draper & Blair, 1996 ; Draper, Kaber, & Usher, 1998). En effet, les deux concepts sont liés par des corrélations (Novak, Hoffman, & Yung, 2000 ; Weibel et al., 2008) et dans le cas de notre étude, les participants *ChargeMentale_{Low}* ont un sentiment de présence plus élevé (total des sous-échelles et sous-échelle performance). Ainsi, les participants qui ont ressenti le moins de charge mentale, au cours de la réalisation du VMT, semblent avoir été plus impliqués dans la tâche et plus motivés, ce qui explique qu'ils aient une performance meilleure. De plus, si la technique d'interaction induit une charge mentale, celle-ci est minimale, et sera principalement ressentie dans les cas où la tâche à réaliser est facile. Notre hypothèse est que l'utilisateur est peu motivé par la tâche.

Les résultats précédents indiquent que la charge mentale est un concept difficile à appréhender. En effet, plusieurs facteurs viennent l'influencer : le profil de l'utilisateur, la difficulté de la tâche et l'environnement (Chanquoy et al., 2007). Dans le cadre de nos travaux, la charge mentale n'est pas clairement modulée par la technique d'interaction, ni par la difficulté de la tâche, ni par le profil du participant. En effet, les groupes de participants *ChargeMentale_{Low}* et

ChargeMentale_{High} comprennent aussi bien des individus experts et novices avec la technique d'interaction, des utilisateurs qui ont participé à la condition de charge mentale élevée et faible et qui utilisent les différentes techniques d'interaction. La charge mentale ne semble donc pas être ressentie par tous de la même manière en fonction de l'engagement dans la tâche.

La charge mentale du participant semble plus être due à des caractéristiques intrinsèques qu'à l'utilisation de la technique d'interaction et qu'aux connaissances antérieures du participant sur l'utilisation des techniques ou à la difficulté de la tâche du moins dans le cas du *VMT*. La charge mentale et le sentiment de présence apparaissent être deux concepts entremêlés. En effet, un EV qui génère beaucoup de charge mentale est associé à un sentiment de présence faible et vice versa (George et al., 2018) et les deux variables sont corrélées entre elles (Draper & Blair, 1996 ; George et al., 2018 ; Ma & Kaber, 2006). Bien que nous ne trouvions pas de corrélation positive entre le sentiment de présence totale et le score au NASA-TLX, les résultats de George et al. (2018) sont concordants avec ceux de notre étude où les techniques qui ont une sélection en 2D génèrent moins de charge mentale et plus de présence mais aussi au niveau du profil du participant où les *ChargeMentale_{Low}* ont un sentiment de présence plus élevé que les *ChargeMentale_{High}*. Dans le cas où les utilisateurs ressentent une charge mentale élevée, ils n'ont pas suffisamment de ressources cognitives disponibles pour porter leur attention sur l'EV, et donc se sentir présents, la réalisation de la tâche et l'utilisation de la technique d'interaction consomment la majorité des ressources cognitives disponibles à ce moment donné. Alors que les utilisateurs qui ont ressenti une charge mentale faible ont un haut sentiment de présence. La réalisation de la tâche et la technique d'interaction ne consomment pas l'intégralité de leurs ressources cognitives et les utilisateurs peuvent porter attention à la qualité graphique de l'EV et au plaisir ludique d'interagir dedans. Dans le *VMT*, la charge mentale ressentie par l'utilisateur, semble donc être majoritairement due à la quantité de ressources attentionnelles disponible au moment de réaliser la tâche. Ainsi, la charge mentale semble être modulée en grande partie par le profil de l'individu, et la taille de sa réserve attentionnelle. Puis, de façon moins importante, la charge mentale est modulée par la technique d'interaction et la difficulté de la tâche.

Kahneman (1973), indique que les capacités cognitives sont limitées. Il y aurait une réserve de ressources qui peut être exploitée jusqu'à épuisement. Cela laisse supposer que plus la personne vient puiser dans ses ressources, plus sa charge mentale est élevée et plus elle risque de réaliser des erreurs et/ou de manquer des informations. Ainsi, la réussite à une tâche serait mobilisée par 2 aspects : le nombre de ressources demandées pour la réalisation de la tâche

(charge mentale) et la taille de la réserve (réserve attentionnelle) (Jaquess et al., 2017). Une grande réserve attentionnelle pourrait permettre à l'utilisateur de mobiliser peu de charge mentale et de compléter la tâche avec peu d'erreur et en peu de temps. Tandis que dans le cas où la réserve attentionnelle est déjà sollicitée, la charge mentale du participant augmenterait et il obtiendrait un score moins élevé et mettrait plus de temps pour compléter la tâche. Par exemple, lors d'une tâche de cuisine, des ressources attentionnelles réduites entraînent un temps de complétion plus long et plus d'erreurs d'omissions (Hendry et al., 2016). De plus, (Jaquess et al., 2017) ont montré, avec des signaux EEG, que la charge mentale était négativement corrélée aux ressources attentionnelles. Ainsi la charge mentale est corrélée avec les potentiels évoqués qui sont la réponse immédiate du cerveau, désignée par une modulation du potentiel électrique, à un stimulus. La taille de la réserve attentionnelle peut donc venir directement influencer la charge mentale ressentie par l'utilisateur.

8.3 Application à un contexte d'évaluation neuropsychologique

8.3.1 Ecologie de la tâche

Les EVs permettent d'évaluer le participant lors de situations issues de la vie réelle et sont suffisamment sensibles pour détecter les troubles cognitifs et permettre de différencier un groupe de sujets contrôles d'un groupe de sujets avec un ralentissement ou des troubles cognitifs (Negut, Matu, Sava, & David, 2016). Leur utilisation permettrait d'évaluer le participant dans des situations écologiques puisqu'ils peuvent simuler la réalité (Rizzo et al., 2004). Si certains auteurs actent la validité écologique des EVs par des corrélations entre les mesures obtenues dans le test virtuel et celles aux tests papier-crayons (Larøi et al., 2010 ; Okahashi et al., 2013 ; T. Parsons & McMahan, 2017), nous ne pouvons pas l'affirmer avec des convictions intègres car, comme nous l'avons discuté dans le chapitre 3, la performance entre réelle et virtuelle à une même tâche est dans certains cas similaire (McGeorge et al., 2001 ; Renison et al., 2012 ; Vallejo et al., 2017) et dans d'autres différente où la performance peut être meilleure en virtuelle (Greenwood et al., 2016) ou moins bonne (Allain et al., 2014 ; Besnard et al., 2016).

Pour rappel, la validité écologique d'une tâche comprend deux versants : la véridicité et la verisimilitude (Franzen & Wilhelm, 1996). Les EVs apparaissent être valides d'un point de vue

véridicité. En effet, la procédure effectuée et la performance obtenue par les participants pour réaliser une recette de cuisine est la même en réel et en virtuel (Vallejo et al., 2017). La validité écologique des EVs, sur un versant verisimilitude, suggère que la charge mentale générée par la tâche soit la même en réel et en virtuel. Les résultats de notre étude indiquent que, dans le cas d'une tâche simple, les techniques d'interaction pouvaient ajouter de la charge mentale et en particulier sur l'implication physique demandée.

Il est donc nécessaire de proposer une tâche suffisamment complexe aux participants pour permettre aux utilisateurs de s'engager dans l'activité. Les résultats de notre étude indiquent qu'il n'y a pas de différence de charge mentale entre la tâche facile et la tâche difficile chez les jeunes adultes. Ils semblent donc plus concentrés sur la tâche difficile et ne ressentent pas de charge mentale contrairement à la tâche facile. Cette recommandation est primordiale d'un point de vue motivationnel mais aussi lors des activités de rééducation puisque les tâches virtuelles doivent refléter la demande cognitive des mêmes tâches en réel. En effet, les IADL requièrent l'implication de nombreuses fonctions cognitives. Par exemple, la réussite à la tâche de la cuisine est liée à aux processus attentionnels, de mémoire et des fonctions exécutives (Hendry et al., 2016). Il s'agit donc de proposer des tâches dont la réalisation demande l'implication de processus cognitifs multiples et non pas uniques. Afin que la tâche ne soit pas trop simple et donc suffisamment sensible pour détecter les difficultés des individus, nous recommandons d'utiliser au moins deux tâches. En effet, Martínez-Pernía et al. (2017) ont demandé à des utilisateurs cérébrolésés d'effectuer une tasse de thé en virtuel, sur un écran tactile, et ceux-ci ont renseigné que la tâche était trop facile.

Dans le cadre de l'évaluation neuropsychologique, les outils de mesures vont être principalement destinés à des populations de personnes âgées et à des individus ayant une lésion cérébrale, qu'elle soit subitement acquise ou non. Ces populations performant moins bien dans les EVs car elles ont un ralentissement cognitif et/ou des déficits ciblés. Afin de permettre la mesure neuropsychologique la plus sensible et écologique, il est nécessaire d'utiliser une technique d'interaction facile à utiliser et qui ne génère pas de charge mentale. En effet, s'il ne semble pas y avoir ou peu de charge mentale générée par la technique d'interaction, des individus ont ressenti une charge mentale élevée lors du VMT. Les personnes âgées ou les patients risquent de ressentir d'autant plus cette charge mentale du fait de leurs difficultés. Il s'agit donc d'adapter les EVs et les techniques d'interaction afin de leur proposer une expérience la plus écologique possible.

8.3.2 Quelle technique d'interaction utiliser ?

La multitude des techniques d'interaction existante fait qu'il peut être difficile de savoir laquelle privilégier au détriment des autres et pour quelles raisons. Bien que plusieurs techniques d'interaction aient déjà été comparées entre elles au cours de diverses situations comme une tâche de sélection (Ramcharitar & Teather, 2017) ou lors d'une tâche spécifique dans un supermarché virtuel (Renner, Dankert, Schneider, Mattar, & Pfeiffer, 2010), les résultats sont difficiles à généraliser car ils dépendent des techniques d'interaction employées et de la tâche effectuée. La comparaison des techniques d'interaction peut être faite entre les différentes techniques ou bien en regroupant celles-ci selon leurs caractéristiques. La deuxième option permet de mieux comprendre comment elles fonctionnent et de faire ressortir plus d'informations.

Comparer les techniques d'interaction entre elles aide à comprendre les points forts et les faiblesses de chacune afin de proposer la technique la plus efficace selon les conditions. On observe que la souris est la technique d'interaction qui a la meilleure utilisabilité (score au SUS et au *throughput*).

Pour mieux comprendre la façon dont les techniques d'interaction peuvent influencer le participant, elles ont été regroupées selon leurs caractéristiques. On observe alors qu'une technique d'interaction avec une sélection en 3D entraîne plus de charge mentale, une moins bonne performance, un score plus élevé en cybermalaise, un score d'utilisabilité plus faible, et un sentiment de présence moins fort qu'une technique avec une sélection en 2D. Les variables qui sont significatives aussi bien en condition de charge mentale faible et élevée que dans l'ensemble des deux conditions sont le score au VMT et la sous-échelle performance du questionnaire de présence.

Une technique d'interaction avec des rotations effectuées via les mouvements de la tête entraîne une charge mentale plus élevée, une moins bonne performance, un score de présence et d'utilisabilité plus faible, et plus de symptômes de cybermalaise. Les variables qui sont significatives aussi bien en condition de charge mentale faible et élevée que dans l'ensemble des deux conditions sont le score au VMT et le score au SUS.

Avec une technique d'interaction présentée en immersif l'utilisateur a un meilleur score au VMT, parcourt moins de distance et ressent plus de cybermalaise qu'avec une technique équivalente mais non-immersive.

Les résultats obtenus sont cohérents avec ceux indiqués dans la littérature. En effet, la souris a le meilleur *throughput* (Natapov et al., 2009 ; Ramcharitar & Teather, 2017) et les techniques

avec le Razer Hydra ont de moins bon résultats que le gamepad et la souris (Farmani & Teather, 2017). De plus, la condition immersive est plus appréciée des participants et génère moins de charge mentale que la condition non-immersive (Speicher et al., 2017). Bien que nous n'ayons pas trouvé de différence au NASA-TLX entre la condition immersive et non-immersive, le score au VMT est tout de même plus élevé en immersif. Ainsi, la condition immersive semble engendrer une charge mentale plus réduite qu'une condition équivalente mais en non-immersive.

Les différences entre les techniques d'interaction sont donc principalement dues à la différence entre une sélection en 2D et en 3D. En effet, les résultats sont plus généralisables selon les conditions. Dans le cas de la sélection en 2D, l'utilisateur doit gérer moins de DDL que lors d'une sélection en 3D puisqu'il n'a pas à se soucier de l'axe z . La restriction des DDL lors de l'interaction dans l'EV favorise une meilleure performance, moins de charge mentale, une meilleure utilisabilité et un sentiment de présence plus fort. Le type de *steering* utilisé semble être moins générateur de différence entre les techniques d'interaction. Ainsi, il peut en être déduit que l'utilisation d'une tâche de sélection telle que celle proposée par la norme ISO 9241-9 est suffisante et ne demande pas de la coupler avec une tâche de navigation. En effet, c'est la façon dont l'utilisateur peut sélectionner les items dans l'EV qui génère le plus de différence entre les techniques d'interaction.

Au cours de expérimentation, les variables liées à l'utilisation de la technique d'interaction ont été enregistrées. Toutes conditions confondues, le RazerHT puis la souris sont les techniques avec lesquelles les utilisateurs utilisent le plus l'*input* de locomotion, ce sont d'ailleurs avec ces techniques que les utilisateurs parcourent le plus de distance. Pour le RazerHT cela semble être dû à une mauvaise utilisabilité de la technique d'interaction tandis que dans le second cas, pour la souris, cela semble plutôt refléter une bonne utilisation de la technique d'interaction avec un comportement d'exploration de l'EV. Les résultats sont sensiblement équivalents lors de la condition facile mais on ne trouve pas de différence au niveau de l'*input* de locomotion lors de la condition difficile. La technique d'interaction qui utilise le moins l'*input* de locomotion est le gamepad. Toutes conditions confondues, on observe que le RazerHT et le gamepad sont les techniques qui utilisent le plus le *steering* et qu'il a plus d'activation des deux *input* (locomotion et rotation de la caméra) en même temps avec le RazerHMD et le RazerHT.

Les données recueillies en fonction de l'activation des *input* de chaque technique d'interaction ne permettent pas d'établir des liens entre la façon dont les techniques sont utilisées et la performance du participant dans un EV. Cependant ces résultats peuvent être exploré, en particulier auprès de populations d'âge divers tels que les enfants et les personnes âgées car la

mobilisation des *input* de différentes natures peuvent rendre plus difficile l'interaction. En effet, chez les populations avec une bonne coordination entre les gestes, la mobilisation du *steering* et de la locomotion en même temps peut être plus difficile. Cela peut être majoritairement le cas lorsque les membres de différentes parties du corps sont impliquées. En effet, l'alternance entre différentes techniques d'interaction a été étudiée. Les sujets pouvaient naviguer en marche réelle et avec le WIP ou bien avec le WIP et le gamepad. Il est plus facile de changer entre deux *input* liés aux mêmes membres, dans le cas de l'étude WIP-marche réelle, que de passer d'un membre inférieur à un membre supérieur (WIP-gamepad) (Bhandari, Tregillus, & Folmer, 2017). Le RazerHT et le RazerHMD demande l'implication des membres supérieurs mais aussi de la tête. L'utilisation des deux peut être plus difficile pour un utilisateur que lorsqu'il n'a que ses mains à utiliser.

Au regard des résultats précédents et de la littérature, nous pouvons avancer les recommandations suivantes pour choisir une technique d'interaction liée à un EV.

Dans un premier temps, il s'agit de décrire l'EV dans son ensemble. Si c'est EV 2D, qui ne nécessite pas de navigation, l'écran tactile est préconisé. D'une part car cette technique est intuitive et les utilisateurs sont rapides avec (Forlines et al., 2007). Si c'est EV 3D, qui nécessite de la navigation, l'utilisation d'une technique d'interaction avec une sélection en 2D est recommandée. La souris est un bon exemple de technique d'interaction d'une part car elle est fréquemment utilisée dans la vie quotidienne et d'autre part car elle a un meilleur *throughput* que le gamepad et donc une meilleure utilisabilité (Natapov et al., 2009). Cependant dans nos résultats, nous pouvons constater que les utilisateurs du gamepad ont un meilleur score. En effet, on observe qu'ils mettent significativement moins de temps et ont un meilleur score à la tâche du poulet que les techniques avec le Razer Hydra alors qu'il n'y a pas de différence significative avec la souris. Bien que les utilisateurs de la souris aient un meilleur *throughput*, la supériorité de cette technique sur le gamepad n'est pas admise par nos résultats. Le choix entre ces deux techniques n'est pas facile à effectuer et d'autres données sont nécessaires pour comprendre comment les utilisateurs interagissent avec ces techniques pour ensuite proposer la plus pertinente des deux selon les contextes. En effet, nous pouvons supposer que dans les cas où l'utilisateur doit être rapide dans la réalisation de la tâche, la souris semble appropriée. Alors que dans le cas où l'utilisateur doit beaucoup naviguer et orienter la position de la caméra, le gamepad semble plus approprié. C'est en effet une technique qui utilise majoritairement du *steering*. Pour permettre ces distinctions entre techniques d'interaction, il est primordial que les *input* liés aux techniques soient tous renseignés.

Dans un second temps, il s'agit d'adapter la technique d'interaction en fonction du profil de l'individu. En effet, les résultats de notre étude ont montré que certains individus ressentaient beaucoup de charge mentale et d'autres peu. Les populations âgées et pathologiques peuvent aussi présenter ces différences interindividuelle. Par exemple, à une tâche de cuisine virtuelle, des participants cérébrolésés ont eut des difficultés pour réaliser la tâche et d'autres ont eut une performance de la même qualité que le groupe de sujets contrôles (Tanguay, Davidson, Guerrero Nuñez, & Ferland, 2014). Nous supposons que les utilisateurs vont ressentir beaucoup de charge mentale s'ils ont une réserve attentionnelle peu importante au moment de la tâche. Afin que la tâche puisse être écologique, les participants susceptibles de générer beaucoup de charge mentale doivent interagir avec une technique d'interaction qui génère peu de charge mentale tandis que les autres peuvent interagir avec une technique qui en demande un peu plus. Ainsi, les prochains EVs destinés à l'évaluation et à la rééducation cognitive vont devoir s'adapter aux différents utilisateurs. Par exemple, déceler avant le début du test virtuel le niveau d'expertise de l'utilisateur avec la technique d'interaction qui sera ensuite employée grâce à une tâche de sélection peut permettre ensuite d'adapter la durée de la phase de familiarisation en fonction de l'expertise de l'utilisateur. Ainsi une phase de familiarisation plus longue où des exercices de navigation et de sélection devrait être proposée à des utilisateurs non experts, en plus de la visite de l'EV et de l'apprentissage des interactions avec les items. Tandis que ceux avec un degré d'expertise plus élevé pourraient se contenter de la phase de familiarisation sans exercices additionnels.

8.3.3 Quelles variables privilégier ?

Lors de la création de l'EV il est aussi nécessaire de renseigner pleinement les variables prises en compte, la façon dont les données sont enregistrées et la technique d'interaction qu'utilise le participant et les *input* de celle-ci telle que les méthodes de sélection et de navigation. Cela permet de faciliter les comparaisons entre études qui utilisent des techniques similaires et de continuer à mieux comprendre comment la technique d'interaction peut venir moduler le comportement et la performance du participant.

Nombre d'actions par tâche

Les résultats indiquent que la tâche de l'épicerie est effectuée en plus d'actions avec les techniques d'interaction avec une sélection en 2D. La tâche de la table est aussi effectuée en

plus d'actions pour les participants *ChargeMentale_{High}* et les experts. La tâche du poulet est aussi effectuée en plus d'actions par les experts.

Pour expliquer ces résultats, une première hypothèse peut être avancée : les participants effectuent plus d'actions pour certaines tâches que d'autres car ils maîtrisent mieux la technique d'interaction et peuvent réaliser des actions superflues sans que cela vienne les pénaliser. Par exemple, lors de la réalisation de la tâche du poulet, certains participants vont rapprocher les ingrédients nécessaires à la recette de cuisine près du plat de cuisson, tandis que d'autres vont laisser les ingrédients sur le comptoir et ne s'en saisir que lorsque cela sera nécessaire. Ainsi, les participants experts réalisent la tâche du poulet et de la table en plus d'actions que les participants novices. Cependant, les participants *ChargeMentale_{High}* réalisent plus d'actions à la tâche de la table.

La réussite de cette tâche est principalement liée à des critères d'utilisabilité de la technique d'interaction et de compréhension de l'EV. En effet, la mise en place des couverts se faisait automatiquement si les participants les déposaient sur la zone attendue, c'est-à-dire avec les couteaux à droites et les fourchettes à gauche de l'assiette. Lorsqu'ils ne respectaient pas cette mise en place, qui correspond à des normes culturelles française, ils pouvaient avoir plus de difficultés à mettre les items correctement et devaient s'y reprendre à plusieurs fois.

Le nombre d'actions réalisées par tâche, dans le cas d'une situation de multitâche ou du moins du VMT, ne semble donc pas être une variable pertinente. En effet, elle semble refléter majoritairement des actions superflues, et sans conséquence sur la réussite de la tâche, des participants experts avec une technique d'interaction. Elle n'a, de plus, pas d'influence sur le score à la tâche. Cependant dans le cas d'une tâche isolée ou de peu de tâches, la variable nombre d'actions peut être un indice d'efficacité. Par exemple dans une tâche de cuisine où le participant doit faire une tasse de café, le nombre d'actions mis pour effectuer la tâche peut être un indice d'efficacité de l'utilisateur. Un utilisateur qui réalisera la tâche en plus d'actions que la moyenne aura sûrement effectué des erreurs dans la séquence d'action et/ou des actions redondantes.

Distance parcourue

Nir-Hadad et al. (2017) proposent que la distance parcourue dans un EV soit liée à une mesure d'efficacité. En effet, dans le cas du super marché virtuel une grande distance parcourue indique que le participant ait fait des allers-retours dans les rayons. Dans le cas du VMT, les

participants *ChargeMentale_{High}* parcourent plus de distance que ceux *ChargeMentale_{Low}*. Cela laisse supposer qu'un participant non-efficient parcourra plus de distance.

Si l'on s'intéresse à la distance parcourue en fonction des techniques d'interaction, nous ne trouvons des différences dans la distance parcourue que dans la condition de charge mentale faible. Au cours de cette condition, les utilisateurs du RazerHT parcourent significativement plus de distance que les utilisateurs avec une autre technique d'interaction. Nous pouvons supposer que les participants parcourent plus de distance avec le RazerHT car ils ont des difficultés pour interagir avec cette technique d'interaction. En effet, ils ont aussi le score le plus faible au SUS dans cette même condition de charge mentale faible. Lorsque nous comparons le RazerHT avec le RazerHMD, autrement dit lorsque nous comparons une situation immersive à une situation non-immersive, nous observons que les utilisateurs parcourent plus de distance. Ces résultats semblent être l'illustration des symptômes de cybermalaise. En effet, les rotations de la caméra avec le visiocasque entraînent, dans le cas du VMT, des cybermalaise. Les utilisateurs ont donc limité les moments où ils devaient tourner et ont préféré faire de la marche arrière.

La variable distance parcourue semble donc particulièrement liée à des critères d'utilisabilité de la technique d'interaction, dans le cas où l'utilisateur effectue des tâches dans un espace réduit. Du moins lorsque la difficulté de la tâche est faible et que l'ensemble des activités à réaliser se déroulent dans la même zone. La variable distance parcourue apparaît être surtout pertinente lorsque l'utilisateur doit suivre un trajet particulier ou doit évoluer dans une grande zone, comme dans le cas du supermarché virtuel.

Le choix des variables pour mesurer la performance des participants doit donc correspondre au type d'EV. Bien que les variables liées au temps de complétion et aux erreurs (ou au score) sont généralisées à toutes sortes d'EVs, les autres ne semblent pas pouvoir l'être et demandent à ce que les expérimentateurs et les concepteurs des tests soient vigilants quant à la signification des variables employées pour qualifier la performance du participant.

8.4 Limites

L'expérimentation que nous avons menée comporte plusieurs limites. D'une part le faible échantillon de participants et le fait que nous nous sommes restreint à des techniques d'interaction majoritairement utilisées en condition non-immersive. Les résultats obtenus ne permettent donc pas d'émettre des suppositions sur les techniques d'interaction créées pour interagir en RV

comme la téléportation. Ce qui est dommage car RV propose des applications appréciées par les utilisateurs (Speicher et al., 2017). En effet, ceux-ci sont curieux et motivés par les environnements de RV. Les applications de RV pourraient donc permettre d'évaluer des utilisateurs lorsqu'ils sont motivés, au moins par la découverte de la technologie.

D'autre part, certains utilisateurs n'ont pas compris le fonctionnement des tâches, en particulier celle des billets et celle de mettre la table. Dans le cas de la tâche des billets, certains n'ont pas trouvé le billet et sont revenus plusieurs fois entre la zone de la tâche et le téléphone pour voir si un évènement allait se déclencher. Il serait donc important d'ajouter un *feedback* aussi bien dans le cas où l'utilisateur a trouvé les billets que dans le cas où il ne les a pas trouvés. Par exemple, lorsque l'utilisateur veut prendre le téléphone, une voix peut remercier pour le billet. Avec la tâche de la table, l'orientation des items déposés sur la table est automatiquement prise en compte par le système si le bon objet est posé sur la bonne zone. Par exemple, si le couteau est déposé à droite de l'assiette, comme c'est culturellement l'habitude, le couteau vient se positionner correctement. Cependant, lorsque les utilisateurs inversent l'ordre attendu et mettent le couteau à gauche, la rotation de l'objet est moins bien prise en compte et les utilisateurs doivent souvent s'y reprendre à plusieurs fois. La conception d'un EV doit donc prendre en compte les habitudes et routines de chaque individu pour que celui-ci ne contienne pas de problème d'heuristique. La tâche de la table est donc un bon exemple de la difficulté de créer des EVs pour les IADL puisque, bien que nous voulions faciliter au maximum les interactions, les utilisateurs ont chacun leurs propres représentations de la façon dont il est souhaitable de mettre la table. De plus, il peut arriver que lorsque l'utilisateur n'arrive pas à faire quelque chose dans un EV, il suppose que cela soit un bug et il n'essaye donc pas de mettre en place une autre tactique.

Une autre limite d'importance à prendre en compte est que les conditions de difficulté utilisée lors de l'expérimentation n'ont pas fait l'objet d'une procédure de pré-test. Ainsi, la charge mentale et la difficulté associée à ces deux conditions n'est que théorique. Pour être sûr que les deux conditions avaient bien deux degrés de difficulté et de charge mentale différents, des participants auraient dû passer les deux conditions et rendre leurs avis subjectifs sur la difficulté de la tâche ainsi que sur la charge mentale générée. En effet, il est possible que les tâches à effectuer ne demandent pas suffisamment de charge mentale même lorsque des évènements inattendus apparaissent. Il est aussi difficile de prendre en compte la performance du participant lorsque les routines et habitudes de celui-ci sont directement concernées, comme dans le cas de tâches à réaliser dans un appartement virtuel. D'autres types d'EVs pourraient être envisagés pour mesurer la cognition d'un participant comme gérer des finances et prendre les transports

en commun. Ces tâches sont celles qui demandent le plus de ressources cognitives parmi les IADL (Gold, 2012).

8.5 Synthèse

L'étude visait à comprendre la charge mentale associée à la technique d'interaction et si celle-ci pouvait venir moduler la performance du participant. Pour répondre à cet objectif, cinq techniques d'interaction ont été comparées au cours de deux conditions : une facile générant une faible charge mentale et une difficile générant une forte charge mentale.

Trois axes d'analyses de la charge mentale ont été empruntés pour appréhender la modulation de la charge mentale en fonction de l'environnement, du profil du participant et de la difficulté de la tâche.

Quant à la modulation de la charge mentale en fonction de l'environnement, dans ce cas en fonction de la technique d'interaction, bien que la souris soit la technique d'interaction avec la meilleure utilisabilité, nous ne trouvons pas de réelle modulation de la charge mentale. Pour mieux comprendre l'implication de la charge mentale selon les techniques d'interaction, nous les avons regroupés selon leurs caractéristiques. Ainsi les techniques d'interaction ont été étudiées selon leur type de sélection (en 2D ou en 3D), leur type de *steering* (avec les mouvements de la tête ou de la main) et le type d'immersion (immersif ou non-immersif). La majeure différence se situe au niveau du type de sélection employée : les techniques d'interaction avec une sélection en 2D ont une meilleure performance que ceux avec une sélection en 3D. De plus, l'impact de la technique d'interaction ne va pas s'exprimer de la même façon en fonction des conditions. En effet, des difficultés d'interaction vont plutôt être renseignées dans le cas où la tâche est difficile alors que lorsque la tâche est facile, les participants seront plus à même de ressentir une demande physique plus importante dans le cas des techniques avec une sélection en 3D. Les techniques avec une sélection en 2D devraient donc être privilégiées dans le cas où les EV ne sont pas immersifs.

Pour étudier la charge mentale en fonction du profil du participant, nous avons commencé par proposer une méthode pour catégoriser le profil du participant entre expert et novice à partir *throughput*. Les résultats n'indiquent pas de lien entre le degré d'expertise d'une technique d'interaction et la charge mentale de l'utilisateur.

Lorsque la charge mentale est comparée en fonction de la difficulté de la tâche, nous pouvons constater que le participant va ressentir plus de charge mentale dans la condition facile que dans

la condition difficile. Nous supposons que lorsque le participant réalise une tâche facile, il va être plus à même de porter son attention sur l'outil d'interaction et les aspects qui peuvent venir le gêner. Tandis que, lorsque le participant réalise une tâche difficile, il va porter son attention sur la tâche et non plus sur la technique d'interaction.

Lorsque la charge mentale du participant est élevée, sa performance va baisser par rapport à un participant qui a une charge mentale faible. Cependant il est difficile de prédire à quoi est due la hausse de charge mentale du participant. Celle-ci semble être dûe à des caractéristiques intrinsèques comme la motivation envers la tâche.

Pour conclure, nous pouvons avancer que l'utilisation de la technique d'interaction ne semble donc pas engendrer de charge mentale spécifique mais que celle-ci est majoritairement engendrée par des caractéristiques internes à l'utilisateur. Il est toutefois préférable d'en utiliser une avec moins de DDL autorisé comme dans le cas de la sélection en 2D.

Conclusion et perspectives

9.1 Conclusions générales

L'usage des EVs et de la RV lors des tâches d'évaluation cognitive et des IADL est de plus en plus fréquent. Les systèmes d'interaction et d'affichage peuvent être variés et apporter chacun leurs avantages et inconvénients. La diversité des technologies demande aux expérimentateurs d'être conscients de la façon dont ils mesurent les fonctions cognitives dans l'EV et si cette mesure est suffisamment sensible, et n'est donc pas biaisée par l'usage d'une technique d'interaction. En effet, les outils d'interaction peuvent induire une charge mentale chez l'utilisateur et cette charge mentale additionnée à celle générée par la tâche peut venir moduler négativement la performance du participant. L'objectif des travaux de cette thèse visait à comprendre si l'usage d'une technique d'interaction impliquait de la charge mentale et si celle-ci peut venir impacter la performance de l'utilisation lors de la réalisation d'IADL dans un EV. Pour répondre à ce but, une synthèse de littérature sur les concepts d'interaction, de charge mentale et de performance dans un EV a été effectuée. Puis une expérimentation est mise en place pour répondre à l'objectif de la thèse.

L'EV développé pour les besoins de l'expérimentation est un appartement virtuel : le VMT. Pour répondre au but, 5 techniques d'interaction ont été comparées au cours de 2 conditions de difficulté d'un scénario de réalisation d'IADL : un facile et un difficile. Un large panel de techniques d'interaction peut être utilisé. Pour conduire l'expérimentation nous avons décidé de nous centrer dans un premier temps sur celles qui sont le plus utilisées lors des EVs destinés à l'évaluation neuropsychologiques, à savoir la souris et le clavier puis les joysticks/gamepad (Valladares-Rodríguez et al., 2016). Le Razer Hydra avec trois modalités de *steering* (via le joystick, via les mouvements de la tête reconnus par une kinect et via les mouvements de la tête reconnus les mouvements du casque de RV HTC Vive) a ensuite été inclus dans le choix des

techniques d'interaction afin de pouvoir comparer les différentes techniques deux à deux selon les axes suivants :

- Comparaison des techniques selon le type de sélection 2D-3D (gamepad et souris vs Razer Hydra, Razer Hydra et *tracking* de tête)
- Comparaison des techniques selon le type de *steering* main-tête (gamepad, souris et Razer Hydra vs RazerHT et RazerHMD)
- Comparaison des techniques selon le type d'environnement : immersif ou non-immersif (RazerHT vs RazerHMD)

Le scénario facile était composé de 3 tâches à réaliser tandis que, lors d'un scénario difficile, 5 tâches additionnelles sont proposées au participant par des coups de téléphone. Pendant la réalisation des tâches le rythme cardiaque de l'utilisateur est enregistré. Celui-ci termine l'expérimentation en remplissant des questionnaires à propos de de la charge mentale ressenti, du sentiment de présence, de l'utilisabilité de la technique d'interaction et des symptômes de cybermalaise. Le protocole expérimental permet donc de mesurer la charge mentale à l'aide de signaux physiologiques, d'un questionnaire et par la mesure de la performance entre une condition facile et une condition difficile.

L'expérimentation a permis de mettre en avant que la différence de charge mentale entre les différentes technique d'interaction est faible et que celle-ci se manifeste plus sensiblement lorsque les techniques sont regroupées par leurs caractéristiques. Ainsi les techniques avec une sélection en 3D et un *steering* effectué par les mouvements de la tête demande plus de charge mentale. Le degré d'expertise avec l'utilisation d'une technique d'interaction n'est pas lié à une modulation de la charge mentale. Cependant les utilisateurs qualifiés d'experts sont plus à même d'explorer le VMT que les novices. De plus, la charge mentale est plus perçue dans les conditions où la difficulté de la tâche est faible alors que dans le cas où la difficulté de la tâche est élevée, ce sont les problèmes d'utilisabilité des techniques d'interaction qui s'expriment. Les participants qui ont une charge mentale élevée ont un score plus faible que les participants qui ont une faible charge mentale, en particulier lors des tâches qui réclament des fonctions cognitives variées.

La charge mentale est donc un concept multifactoriel qui n'est pas aisé à isoler ni à mesurer. En effet, plusieurs types de mesures peuvent être réalisés pour avoir une idée plus précise du degré de charge mentale du participant. De plus, les mesures sont imbriquées dans la performance

du participant et il ne semble pas pertinent de suggérer qu'une bonne performance est due à une faible charge mentale dans le cas de la réalisation d'IADL. En effet, il est difficile d'isoler la raison de la hausse de la charge mentale chez les utilisateurs, est-elle dûe à des caractéristiques internes ? A la technique d'interaction ? Ou à la difficulté de la tâche ? Nous supposons que la charge mentale s'exprime plus ou moins faiblement chez les utilisateurs en fonction de facteurs internes et que la performance au IADL dans un EV est aussi modulée par des aspects de motivation.

9.2 Perspectives envisagées

Bien que la RV génère des symptômes de cybermalaise, ceux-ci ne touche pas tout le monde de la même façon. Par exemple, ils pourraient être liés de façon inconsciente aux sujets de sorte à ce qu'un sujet qui s'attend à expérimenter des symptômes néfastes, les verra plus facilement apparaître qu'une personne qui ne s'y attend pas. Cette hypothèse est étayée par les travaux de Almeida, Rebelo, Noriega, et Vilar (2017) qui ont montré que lorsque le groupe de sujets qui avaient lu avec attention le formulaire de consentement précisant le risque de cybermalaise liée à une tâche en RV avait plus de symptômes que le groupe qui avait porté peu d'attention au formulaire de consentement. On retrouve en particulier une différence significative pour les échelles fatigue et transpiration. De plus, les avancées technologiques laissent supposer que les symptômes vont être de moins en présents et permettent de pallier aux aspects négatifs d'interaction liés à l'usage des visiocasques antérieurs. En effet, auparavant les utilisateurs pouvaient être gênés car ils ne voyaient pas l'outils d'interaction avec le port du HMD. Maintenant les manettes fournies avec les nouveaux visiocasques contiennent des capteurs de position si bien qu'une réplique virtuelle de la manette peut être visible dans l'EV.

Ainsi, les applications de RV proposent une alternative intéressante aux EVs *desktop* car elles intéressent et sont plus agréables pour les participants (Speicher et al., 2017). De plus, les résultats que nous avons obtenus indiquent que la performance du participant est meilleure dans une condition immersive que dans une condition équivalente mais en non-immersive. Nos travaux devraient être poursuivis en se centrant sur la réalisation d'IADL en immersif avec différentes méthodes d'interaction afin de distinguer différents profils d'utilisateurs.

Les résultats présentés dans cette thèse sont issus de travaux réalisés chez le jeune adulte sain. Les résultats montrent qu'ils ne semblent pas avoir ressenti de charge mentale associée à la technique d'interaction, ou alors que celle-ci était minime. Ces travaux devraient être éten-

dus chez la personne âgée et auprès d'individus avec des troubles cognitifs, par exemple dus à une lésion cérébrale. Ces populations sont, en effet, le public cible des EVs évaluant les IADL. La performance des personnes âgées peut diminuer dès 50 ans au cours d'IADL (Sakai et al., 2018). L'utilisation de la technique d'interaction pourrait venir moduler leur performance d'une façon différente que chez les jeunes adultes. De plus, dans le but de répondre aux hypothèses soulevées sur la réserve attentionnelle lors de la discussion, il est nécessaire de mesurer la réserve attentionnelle des utilisateurs pour comprendre si les participants expérimentant beaucoup de charge mentale sont aussi ceux qui ont une plus faible quantité de ressources attentionnelles. Ainsi, les EVs pourraient être simplifié dans le cas où le participant a moins de ressources disponibles lors du test.

Toujours dans l'optique d'adapter les EVs en fonction du profil des utilisateurs et des tâches à réaliser, la récolte de données liées à l'utilisation des techniques d'interaction devraient être approfondies. En effet, en fonction des applications il est possible que les utilisateurs utilisent différemment les techniques d'interaction. Par exemple, plus de *steering* dans un type de tâche plutôt qu'un autre. Une meilleure compréhension de l'utilisation des techniques d'interaction en terme de temps d'utilisation des *inputs* pourrait permettre de préférer une technique pour une tâche particulière. Ainsi, une tâche qui ne demande pas ou peu d'exploration visuelle pourrait être réalisée plus efficacement avec une technique d'interaction où les *input* de sélection sont activés rapidement. A contrario, une tâche qui nécessite beaucoup d'exploration visuelle serait peut être mieux réalisée avec une technique d'interaction où l'orientation de la caméra est effectuée facilement comme avec le gamepad.

Références

- Allain, P., Foloppe, D. A., Besnard, J., Yamaguchi, T., Etcharry-Bouyx, F., Le Gall, D., ... Richard, P. (2014). Detecting everyday action deficits in alzheimer's disease using a nonimmersive virtual reality kitchen. *Journal of the International Neuropsychological Society*, 20(5), 468–477.
- Almanji, A., Davies, T. C., & Stott, N. S. (2014). Using cursor measures to investigate the effects of impairment severity on cursor control for youths with cerebral palsy. *International Journal of Human-Computer Studies*, 72(3), 349–357.
- Almanji, A., Payne, A. R., Amor, R., & Davies, T. C. (2015). A nonlinear model for mouse pointing task movement time analysis based on both system and human effects. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(6), 1003–1011.
- Almeida, A., Rebelo, F., Noriega, P., & Vilar, E. (2017). Virtual reality self induced cybersickness : An exploratory study. In *International conference on applied human factors and ergonomics* (pp. 26–33).
- Argelaguet, F., & Andujar, C. (2013). A survey of 3d object selection techniques for virtual environments. *Computers & Graphics*, 37(3), 121–136.
- Aubin, G., Béliveau, M.-F., & Klinger, E. (2015). An exploration of the ecological validity of the virtual action planning–supermarket (vap-s) with people with schizophrenia. *Neuropsychological rehabilitation*, 1–20.
- Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., & Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage*, 59(1), 36–47.
- Bailey, H. R., Kurby, C. A., Giovannetti, T., & Zacks, J. M. (2013). Action perception predicts action performance. *Neuropsychologia*, 51(11), 2294–2304.
- Balakrishnan, R. (2004). “beating” fitts’ law : virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies*, 61(6), 857–874.
- Baldwin, C. L., & Penaranda, B. (2012). Adaptive training using an artificial neural network

- and eeg metrics for within-and cross-task workload classification. *NeuroImage*, 59(1), 48–56.
- Balk, S. A., Bertola, M. A., & Inman, V. W. (2013). Simulator sickness questionnaire : twenty years later.
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual sus scores mean : Adding an adjective rating scale. *Journal of usability studies*, 4(3), 114–123.
- Banville, F., Couture, J.-F., Verhulst, E., Besnard, J., Richard, P., & Allain, P. (2017). Using virtual reality to assess the elderly : The impact of human-computer interfaces on cognition. In *International conference on human interface and the management of information* (pp. 113–123).
- Banville, F., Nolin, P., Lalonde, S., Henry, M., Dery, M.-P., & Villemure, R. (2010). Multitasking and prospective memory : can virtual reality be useful for diagnosis? *Behavioural neurology*, 23(4), 209–211.
- Barrouillet, P. (1996). Ressources, capacités cognitives et mémoire de travail : postulats, métaphores et modèles : La charge mentale. *Psychologie française*, 41(4), 319–338.
- Besançon, L., Issartel, P., Ammi, M., & Isenberg, T. (2016). Usability comparison of mouse, touch and tangible inputs for 3d data manipulation. *arXiv preprint arXiv :1603.08735*.
- Besnard, J., Richard, P., Banville, F., Nolin, P., Aubin, G., Le Gall, D., ... Allain, P. (2016). Virtual reality and neuropsychological assessment : The reliability of a virtual kitchen to assess daily-life activities in victims of traumatic brain injury. *Applied Neuropsychology : Adult*, 23(3), 223–235.
- Bhandari, J., Tregillus, S., & Folmer, E. (2017). Legomotion : scalable walking-based virtual locomotion. In *Proceedings of the 23rd acm symposium on virtual reality software and technology* (p. 18).
- Bier, N., & Macoir, J. (2010). How to make a spaghetti sauce with a dozen small things i cannot name : A review of the impact of semantic-memory deficits on everyday actions. *Journal of Clinical and Experimental Neuropsychology*, 32(2), 201–211.
- Bilder, R. M. (2011). Neuropsychology 3.0 : Evidence-based science and practice. *Journal of the international Neuropsychological Society*, 17(1), 7–13.
- Boletsis, C. (2017). The new era of virtual reality locomotion : A systematic literature review of techniques and a proposed typology. *Multimodal Technologies and Interaction*, 1(4), 24.
- Bowman, D. A., & Hodges, L. F. (1997). An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the*

1997 symposium on interactive 3d graphics (pp. 35–ff).

- Bowman, D. A., & Hodges, L. F. (1999). Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *Journal of Visual Languages & Computing*, *10*(1), 37–53.
- Bowman, D. A., Kruijff, E., LaViola Jr, J. J., & Poupyrev, I. (2001). An introduction to 3-d user interface design. *Presence : Teleoperators and virtual environments*, *10*(1), 96–108.
- Brooke, J., et al. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4–7.
- Burdea, G., & Coiffet, P. (1993). *La réalité virtuelle*. Paris : Hermès Science.
- Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the international neuropsychological society*, *4*(6), 547–558.
- Burgess, P. W., Alderman, N., Forbes, C., Costello, A., LAURE, M. C., Dawson, D. R., ... Channon, S. (2006). The case for the development and use of “ecologically valid” measures of executive function in experimental and clinical neuropsychology. *Journal of the international neuropsychological society*, *12*(2), 194–209.
- Burgess, P. W., Gonen-Yaacovi, G., & Volle, E. (2011). Functional neuroimaging studies of prospective memory : what have we learnt so far ? *Neuropsychologia*, *49*(8), 2246–2257.
- Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, *38*(6), 848–863.
- Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological psychology*, *42*(3), 249–268.
- Caywood, M. S., Roberts, D. M., Colombe, J. B., Greenwald, H. S., & Weiland, M. Z. (2017). Gaussian process regression for predictive but interpretable machine learning models : An example of predicting mental workload across tasks. *Frontiers in human neuroscience*, *10*, 647.
- Cegarra, J., & Chevalier, A. (2008). The use of tholos software for combining measures of mental workload : Toward theoretical and methodological improvements. *Behavior Research Methods*, *40*(4), 988–1000.
- Chanquoy, L., Tricot, A., & Sweller, J. (2007). *La charge cognitive : Théorie et applications*. Armand Colin.
- Chapman, P., Clinton, J., Khabaza, T., Reinartz, T., & Wirth, R. (1999). The crisp-dm process model. *The CRIP-DM Consortium*, 310.
- Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological

- tests : A review of the literature on everyday cognitive skills. *Neuropsychology review*, 13(4), 181–197.
- Chevignard, M., Taillefer, C., Picq, C., Poncet, F., Noulhiane, M., & Pradat-Diehl, P. (2008). Ecological assessment of the dysexecutive syndrome using execution of a cooking task. *Neuropsychological Rehabilitation*, 18(4), 461–485.
- Chittaro, L., & Ranon, R. (2007). Web3d technologies in learning, education and training : Motivations, issues, opportunities. *Computers & Education*, 49(1), 3–18.
- Chittaro, L., Ranon, R., & Ieronutti, L. (2006). Vu-flow : A visualization tool for analyzing navigation in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 12(6), 1475–1485.
- Cinaz, B., Arnrich, B., La Marca, R., & Tröster, G. (2013). Monitoring of mental workload levels during an everyday life office-work scenario. *Personal and ubiquitous computing*, 17(2), 229–239.
- Couclelis, H., Golledge, R. G., Gale, N., & Tobler, W. (1987). Exploring the anchor-point hypothesis of spatial cognition. *Journal of Environmental Psychology*, 7(2), 99–122.
- Courage, M. L., Bakhtiar, A., Fitzpatrick, C., Kenny, S., & Brandeau, K. (2015). Growing up multitasking : The costs and benefits for cognitive development. *Developmental Review*, 35, 5–41.
- Csikszentmihalyi, M. (1990). *The psychology of optimal experience new york*. Harper & Row.
- Csikszentmihalyi, M. (1997). *Finding flow : The psychology of engagement with everyday life*. Basic Books.
- Deci, E. L., & Ryan, R. M. (2000). The " what " and " why " of goal pursuits : Human needs and the self-determination of behavior. *Psychological inquiry*, 11(4), 227–268.
- Dehn, L. B., Kater, L., Piefke, M., Botsch, M., Driessen, M., & Beblo, T. (2018). Training in a comprehensive everyday-like virtual reality environment compared to computerized cognitive training for patients with depression. *Computers in Human Behavior*, 79, 40–52.
- Denisova, A., & Cairns, P. (2015). First person vs. third person perspective in digital games : do player preferences affect immersion ? In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 145–148).
- de Winter, J. C. (2014). Controversy in human factors constructs and the explosive use of the nasa-tlx : a measurement perspective. *Cognition, technology & work*, 16(3), 289–297.
- Draper, J. V., & Blair, L. M. (1996). Workload, flow, and telepresence during teleoperation. In *Robotics and automation, 1996. proceedings., 1996 ieee international conference on*

(Vol. 2, pp. 1030–1035).

- Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human factors*, *40*(3), 354–375.
- D’Souza, L., Pathirana, I., McMeel, D., & Amor, R. (2011). Kinect to architecture. *IVCNZ 2011*, 155–160.
- Dux, P. E., Tombu, M. N., Harrison, S., Rogers, B. P., Tong, F., & Marois, R. (2009). Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron*, *63*(1), 127–138.
- Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance : The processing efficiency theory. *Cognition & Emotion*, *6*(6), 409–434.
- Fallahi, M., Motamedzade, M., Heidarimoghadam, R., Soltanian, A. R., Farhadian, M., & Miyake, S. (2016). Analysis of the mental workload of city traffic control operators while monitoring traffic density : A field study. *International Journal of Industrial Ergonomics*, *54*, 170–177.
- Fallahi, M., Motamedzade, M., Heidarimoghadam, R., Soltanian, A. R., & Miyake, S. (2016). Assessment of operators’ mental workload using physiological and subjective measures in cement, city traffic and power plant control centers. *Health promotion perspectives*, *6*(2), 96.
- Farmani, Y., & Teather, R. J. (2017). Player performance with different input devices in virtual reality first-person shooter games. In *Proceedings of the 5th symposium on spatial user interaction* (pp. 165–165).
- Fenouillet, F., Kaplan, J., & Yennek, N. (2009). Serious games et motivation. In *4eme conference francophone sur les environnements informatiques pour l’apprentissage humain (eiah’09)*, vol. actes de l’atelier” jeux serieux : conception et usages.
- Ferracani, A., Pezzatini, D., Bianchini, J., Biscini, G., & Del Bimbo, A. (2016). Locomotion by natural gestures for immersive virtual environments. In *Proceedings of the 1st international workshop on multimedia alternate realities* (pp. 21–24).
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, *47*(6), 381.
- Forlines, C., Wigdor, D., Shen, C., & Balakrishnan, R. (2007). Direct-touch vs. mouse input for tabletop displays. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 647–656).
- Franzen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment.

- Frommel, J., Sonntag, S., & Weber, M. (2017). Effects of controller-based locomotion on player experience in a virtual reality exploration game. In *Proceedings of the 12th international conference on the foundations of digital games* (p. 30).
- Fuchs, P. (2016). *Les casques de réalité virtuelle et de jeux vidéo*. Presses des Mines-Transvalor.
- Fuchs, P., & Moreau, G. (2006). *Le traité de la réalité virtuelle* (Vol. 2). Presses des MINES.
- Gamito, P., Oliveira, J., Caires, C., Morais, D., Brito, R., Lopes, P., ... others (2015). Virtual kitchen test. *Methods of Information in Medicine*, 54(2), 122–126.
- George, C., Demmler, M., & Hussmann, H. (2018). Intelligent interruptions for ivr : Investigating the interplay between presence, workload and attention. In *Extended abstracts of the 2018 chi conference on human factors in computing systems* (p. LBW511).
- Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, 4(1-2), 113–131.
- Giovannetti, T., Bettcher, B. M., Brennan, L., Libon, D. J., Burke, M., Duey, K., ... Wambach, D. (2008). Characterization of everyday functioning in mild cognitive impairment : A direct assessment approach. *Dementia and geriatric cognitive disorders*, 25(4), 359–365.
- Giovannetti, T., Bettcher, B. M., Brennan, L., Libron, D. J., Kessler, R. K., & Duey, K. (2008). Coffee with jelly or unbuttered toast : Commissions and omissions are dissociable aspects of everyday action impairment in alzheimer's disease. *Neuropsychology*, 22(2), 235.
- Giovannetti, T., Britnell, P., Brennan, L., Siderowf, A., Grossman, M., Libon, D. J., ... Seidel, G. A. (2012). Everyday action impairment in parkinson's disease dementia. *Journal of the International Neuropsychological Society*, 18(5), 787–798.
- Gold, D. A. (2012). An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. *Journal of clinical and experimental neuropsychology*, 34(1), 11–34.
- Goldin, S. E., & Thorndyke, P. W. (1982). Simulating navigation for spatial knowledge acquisition. *Human factors*, 24(4), 457–471.
- Gonçalves, A., & Cameirão, M. (2016). Evaluating body tracking interaction in floor projection displays with an elderly population. In *Phycs* (pp. 24–32).
- Greenwood, K. E., Morris, R., Smith, V., Jones, A.-M., Pearman, D., & Wykes, T. (2016). Virtual shopping : A viable alternative to direct assessment of real life function? *Schizophrenia research*, 172(1), 206–210.
- Gregg, L., & Tarrier, N. (2007). Virtual reality in mental health. *Social psychiatry and psychiatric epidemiology*, 42(5), 343–354.
- Gronwall, D., & Wrightson, P. (1974). Delayed recovery of intellectual function after minor

- head injury. *The Lancet*, 304(7881), 605–609.
- Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index) : Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier.
- Hartmann, K., Goldenberg, G., Daumüller, M., & Hermsdörfer, J. (2005). It takes the whole brain to make a cup of coffee : the neuropsychology of naturalistic actions involving technical devices. *Neuropsychologia*, 43(4), 625–637.
- Hendry, K., Ownsworth, T., Beadle, E., Chevignard, M. P., Fleming, J., Griffin, J., & Shum, D. H. (2016). Cognitive deficits underlying error behavior on a naturalistic task after severe traumatic brain injury. *Frontiers in behavioral neuroscience*, 10, 190.
- Hintzman, D. L., O'Dell, C. S., & Arndt, D. R. (1981). Orientation in cognitive maps. *Cognitive Psychology*, 13(2), 149–206.
- Hourcade, J. P. (2006). Learning from preschool children's pointing sub-movements. In *Proceedings of the 2006 conference on interaction design and children* (pp. 65–72).
- Hourcade, J. P., Nguyen, C. M., Perry, K. B., & Denburg, N. L. (2010). Pointassist for older adults : analyzing sub-movement characteristics to aid in pointing tasks. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1115–1124).
- Hwang, Y., & Jeong, S.-H. (2018). Multitasking and task performance : Roles of task hierarchy, sensory interference, and behavioral response. *Computers in Human Behavior*, 81, 161–167.
- Inkpen, K. M. (2001). Drag-and-drop versus point-and-click mouse interaction styles for children. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 8(1), 1–33.
- Jankowski, J., & Hachet, M. (2013). A survey of interaction techniques for interactive 3d environments. In *Eurographics 2013-star*.
- Jaquess, K., Lo, L., Oh, H., Domingues, C., Neiman, S., Rietschel, J., ... Hatfield, B. (2017). Empirical evidence for the relationship between cognitive workload and attentional reserve. In *International journal of exercise science : Conference proceedings* (p. 51).
- Jasper, P., Sibley, C., & Coyne, J. (2016). Using heart rate variability to assess operator mental workload in a command and control simulation of multiple unmanned aerial vehicles. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 60, pp. 1125–1129).
- Jeong, S.-H., & Hwang, Y. (2016). Media multitasking effects on cognitive vs. attitudinal outcomes : A meta-analysis. *Human Communication Research*, 42(4), 599–618.
- Jimenez-Molina, A., & Lira, H. (2017). Towards a continuous assessment of cognitive workload

- for smartphone multitasking users. In *The first international symposium on human mental workload*. Dublin Institute of Technology.
- Josman, N., Klinger, E., & Kizony, R. (2008). Performance within the virtual action planning supermarket (vap-s) : an executive function profile of three different populations suffering from deficits in the central nervous system. *Maia, Portugal : Proc 7th ICDVRAT*.
- Jung, J., Park, H., Hwang, D., Son, M., Beck, D., Park, J., & Park, W. (2014). A review on interaction techniques in virtual environments. In *Proc. 2014 international conference on industrial engineering and operations management* (pp. 1582–1590).
- Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ.
- Kang, N. E., & Yoon, W. C. (2008). Age-and experience-related user behavior differences in the use of complicated electronic devices. *International Journal of Human-Computer Studies*, *66*(6), 425–437.
- Kang, Y. J., Ku, J., Han, K., Kim, S. I., Yu, T. W., Lee, J. H., & Park, C. I. (2008). Development and clinical trial of virtual reality-based cognitive assessment in people with stroke : preliminary study. *CyberPsychology & Behavior*, *11*(3), 329–339.
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire : An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, *3*(3), 203–220.
- Kessler, R. K., Giovannetti, T., & MacMullen, L. R. (2007a). Everyday action in schizophrenia : Performance patterns and underlying cognitive mechanisms. *Neuropsychology*, *21*(4), 439.
- Kessler, R. K., Giovannetti, T., & MacMullen, L. R. (2007b). Everyday action in schizophrenia : Performance patterns and underlying cognitive mechanisms. *Neuropsychology*, *21*(4), 439.
- Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education*, *67*, 135–142.
- Kliegel, M., & Martin, M. (2003). Prospective memory research : Why is it relevant ? *International Journal of Psychology*, *38*(4), 193–194.
- Klinger, E., Chemin, I., Lebreton, S., & Marié, R.-M. (2006). Virtual action planning in parkinson's disease : A control study. *Cyberpsychology & Behavior*, *9*(3), 342–347.
- Knaepen, K., Marusic, U., Crea, S., Guerrero, C. D. R., Vitiello, N., Pattyn, N., ... Meeusen, R. (2015). Psychophysiological response to cognitive workload during symmetrical, asymmetrical and dual-task walking. *Human movement science*, *40*, 248–263.
- Kramer, A. F., & Parasuraman, R. (2007). Neuroergonomics : Applications of neuroscience to

human factors.

- Kulik, A. (2009). Building on realism and magic for designing 3d interaction techniques. *IEEE Computer Graphics and Applications*, 29(6), 22–33.
- Kurtz, M. M., Baker, E., Pearlson, G. D., & Astur, R. S. (2006). A virtual reality apartment as a measure of medication management skills in patients with schizophrenia : a pilot study. *Schizophrenia bulletin*, 33(5), 1162–1170.
- Kvavilashvili, L., & Ellis, J. (1996). Varieties of intention : Some distinctions and classifications. *Prospective memory : Theory and applications*, 6, 183–207.
- Larøi, F., Canlaire, J., Mourad, H., & Van der Linden, M. (2010). Relations between a computerized shopping task and cognitive tests in a group of persons diagnosed with schizophrenia compared with healthy controls. *Journal of the International Neuropsychological Society*, 16(1), 180–189.
- Law, A. S., Logie, R. H., & Pearson, D. G. (2006). The impact of secondary tasks on multi-tasking in a virtual environment. *Acta psychologica*, 122(1), 27–44.
- Lee, J., Kim, B., Suh, B., & Koh, E. (2016). Exploring the front touch interface for virtual reality headsets. In *Proceedings of the 2016 chi conference extended abstracts on human factors in computing systems* (pp. 2585–2591).
- Levy, R. (2013). Psychometric and evidentiary advances, opportunities, and challenges for simulation-based assessment. *Educational Assessment*, 18(3), 182–207.
- Lewis, M. W., Babbage, D. R., & Leathem, J. M. (2011). Assessing executive performance during cognitive rehabilitation. *Neuropsychological rehabilitation*, 21(2), 145–163.
- Lezak, M. (1995). *Neuropsychological assessment (3rd ed.)*. New York : Oxford.
- Liang, J., Wilkinson, K., & Sainburg, R. L. (2017). Is hand selection modulated by cognitive-perceptual load? *Neuroscience*.
- Long, C., & Kibby, M. (1995). Ecological validity of neuropsychological tests : A look at neuropsychology's past and the impact that ecological issues may have on its future. *Advances in Medical Psychotherapy*, 8, 59–78.
- Longo, L., & Barrett, S. (2010). A computational analysis of cognitive effort. In *Asian conference on intelligent information and database systems* (pp. 65–74).
- Luximon, A., & Goonetilleke, R. S. (2001). Simplified subjective workload assessment technique. *Ergonomics*, 44(3), 229–243.
- Ma, R., & Kaber, D. B. (2006). Presence, workload and performance effects of synthetic environment design factors. *International Journal of Human-Computer Studies*, 64(6), 541–552.

- MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*, 7(1), 91–139.
- MacKenzie, I. S., Sellen, A., & Buxton, W. A. (1991). A comparison of input devices in element pointing and dragging tasks. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 161–166).
- Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., & Schwartz, P. J. (1996). Heart rate variability : Standards of measurement, physiological interpretation, and clinical use. *European heart journal*, 17(3), 354–381.
- Martel, E., & Muldner, K. (2017). Controlling vr games : control schemes and the player experience. *Entertainment computing*, 21, 19–31.
- Martínez-Pernía, D., Núñez-Huasaf, J., Del Blanco, Á., Ruiz-Tagle, A., Velásquez, J., Gomez, M., ... Slachevsky, A. (2017). Using game authoring platforms to develop screen-based simulated functional assessments in persons with executive dysfunction following traumatic brain injury. *Journal of biomedical informatics*, 74, 71–84.
- Martono, N. P., Yamaguchi, T., & Ohwada, H. (2016). Utilizing finger movement data to cluster patients with everyday action impairment. In *Cognitive informatics & cognitive computing (icci* cc), 2016 ieee 15th international conference on* (pp. 459–464).
- McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory : An overview and synthesis of an emerging field*. Sage Publications.
- McGeorge, P., Phillips, L. H., Crawford, J. R., Garden, S. E., Sala, S. D., Milne, A. B., ... Callender, J. S. (2001). Using virtual environments in the assessment of executive dysfunction. *Presence : Teleoperators & Virtual Environments*, 10(4), 375–383.
- Miyake, S. (2001). Multivariate workload evaluation combining physiological and subjective measures. *International journal of psychophysiology*, 40(3), 233–238.
- Monsell, S. (2003). Task switching. *Trends in cognitive sciences*, 7(3), 134–140.
- Moustafa, K., Luz, S., & Longo, L. (2017). Assessment of mental workload : a comparison of machine learning methods and subjective assessment techniques. In *International symposium on human mental workload : Models and applications* (pp. 30–50).
- Murata, A., & Iwase, H. (2001). Extending fitts' law to a three-dimensional pointing task. *Human movement science*, 20(6), 791–805.
- Muthukrishnan, S. P., Gurja, J. P., & Sharma, R. (2017). Does heart rate variability predict human cognitive performance at higher memory loads? *Indian J Physiol Pharmacol*, 61(1), 14–22.
- Nabiyouni, M., Saktheeswaran, A., Bowman, D. A., & Karanth, A. (2015). Comparing the

- performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality. In *3d user interfaces (3dvi), 2015 ieee symposium on* (pp. 3–10).
- Natapov, D., Castellucci, S. J., & MacKenzie, I. S. (2009). Iso 9241-9 evaluation of video game controllers. In *Proceedings of graphics interface 2009* (pp. 223–230).
- Negut, A., Matu, S.-A., Sava, F. A., & David, D. (2016). Virtual reality measures in neuropsychological assessment : a meta-analytic review. *The Clinical Neuropsychologist*, *30*(2), 165–184.
- Nir-Hadad, S. Y., Weiss, P. L., Waizman, A., Schwartz, N., & Kizony, R. (2017). A virtual shopping task for the assessment of executive functions : Validity for people with stroke. *Neuropsychological rehabilitation*, *27*(5), 808–833.
- Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1–18). Springer.
- Novak, T. P., Hoffman, D. L., & Yung, Y.-F. (2000). Measuring the customer experience in online environments : A structural modeling approach. *Marketing science*, *19*(1), 22–42.
- Okahashi, S., Seki, K., Nagano, A., Luo, Z., Kojima, M., & Futaki, T. (2013). A virtual shopping test for realistic assessment of cognitive function. *Journal of neuroengineering and rehabilitation*, *10*(1), 59.
- Olson, J. R., & Olson, G. M. (1990). The growth of cognitive modeling in human-computer interaction since goms. *Human-computer interaction*, *5*(2-3), 221–265.
- Ouellet, É., Boller, B., Corriveau-Lecavalier, N., Cloutier, S., & Belleville, S. (2018). The virtual shop : A new immersive virtual reality environment and scenario for the assessment of everyday memory. *Journal of neuroscience methods*.
- Oviatt, S. (2006). Human-centered design meets cognitive load theory : designing interfaces that help people think. In *Proceedings of the 14th acm international conference on multimedia* (pp. 871–880).
- Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational psychology review*, *6*(4), 351–371.
- Parasuraman, R. (2003). Neuroergonomics : Research and practice. *Theoretical issues in ergonomics science*, *4*(1-2), 5–20.
- Park, B., & Brünken, R. (2015). The rhythm method : A new method for measuring cognitive load—an experimental dual-task study. *Applied Cognitive Psychology*, *29*(2), 232–243.
- Parsons, T., & McMahan, T. (2017). An initial validation of the virtual environment grocery store. *J Neurosci Methods*, *291*(1), 13-19.
- Parsons, T. D. (2015). Ecological validity in virtual reality-based neuropsychological assess-

- ment. In *Encyclopedia of information science and technology, third edition* (pp. 1006–1015). IGI Global.
- Patel, K. K., & Vij, S. K. (2010). Spatial navigation in virtual world. *Advanced knowledge based systems : model, applications and research, TMRF e-Book*, 101–125.
- Pincioli, F. (2016). An hci quality attributes taxonomy for an impact analysis to interactive systems design and improvement. In *Xxiii congreso argentino de ciencias de la computación (cacic 2016)*.
- Poupyrev, I., Ichikawa, T., Weghorst, S., & Billinghamurst, M. (1998). Egocentric object manipulation in virtual environments : empirical evaluation of interaction techniques. In *Computer graphics forum* (Vol. 17, pp. 41–52).
- Ramcharitar, A., & Teather, R. J. (2017). A fitts' law evaluation of video game controllers : Thumbstick, touchpad and gyrosensor. In *Proceedings of the 2017 chi conference extended abstracts on human factors in computing systems* (pp. 2860–2866).
- Ramírez-Fernández, C., Morán, A. L., & García-Canseco, E. (2015). Haptic feedback in motor hand virtual therapy increases precision and generates less mental workload. In *Pervasive computing technologies for healthcare (pervasivehealth), 2015 9th international conference on* (pp. 280–286).
- Read, J. M., & Saleem, J. J. (2017). Task performance and situation awareness with a virtual reality head-mounted display. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 61, pp. 2105–2109).
- Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique : A scaling procedure for measuring mental workload. *Advances in psychology*, 52, 185–218.
- Reinhardt, D., & Hurtienne, J. (2017). Interaction under pressure : Increased mental workload makes issues of intuitive interaction visible. In *Proceedings of the 2016 acm conference companion publication on designing interactive systems* (pp. 67–71).
- Reinhardt, D., & Hurtienne, J. (2018). The impact of tangible props on gaming performance and experience in gestural interaction. In *Proceedings of the twelfth international conference on tangible, embedded, and embodied interaction* (pp. 638–646).
- Renison, B., Ponsford, J., Testa, R., Richardson, B., & Brownfield, K. (2012). The ecological and construct validity of a newly developed measure of executive function : the virtual library task. *Journal of the International Neuropsychological Society*, 18(3), 440–450.
- Renner, P., Dankert, T., Schneider, D., Mattar, N., & Pfeiffer, T. (2010). Navigating and selecting in the virtual supermarket : Review and update of classic interaction techniques. In *Virtuelle und erweiterte realität : 7. workshop der gi-fachgruppe vr/ar*.

- Rizzo, A. A., Schultheis, M., Kerns, K. A., & Mateer, C. (2004). Analysis of assets for virtual reality applications in neuropsychology. *Neuropsychological rehabilitation*, *14*(1-2), 207–239.
- Rosa, P. J., Morais, D., Gamito, P., Oliveira, J., & Saraiva, T. (2016). The immersive virtual reality experience : A typology of users revealed through multiple correspondence analysis combined with cluster analysis technique. *Cyberpsychology, Behavior, and Social Networking*, *19*(3), 209–216.
- Roupé, M., Bosch-Sijtsema, P., & Johansson, M. (2014). Interactive navigation interface for virtual reality using the human body. *Computers, Environment and Urban Systems*, *43*, 42–50.
- Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology : Human Perception and Performance*, *27*(4), 763.
- Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload : A comparison of swat, nasa-tlx, and workload profile methods. *Applied Psychology*, *53*(1), 61–86.
- Rupp, A., Levy, R., Dicerbo, K. E., Sweet, S. J., Crawford, A. V., Calico, T., . . . others (2012). Putting ecd into practice : The interplay of theory and data in evidence models within a digital learning environment. *JEDM/ Journal of Educational Data Mining*, *4*(1), 49–110.
- Sakai, H., Nagano, A., Seki, K., Okahashi, S., Kojima, M., & Luo, Z. (2018). Development of a cognitive function test using virtual reality technology : examination in healthy participants. *Aging, Neuropsychology, and Cognition*, *25*(4), 561–575.
- Salvucci, D. D., & Taatgen, N. A. (2010). *The multitasking mind*. Oxford University Press.
- Sangani, S., Fung, J., Kizony, R., Koenig, S., & Weiss, P. (2013). Navigating and shopping in a complex virtual urban mall to evaluate cognitive functions. In *Virtual rehabilitation (icvr), 2013 international conference on* (pp. 9–14).
- Santos, B. S., Dias, P., Pimentel, A., Baggerman, J.-W., Ferreira, C., Silva, S., & Madeira, J. (2009). Head-mounted display versus desktop for 3d navigation in virtual reality : a user study. *Multimedia Tools and Applications*, *41*(1), 161.
- Sayers, H. (2004). Desktop virtual environments : a study of navigation and age. *Interacting with Computers*, *16*(5), 939–956.
- Seelye, A., Hagler, S., Mattek, N., Howieson, D. B., Wild, K., Dodge, H. H., & Kaye, J. A. (2015). Computer mouse movement patterns : A potential marker of mild cognitive impairment. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*,

1(4), 472–480.

- Seligman, S. C., Giovannetti, T., Sestito, J., & Libon, D. J. (2014). A new approach to the characterization of subtle errors in everyday action : implications for mild cognitive impairment. *The Clinical Neuropsychologist*, 28(1), 97–115.
- Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, 114(2), 727–741.
- Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in child development and behavior*, 10, 9–55.
- Skalski, P., Tamborini, R., Shelton, A., Buncher, M., & Lindmark, P. (2011). Mapping the road to fun : Natural video game controllers, presence, and game enjoyment. *New Media & Society*, 13(2), 224–242.
- Slater, M., Usoh, M., & Steed, A. (1995). Taking steps : the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(3), 201–219.
- Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory : investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 29(3), 347.
- Soukoreff, R. W., & MacKenzie, I. S. (2004). Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. *International journal of human-computer studies*, 61(6), 751–789.
- Speicher, M., Daiber, F., Kiefer, G.-L., & Krüger, A. (2017). Exploring task performance and user's preference of mid-air hand interaction in a 3d docking task experiment. In *Proceedings of the 5th symposium on spatial user interaction* (pp. 160–160).
- Stanney, K. M. (2002). *Handbook of virtual environments : Design, implementation, and applications*. Lawrence Erlbaum Associates Publishers.
- Stevens, J., Mondesire, S. C., Maraj, C. S., Badillo-Urquiola, K. A., & Maxwell, D. B. (2016). *Workload analysis of virtual world simulation for military training*.
- Stoet, G., O'Connor, D. B., Conner, M., & Laws, K. R. (2013). Are women better than men at multi-tasking? *BMC Psychology*, 1(1), 18.
- Sutcliffe, A., & Gault, B. (2004). Heuristic evaluation of virtual reality applications. *Interacting with computers*, 16(4), 831–849.
- Tan, D. S., Robertson, G. G., & Czerwinski, M. (2001). Exploring 3d navigation : combining speed-coupled flying with orbiting. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 418–425).

- Tanguay, A. N., Davidson, P. S., Guerrero Nuñez, K. V., & Ferland, M. B. (2014). Cooking breakfast after a brain injury. *Frontiers in behavioral neuroscience*, *8*, 272.
- Teather, R. J., Stuerzlinger, W., & Pavlovych, A. (2014). Fishtank fits : a desktop vr testbed for evaluating 3d pointing techniques. In *Chi'14 extended abstracts on human factors in computing systems* (pp. 519–522).
- Teo, G., Reinerman-Jones, L., Matthews, G., Szalma, J., Jentsch, F., Hudson, I., & Hancock, P. (2017). Selecting workload and stress measures for performance prediction. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 61, pp. 2042–2046).
- Thompson, S. G., McConnell, D. S., Slocum, J. S., & Bohan, M. (2007). Kinematic analysis of multiple constraints on a pointing task. *Human movement science*, *26*(1), 11–26.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, *55*(4), 189.
- Tong, T., & Chignell, M. (2014). Developing serious games for cognitive assessment : aligning game parameters with variations in capability. In *Proceedings of the second international symposium of chinese chi* (pp. 26–27).
- Tong, T., Chignell, M., Tierney, M. C., & Lee, J. (2016). A serious game for clinical assessment of cognitive status : validation study. *JMIR serious games*, *4*(1).
- Tricot, A., & Chanquoy, L. (1996). La charge mentale : “vertu dormitive » ou concept opérationnel? *Psychologie française*, *41*(4), 313–318.
- Tsang, P. S., & Velazquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, *39*(3), 358–381.
- Tucker, A. B. (2004). *Computer science handbook*. CRC press.
- Valladares-Rodríguez, S., Pérez-Rodríguez, R., Anido-Rifón, L., & Fernández-Iglesias, M. (2016). Trends on the application of serious games to neuropsychological evaluation : a scoping review. *Journal of biomedical informatics*, *64*, 296–319.
- Vallejo, V., Wyss, P., Chesham, A., Mitache, A. V., Müri, R. M., Mosimann, U. P., & Nef, T. (2017). Evaluation of a new serious game based multitasking assessment tool for cognition and activities of daily living : Comparison with a real cooking task. *Computers in human behavior*, *70*, 500–506.
- Van der Ham, I. J., Faber, A. M., Venselaar, M., van Kreveld, M. J., & Löffler, M. (2015). Ecological validity of virtual environments to assess human navigation ability. *Frontiers in psychology*, *6*, 637.
- Verhulst, E., Banville, F., Richard, P., Tabet, S., Lussier, C., Massicotte, É., & Allain, P. (2017). Navigation patterns in elderly during multitasking in virtual environment. In *International conference on human interface and the management of information* (pp.

176–188).

- Verhulst, E., Richard, P., Richard, E., Allain, P., & Nolin, P. (2016). 3d interaction techniques for virtual shopping : Design and preliminary study. In *Proceedings of the 11th joint conference on computer vision, imaging and computer graphics theory and applications : Volume 1 : Grapp* (pp. 271–279).
- Vlakoveld, W. P., Twisk, D., Christoph, M., Boele, M., Sikkema, R., Remy, R., & Schwab, A. L. (2015). Speed choice and mental workload of elderly cyclists on e-bikes in simple and complex traffic situations : A field experiment. *Accident Analysis & Prevention, 74*, 97–106.
- Wang, C., Zhang, J., Yu, H., Wang, D., Yang, X., Wang, Q., & Li, F. (2015). Experimental study of optimized interface displays of navigation information system. In *Proceedings of the 15th international conference on man-machine-environment system engineering* (pp. 373–379).
- Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., & Groner, R. (2008). Playing online games against computer-vs. human-controlled opponents : Effects on presence, flow, and enjoyment. *Computers in human behavior, 24*(5), 2274–2291.
- Werner, P., Rabinowitz, S., Klinger, E., Korczyn, A. D., & Josman, N. (2009). Use of the virtual action planning supermarket for the diagnosis of mild cognitive impairment. *Dementia and geriatric cognitive disorders, 27*(4), 301–309.
- Wild, K. V., Mattek, N. C., Maxwell, S. A., Dodge, H. H., Jimison, H. B., & Kaye, J. A. (2012). Computer-related self-efficacy and anxiety in older adults with and without mild cognitive impairment. *Alzheimer's & Dementia, 8*(6), 544–552.
- Williams, B., Bailey, S., Narasimham, G., Li, M., & Bodenheimer, B. (2011). Evaluation of walking in place on a wii balance board to explore a virtual environment. *ACM Transactions on Applied Perception (TAP), 8*(3), 19.
- Wilson, G. F., & Eggemeier, F. T. (1991). Psychophysiological assessment of workload in multi-task environments. *Multiple-task performance, 329360*.
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments : A presence questionnaire. *Presence, 7*(3), 225–240.
- Wong, A. W., Chan, C. C., Li-Tsang, C. W., & Lam, C. S. (2009). Competence of people with intellectual disabilities on using human-computer interface. *Research in Developmental Disabilities, 30*(1), 107–123.
- Wood, R. T., Griffiths, M. D., Chappell, D., & Davies, M. N. (2004). The structural characteristics of video games : A psycho-structural analysis. *CyberPsychology & behavior, 7*(1),

1–10.

- Yang, C., Lin, Y., Cai, M., Qian, Z., Kivol, J., & Zhang, W. (2017). Cognitive fatigue effect on rehabilitation task performance in a haptic virtual environment system. *Journal of Rehabilitation and Assistive Technologies Engineering*, 4, 2055668317738197.
- Yousef Zade, A., Mazloumi, A., Abbasi, M., & Akbar Zade, A. (2016). Investigating the relationship between cognitive failures and workload among nurses of imam khomeini and vali-e-asr hospitals in tehran. *Health and Safety at Work*, 6(2), 57–68.
- Zanbaka, C. A., Lok, B. C., Babu, S. V., Ulinski, A. C., & Hodges, L. F. (2005). Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics*, 11(6), 694–705.
- Zaranek, A., Ramoul, B., Yu, H. F., Yao, Y., & Teather, R. J. (2014). Performance of modern gaming input devices in first-person shooter target acquisition. In *Chi'14 extended abstracts on human factors in computing systems* (pp. 1495–1500).
- Zhao, X., Shi, C., You, X., & Zong, C. (2017). Analysis of mental workload in online shopping : Are augmented and virtual reality consistent ? *Frontiers in psychology*, 8.
- Zucchella, C., Sinforiani, E., Tassorelli, C., Cavallini, E., Tost-Pardell, D., Grau, S., ... others (2014). Serious games for screening pre-dementia conditions : from virtuality to reality ? a pilot project. *Functional neurology*, 29(3), 153.
- Zygouris, S., Giakoumis, D., Votis, K., Doumpoulakis, S., Ntovas, K., Segkouli, S., ... Tsolaki, M. (2015). Can a virtual reality cognitive training application fulfill a dual role ? using the virtual supermarket cognitive training application as a screening tool for mild cognitive impairment. *Journal of Alzheimer's Disease*, 44(4), 1333–1347.
- Zygouris, S., & Tsolaki, M. (2015). Computerized cognitive testing for older adults : a review. *American Journal of Alzheimer's Disease & Other Dementias*® , 30(1), 13–28.

Annexe : Journal d'information issu du VMT

Logs des evenements du 2017-10-26_16-47

Debut de la simulation

0:09 : La porte d'entrée a été ouverte
0:25 : Ramasser Cereale1
0:32 : Un placard de la cuisine a été ouverte
0:35 : Déposer Cereale1 -> Placard
0:41 : Ramasser Poulet
0:42 : Poulet Rôti: Ajouter Poulet_Poulet +1 point
0:42 : Déposer Poulet -> Rotissoire
0:48 : Appel: Appel defroisser la chemise: cuisine_téléphone_combiné (VMT.interactions.Telephone)
1:14 : Ramasser Sirop d'erable
1:18 : Poulet Rôti: Ajouter sirop_erable +1 point
1:18 : Déposer Sirop d'erable -> Versage
1:28 : Déposer Sirop d'erable -> Placard
1:35 : Ramasser Poivre
1:37 : Poulet Rôti: Ajouter Poivriere +1 point
1:37 : Déposer Poivre -> Versage
1:43 : Déposer Poivre -> Placard
1:46 : Ramasser Sel
1:49 : Poulet Rôti: Ajouter saliere +1 point
1:49 : Poulet Rôti: Debut Marinade
1:49 : Déposer Sel -> Versage
1:55 : Déposer Sel -> Placard
2:01 : Le four est allumé
2:07 : La minuterie du four est regler a 10minutes
2:08 : La minuterie du four est regler a 0minutes
2:11 : Appel: Appel des Billets: cuisine_téléphone_combiné (VMT.interactions.Telephone)
2:38 : Ramasser Lettre2
2:40 : Déposer Lettre2 -> Table a l'entré du salon (billets)
2:42 : Ramasser Lettre4
2:44 : Déposer Lettre4 -> Table a l'entré du salon (billets)
2:45 : Le billet a été vérifié
2:59 : Appel: Tache Billet: cuisine_téléphone_combiné (VMT.interactions.Telephone)
3:14 : La porte de la grande chambre a été ouverte
3:17 : Ramasser Chemise
3:23 : La porte droite du placard de la sècheuse a été ouverte

FIGURE A.1 – Exemple de journal d'informations issu du VMT.

Annexe : Questionnaire Général

Rubrique 1 sur 3

Avis général

Description du formulaire

Quel est votre sexe ? *

Masculin

Féminin

Quelle est votre date de naissance ? *

Réponse courte

Quel est votre nationalité ? *

Réponse courte

Quel est votre dernier diplôme obtenu ? *

Réponse courte

Etes-vous droitier ou gaucher ? *

Je suis droitier(e)

Je suis gaucher(e)

Je suis ambidextre



Vous et les ordinateurs

Description (facultative)

A quelle fréquence utilisez-vous un ordinateur ? *

- Jamais
- Rarement
- plusieurs fois par mois
- plusieurs fois par semaines
- tous les jours
- plusieurs fois par jours

Depuis combien de temps utilisez-vous un ordinateur ?

- Quelques mois
- 1 an
- 3 ans
- 5 ans
- 8 ans et plus

Rubrique 2 sur 3



Avec quels matériels utilisez-vous un ordinateur ?

- Souris
- Clavier
- Pad
- Ecran tactile

Rubrique 3 sur 3



Vous et les jeux vidéos

Description (facultative)

A quelle fréquence jouez-vous aux jeux vidéos ?

- Jamais
- Rarement
- plusieurs fois par mois
- plusieurs fois par semaines
- tous les jours
- plusieurs fois par jours

Depuis combien de temps jouez-vous aux jeux vidéos ?

- Quelques mois
- 1 an
- 3 ans
- 5 ans
- 8 ans et plus

Sur/avec quels matériels jouez-vous ?

- PC
- Wii
- Wii U
- Xbox / PS
- Téléphone
- Tablette
- Jeux de Réalité Virtuelle (ex., casque HTC Vive, Oculus)
- RazerHydra

Annexe : *System Usability Scale* - SUS

1. J'aimerais utiliser ce système fréquemment *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

2. Je trouve ce système inutilement complexe *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

3. Je pense que ce système est facile à utiliser *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

4. J'aurais besoin d'un support technique pour utiliser ce système *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

5. Les différentes fonctionnalités de ce système sont bien intégrées *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

6. Ce système est truffé d'incohérences *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

7. Le grand public peut apprendre à utiliser ce système rapidement *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

8. Ce système est lourd à utiliser *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

9. J'ai confiance en ce système *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

10. J'ai dû apprendre beaucoup avant de pouvoir utiliser le système *

	1	2	3	4	5	
Pas du tout d'accord	<input type="radio"/>	Tout à fait d'accord				

Annexe : *Presence Questionnaire* - PQ

5. Dans quelle mesure les mécanismes permettant votre mouvement dans l'environnement vous semblaient-ils naturels?

|_____||_____||_____||_____||_____||
EXTRÊMEMENT À MI-CHEMIN COMPLÈTEMENT
ARTIFICIELS NATURELS

6. Dans quelle mesure vos sens étaient-ils trompés par le réalisme du mouvement des objets à travers l'espace?

|_____||_____||_____||_____||_____||
PAS DU TOUT MODÉRÉMENT TRÈS
TROMPÉS TROMPÉS

7. Dans quelle mesure les expériences que vous avez vécues dans l'environnement virtuel ressemblaient-elles à celles de l'environnement réel ?

|_____||_____||_____||_____||_____||
PAS MODÉRÉMENT TRÈS
RESSEMBLANT RESSEMBLANT RESSEMBLANT

8. Étiez-vous capable d'anticiper les conséquences des mouvements que vous faisiez?

|_____||_____||_____||_____||_____||
PAS DU TOUT ASSEZ COMPLÈTEMENT

9. Jusqu'à quel point étiez-vous en mesure d'explorer activement l'environnement de façon visuelle?

|_____||_____||_____||_____||_____||
PAS DU TOUT ASSEZ COMPLÈTEMENT

10. Jusqu'à quel point la sensation de déplacement à l'intérieur de l'environnement virtuel était-elle confondante (réaliste)?

|_____||_____||_____||_____||_____||
PAS MODÉRÉMENT TRÈS
CONFONDANTE CONFONDANTE CONFONDANTE

11. À quelle distance pouviez-vous examiner les objets?

|_____| |_____| |_____| |_____| |_____| |_____| |_____|
PAS PROCHE PLUTÔT TRÈS
DU TOUT PROCHE PROCHE

12. Jusqu'à quel point pouviez-vous examiner les objets sous différents angles?

|_____| |_____| |_____| |_____| |_____| |_____| |_____|
PAS DU TOUT ASSEZ COMPLÈTEMENT

13. Jusqu'à quel point étiez-vous impliqué(e) dans l'expérience vécue dans l'environnement virtuel?

|_____| |_____| |_____| |_____| |_____| |_____| |_____|
PAS DU TOUT MOYENNEMENT COMPLÈTEMENT
ENGAGÉ(E) ENGAGÉ(E) ABSORBÉ(E)

14. Jusqu'à quel point avez-vous ressenti un délai séparant vos actions de leurs conséquences?

|_____| |_____| |_____| |_____| |_____| |_____| |_____|
AUCUN DÉLAI LONG
DÉLAI MODÉRÉ DÉLAI

15. À quel rythme vous êtes-vous adapté(e) à l'expérience vécue dans l'environnement virtuel?

|_____| |_____| |_____| |_____| |_____| |_____| |_____|
PAS ADAPTÉ(E) LENTEMENT EN MOINS
DU TOUT D'UNE MINUTE

16. En termes d'interactions et de déplacements dans l'environnement virtuel, jusqu'à quel point vous sentiez-vous compétent(e) à la fin de l'expérience?

|_____| |_____| |_____| |_____| |_____| |_____| |_____|
PAS RAISONNABLEMENT TRÈS
COMPÉTENT(E) COMPÉTENT(E) COMPÉTENT(E)

RÉPONDEZ SI L'ENVIRONNEMENT VIRTUEL COMPRENAIT DES ASPECTS HAPTIQUE (TOUCHER) :

23. Dans quelle mesure pouviez-vous explorer activement et de façon tactile (par le toucher) l'environnement?

|_____| |_____| |_____| |_____| |_____| |_____|
PAS DU TOUT ASSEZ COMPLÈTEMENT

24. Jusqu'à quel point pouviez-vous déplacer ou manipuler les objets dans l'environnement virtuel?

|_____| |_____| |_____| |_____| |_____| |_____|
PAS DU TOUT ASSEZ COMPLÈTEMENT

Dernière version: Mars 2013

*Version originale : Witmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence : Teleoperators and Virtual Environments*, 7(3), 225-240. Structure factorielle révisée: Witmer, B.J., Jerome, C.J., & Singer, M.J. (2005). The factor structure of the Presence Questionnaire. *Presence*, 14(3) 298-312.

Annexe : NASA-TLX

Annexe : Simulator Sickness Questionnaire
- (SSQ)



Questionnaire sur les cybermalaises*

Laboratoire de Cyberpsychologie de l'UQO
(Traduit de Kennedy, R.S. et al., 1993)

Numéro _____

Date _____

Consignes : Encerclez à quel point chaque symptôme ci-dessous vous affecte présentement.

1. Inconfort général	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
2. Fatigue	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
3. Mal de tête	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
4. Fatigue des yeux	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
5. Difficulté à faire le focus	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
6. Augmentation de la salivation	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
7. Transpiration	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
8. Nausées	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
9. Difficulté à se concentrer	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
10. Impression de lourdeur dans la tête	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
11. Vision embrouillée	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
12. Étourdissement les yeux ouverts	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
13. Étourdissement les yeux fermés	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
14. *Vertiges	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
15. **Conscience de l'estomac	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>
16. Rots	<u>Pas du tout</u>	<u>Un peu</u>	<u>Modérément</u>	<u>Sévèrement</u>

* Les vertiges sont vécus comme une perte de l'orientation par rapport à la position verticale.

** L'expression « conscience de l'estomac » est habituellement utilisée pour désigner un sentiment d'inconfort sans nausée.

Dernière version : Mars 2013

***Version originale : Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology*, 3(3), 203-220

Annexe : Recette de cuisine

Poulet rôti aux légumes

Ingrédients

- Sirop d'érable
- Huile d'olive
- Poulet
- Sel et
- poivre
- carottes moyennes
- navet
- oignons moyens
- pommes de terre grelots
- Sauce préparée : bouillon de poulet et ail

Préparation

1. Déposer le poulet dans le plat de cuisson. Ajouter le sirop d'érable. Saler et poivrer. Laisser reposer environ 5 minutes.
2. Préchauffer le four à 190 °C (375 °F).
3. Par la suite, ajouter les légumes (navets, oignons, carottes et pomme de terre), la sauce préparée et l'huile d'olive sur le poulet. Mettre la préparation et le poulet au four pour 1h30 minutes (programmer le minuteur sur le four).

FIGURE G.1 – Recette de la tâche de poulet.

Titre : Contribution de l'étude de l'interaction en environnement virtuel : intérêt de la charge mentale

Mots clés : Environnement virtuel, Technique d'interaction, Charge mentale, Evaluation de la performance

Résumé : Les environnements virtuels (EVs) sont de plus en plus utilisés dans le domaine de la recherche et de la clinique avec pour avantage que chaque environnement peut être adapté au besoin du participant mais avec pour inconvénient que la pluralité des variables mesurées lors des tests virtuels rend la performance du sujet de plus en plus difficile à expertiser. L'objectif de la thèse est de proposer une approche permettant de qualifier, de façon la plus pertinente possible, la performance du participant dans un EV en tentant de comprendre la charge mentale associée à chaque technique d'interaction et comment celle-ci peut venir influencer la performance du participant. Cinquante-trois participants ont réalisé des tâches dans un EV avec une des cinq techniques d'interaction disponible : la souris, le gamepad, le Razer Hydra, le Razer Hydra avec un *tracking* de tête et le Razer Hydra avec un casque de RV. Ils étaient attribués à une des deux conditions de difficulté de la tâche : une facile et une difficile avec des tâches additionnelles. Le rythme cardiaque du participant était enregistré et ceux-ci ont renseigné des questionnaires de charge mentale, présence et cybermalaise.

L'expérimentation a permis de mettre en avant que la différence de charge mentale entre les différentes techniques d'interaction est faible et que celle-ci se manifeste plus sensiblement lorsque les techniques sont regroupées par leurs caractéristiques. Ainsi les techniques avec une sélection en 3D et un *steering* effectué par les mouvements de la tête demande plus de charge mentale. Le degré d'expertise avec l'utilisation d'une technique d'interaction n'est pas lié à une modulation de la charge mentale. Cependant les utilisateurs qualifiés d'experts sont plus à même d'explorer le VMT que les novices. De plus, la charge mentale est plus perçue dans les conditions où la difficulté de la tâche est faible alors que dans le cas où la difficulté de la tâche est élevée, ce sont les problèmes d'utilisabilité des techniques d'interaction qui s'expriment. Les participants qui ont une charge mentale élevée ont un score plus faible que les participants qui ont une faible charge mentale, en particulier lors des tâches qui réclament des fonctions cognitives variées.

Title: Study of interaction in virtual environment : focus on workload

Keywords: Virtual Environment, Interaction Technique, Workload, Performance assessment

Abstract: Virtual environments (Ves) can be modulated and be adapted to the need of each user but the several variables used during cognitive assessment during virtual tests give the user's performance uneasy to expertise. The objective of the thesis is to propose an approach to qualify user performance taking account of his/her workload during the task realization. Fifty-three participants realized tasks in a VE with one of five interaction technique: mouse, gamepad, Razer Hydra, Razer Hydra and head tracking and Razer Hydra and HMD. They participate in one of the two conditions: the easy one or the more difficult with additional tasks. Heart rate, workload, presence and cybersickness were measured.

Results showed that there is only a weak difference of workload across the several interaction technique and the difference is stronger when interaction technique are grouped according their similarities. Interaction with 3D selection and with head steering ask for more workload. User abilities with the use of an interaction technique is not related to workload but experts are more likely to explore the VE. What's more the perceived workload is higher in easy condition whereas usability issues are perceived higher in hard condition. Users with high workload have lower score than those with low workload especially for tasks with high cognitive demand.