

École doctorale n° 432 : Science des Métiers de l'ingénieur

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École Nationale Supérieure d'Arts et Métiers

Spécialité « Génie industriel – Conception & Innovation »

présentée et soutenue publiquement par

Rémy EYNARD

le 29 Avril 2016

**Investigation de l'impact de l'interaction sociale vocale sur l'expérience
utilisateur dans les environnements 3D temps-réel immersifs**

Directeur de thèse : **Pr. Simon RICHIR**

Co-encadrement de la thèse : **Dr. Marc PALLOT**

Jury

M. Sébastien GEORGE, Professeur, IEIAH du LIUM, Université du Maine

M. Pierre CHEVAILLIER, Professeur, Lab-STICC, ENIB

M. Stéphane NATKIN, Professeur, CEDRIC, CNAM de Paris

M. Guillaume MOREAU, Professeur, dép. Informatique et Mathématiques, Centrale Nantes

M. Olivier CHRISTMANN, Docteur, LAMPA, Arts & Métiers ParisTech

M. Marc PALLOT, Docteur, LAMPA, Arts & Métiers ParisTech

M. Simon RICHIR, Professeur, LAMPA, Arts & Métiers ParisTech

Président

Rapporteur

Rapporteur

Examineur

Examineur

Examineur

Examineur

**T
H
È
S
E**

Résumé

L'expérience utilisateur est récemment devenue un sujet central dans la conception de produits ou services. Avec la démocratisation de la Réalité Virtuelle (RV), de plus en plus d'acteurs économiques prennent conscience du potentiel de ces technologies dans leurs secteurs respectifs. Dans ce contexte de plus en plus dynamique, l'expérience utilisateur (UX) permet d'anticiper les attentes des utilisateurs et des concepteurs afin de proposer l'expérience la plus satisfaisante possible. De plus, au travers de l'anticipation de l'expérience utilisateur, nous sommes en mesure de prédire de plus en plus précisément le degré d'adoption d'une solution technologique par les futurs utilisateurs. Nous avons aujourd'hui compris que la seule dimension technologique n'est plus suffisante pour assurer la réussite d'un produit matériel et/ou logiciel. En revanche, ce produit doit correspondre à l'usage que l'utilisateur compte en faire, et plus encore, il doit offrir une expérience unique, au travers de sensations, de perceptions, d'émotions et de réflexion, induisant un degré de satisfaction le plus élevé possible.

Un consensus est depuis longtemps établi sur le fait que jouer, créer, apprendre ensemble est plus efficace et entraînant qu'une activité solitaire. Partager et communiquer est primordial, c'est au travers des interactions sociales avec nos semblables que nous sommes capables de combler nombre de nos besoins. Nous avons, au travers de l'étude empirique conduite dans le cadre de cette thèse, apporté une meilleure compréhension de l'impact des interactions sociales par la modalité vocale dans le processus d'expérience utilisateur en 3D-IVE (3D Immersive Virtual Environment). À l'aide d'un état de l'art s'appuyant sur les études existantes de l'impact de l'interaction sociale sur la présence virtuelle (sentiment d'être là) ou l'engagement de l'utilisateur (flow), nous proposons une vision plus holistique de cet impact par une étude centrée sur l'expérience utilisateur globale.

Pour ce faire, nous avons conduit trois expérimentations immergeant des binômes d'utilisateurs dans un 3D-IVE. Par la répartition des participants en deux échantillons, l'un étant en mesure de communiquer vocalement, l'autre ne l'étant pas, nous avons pu comparer des données quantitatives et qualitatives afin de mettre en exergue l'impact des interactions sociales vocales sur l'expérience vécue. Les résultats quantitatifs obtenus ont montré que dans le cas de tâches partagées, la coordination par interaction vocale rend plus efficace le binôme. En revanche, dans le cas de tâches spécifiquement allouées à chaque membre du binôme, l'effet de translucence est comparable à celui de l'interaction sociale vocale sur le plan de l'efficacité. Fait surprenant, nous avons aussi observé que la présence virtuelle ressentie tendait à être plus forte lorsque les participants, ne pouvant pas communiquer vocalement, se complétaient par le phénomène naturel de translucence. Toutefois, nous avons dû pondérer ces résultats, eu égard aux groupes de discussion qui ont fait émerger des différences d'appréhension de la performance, de la présence virtuelle et sociale. Cette triangulation entre quantitatif objectif, quantitatif subjectif et qualitatif a confirmé la nature variable de l'UX (en fonction de l'utilisateur et du temps) et l'influence forte de la nature de l'activité de l'expérimentation.

Nos travaux ont permis de valider la capacité des environnements immersifs simples (logiciel de jeu Minecraft) à générer un bon niveau de présence virtuelle et d'engagement et, de manière générale, un haut niveau de satisfaction chez les utilisateurs. Nous apportons aussi un éclairage sur les méthodes à employer dans ce type de contexte vis-à-vis de la conception des tâches à accomplir ; en effet, une même tâche présentée différemment peut induire une expérience assez dissemblable. Malgré la difficulté à généraliser un résultat dû à la nature très variable de l'expérience utilisateur, nous pensons que nos travaux contribueront à une plus grande compréhension de l'impact des interactions sociales vocales sur l'UX dans les 3D-IVEs.

Mots clés : *Réalité Virtuelle ; Environnements 3D Immersifs ; Expérience Utilisateur ; Interactions Sociales ; Communication Verbale ; Présence Sociale.*

Summary

User experience has recently become a central issue regarding to the design of products or services. With the democratization of Virtual Reality (VR), more and more economic actors start to being aware of the potential of these technologies in their respective sectors.

In this increasingly dynamic context, user experience (UX) allows the anticipation of users' and designers' expectations in order to offer the most satisfying experience as possible. In addition, through the anticipation of user experience, we are able to predict more precisely the adoption level of a technological solution by future users. We have understood today that the only technological dimension is no longer enough to ensure the success of a hardware and/or software product. However, this must fit to the use that user will have of it, and more, it must provide a unique experience, through sensations, perceptions, emotions and thinking, inducing the highest satisfaction possible.

A consensus has been established since a long time about the fact that play, create, learn together are more effective and catchy than a lonely activity. Share and communicate is a primary need, it is through social interactions with our peers that we are able to fulfill many of our needs, from the most pragmatic to the most ethereal. We have, through the empirical study conducted as part of this thesis, brought a better understanding of the impact of social interactions by vocal mode on the user experience process in a 3D-IVE (3D Immersive Virtual Environment). Through a literature review based on existing studies of the impact of social interaction on virtual presence (feeling of being there) or on users' engagement (flow), we propose a more holistic view of this impact by a study focused on the overall user experience.

With this aim, we have conducted three experiments immersing pairs of users in a 3D-IVE. By the division of participants into two samples, one being able to vocally communicate, the other not being so, we were able to compare quantitative and qualitative data to highlight the impact of social interactions (voice) on the user experience. Obtained quantitative results showed that in the case of shared tasks, coordination by vocal interaction make the pair more efficient. However, in the case of tasks specifically allocated to each member of the pair, the effect of translucence is comparable to vocal social interaction in terms of efficiency. Surprisingly, we also observed that the perceived virtual presence tended to be higher when participants that were not able to communicate vocally, relied on the natural phenomenon of translucence. However, we had to balance these results, given to focus groups that have arisen differences regarding to the apprehension of performances, virtual and social presence. This triangulation between objective quantitative, subjective quantitative and qualitative data has confirmed the variable nature of the UX (depending on the user and time) and the strong influence of the nature of the activity of the experimentation.

Our studies have validated the ability of simple immersive environments (Minecraft game software) to generate a high level of virtual presence and engagement, and in general, a high level of satisfaction among users. We also provide an insight into methods to be used in this type of context regarding to the design of tasks. Indeed, the same task presented differently can induce a quite dissimilar experience. Despite the difficulty to generalize a result due to the highly variable nature of user experience, we believe that our work will contribute to a greater understanding of the impact of vocal social interactions on UX in 3D-IVEs.

Keywords : *Virtual Reality ; 3D Immersive Environment ; User Experience ; Social Interactions ; Verbal Communication ; Social Presence.*

Remerciements

Je souhaiterais en premier lieu adresser mes remerciements au Pr Simon Richir et au Dr Marc Pallot pour m'avoir permis d'entreprendre ce travail et pour m'avoir accompagné durant ces trois années. Merci à ceux qui m'ont apporté leur aide, en particulier le Dr Olivier Christmann pour ces conseils avisés et la patience dont il a fait preuve à mon égard et sans laquelle je n'en serais pas là aujourd'hui. Benjamin Poussard pour son aide à la conception du « auraMod », le Dr Grégoire Cliquet qui a rendu possible ma première expérimentation et bien-sûr tous les participants nous ayant offert un peu de leur temps dans le cadre de notre étude empirique.

Je désire remercier chaleureusement ceux qui m'ont soutenu durant ces trois années par leur amitié et leur aide. Évidemment, mes parents, qui malgré la distance, ont toujours été aussi présents que le permettent les truchements de la téléphonie. Louise, pour son attention de chaque instant et son soutien indéfectible. Pascaline, Alex, Kévin, Le Bix et les autres pour avoir été de si bons compagnons, intéressés par mon travail et toujours enclins à me faire oublier les vicissitudes de ces trente-six mois de labeur.

À tous, sincèrement, merci.



Publications

Journaux à comité de relecture

Rémy EYNARD, Marc PALLOT, Olivier CHRISTMANN, & Simon RICHIR (2015). The Power of Vocal Social Interaction within 3D-IVE, Myth or Reality? Soumis à *Presence: Teleoperators and Virtual Environments*, mars 2016.

Actes de conférence

Rémy EYNARD, Marc PALLOT, Olivier CHRISTMANN, & Simon RICHIR (2015). Impact of Verbal Communication on User Experience in 3D Immersive Virtual Environments. Présenté à la *21th ICE/IEE International Technology Management Conference*, Belfast.

Benjamin POUSSARD, Guillaume LOUP, Olivier CHRISTMANN, Rémy EYNARD, Marc PALLOT, Simon RICHIR et al. (2014). Investigating the main characteristics of 3D real time tele-immersive environments through the example of a computer augmented golf platform (pp. 1–8). Présenté à la *2014 Virtual Reality International Conference*, Laval: ACM Press.

Marc PALLOT, Rémy EYNARD, Benjamin POUSSARD, Olivier CHRISTMANN, & Simon RICHIR (2013). Augmented Sport: Exploring Collective User Experience. Présenté à la *2013 Virtual Reality International Conference*, Laval.

Glossaire

Terme	Définitions
3D-IVEs	De IVE : abréviation de l'anglais <i>Immersive Virtual Environment</i> , voir Garau (2003), acronyme auquel nous avons ajouté 3D pour plus de précision.
ACV	Abréviation de Avec Communication Vocale.
ACV-AC	Abréviation de Avec Communication Vocale, Avec Cube émotionnel.
ACV-SC	Abréviation de Avec Communication Vocale, Sans Cube émotionnel.
Agent	Représentation virtuelle d'un être intelligent contrôlée par un ordinateur.
Avatar	Représentation virtuelle d'un être intelligent contrôlée par un humain.
DIME	De l'anglais « <i>Distributed Interactive Multimedia Environment</i> . » (Wu et al., 2009)
ECG	Abréviation d'électrocardiogramme, mesure de l'activité cardiaque.
Échelle bipolaire sémantique	Échelle d'évaluation dont les pôles sont sémantiquement décrits et qui peuvent donc revêtir n'importe quel aspect.
Échelle de Likert	Échelle de jugement très répandue dans les questionnaires psychométriques par laquelle la personne interrogée exprime son degré d'accord ou de désaccord vis-à-vis d'un énoncé sous la forme d'une affirmation.
EEG	Abréviation d' électroencéphalogramme, mesure de l'activité électrique du cerveau.
GSR	Abréviation de l'anglais <i>Galvanic Skin Response</i> , mesure des micro variations de la sudation par la conductance de la peau.
HMD	Abréviation de l'anglais « <i>head-mounted display</i> . »
MMOG	Abréviation de l'anglais « <i>Mass Multiplayer Online Game</i> . »
MMORPG	Abréviation de « <i>Massively Multiplayer Online Role-Playing Game</i> », en français « <i>Jeu de rôle en ligne massivement multijoueur</i> . »
MOBA	Abréviation de « <i>Multiplayer Online Battle Arena</i> » en français « <i>Arène de bataille en ligne multijoueur</i> . »
MOOCs	Abréviation de l'anglais « <i>Massive Open Online Course</i> », en français « <i>Formation en Ligne Ouverte à Tous</i> » (FLOT).
SCV	Abréviation de Sans Communication Vocale.
SCV-AC	Abréviation de Sans Communication Vocale, Avec Cube émotionnel.
SCV-SC	Abréviation de Sans Communication Vocale, Sans Cube émotionnel.

Table des matières

1	Introduction	1
1.1	Motivations	2
1.2	Questions de recherche	3
1.3	Objectifs, contributions et stratégie de la recherche	6
1.4	Plan du mémoire	7
2	État de l’art	9
2.1	Les interactions sociales	11
2.1.1	Les interactions sociales verbales	14
2.1.2	Les interactions sociales non-verbales	16
2.2	L’interaction sociale dans les environnements virtuels	21
2.2.1	Agents et avatars, les outils de l’interaction sociale en environnement virtuel	22
2.2.2	La présence, la co-présence et la présence-sociale	23
2.2.3	Une interaction sociale convaincante en environnement virtuel ?	37
2.3	L’expérience utilisateur	40
2.3.1	De l’utilisabilité à l’expérience utilisateur, évolution d’un concept multimodal	40
2.3.2	Modèles d’expérience utilisateur retenus	44
2.4	Résumé de l’état de l’art	55
2.5	Hypothèses et poursuite des travaux	57
3	Méthodologie	61
3.1	Typologies des méthodes	62
3.1.1	Méthodes quantitatives	62
3.1.2	Méthodes qualitatives	64
3.2	Croiser les données par triangulation	66
3.3	Facteurs mesurés et moyens de mesure	67
3.4	Méthodologie d’étude empirique, comparaisons et itérations	75
3.5	Résumé du chapitre Méthodologie	76

4	Expérimentations	79
4.1	Choix de l'environnement virtuel	81
4.1.1	Minecraft, bac à sable virtuel	81
4.1.2	Pourquoi Minecraft ?	84
4.2	Première expérimentation	87
4.2.1	Objectif et activité proposée	87
4.2.2	Environnement Virtuel	89
4.2.3	Participants	93
4.2.4	Mesures	94
4.2.5	Équipement des participants	97
4.2.6	Procédure	99
4.2.7	Résultats	100
4.2.8	Analyse et synthèse	111
4.3	Seconde expérimentation	117
4.3.1	Objectif et activité proposée	117
4.3.2	Environnement Virtuel	120
4.3.3	Participants	122
4.3.4	Mesures	122
4.3.5	Équipement des participants	125
4.3.6	Procédure	126
4.3.7	Résultats	127
4.3.8	Analyse et synthèse	145
4.4	Expérimentation complémentaire	151
4.4.1	Objectifs de l'expérimentation	151
4.4.2	Environnement virtuel	151
4.4.3	Participants	151
4.4.4	Mesures	152
4.4.5	Équipement des participants	153
4.4.6	Procédure	153
4.4.7	Résultats	154
4.4.8	Analyse et synthèse	168
5	Discussion	171
5.1	Présence sociale	172
5.2	<i>Flow</i> Psychologique (Engagement)	174
5.3	Présence Virtuelle	176

5.4	Performances & Efficacité	177
5.5	Adoption technologique (ressentit hédonique global)	179
5.6	Schémas récapitulatifs des résultats	179
5.7	Tableau récapitulatif des résultats	181
6	Conclusion	185
6.1	Introduction	186
6.2	Réponses aux questions de recherche et validation des hypothèses	187
6.3	Contributions à la connaissance	191
6.4	Limites	192
6.5	Travaux futurs	193
A	Figures supplémentaires	I
B	Annexes à l'expérimentation 1	III
B.1	Environnement virtuels	III
B.2	Configuration Minecraft	IV
B.3	Livret d'informations aux participants	VIII
B.4	Questionnaire Pré Expérimentation	IX
B.5	Questionnaire Post Expérimentation	XI
B.6	Questionnaire a posteriori de l'expérimentation	XIX
C	Annexes à l'expérimentation 2	XXI
C.1	Configuration Minecraft	XXI
C.2	Livret d'informations aux participants	XXIII
C.3	Questionnaire Post-Expérimentation	XXIII
D	Annexes à l'expérimentation complémentaire	XXXIII
D.1	Questionnaire Post-Expérimentation - Echantillons 1 & 2	XXXIII
D.2	Questionnaire Post-Expérimentation - Echantillon 3	XXXVII
D.3	Guide pour groupe de discussion	XLII

Liste des figures

2.1	Modèle de communication de Wilbur Schramm (1954).	12
2.2	Pyramide des besoins humains de Abraham Maslow (1943).	12
2.3	Modulation et perturbations de l'interaction sociale vocale.	15
2.4	Diagramme des sphères proxémiques d'après Edward T.Hall (1966).	18
2.5	Les six émotions fondamentales décrites par Ekman.	19
2.6	Carte de la présence d'après Jayesh Pillai.	24
2.7	Niveaux et dimensions de la présence sociale de Biocca et Harms (2002).	29
2.8	Les symétries de l'interaction pour la présence sociale.	32
2.9	Courbe de la « Uncanny Valley ».	34
2.10	Exemples de représentations tombant dans la « <i>Uncanny Valley</i> ».	37
2.11	Modèle théorique d'évaluation d'un système interactif de Hassenzahl <i>et al.</i> (2000). . .	42
2.12	Ulna, poignée de porte anti-contamination.	44
2.13	Modèle d'expérience utilisateur en nid d'abeille (Morville, 2004).	45
2.14	Chaîne de causalité entre QoS & QoE (Wu <i>et al.</i> , 2009).	45
2.15	Modèle étendu de qualité des DIMEs de Wu <i>et al.</i> (2009).	47
2.16	Modèle de communication en DIMEs (Wu <i>et al.</i> , 2009).	48
2.17	Modèle d'UX théorique de Pallot <i>et al.</i> (2013) d'après Wu <i>et al.</i> (2009).	50
2.18	Modèle du <i>flow</i> psychologique de Csikszentmihalyi (1997).	53
3.1	Méthodologie par triangulation pour l'analyse des données.	67
4.1	Vue d'un monde <i>Minecraft</i> à l'aube.	82
4.2	Différents types de biomes dans <i>Minecraft</i>	83
4.3	Evolution du nombre de publications <i>Minecraft</i>	85
4.4	Exemple de modifications de la qualité visuelle de <i>Minecraft</i>	87
4.5	Version <i>chibi pixel art</i> du personnage <i>Loki</i> par <i>Mattmadeacomc</i>	89
4.6	Différence entre les packs de textures par défaut et <i>Sphax PureBDcraft 64x64</i>	90
4.7	Éléments de l'interface utilisateur de <i>Minecraft</i>	91
4.8	Environnement participant de l'expérimentation 1.	91

4.9	Éléments de la seconde partie de l'expérimentation 1.	92
4.10	Éléments de la partie réservée à l'expérimentateur.	93
4.11	Comparaison entre échelles de Likert à sept et cinq niveaux.	95
4.12	Schéma de la configuration globale de l'expérimentation 1.	99
4.13	Moyennes des évaluations de l'attention allouée.	103
4.14	Moyennes des évaluations de la compréhension mutuelle.	103
4.15	Moyennes des évaluations de l'interdépendance comportementale.	104
4.16	Moyennes des évaluations de l'assistance mutuelle.	104
4.17	Moyennes des évaluations de l'état émotionnel général.	105
4.18	Moyennes des évaluations du <i>Flow</i> Psychologique.	105
4.19	Moyennes des évaluations du sentiment de présence virtuelle.	106
4.20	Moyennes des évaluations de la vivacité sensorielle.	106
4.21	Moyennes des évaluations du degré de contrôle.	107
4.22	Moyennes des évaluations de l'efficacité perçue.	109
4.23	Comparaison du nombre de blocs placés entre les contextes SCV & ACV.	110
4.24	Moyennes des évaluations de l'occurrence du stress.	111
4.25	Moyennes des réponses obtenues au questionnaire a posteriori.	111
4.26	Modèle du personnage à construire.	118
4.27	<i>Mod</i> cube émotionnel.	119
4.28	Vision de l'environnement <i>Minecraft</i> depuis la plateforme d'observation.	120
4.29	Éléments de l'environnement <i>Minecraft</i>	121
4.30	Comparaisons statistiques menées pour l'expérimentation 2.	128
4.31	Moyennes des évaluations de l'attention allouée.	130
4.32	Moyennes des évaluations de la compréhension mutuelle - Q.14.	131
4.33	Moyennes des évaluations de la compréhension mutuelle - Q.15.	131
4.34	Moyennes des évaluations de l'interdépendance comportementale - Q.16.	132
4.35	Moyennes des évaluations de l'interdépendance comportementale - Q.17.	133
4.36	Moyennes des évaluations de l'assistance mutuelle.	133
4.37	Moyennes des évaluations de l'empathie.	134
4.38	Moyennes des évaluations de l'état émotionnel - Q.1.	135
4.39	Moyennes des évaluations du niveau de défi perçu - Q.2.	135
4.40	Moyennes des évaluations de l'amusement - Q.6.	136
4.41	Moyennes des évaluations de la concentration - Q.3.	136
4.42	Moyennes des évaluations du sentiment de présence virtuelle - Q.7.	137
4.43	Moyennes des évaluations du sentiment de présence virtuelle - Q.8.	137
4.44	Moyennes des évaluations du degré de contrôle - Q.4.	138

4.45	Moyennes des évaluations du degré de contrôle - Q.5.	138
4.46	Moyennes des évaluations de la concentration.	139
4.47	Moyennes des évaluations de l'utilité perçue du cube émotionnel.	139
4.48	Moyennes des évaluations de l'efficacité perçue - Q.10.	142
4.49	Moyennes des évaluations de l'efficacité perçue - Q.11.	142
4.50	Moyennes du nombre de blocs placés vis-à-vis de l'expérience préalable de <i>Minecraft</i> . .	143
4.51	Moyennes des évaluations de l'intention d'utilisation - Q.37.	144
4.52	Moyennes des évaluations de l'intention d'utilisation - Q.38.	145
4.53	Moy.éval. de l'attention allouée et de l'imp. du comp., éch. 1 & 2.	156
4.54	Moy. éval. de l'attention allouée et de l'imp. du comp., éch.3.	156
4.55	Moy. éval. de l'attention allouée et de l'imp. de la comm. verb., éch. 1 & 2.	157
4.56	Moy. éval. de l'attention allouée et de l'imp. de la comm. verb., éch. 3.	157
4.57	Moy. éval. de l'attention allouée et de l'imp. de la comm. verb. et du comp., éch. 1 & 2	159
4.58	Moy. éval. de l'attention allouée et de l'imp. de la comm. verb. et du comp., éch. 3 . .	159
4.59	Moyennes des évaluations de l'amusement pour les échantillons 1 & 2.	162
4.60	Moyennes des évaluations de l'amusement pour l'échantillon 3.	162
4.61	Moyennes des évaluations de la présence virtuelle pour les échantillons 1 & 2.	164
4.62	Moyennes des évaluations de la présence virtuelle pour l'échantillon 3.	164
4.63	Moyennes des évaluations de l'efficacité pour les échantillons 1 & 2.	166
4.64	Moyennes des évaluations de l'efficacité pour l'échantillon 3.	166
5.1	Résumé schématique de l'impact de la communication vocale sur l'UX en 3D-IVE. . . .	180
5.2	Résumé schématique des impacts inter facteur de l'expérience utilisateur en 3D-IVE. .	181
A.1	Différentes démarches ou attitudes de marche.	I
A.2	Différentes attitudes en position assise.	I
A.3	Différents gestes associés à leur état d'esprit.	II
A.4	Le questionnaire <i>Self-Assessment Manikin</i> (SAM) (Bradley et Lang, 1994).	II
B.1	Panneaux d'informations dans l'environnement <i>Minecraft</i> de l'expérimentation 1. . . .	III
B.2	Système de blocage du temps et des intempéries.	III
B.3	<i>Minecraft launcher</i> , choix de la version	V
B.4	Livret d'informations aux participants - Expérimentation 1 (partie 1).	VIII
B.5	Livret d'informations aux participants - Expérimentation 1 (partie 2).	IX
C.1	Livret d'informations aux participants - Expérimentation 2.	XXIII

Liste des tableaux

3.1	Différence de paradigme entre mesures quantitatives et qualitatives (Creswell, 2009).	66
3.2	Récapitulatif des moyens de mesures des facteurs de l'UX en 3D-IVE - partie 1.	77
3.3	Récapitulatif des moyens de mesures des facteurs de l'UX en 3D-IVE - partie 2.	78
4.1	Évaluation de l'intrusivité des matériels utilisés.	102
4.2	Répartition des participants vis-à-vis de leur expérience préalable de <i>Minecraft</i>	109
4.3	Intervalles de mesure pour les capteurs ECG et GSR pour la définition des émotions.	119
4.4	Évaluation de l'intrusivité des matériels utilisés.	129
4.5	Répartition des participants vis-à-vis de leur expérience préalable de <i>Minecraft</i>	143
5.1	Récapitulatif des résultats de l'étude - 1	182
5.2	Récapitulatif des résultats de l'étude - 2	183
5.3	Récapitulatif des résultats de l'étude - 3	184

Chapitre 1

Introduction

*« La réalité, ça ne m'enthousiasme pas plus
que ça, mais c'est encore le seul endroit où
l'on peut faire un repas correct. »*

Groucho Marx - Comédien américain

(1890 - 1977)

Sommaire

1.1	Motivations	2
1.2	Questions de recherche	3
1.3	Objectifs, contributions et stratégie de la recherche	6
1.4	Plan du mémoire	7

1.1 Motivations

Bien que très largement étudiée par la communauté scientifique, la compréhension de l'expérience utilisateur¹ soulève encore de nombreuses questions. En raison de sa nature variable propre à l'individu et extrêmement dépendante du contexte, il est assez difficile d'anticiper les réactions des utilisateurs et de savoir sur quels leviers agir afin de rendre l'expérience utilisateur positive. Le but est de proposer une expérience toujours plus satisfaisante (p.ex. intéressante, enrichissante et ludique), tout en restant en adéquation avec les objectifs des concepteurs de produits ou de services. Assurément, une expérience satisfaisante n'est pas uniquement dépendante des qualités ou des fonctionnalités inhérentes à un produit ou à un service, elle implique aussi de nombreux facteurs propres à l'utilisateur. C'est dans la prise en compte de ces nombreux facteurs que se situe l'évolution la plus notable entre expérience utilisateur et utilisabilité (Robert et Lesage, 2010). Une démarche centrée utilisateur (étude, conception, empirique ou théorique) se doit d'avoir une vision holistique, de prendre en compte l'engagement de l'utilisateur ainsi que ses attributs culturels, émotionnels et sociaux.

Les interactions sociales sont de plus en plus souvent intégrées dans les produits ou services proposés aujourd'hui. Depuis que le monde entier est connecté (internet), les possibilités de communication et d'augmentation de l'expérience utilisateur par les interactions sociales se sont fortement accrues, et sont largement plébiscitées. En témoigne le nombre d'utilisateurs des jeux vidéo multijoueurs à univers persistants de type MMORPG² ou MOBA³ qui sont devenus de loin les *leaders* du marché⁴. Les jeux vidéo n'intégrant pas de fonctionnalité sociale sont devenus aujourd'hui assez rares, même ceux qualifiés de « *casual games* » (jeux occasionnels pour la détente) se sont vus pourvus de fonctionnalités qui permettent de solliciter l'aide de ses amis au travers de plateformes sociales comme *Facebook*. Bien que le jeu vidéo soit le champ d'application le plus visible, les interactions sociales médiées se montrent très utiles dans d'autres domaines d'expérience multi-utilisateurs. Par exemple, des applications de réalité virtuelle impliquant les interactions sociales sont déjà fréquemment utilisées pour le travail à distance et la collaboration (Pallot, 2011), la formation pour de nombreux types de métiers ou pour de nouvelles tâches (Rhienmora *et al.*, 2011 ; Rezazadeh *et al.*, 2011), ou encore pour l'apprentissage dans l'éducation au travers des MOOCs⁵ ou de mondes virtuels tel que *Second Life*⁶ (Steel et Jones, 2013). Ces technologies sont aussi utilisées pour tenter d'aider les personnes autistes (Georgescu *et al.*, 2014 ; Smith *et al.*, 2014), les victimes de chocs post-traumatiques (Aiken et Berry, 2015 ; Gahm *et al.*, 2013) ou encore les personnes atteintes de phobies sociales (Brinkman *et al.*, 2012) ou d'autres natures (Cavrag *et al.*, 2014 ; Garcia-Palacios *et al.*, 2002). Des plateformes spécifiquement dédiées aux soins psychologiques commencent à voir le jour, par exemple le projet *Phobos* (PsyTech, 2015). Faciliter la conception de telles applications par une meilleure compréhension des processus de l'expérience

1. Expérience utilisateur : ensemble des perceptions et des réponses d'une personne à l'égard de l'utilisation ou l'anticipation de l'utilisation d'un produit ou service (ISO 9241-11 (2010))

2. MMORPG : abréviation de « *Massively Multiplayer Online Role-Playing Game* », en français « *Jeu de rôle en ligne massivement multijoueur* ».

3. MOBA : abréviation de « *Multiplayer Online Battle Arena* » en français « *Arène de bataille en ligne multijoueur* ».

4. Voir étude statistique sur la plateforme de communication en jeu *Raptr* par la société *Statista* (Sta, 2015)

5. MOOCs : abréviation de l'anglais « *Massive Open Online Course* », en français « *Formation en Ligne Ouverte à Tous* » (FLOT).

6. Voir : *secondlife.com*

utilisateur, permettrait au fil du temps de les rendre plus abordables et ainsi d'aider un plus grand nombre de personnes quelque soit le besoin exprimé.

La réalité virtuelle s'est vue, depuis l'apparition de l'*Oculus Rift* (premier HMD⁷ abordable pour le grand public), propulsée sur le devant de la scène. Depuis s'est engagée une course à l'innovation entre de nombreuses entreprises, géants du numérique ou petites *startups* afin de savoir qui sortira le meilleur HMD et au prix le plus bas. La compétition se joue ici comme ailleurs sur le terrain du marketing ou par la prise de contrôle stratégique d'acteurs influents du marché : rachat de l'entreprise *Oculus Rift* par *Facebook*, partenariat entre *Steam* — géant du jeu vidéo dématérialisé — et *HTC* pour la création du *HTC Vive*. Cette belligérance, en raison de sa médiatisation, a l'avantage de susciter la curiosité de nombreux corps de métiers ne s'étant pas encore posé la question des apports de la réalité virtuelle dans leurs domaines. De nouveaux défis quant à la conception d'applications de réalité virtuelle dédiées à ces besoins émergent. Alors, force est de constater qu'une majorité des applications à concevoir pour ces nouveaux besoins impliquent peu ou prou les interactions sociales.

Le design de produits ou de services, suivant les préconisations actuelles de l'expérience utilisateur et l'orientation de cette dernière, requiert une bonne compréhension des phénomènes impliqués. Nous pensons que c'est uniquement par cette compréhension que nous serons capables de proposer des applications pertinentes, centrées sur le besoin sans ajouts de fonctionnalités superflues en jouant sur les bons leviers de l'expérience utilisateur. Par exemple, a-t-on besoin d'un réalisme extrême pour désensibiliser une personne atteinte d'agoraphobie ? Nous décrirons au cours de ce manuscrit l'importance de l'interaction sociale pour l'expérience utilisateur, nous verrons que, malgré de nombreuses études focalisées sur les interactions sociales et sur leur impact sur un ou plusieurs facteurs de l'expérience utilisateur (p.ex. le *flow* psychologique ou la présence virtuelle), nous avons constaté un manque de travaux proposant un point de vue holistique de cette problématique. Ces raisons nous ont encouragé dans le cadre de ce travail de thèse à nous attacher à la compréhension de l'impact des interactions sociales vocales sur l'expérience utilisateur dans les environnements virtuels 3D temps réel immersif (3D-IVEs⁸). Impact que nous serons en mesure de déceler au travers de facteurs de l'expérience utilisateur tels que la présence virtuelle, l'efficacité de l'utilisateur ou les comportements sociaux.

1.2 Questions de recherche

Notre problématique de thèse dans sa version initiale reposait sur la question suivante :

les interactions sociales ont-elles un impact sur l'expérience utilisateur en environnements virtuels ?

Nous nous sommes rapidement aperçus lors de nos premières recherches bibliographiques que ce champ d'investigation était très, trop large. En effet, les interactions sociales comme nous le verrons

7. HMD : abréviation de l'anglais « *head-mounted display* » désigne les casques immersifs ou visiocasques. Système de visualisation se portant sur la tête et muni d'un petit écran.

8. De IVE : abréviation de l'anglais *Immersive Virtual Environment*, voir Garau (2003), acronyme auquel nous avons ajouté 3D pour plus de précision.

plus en détail dans notre état de l'art suivent la dichotomie suivante : interactions sociales verbales et non verbales. Et au sein même de ces catégories, d'autres divisions apparaissent, ainsi, les interactions sociales non verbales peuvent être étudiées sous l'angle du corps, du visage, du comportement social ou encore des réactions physiologiques incontrôlables (rougissement, sudation). Les interactions sociales verbales quant à elles, comprennent le langage textuel ainsi que le langage vocal. Conscients que nous ne serions pas — dans le temps alloué à cette thèse — en mesure d'étudier les interactions sociales dans leur ensemble, nous avons dû faire un choix cohérent avec nos contraintes temporelles mais aussi avec celles inhérentes à l'étude des environnements virtuels immersifs. À la lumière de travaux tels que ceux de Delamarre (2011) nous avons compris que les interactions verbales et non verbales n'étaient pas égales et n'étaient pas traitées de la même manière ni avec la même priorité par notre cerveau. Le verbal étant traité de manière prioritaire et le non verbal comme un complément (dans la plupart des cas). Nous avons découvert le nombre important d'études portant sur la restitution du non verbal dans les environnements virtuels et de son impact sur des concepts fondamentaux de l'expérience utilisateur en environnement virtuel tel que la présence. Puis nous avons constaté le peu d'études portant sur ce que nous considérons comme un acquis, quelque chose de naturel : notre capacité à parler.

Pour ces raisons, nous avons pris la décision de focaliser nos travaux sur l'impact de l'interaction sociale vocale sur l'expérience utilisateur en 3D-IVEs. Cette focalisation nous permet une étude non pas centrée sur un ou une partie des facteurs de l'expérience utilisateur, mais sur son ensemble. À partir de ce point, nous avons posé cette problématique sous forme de questions opérationnelles, auxquelles nous pourrions apporter des réponses d'une part par notre revue de la littérature, puis d'autre part grâce à un ensemble d'expérimentations. C'est en raison de la nature de l'expérience utilisateur que nous avons fait le choix d'une démarche empirique. En effet, conformément à sa subjectivité, à sa variabilité d'une personne à l'autre, il est nécessaire de produire et d'accumuler les observations en vue d'une analyse permettant d'en extraire des constats plus généraux. Devant l'impossibilité de proposer une formulation mathématique de l'expérience utilisateur, seuls les résultats d'une étude empirique (observations et raisonnement synthétique) permettront d'induire une nouvelle connaissance (ou compréhension du phénomène observé) dans le contexte étudié. Nous allons à présent exposer ces questions ainsi que les différents points nécessaires à l'apport de réponses.

Q1 : Les interactions sociales vocales ont-elles un impact sur l'expérience utilisateur lors de l'immersion de plusieurs utilisateurs dans un 3D-IVE partagé ?

- **Q.1.1 :** Quels sont les facteurs mesurables des modèles existants de l'expérience utilisateur ?
 - **Q.1.1.1 :** Existe-t-il une relation d'ordre entre facteurs ? Certains facteurs ont-ils un impact plus fort sur l'expérience utilisateur que d'autres ?
 - **Q.1.1.2 :** Existe-t-il des corrélations entre facteurs (impact d'un facteur sur l'autre) ?
- **Q.1.2 :** D'après la littérature et la norme ISO 9241-11, un consensus existe sur le fait que l'expérience utilisateur est subjective puisqu'elle est principalement caractérisée par les ressentis de l'utilisateur vis-à-vis de l'expérience qu'il a vécu, vit ou vivra. Dans ce cas, est-il possible de mesurer certains de ces facteurs de manière objective ? Pour donner un point de comparaison, les émotions sont un phénomène purement subjectif. Néanmoins, il est possible de les mesurer

par la physiologie de la personne (rythme cardiaque, conductance de la peau), donc de manière plus objective.

— **Q.1.3 :** Quelles sont les composantes de l’interaction sociale vocale, les stimulations induites et les réponses produites vis-à-vis de l’interlocuteur ?

— **Q.1.3.1 :** Existe-t-il une relation d’ordre entre les composantes de l’interaction sociale vocale (composante plus importante dans la compréhension, la force de la réponse obtenue) ?

Les réponses aux points sus-cités seront apportées par la revue de littérature. Ainsi, nous devons identifier les facteurs de l’expérience utilisateur que nous serons capable d’évaluer, les possibilités de mesures de facteurs comme la présence virtuelle dans un 3D-IVE afin de proposer une évaluation plus robuste de l’expérience utilisateur. Nous devons aussi mettre en exergue les composantes de l’interaction sociale vocale : intonation, débit, volume ainsi que leur fonctionnement au sein d’un 3D-IVE. Enfin, à l’égard des constats produits par cette revue de littérature, nous pourrions nous questionner sur les possibilités d’évaluation de l’impact des interactions sociales vocales sur l’expérience utilisateur lors de l’immersion de plusieurs utilisateurs dans un 3D-IVE partagé.

Q.2 : Cet impact est-il fortement soumis à la nature de la tâche, de l’activité proposée ?

En raison de sa nature subjective et évolutive, il est difficile de généraliser des résultats quant à l’impact d’un facteur sur l’expérience utilisateur. En effet, selon la personne et le contexte, l’expérience vécue peut se voir largement modifiée. C’est pourquoi afin d’apporter une vision plus générale du potentiel impact des interactions sociales vocales sur l’expérience utilisateur en 3D-IVE, il est important de produire les mêmes évaluations suivant des protocoles très proches et à la fois différents. Afin de tester cette interrogation, nous nous sommes attachés à la nature de l’activité proposée et en accord avec un domaine dans lequel les interactions sociales sont fortement sollicitées, le travail à distance : la tenue d’une activité à plusieurs. Néanmoins, la difficulté ici était de proposer des activités différentes sans pour autant l’être au point de totalement changer l’expérience vécue. Nous nous sommes alors tournés vers deux modes de travail : la collaboration et la coopération. Ces deux modes présentent une forte similarité dans le sens où leur objectif est d’accomplir une tâche commune. En revanche, leur différence réside dans la manière d’accomplir cette tâche. En collaboration, les personnes impliquées travaillent de concert sur la tâche qui leur est soumise, sans énoncé particulier quant à la division du travail. A contrario, en coopération, le travail est clairement divisé entre les personnes impliquées et la réussite de la tâche globale implique la réussite de chacune des parties impliquées (Howard *et al.*, 2014). L’objectif est alors d’évaluer les différences et similarités entre ces deux méthodes, si similitude il y a, en raison de leur récurrence, les impacts relevés seront plus enclin à apparaître dans d’autres types de tâches, de contextes. Nous apporterons des éléments de réponse grâce à une étude empirique par la comparaison de nos expérimentations proposant des protocoles très proches, mais comportant cette différence sensible au niveau de la tâche à exécuter en mode collaboration ou coopération.

1.3 Objectifs, contributions et stratégie de la recherche

Malgré les nombreuses recherches déjà menées dans le domaine de la compréhension de l'expérience utilisateur, les questions soulevées au croisement de la réalité virtuelle, des interactions sociales et de l'expérience utilisateur n'ont pas reçu aujourd'hui de réponses consensuelles. Des travaux ont été menés dans plusieurs directions que nous allons explorer afin de construire une vision claire et holistique de notre problématique, préalablement aux expérimentations que nous souhaitons mener.

Dans nos travaux, qui seront décrits dans les parties suivantes, nous parlerons d'immersion partagée ou commune dans un 3D-IVE, nécessitant une communication parfaitement synchrone sans laquelle les phénomènes de co-présence et de présence sociale ne peuvent apparaître. Nous faisons ici une différence entre la présence physique de l'autre, le fait de se trouver spatialement immergé dans le même environnement virtuel, et sa présence sociale qui implique non seulement la présence virtuelle, mais aussi une forme de communication entre les utilisateurs.

Dans le cadre d'une communication synchrone dans un 3D-IVE partagé, de nombreuses études ont été menées. Toutefois, ces travaux sont cloisonnés, s'intéressant principalement à l'expérience utilisateur ou aux interactions sociales, mais peu, voir pas du tout aux liens, aux intrications qu'ils présentent l'une envers l'autre. C'est cette lacune que nous désirons commencer à combler par ce travail de thèse. Du côté des plateformes de jeux en ligne précédemment citées, par exemple *World of Warcraft (Wow)*, les travaux menés donnent de nombreux indices sur l'engagement des utilisateurs, sur leurs manières d'interagir et sur les modèles sociaux se dessinant dans ce type de contexte (Hughes, 2015 ; Siitonen, 2007 ; Yee, 2006). Bien que le contexte des MMOG⁹ soit quelque peu différent de celui que nous étudions (les 3D-IVEs), nous pouvons trouver ici de nombreuses indications afin d'orienter notre travail. Dans un contexte similaire au notre, les travaux de Bailenson *et al.* (2008a;b; 2006; 2005a;b; 2004; 2003; 2002; 2001a;b) sont très complets à propos des interactions sociales en 3D-IVE. Ils démontrent d'une part le besoin de réalisme comportemental de la restitution virtuelle des interactions sociales, mais surtout que les comportements sociaux des utilisateurs dans les environnements immersifs sont cohérents avec ceux produits dans le monde réel. Les travaux de Bailenson *et al.* ne présentent néanmoins pas de liens explicites avec l'expérience utilisateur bien que les facteurs évalués soient comme nous le verrons, impliqués dans l'expérience utilisateur selon la vision que nous en adoptons.

Cette vision de l'expérience utilisateur en 3D-IVEs sur laquelle s'appuient nos travaux repose sur des modèles existants, dont un présentant une dichotomie particulièrement intéressante de l'expérience utilisateur en deux branches (Wu *et al.*, 2009):

- **La qualité de l'expérience (QoE)** comme décrite par Gaggioli *et al.* (2003), est la partie subjective de l'expérience utilisateur. L'une de ses principales composantes est le sentiment de présence également appelé téléprésence ou présence virtuelle selon le contexte de la recherche, synthétisé comme « *le sentiment d'être là* » (« *Being there* ») dans un 3D-IVE.
- **La qualité de service (QoS)** correspond quant à elle aux performances techniques de l'envi-

9. MMOG : abréviation de l'anglais « *Mass Multiplayer Online Game* », jeu vidéo en ligne impliquant un très grand nombre de joueurs simultanément.

ronnement virtuel ou de l’application, elle conditionne par relation directe la QoE. Elle inclut par exemple la légère latence entre le mouvement de la tête d’un utilisateur et sa reproduction virtuelle (latence en partie coupable de la « *cyber sickness* »¹⁰).

Cette construction multimodale de l’expérience utilisateur dans les 3D-IVEs nous est apparue comme très pertinente, non seulement dans les facteurs qu’elle propose — que nous détaillerons par la suite — mais aussi dans les liens de causalité entre QoS et QoE permettant une modélisation de l’impact de facteurs pouvant sembler mineurs de prime abord, mais ayant un impact non négligeable. Cette modélisation permet aussi une meilleure compréhension de l’expérience utilisateur en elle-même, vue non plus comme un système de facteurs rendus quelque peu nébulaire par sa subjectivité ; mais comme un ensemble cohérent proposant des liens, des mises en relation interfacteurs mesurables et, dans une certaine mesure, contrôlables.

Les objectifs que nous nous proposons d’atteindre par ce travail de thèse sont les suivants :

- **O.1** : mettre en exergue les leviers liés aux interactions sociales vocales à mettre en place pour offrir une expérience utilisateur satisfaisante dans les 3D-IVEs. Pour ce faire nous devons :
- **O.2** : apporter une meilleure compréhension des facteurs liés aux interactions sociales vocales impactant l’expérience utilisateur ;
- **O.3** : mesurer certains facteurs de l’expérience utilisateur de manière plus objective.

Par l’atteinte de ces objectifs, nous espérons non seulement contribuer à améliorer la connaissance et la compréhension de l’expérience utilisateur en 3D-IVE, mais aussi permettre aux concepteurs de faire les bons choix quant à l’utilisation des interactions sociales dans leurs applications en accord avec leurs propres besoins. Pour ce faire, nous nous proposons de construire plusieurs expérimentations comparant chacune deux contextes : l’un sans communication vocale, l’autre avec. Dans ce cadre, nous établirons des mesures quantitatives et qualitatives des facteurs subjectifs et objectifs de l’expérience utilisateur afin d’avoir une vision holistique de nos résultats par la triangulation des données quantitatives et qualitatives qui seront collectées.

1.4 Plan du mémoire

Ce travail de thèse, et par extension l’étude empirique qu’il relate, a pour objectif d’apporter une meilleure compréhension de l’impact des interactions sociales vocales sur l’expérience utilisateur au sein des 3D-IVEs. Dans ce but, le **chapitre 2** présente la revue de la littérature menée premièrement sur l’interaction sociale sous ses aspects verbaux (comprenant la communication vocale) et non verbaux. Deuxièmement, elle propose un résumé des publications scientifiques ayant étudié les interactions sociales dans le contexte de la réalité virtuelle ainsi qu’un exposé des concepts fondamentaux de ce champ d’investigation. Cela, avec pour finalité d’attester de la possibilité d’une interaction sociale convaincante en environnement virtuel. Cette revue de littérature traite par la suite l’expérience utili-

10. Cyber sickness : en français mal de la réalité virtuelle est un trouble nauséeux présumément causé par une discordance entre la perception visuelle et le système vestibulaire (oreille interne). Les symptômes du mal de la réalité virtuelle sont quasi identiques à ceux que ressentent les personnes atteintes de mal de mer ou du mal de l’air.

sateur en environnement virtuel, afin de comprendre les taxonomies de facteurs des modèles existants et leurs liens avec les concepts fondamentaux de la réalité virtuelle précédemment exposés. Enfin grâce à un résumé de la revue de littérature effectuée, nous introduirons la suite des travaux à mener (étude empirique) afin d'éprouver les hypothèses qui seront alors exposées.

Grâce à cette étude bibliographique, nous serons dans le **chapitre 3** en mesure de proposer une méthodologie d'étude empirique basée sur la triangulation de données quantitatives et qualitatives. Pour ces types de données seront présentés des méthodes d'évaluation en adéquation avec les taxonomies de facteurs précédemment mises en exergue par l'étude de l'expérience utilisateur. Et pour finir, nous proposerons une méthodologie d'étude empirique basée sur l'itération et la comparaison des expérimentations.

De ces recherches bibliographiques et méthodologiques découlera, dans le **chapitre 4**, l'exposé de notre étude empirique composée de deux expérimentations principales ainsi que d'une expérimentation complémentaire. Chacune d'entre elles sera décrite selon une trame commune présentant : ses objectifs, les choix faits quant à la conception de l'activité proposée aux participants, la conception de l'environnement virtuel, les données afférentes aux participants, l'équipement de ces derniers, le protocole expérimental mis en place et enfin les résultats obtenus ainsi que l'analyse les concernant.

Le **chapitre 5** présentera les discussions générales liées à notre étude et en proposera une vision analytique et critique par la comparaison des trois expérimentations produites, ce afin de proposer une vision globale des résultats obtenus et ainsi de la validation de nos hypothèses. Au sein de ce chapitre, nous proposerons également un diagramme récapitulatif de nos résultats, apport important de nos travaux.

Le **chapitre 6** clôturera ce manuscrit en répondant aux questions de recherches, il décrira la contribution de ce travail de thèse à l'ensemble de la connaissance existante, énoncera les leçons apprises et présentera des recommandations quant aux résultats annoncés ultérieurement. Enfin, il reprendra les limites de nos travaux d'un point de vue général ainsi que les prérequis en vue de futurs travaux de recherches.

Chapitre 2

État de l'art

*« Quand nous débattons, vous et moi,
chacun de nous deux devient l'autre. Car
lorsque je comprends ce que vous
comprenez, je deviens votre compréhension
et je suis vous, d'une certaine ineffable
façon. »*

Érigène - Clerc et philosophe
irlandais (815 - 877)

Sommaire

2.1	Les interactions sociales	11
2.1.1	Les interactions sociales verbales	14
2.1.1.1	Qu'est-ce que l'interaction sociale verbale ?	14
2.1.1.2	L'interaction sociale verbale médiate	15
2.1.2	Les interactions sociales non-verbales	16
2.1.2.1	Qu'est ce que l'interaction sociale non-verbale ?	16
2.1.2.2	L'interaction sociale non-verbale médiate	19
2.1.2.3	La translucence sociale, comportement non verbal	20
2.2	L'interaction sociale dans les environnements virtuels	21
2.2.1	Agents et avatars, les outils de l'interaction sociale en environnement virtuel	22
2.2.2	La présence, la co-présence et la présence-sociale	23
2.2.2.1	La présence, concept fondamental	23
2.2.2.2	La co-présence, et la présence sociale « Being there with others » .	28
2.2.2.3	La théorie de la « Uncanny Valley » et l'importance de la cohérence	33
2.2.3	Une interaction sociale convaincante en environnement virtuel ?	37
2.3	L'expérience utilisateur	40
2.3.1	De l'utilisabilité à l'expérience utilisateur, évolution d'un concept multimodal	40

2.3.2	Modèles d'expérience utilisateur retenus	44
2.3.2.1	Wu <i>et al.</i> (2009), dichotomie entre QoS et QoE	44
2.3.2.2	Pallot <i>et al.</i> (2013), Expérience Immersive & QoE	48
2.4	Résumé de l'état de l'art	55
2.5	Hypothèses et poursuite des travaux	57

Les notions et concepts décrits dans cet état de l'art sont très souvent dépendants d'autres notions ou concepts. En effet, comme nous l'avons déjà évoqué, l'expérience utilisateur, conformément à sa nature fluctuante et subjective, fait appel à plusieurs éléments en même temps. Cette nature, aux antipodes de la linéarité inhérente à tout document écrit, a rendu la rédaction de ce chapitre délicate. Une attention particulière a été apportée à la facilité de compréhension de la globalité du chapitre en essayant de minimiser les retours en arrière ou les évocations anticipées de notions ou de concepts non explicités. Nous commencerons donc ce chapitre par une revue de ce que sont les interactions sociales ainsi que par la définition du concept de « présence », un fondamental en réalité virtuelle. Puis nous nous intéresserons aux modèles d'expérience utilisateur sur lesquels se base notre étude empirique, nous en définirons chacun des facteurs, ce qui nous permettra de comprendre ceux qui font partie du modèle des 3D-IVEs.

2.1 Les interactions sociales

Les interactions sociales¹ sont le fondement de tout échange entre humains, elles nous ont permis depuis les prémices de notre espèce d'échanger sur des concepts simples, complexes, abstraits, de transmettre à nos semblables de manière consciente ou inconsciente nos émotions et notre état d'esprit. Elles sont « *une condition sine qua non de la vie humaine et de l'ordre social* » [notre traduction] (Watzlawick *et al.*, 1972). En effet, sans expressions de nos désirs, préoccupations et aspirations, comment faire valoir sa position sociale ? Comment mettre en commun des objectifs sommaires tels qu'une répartition de tâche ou simplement demander de l'aide à un semblable ?

L'interaction sociale est un phénomène dynamique et transactionnel, chaque action de l'une des parties impliquées sera perçue, décodée puis engendrera une réponse (il est admis que le silence ou l'absence de réaction est une réponse) qui elle-même sera perçue, décodée et ainsi de suite jusqu'à l'apparition d'un protocole de rupture de l'interaction sociale (Picard, 2001). Ce processus s'illustre par le modèle de Schramm (1954) qui, bien que datant de 1954, reste en raison de sa simplicité une bonne illustration du processus d'interaction sociale (figure 2.1). Sur ce schéma, les rôles d'émetteur et de récepteur ne sont pas explicitement inscrits puisque les protagonistes assument chacun des rôles tour à tour au fil des boucles de communication.

L'interaction sociale a toujours un but : **influencer**¹. Ici, la notion d'influence a de multiples facettes : enseigner un savoir, transmettre un sentiment, prévenir d'un danger, féliciter, encourager, se moquer, déprécier, interroger, requérir, ponctuer la conversation pour montrer notre approbation ou encore notre engagement. Ces tentatives d'influence produiront en retour différents types de réactions : une parole, un mouvement, une absence de réaction. Réactions qui, à leur tour auront une influence sur l'émetteur du premier message (Figure 2.1). Bien sûr, ces interactions ne sont pas anodines et sont émises en réaction à un besoin.

1. Interaction sociale : Relation interpersonnelle entre deux individus au moins, par laquelle les comportements de ces individus s'influencent mutuellement et se modifient chacun en conséquence (Larousse 2015).

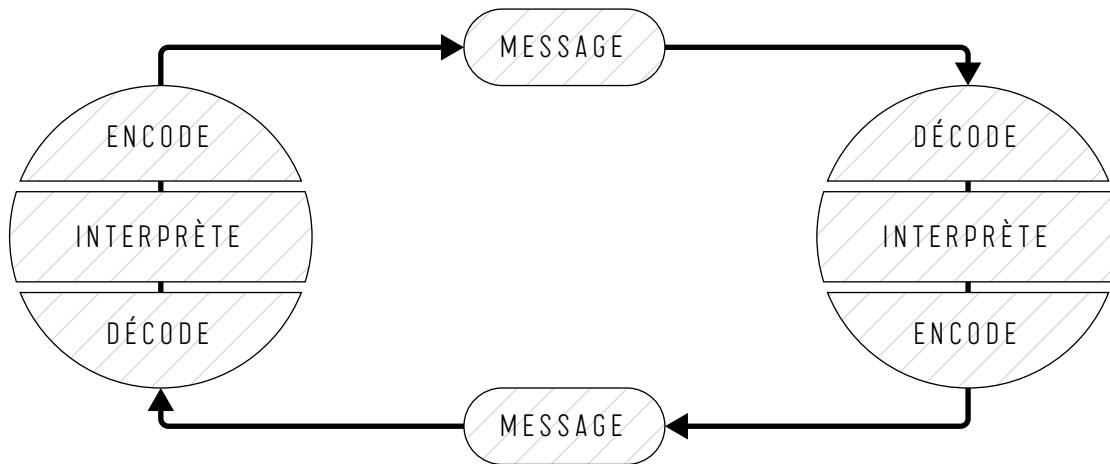


FIGURE 2.1 – Modèle de communication de Wilbur Schramm (1954).

Une description communément admise de ces besoins est celle proposée par le psychologue Abraham Maslow (1943) au travers de sa pyramide des besoins. Cette pyramide hiérarchise les besoins de l'homme du plus élémentaire au plus spirituel (du corps à l'âme) (Figure 2.2). Cette hiérarchisation trouve sa logique dans le fait qu'il est difficile de se préoccuper de l'accomplissement de soi quand on n'est pas certains de pouvoir se nourrir ou de savoir si ses proches seront en sécurité pour les heures, les jours ou les mois à venir.

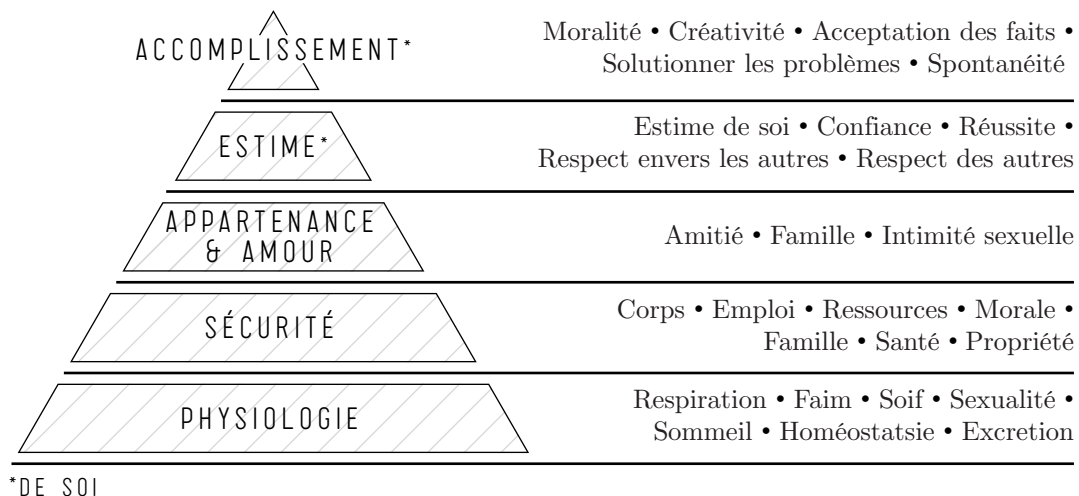


FIGURE 2.2 – Pyramide des besoins humains de Abraham Maslow (1943).

Nous avons, pour combler ces besoins, une absolue nécessité des autres, par exemple comment commercer pour obtenir de la nourriture sans interagir socialement avec un autre individu ? Comment séduire un futur partenaire sans lui exprimer notre intérêt ? Bien que la parole ne soit pas le seul canal de communication impliqué dans la séduction, il reste primordial. Mais encore faut-il que nous partagions les mêmes modalités d'interactions : le langage, la culture ainsi que les normes sociales. C'est pourquoi, de manière bien souvent inconsciente, le premier message que nous envoyons à un nouvel interlocuteur tend à tester le partage ou non de ces modalités et en fonction de cela, nous permet de définir quelle stratégie adopter pour que l'interaction soit fructueuse. Un simple « *Bonjour* » nous permet de savoir instantanément ou presque si notre interlocuteur partage notre langue et s'il

comprend que ce terme est un rituel d'ouverture d'interaction sociale (Picard, 2001).

La complexité de ce processus peut facilement être expérimentée lors d'un voyage dans un pays étranger dont nous ne parlons pas la langue. Face à ce genre de problème, nous essayons alors de créer un langage partagé par la gestuelle, par exemple en montrant l'objet convoité. Cela implique alors l'utilisation d'une communication non verbale ainsi que l'utilisation contextuelle de l'environnement. En effet, toute communication interpersonnelle est multimodale. Elle n'est pas seulement un alignement de sons produits par nos cordes vocales, elle comporte de nombreux autres paramètres non verbaux que nous sommes, grâce à notre éducation et notre programmation génétique, capables de décoder la plupart du temps d'une manière plus ou moins adéquate et perspicace.

La grande variété de facteurs impliqués dans les communications verbales et non verbales fait que l'interaction sociale à travers les technologies numériques constitue un défi. Sans une transmission globale et fidèle de l'interaction sociale, la communication devient plus complexe et peut mener à une incompréhension mutuelle; tout le monde a un jour mal compris un SMS (Small Message System) ou un message instantané. Assurément, l'ironie est une figure de style bien difficile à faire passer sans indice supplémentaire que le simple sens des mots.

Il nous faut donc retenir de l'interaction sociale qu'elle est générée pour répondre à nos besoins, qu'elle est dynamique puisque chaque nouvelle boucle de perception-décodage en entraîne une autre et qu'elle implique de multiples modalités verbales et non verbales. Afin de mieux comprendre cette multimodalité, il est nécessaire de faire une dichotomie des interactions sociales suivant les deux canaux principaux de communication : le verbal et le non verbal.

La dichotomie de l'interaction sociale commence dans notre cerveau. En effet, l'organisation de nos neurones n'est pas étrangère à la perception que nous avons du langage verbal et du langage non verbal. Comme l'explique Delamarre (2011), notre cheminement cérébral vis-à-vis de la communication est hiérarchisé par deux types de neurones :

- **Les neurones pyramidaux** (ainsi nommés en référence à leur forme) sont à l'origine des mouvements volontaires et sont les acteurs principaux de la communication verbale (discours, lecture).
- **Les neurones extrapyramidaux** sont à l'origine des mouvements involontaires et les acteurs essentiels de la communication non verbale par le biais du système sensorimoteur.

Les cellules pyramidales sont décrites comme étant les instances supérieures de notre cerveau auxquelles il va se référer en premier, tandis que les cellules extrapyramidales se situent plus bas dans l'échelle hiérarchique et ne sont donc pas prioritaires dans le traitement des informations reçues. Cette hiérarchie explique que nous nous référons toujours ou presque en premier lieu au langage verbal, ne lisant les informations non verbales que pour combler un manque de compréhension ou pour obtenir des informations supplémentaires. Cependant, même par l'utilisation d'un seul type de canal, les possibilités de messages différents sont immenses, là encore grâce à notre physiologie.

2.1.1 Les interactions sociales verbales

2.1.1.1 Qu'est-ce que l'interaction sociale verbale ?

En premier lieu, nous donnerons deux définitions qui permettront d'éviter toute confusion vis-à-vis des termes utilisés. La communication verbale est un mode de communication basé sur l'utilisation du **verbe**. Cela signifie qu'elle n'implique pas nécessairement l'emploi de la voix, car l'écrit ou l'usage de la langue des signes sont des formes de communication verbale. La communication vocale, quant à elle, passe forcément par la **voix** et implique, pour la modulation du message transmis, les critères suivants : hauteur, intensité, timbre et intonation.

La voix dépend d'un organe en particulier, le larynx. Placé sur le chemin de l'air inspiré et expiré, il abrite les cordes vocales. Ces dernières, permettent grâce à leurs tensions ou relâchements de produire des sons, de nous donner une voix. On peut distinguer trois modalités de la voix : le cri, la parole, le chant. Mais l'éventail des possibilités de la voix ne se limite pas à ces trois modalités. Lors de l'émission d'un message vocal, nous sommes capables de moduler les sons émis de plusieurs manières, cela entraîne des effets différents sur le sens donné et la compréhension du message.

Les trois principaux facteurs variables de la voix sont :

- **La hauteur**, elle est ce que nous identifions comme la caractéristique grave ou aiguë d'une voix. Définie par la fréquence de l'onde sonore sur un oscillographe et dans la réalité, par le fait qu'un homme possède dans la majorité des cas une voix plus grave que celle d'une femme ou d'un enfant.
- **L'intensité** ou autrement dit le volume, elle représente la puissance du son émis. Physiologiquement, elle dépend en grande partie de l'énergie pulmonaire déployée par l'émetteur du son.
- **Le timbre** de la voix et aussi appelé couleur de la voix, il est ce qui permet de reconnaître une personne sur simple écoute de quelques mots / phrases.
- **l'intonation**, elle est le mouvement musical de la voix et se caractérise par les variations de la hauteur des voyelles. Elle permet d'exprimer la catégorie à laquelle appartient la phrase prononcée (affirmative, interrogative, impérative) ainsi que les modalités expressives de cette dernière (interrogation ironique, doute, dégoût, surprise).

Grâce à ces variateurs vocaux, nous allons pouvoir mettre en place un arsenal de fluctuations afin de transmettre le message désiré et surtout afin qu'il soit compris de la manière voulue. Par exemple, si le temps est pluvieux et que cela m'attriste, j'utiliserai la phrase « *Il pleut aujourd'hui* » avec une intensité et une hauteur assez basse ainsi qu'avec une intonation descendante ce qui donnera à mon discours un ton triste. Celui-ci permettra à mon interlocuteur non seulement de comprendre à quoi je fais référence (grâce au sens des mots), mais aussi l'effet que produit sur moi ce constat et donc mon état émotionnel. A contrario, si je suis agriculteur et que j'attends la pluie depuis plusieurs semaines, je prononcerais cette même phrase avec une intensité et une hauteur plutôt élevée et avec une intonation ascendante qui traduira une émotion positive. Ces modulations de l'intensité, de la

hauteur et de l'intonation de la voix nous permettent l'emploi de très nombreux tons, chacun donnant des indications sur notre état émotionnel et sur la direction dans laquelle nous souhaitons orienter l'interaction sociale en cours.

Nous avons utilisé le terme de besoin, nous parlons bien sûr des besoins de l'humain évoqués au travers de la pyramide de Maslow (Figure 2.2), mais aussi des besoins contextuels qui peuvent être liés à l'émetteur, à l'interlocuteur ou à l'environnement et qui peuvent venir perturber le bon déroulement de l'interaction sociale. Par exemple, nous pouvons citer les problèmes d'élocution tels que le bégaiement ou les problèmes d'articulation pouvant survenir après un AVC². Du côté du récepteur, les problèmes d'audition dus à des traumatismes sonores ou à des maladies telles que l'otospongiose³ rendent l'interaction sociale vocale difficile et demandent un effort pour comprendre le message reçu. Outre les possibles problèmes physiologiques, des facteurs environnementaux impactent fortement l'interaction sociale vocale : la distance entre l'émetteur et le récepteur, les obstacles présents entre les interlocuteurs et surtout le bruit ambiant. Ces environnements différents vont conditionner les modalités vocales que nous allons déployer en prenant en compte que la difficulté de compréhension est facteur du nombre d'obstacles auxquels notre voix devra faire face. Le schéma suivant (Figure 2.3) résume les explications données vis-à-vis de l'interaction sociale vocale.

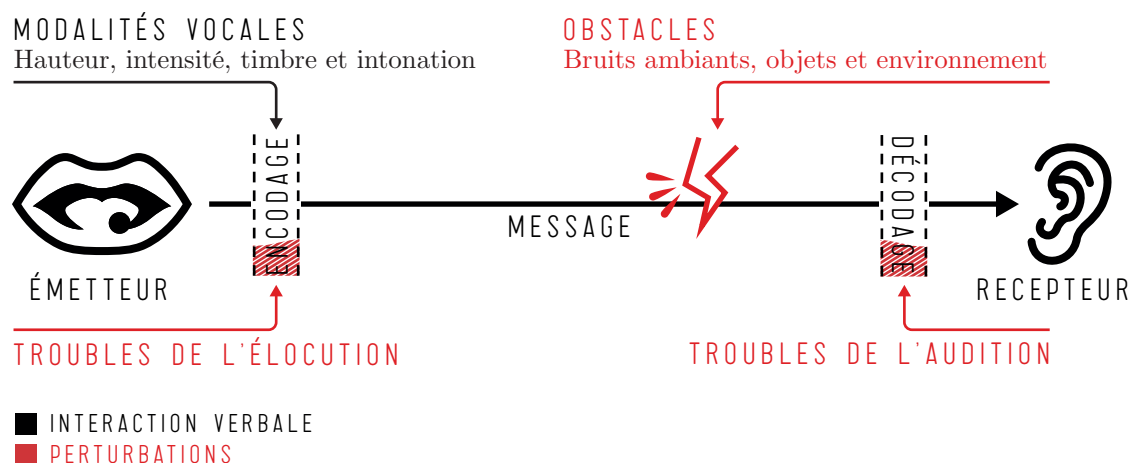


FIGURE 2.3 – Modulation et perturbations de l'interaction sociale vocale.

NOTA : Il est considéré que la hauteur, l'intensité, le timbre et l'intonation font partie des interactions sociales non verbales (Hinde, 1975). Nous pourrions alors résumer l'interaction sociale vocale aux mots et au sens, mais dans le cadre de ce cet état de l'art, il nous est apparu comme plus compréhensible de traiter ces facteurs en corrélation avec l'interaction sociale vocale.

2.1.1.2 L'interaction sociale verbale médiate

Les interactions sociales verbales médiates peuvent être produites de plusieurs manières. Il nous faut distinguer l'utilisation de la parole de celle du texte ainsi que l'utilisation de ces deux modes de manière synchrone (en temps réel) ou asynchrone.

2. AVC : abréviation de Accident Vasculaire Cérébral

3. Otospongiose : maladie générant une croissance anormale des os de l'oreille interne

L'exemple le plus commun de communication vocale synchrone est le téléphone ; depuis l'avènement d'internet de nombreuses applications permettant de communiquer vocalement au travers du réseau ont vu le jour. Toutefois, ces applications restent en soi une forme de téléphone augmenté, ils permettent d'appeler des contacts (un ou plusieurs) et de créer une conversation en temps réel. En plus de cela, il est possible de passer en mode vidéo ce qui permet de voir la personne et ainsi d'ajouter à l'interaction (sociale) vocale une partie du langage non verbal. L'utilisation de la communication vocale asynchrone (p.ex. laisser un message sur un répondeur) reste assez marginale dans le cadre de l'interaction sociale verbale. L'utilisation du MMS (Multimedia Messaging Service) pour la vidéo semble quant à elle plus réservée à montrer quelque chose, un instant, un paysage, une action qu'à transmettre un message demandant une réelle réponse, en résulte un usage proche de la mise en ligne d'une vidéo sur un réseau social.

Dans le cadre de la communication verbale textuelle, la scission synchrone / asynchrone se situe plus de côtés de l'usage que de la technologie elle même. Hormis dans le cadre d'une lettre envoyée par voie postale, le message est délivré instantanément (sauf problèmes techniques de connexion par exemple), que ce soit par courrier électronique, messagerie instantanée, forum ou SMS. C'est donc l'utilisateur qui va rendre la communication synchrone ou asynchrone selon son temps de réponse.

La communication asynchrone présente plusieurs contraintes incompatibles avec son utilisation dans le cadre des 3D-IVEs. L'utilisation du clavier devient beaucoup plus difficile lors du port d'un casque immersif, celui-ci occultant complètement le champ visuel de l'utilisateur. La communication textuelle comme décrite par Berge (1997) est plus réflexive qu'expressive ce qui est en contradiction avec les besoins d'interactivité et d'instantanéité qu'implique une immersion courte (de 10 à 30 min, temps moyen de l'immersion dans le cadre d'études empiriques) d'utilisateurs ne se connaissant pas de prime abord et devant interagir socialement. Ces raisons excluent l'utilisation d'une forme de communication asynchrone dans la conduite de nos travaux.

2.1.2 Les interactions sociales non-verbales

2.1.2.1 Qu'est ce que l'interaction sociale non-verbale ?

Cette partie sera assez brève et synthétique, car le travail présenté dans ce manuscrit repose sur les interactions sociales vocales. Néanmoins, il est important de comprendre le fonctionnement et le rôle des interactions sociales non verbales dans le processus de communication interpersonnel car, comme nous le verrons par la suite dans la section 2.5, interactions sociales verbales et non verbales sont liées et peuvent impacter les mêmes facteurs. Ce lien implique que, même si nous n'utilisons pas sciemment les interactions sociales non verbales dans nos travaux, nous devons les prendre en compte dans notre évaluation de l'expérience utilisateur.

L'interaction sociale non verbale ou paralangage⁴ décrite par Patterson *et al.* (2002) comme « la

4. Paralangage : ensemble des moyens de communication naturels qui ne font pas partie du système linguistique, mais qui peuvent accompagner et renforcer la parole (expressions du visage, gestes, etc. ...) (Larousse 2015).

transmission d'information et d'influence par les indices physiques et comportementaux d'un individu » peut, elle aussi, être découpée en plusieurs éléments qui, de la même manière que pour l'interaction sociale vocale, se combinent afin d'offrir de nombreuses possibilités. Ces facteurs font partie de ce que l'on nomme le comportement social (Argyle, 1988). Il faut distinguer deux types d'interaction sociale non verbale :

L'interaction non verbale intentionnelle : utiliser les gestes pour montrer quelque chose, accompagnant (ou non) une phrase du type « *Tiens ! Regarde, tu as vu ce nuage là-bas ?* », les gestes utilisés pour appuyer un propos en disant « *Sans rire, j'ai pêché un poisson gros comme ça !* », les gestes liés aux normes sociales comme agiter la main pour dire au revoir, opiner de la tête pour signifier son approbation, lever le poing en guise de menace, ouvrir les bras pour accueillir un proche, se racler la gorge pour attirer l'attention (l'utilisation d'onomatopées, de sifflements ou d'autres sons non verbaux fait partie du domaine de la communication non verbale (Argyle, 1988 ; Hinde, 1975)) ou encore par l'utilisation directe de notre environnement, comme frapper à une porte pour signifier notre présence.

L'interaction non verbale instinctive : les éléments instinctifs de l'interaction sociale non verbale sont très intéressants lors d'une communication interpersonnelle. En effet, conformément à leur nature involontaire, ils reflètent le véritable état d'esprit émotionnel de l'émetteur de ces signaux. Malheureusement, nous ne sommes pas tous égaux vis-à-vis du décodage de ces signaux, certaines personnes sont particulièrement douées de naissance. D'autres apprennent à les déchiffrer et certains tentent d'apprendre à les contrôler, ce qui est possible. Il faut donc tenir comme acquis que ce que nous définissons par instinctif ou involontaire peut être contrôlé par le travail, comme la colère peut l'être par la pratique de la méditation.

Afin d'avoir une vision aussi globale que possible de l'interaction sociale non verbale, nous allons expliquer brièvement chaque type de signal ou comportement non verbal.

La proxémie, ou distance interpersonnelle, introduite par l'anthropologue Andrew T.Hall (1966) représente la distance qui nous sépare de notre interlocuteur. On distingue quatre sphères proxémiques comportant chacune deux modes le « proche » et « l'éloigné » (figure 2.4) :

Cette approche de la distance interpersonnelle est très largement partagée par la communauté scientifique, mais l'étude étant datée de 1966, la proxémie a pu évoluer en même temps que notre société. Il faut également remarquer que, comme nous l'avons déjà évoqué, la culture a une grande importance dans les processus de communication. En l'occurrence, cette étude a été menée sur des personnes de la classe moyenne de la côte nord-est des États-Unis.

Les attitudes et les postures que nous adoptons de manière involontaire ou volontaire (quand nous y sommes préparés) peuvent être explicites. Nous parlons ici d'indications non verbales sur un état d'esprit, une émotion, le verbe n'est donc pas le moyen le plus adapté à leur description. Les figures A.1 et A.2 en annexe de ce document illustrent un échantillon d'attitudes et de postures communes et faciles à détecter pour peu que l'on y prête attention.

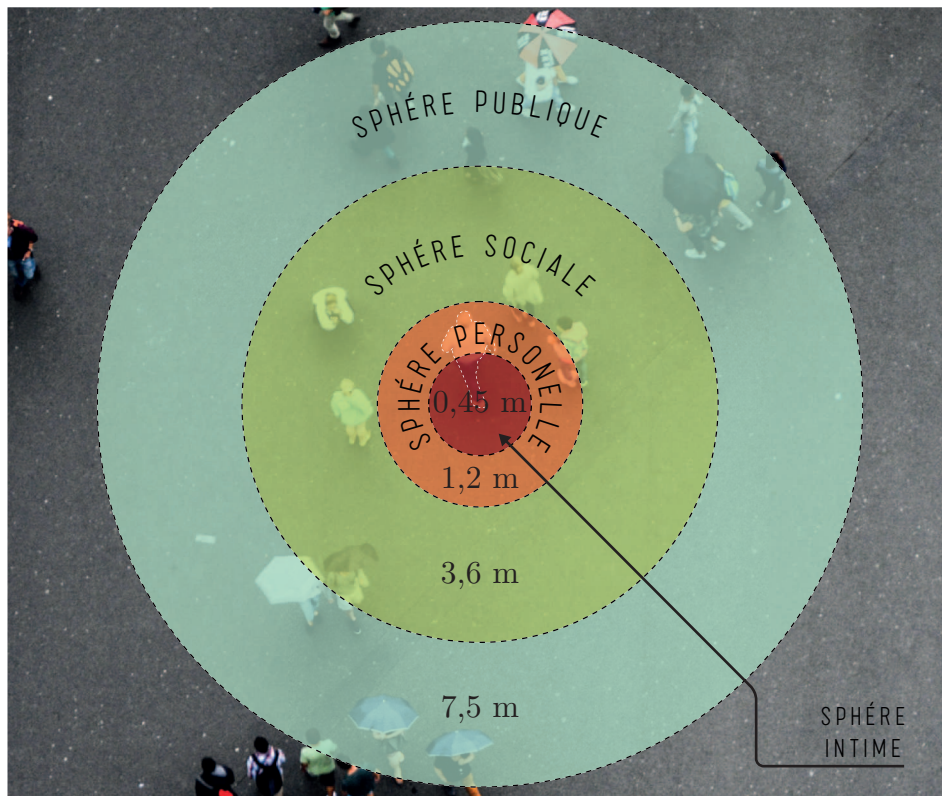


FIGURE 2.4 – Diagramme des sphères proxémiques d'après Edward T. Hall (1966).

Outre les gestes volontaires mentionnés précédemment, nous faisons aussi montre de gestes involontaires (p. ex. croiser les bras, pencher la tête,...) Ces gestes sont détaillés en figure A.3 en annexe de ce document.

Le toucher est la première voie de communication que nous utilisons avec nos semblables. Un nouveau-né n'étant en mesure ni de comprendre les paroles émises ni de voir les expressions sur le visage de ses parents ou de son entourage (la capacité de reconnaître le visage des parents s'acquière entre le second et le troisième mois), son seul moyen de percevoir le monde est le toucher. Une fois adulte, le toucher reste un moyen de communication important, très sensible : une tape dans le dos, une main sur l'épaule semble plus importante et efficace que la parole d'encouragement qu'elle remplace. Il faut aussi distinguer les auto contacts par exemple se triturer les mains ou se passer la main à l'arrière de la tête quand nous sommes gênés, se masser le front ou la racine du nez pour se remotiver, se recentrer (Pierson, 2003).

La tenue vestimentaire, les accessoires ou encore les tatouages sont considérés comme autant d'indications non verbales de la personnalité, de l'état d'esprit d'une personne, mais ils sont surtout une démonstration de son appartenance ou de son désir d'appartenance à un groupe social, une ethnie, une profession : « *Les directeurs de banque ne s'habillent pas comme des hippies* » [notre traduction] (Hinde, 1975).

Le visage est sans commune mesure la partie la plus expressive de notre anatomie. C'est celle qui livre le plus d'indications non verbales quand nous interagissons. Son observation permet la détection des six émotions fondamentales (Ekman *et al.*, 1972) grâce aux mouvements spécifiques de la bouche,

des yeux et des sourcils qu'elles génèrent (figure 2.5). Néanmoins, il faut être vigilant, par exemple, un sourire peut être feint, ce qui n'est pas le cas d'autres signaux incontrôlables tels que la moiteur du visage qui peut indiquer que l'interlocuteur est gêné, mis sous pression (stressé) par l'interaction en cours. Une réaction qui va souvent de pair avec un rougissement des joues, voire du visage entier. La pâleur indique un changement psychologique fort, voire un choc émotionnel. Le tonus musculaire au niveau de la mâchoire, serrer les dents est un signe flagrant de colère ou de stress contenu. Ce qui nous amène au « miroir de l'âme » : (expression qui semblerait ne pas être fondée uniquement sur la subjectivité comme l'arguent Larsson *et al.* (2007)) les yeux et surtout le regard. Le maintien du regard mutuel est un indice très pertinent de la confiance d'une personne en son discours et accroît grandement la portée de celui-ci (Argyle *et al.*, 1974). A contrario, un regard fuyant, fixé sur ses propres mains ou sur une partie du corps autre que le visage de l'interlocuteur signifie souvent un manque de confiance en soi ou dans le discours promulgué et donne donc des informations sur le crédit à apporter à ces propos et à celui qui les émet. Nous verrons un peu plus loin dans cet état de l'art l'impact produit par le regard mutuel dans le cadre des 3D-IVEs (voir section 2.2.3) afin d'éprouver la viabilité et la réalité de l'interaction sociale en environnement virtuel.



FIGURE 2.5 – Les six émotions fondamentales décrites par Ekman.

2.1.2.2 L'interaction sociale non-verbale médiée

La transmission de l'interaction sociale non verbale au travers d'un média n'est pas chose aisée. Dans le cadre d'un échange verbal synchrone, téléphonique par exemple, elle est drastiquement réduite et se résume aux quelques indices vus précédemment que sont l'intonation de la voix, les onomatopées et autres émissions vocales considérées comme non verbales. Lors d'un échange vidéo, la vue du visage de notre interlocuteur permet d'avoir accès aux expressions faciales et autres indications offertes par le visage. Néanmoins, il est encore fréquent lors d'un échange vidéo synchrone que le taux de rafraîchissement de la vidéo ne soit pas suffisant ou du moins pas suffisamment constant pour permettre une bonne lecture non verbale de l'interlocuteur.

Dans le cadre d'échanges textuels, nous ne sommes plus en mesure de capter des indications non verbales. Dans le but de pallier ce manque, Kalra et Karahalios (2005) ont proposé une amélioration des applications de messagerie instantanée nommée *TextTone* qui, par l'affichage du texte en différentes couleurs et tailles, tend à transmettre les émotions. L'autre solution très commune pour apporter du non-verbal dans les conversations textuelles est le « *smiley* » qui a donné lieu aux « *emoticons* » permettant à la fois d'exprimer une émotion, un geste ou un état physique ou mental. Les « *emoticons* » n'ont plus à prouver leur efficacité, le simple constat de leur adoption quasi globale démontre bien leur efficacité non seulement dans la transmission d'indications non verbales, mais aussi dans la ponctuation, le rythme et donc la structuration de la communication (Provine *et al.*, 2007). De plus, et contrairement à la solution proposée par Kalra et Karahalios (2005), leur simplicité d'utilisation est plus à même de traduire la fugacité d'une émotion.

Il est néanmoins une forme de communication non verbale que les 3D-IVEs permettent de transmettre aisément à partir du moment où les personnes immergées possèdent toutes un avatar. En effet, la possibilité de se voir apporte la conscience de la présence de l'autre dans l'environnement, et permet par l'observation de moduler son comportement en fonction des autres protagonistes. Cela fait échos à un comportement nommé la translucence sociale.

2.1.2.3 La translucence sociale, comportement non verbal

Le terme translucence, extrait de translucidité, évoque la possibilité de voir au travers de quelque chose. Appliquée au comportement social et comme décrit par Erickson et Kellogg (2000), elle représente notre capacité à comprendre et à réagir en fonction du comportement des autres. Prenons deux exemples proposés par Erickson : « *Vous faites du shopping afin de trouver du vin pour le dîner. Alors que vous naviguez entre les rayonnages vous entendez un discret « Aha ! » et voyez un autre client prendre deux bouteilles sur un rayon presque vide. Vous prenez l'une de ces mêmes bouteilles pour vous.* » [notre traduction].

Le second exemple permet d'expliquer plus clairement les trois composantes de la translucence, la **visibilité**, la **conscience** et la **responsabilité** : Erickson décrit l'une des portes du bâtiment dans lequel il travaille, cette porte s'ouvre rapidement, elle présente un risque si une personne se présente de chaque côté de la porte que l'une des deux reçoive la porte en plein visage ! La pose d'un hublot sur cette porte permet de la transformer en un système translucent, et ce pour trois raisons :

- Le hublot rend **visibles** les informations socialement importantes, les autres. En tant qu'humain nous sommes plus réactifs vis-à-vis du visage ou de la présence d'un de nos semblables qu'aux indications écrites sur un quelconque panneau.
- Le hublot permet d'avoir **conscience** de l'autre, « *J'ouvre la porte doucement, car je sais que quelqu'un se trouve de l'autre côté* ». Ce comportement fait appel aux normes sociales qui régissent notre vie avec les autres.
- La troisième raison de l'efficacité du hublot est plus subtile : admettons que je ne me préoccupe pas de blesser les autres. Néanmoins, je vais ouvrir la porte doucement parce que, *je sais que*

vous savez que je sais que vous êtes là, et je vais donc être tenu **responsable** de mes actions.

Visibilité, conscience et responsabilité, ces trois composantes de la translucence interviennent chaque jour dans la régulation de nos comportements sociaux, nous entrons dès la croisée d'un regard dans cet état d'adaptabilité, et ce sans avoir à prononcer le moindre mot. Ce comportement présente un intérêt certain pour nos travaux, car, quoi que nous entreprenions afin d'évaluer l'impact de l'interaction sociale vocale sur l'expérience utilisateur, cette évaluation aura pour contexte un 3D-IVE partagés par plusieurs utilisateurs — à n'en pas douter incarnés par des avatars — par cette mise en présence de plusieurs avatars le phénomène de translucence sera à même d'apparaître. Par son observation, nous serons en mesure d'obtenir des indications sur le traitement du phénomène par l'utilisateur. Par traitement nous entendons ici que la translucence pourrait avoir la priorité sur la communication vocale, être perçue comme un complément ou un palliatif à des difficultés ou à l'absence de communication vocale.

2.2 L'interaction sociale dans les environnements virtuels

Nous n'avons jusqu'à présent que peu parlé de réalité virtuelle, il était primordial de donner au lecteur les notions importantes inhérentes aux interactions sociales avant de pouvoir nous pencher sur leur mise en œuvre au travers du médium qu'est la réalité virtuelle. L'interaction sociale dans le monde réel est multimodale, comme nous l'avons déjà évoqué, et il apparaît comme évident qu'il en va de même dans les environnements virtuels. Cette multimodalité est même exacerbée dans le contexte des 3D-IVEs, d'autres facteurs entrant en ligne de compte. Le premier est la perception des représentations d'êtres intelligents communément appelées avatars⁵ ou agents virtuels⁶. Il est nécessaire de comprendre comment ils sont perçus par l'utilisateur et si ces perceptions sont différentes en fonctions du type de représentation. Le réalisme visuel, comportemental ainsi que le niveau d'anthropomorphisme jouent un grand rôle dans cette perception et présentent des interactions complexes difficiles à interpréter (Bailenson *et al.*, 2005b). Il est aisé de se dire que l'utilisation de représentations très réalistes à tous niveaux engendrerait un même niveau de réalisme quant aux interactions sociales produites au travers d'elles. Malheureusement, outre les difficultés techniques à faire « vivre » en temps réel un avatar (plus les modélisations, textures et animations sont complexes, plus les ressources matérielles demandées sont élevées), un réalisme trop poussé peut mener au rejet pur et simple de la représentation virtuelle par l'utilisateur. Ceci, jugulant l'occurrence d'interactions sociales naturelles, de la même manière que dans le monde réel il nous est difficile d'entrer en interactions avec un être dont la vision nous perturbe.

5. Avatar : représentation virtuelle d'un être intelligent contrôlé par un humain.

6. Agent : représentation virtuelle d'un être intelligent contrôlé par un ordinateur.

2.2.1 Agents et avatars, les outils de l'interaction sociale en environnement virtuel

Le choix de l'utilisation d'avatars (plusieurs utilisateurs) ou d'agents (un utilisateur et un ou plusieurs agents) est primordial. En effet, il est souhaitable lors de l'immersion dans un système de réalité virtuelle multiutilisateur que de nombreuses interactions vocales puissent avoir lieu. Or agents et avatars ne sont pas égaux vis-à-vis des réactions qu'ils génèrent chez l'utilisateur. En effet, le simple fait de dire à l'utilisateur qu'il va se trouver en présence d'un agent ou d'un avatar change son comportement vis-à-vis de ce dernier. Afin d'éprouver ce postulat, Guadagno *et al.* (2007), de Melo *et al.* (2013), dans deux études différentes, ont exposé des utilisateurs à des agents, mais ont induit la moitié de leurs participants en erreur en leur disant qu'ils allaient être mis en présence d'avatars. Les résultats de l'évaluation des agents montrent que les utilisateurs qui pensaient être en présence d'avatars ont donné une évaluation plus positive vis-à-vis de la confiance accordée à l'agent et du réalisme de leur comportement (Guadagno *et al.*, 2007). Un effet notable a aussi été relevé sur les intentions de coopération et sur l'engagement social vis-à-vis de la représentation en présence (de Melo *et al.*, 2013). Garau *et al.* (2005) ont eux mis en évidence un effet sur l'intrusion dans les sphères proxémiques avec un meilleur respect des distances « de courtoisie » pour les utilisateurs pensant être en présence d'avatars. Ces résultats confirment les travaux précédents de Bailenson *et al.* (2001b) sur la proxémie et le regard mutuel, les utilisateurs sachant être en présence d'un agent n'hésitant pas à leur passer au travers par exemple. Il apparaît donc que si l'utilisateur est conscient d'être en présence d'un agent, il le considérera comme tel, comme un être non doté de sentiments et n'aura pour lui que peu d'égards. Ce constat est à modérer, car un élément peut totalement changer la perception d'un agent par l'utilisateur, son bon vouloir, si l'utilisateur décide de personnifier un agent. Ce dernier ne sera certes pas perçu comme humain, mais l'utilisateur fera montre d'un comportement plus affectif, moins pragmatique vis-à-vis de lui. Cette personnification est difficile à étudier de manière empirique dans le cadre des 3D-IVEs, on la constate plus souvent dans les jeux vidéos où, à force d'interagir avec un personnage non joueur (PNJ), l'utilisateur va finir par éprouver des sentiments à son égard que ce soit de l'amitié, de la haine, voire même de l'attirance. L'une des premières apparitions mémorables du sentiment d'empathie généré par un contenu vidéoludique remonte à 1997 avec la sortie du jeu vidéo « *Final Fantasy VII* » où l'on voit mourir l'un des personnages principaux sous les coups de la némésis du héros. Cette scène et le ressentiment à l'égard de « *Sephiroth* », l'assassin, sont encore présents dans la mémoire de nombreux joueurs et a contribué à élever ce jeu au rang d'icône du jeu vidéo. Cet exemple démontre le pouvoir émotionnel détenu par un personnage non joueur bien qu'en 1997 le niveau de réalisme visuel était assez bas.

En plus du fait qu'ils soient de bons vecteurs d'émotion, les agents ou PNJ présentent certains intérêts pour les applications impliquant l'interaction sociale. En effet, leur utilisation amène une plus grande propension de l'utilisateur à dévoiler des informations personnelles (Bailenson *et al.*, 2006) ou entraîne une baisse de l'anxiété sociale ressentie (Kwon *et al.*, 2013). En raison de sa nature non humaine, l'agent n'est pas en mesure d'émettre un jugement, ce qui rassure l'utilisateur. Ces spécificités rendent possible leur utilisation par exemple dans le cadre de l'amélioration des compétences sociales (Hoque *et al.*, 2013).

Ces études nous montrent que, pour une interaction sociale convaincante, il est préférable que l'utilisateur ait la conviction qu'il est impliqué avec un autre humain. Pour autant, cette conviction n'est pas suffisante pour l'inciter à interagir socialement de la même manière qu'il le ferait en réalité, une des conditions est qu'il doit se sentir à l'aise avec la ou les représentations qui lui sont présentées. Bien que totalement subjective, cette sensation est très influencée par le réalisme visuel, comportemental ainsi que le niveau d'anthropomorphisme de la représentation jouant le rôle d'interlocuteur et surtout par la cohérence entre ces paramètres, comme nous allons le voir dans la section suivante.

2.2.2 La présence, la co-présence et la présence-sociale

Avant de pousser plus loin notre étude des interactions sociales en 3D-IVE au travers du phénomène de présence sociale, il nous faut comprendre le concept égocentré de présence dans l'environnement virtuel. Comme nous le verrons, la présence égocentrée et la présence sociale sont des phénomènes assez proches comportant des facteurs en commun. La différence se situe dans le fait que, pour la présence égocentrée, nous parlons plutôt du sentiment de présence généré par l'environnement virtuel et les interactions qu'il propose. Tandis que, pour la présence sociale, nous parlons du sentiment de présence généré par les avatars ou agents présents dans l'environnement.

2.2.2.1 La présence, concept fondamental

D'une manière générale, nous vivons la présence à chaque instant et pas seulement lors de l'immersion dans un environnement virtuel. Au moment de l'écriture de ces mots, je suis présent devant un ordinateur : la sensation du clavier sous mes doigts, mon environnement sonore et visuel me permettent de me sentir présent dans la réalité qui est la mienne, je suis dans ce que nous pouvons appeler la réalité primaire, celle que nous considérons comme la « vraie réalité ». Une réalité primaire de laquelle nous sommes capables de nous dissocier, lors de la lecture d'un roman, notre esprit va créer des images, peut-être allons-nous nous mettre à la place des protagonistes, adopter leur vision des choses, voir leur environnement, les personnes avec qui ils interagissent, ressentir leurs émotions en nous sentant présents à leur place ou à leur côté. Néanmoins, nous restons conscients du monde qui nous entoure, oscillant entre la réalité transmise par l'ouvrage (réalité évoquée) et la réalité primaire. Si quelque chose, un bruit, un changement dans la lumière ambiante, attire notre attention, ce sentiment de présence a de fortes chances d'être brisé, les représentations mentales produites à cet instant s'évaporant sous le coup du retour à la réalité primaire. Ce retour peut se faire sciemment ou être causé par un stimulus extérieur, phénomène alors appelé « *Break In Presence* » (BIP) (Slater et Steed, 2000). Cette expérience subjective qu'est la présence apparaît bien évidemment lors de l'immersion dans un 3D-IVE mais aussi lors du visionnage d'un film, d'une partie de jeu vidéo, lorsqu'un ami vous raconte ses vacances à l'étranger, lorsque nous sommes perdus dans nos pensées ou même lorsque nous rêvons. Hormis dans le cadre de la réalité primaire, la présence peut être définie comme : « *l'illusion forte d'être dans un lieu en dépit du fait de savoir que vous n'y êtes pas* » [notre traduction] (Slater, 2009). Ces oscillations entre réalité primaire, évoquée et médiate ont été décrites par Pillai *et al.* (2013)

comme le continuum de la présence et illustrées par sa « Carte de la réalité-présence » présentée en figure 2.6. Cette carte illustre parfaitement le fait que la présence est intimement liée à l'attention que nous prêtons aux autres réalités comme le souligne Draper *et al.* (1999) à propos des environnements virtuels : « *Plus l'utilisateur prête attention aux stimuli présentés par les affichages, plus l'identification à l'environnement virtuel et le sentiment de téléprésence seront grands* » [notre traduction]. Notons ici l'utilisation du terme téléprésence au lieu de présence, autre terme utilisé pour identifier cette expérience subjective.

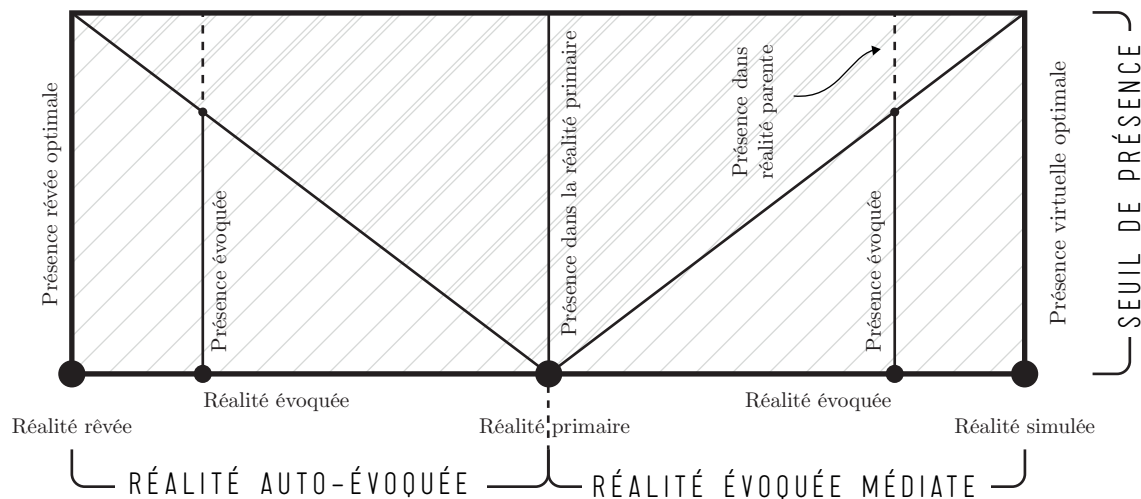


FIGURE 2.6 – Carte de la présence d'après Pillai *et al.* (2013).

Effectivement, le concept de présence, fut et reste encore aujourd'hui un terrain d'investigation très prolifique. Un grand nombre de chercheurs issus de diverses disciplines, communication (Kim et Biocca, 1997 ; Lombard et Ditton, 1997), informatique (Slater *et al.*, 1994), ingénierie industrielle, psychologie, philosophie ou éducation a mené à une multiplication des définitions de la présence. Dans le cadre de nos travaux nous préférons l'utilisation du terme « présence virtuelle » (Sheridan, 1992) pour une question de précision, de localisation dans un contexte technologique précis. Le terme présence à lui seul est général comme nous l'avons vu au travers du continuum de la présence de Pillai *et al.* (2013). Le terme téléprésence est, quant à lui, proposé en premier par Minsky (1980) en faisant référence au sentiment que pouvait ressentir une personne impliquée dans une tâche à distance par un système de téléopération (chirurgie à distance, travaux en environnements hostiles, télétravail), repris par la suite comme définissant le sentiment de présence dans un environnement par le biais d'un système de communication (Steuer, 1992) (ledit environnement pouvant être distant physiquement ou temporellement ou ne pas exister comme c'est le cas des 3D-IVEs), est là encore trop large. Si nous considérons la photo d'un village prise au siècle dernier, cette dernière est un médium, par ce qu'elle présente elle nous permet de nous transposer dans le monde qu'elle reproduit et génère donc un sentiment de « téléprésence ». Dans le cadre de nos travaux, le terme de présence virtuelle est plus adéquat, car représentant le sentiment d'être présent dans un environnement distant (virtuellement) au travers d'un média (l'environnement virtuel) généré par ordinateur.

De même que pour les termes employés pour désigner la présence, plusieurs analyses des facteurs impliqués dans le phénomène existent (Slater *et al.*, 2009 ; Ijsselstein et Riva, 2003 ; Slater, 2002 ;

Schubert *et al.*, 2001 ; Durlach et Slater, 2000 ; Schubert *et al.*, 1999 ; Lombard et Ditton, 1997 ; Steuer, 1992 ; Heeter, 1992). Nous avons voulu, en accord avec notre contexte d'étude et avec la littérature, isoler les facteurs nous semblant les plus pertinents et, comme nous le verrons dans le chapitre *méthodologie* (page 61), les plus aisément mesurables.

Les facteurs présentés ici sont à séparer en deux types, ceux afférents à (1) l'immersion psychologique ; représentant les réponses cognitives de l'utilisateur aux stimuli envoyés par l'environnement médiat, et à (2) l'immersion perceptuelle ; représentant le flux continu desdits stimuli (Lombard et Ditton, 1997 ; Biocca et Delaney, 1995).

1. L'immersion psychologique :

- **Presence spatiale :** d'après Schubert *et al.* (2001), elle est la relation entre l'environnement virtuel comme espace et le corps de l'utilisateur. Un facteur intimement lié non seulement au réalisme et à la cohérence de l'environnement mais aussi à l'incarnation virtuelle de l'utilisateur désignant la représentation du corps de ce dernier dans l'environnement virtuel. L'incarnation du corps est considérée comme un facteur déterminant de la présence virtuelle surtout dans le cas de l'utilisation d'un casque immersif (França et Soares, 2015 ; Haans et Ijsselsteijn, 2012 ; Kiltner *et al.*, 2012 ; Schultze, 2010).
- **Le niveau d'intérêt :** est l'intérêt préalable (envie par anticipation de l'utilisation) que porte l'utilisateur au 3D-IVE mais surtout l'intérêt que saura susciter ce dernier, ce facteur est très fortement influencé par la narration et l'interactivité proposées lors de l'immersion.
- **L'amusement :** est lié au fait que plus l'utilisateur éprouvera de plaisir, plus sa présence virtuelle augmentera. En effet, lorsque que l'on passe un bon moment, le temps semble se contracter, passer plus vite à l'inverse des moments où nous nous ennuyons (Danckert et Allman, 2005). Cette illusion produite par notre cerveau est considérée comme un bon indice du niveau de présence virtuelle ressentie (Ijsselsteijn *et al.*, 2001). Néanmoins, créer du plaisir n'est pas chose aisée, dans le contexte des jeux vidéo par exemple le « *GameFlow model* » de Sweetser et Wyeth (2005) montre que pour ressentir du plaisir, le joueur doit être concentré et engagé dans le déroulement de l'action et que l'apparition de ces deux phénomènes demande un équilibre entre le défi proposé et l'habileté du joueur à relever ledit défi. Dans le même contexte, il est aussi démontré que l'aspect social, le fait de pouvoir communiquer avec d'autres joueurs accroît de manière significative l'amusement (Cole et Griffiths, 2007).
- **La concentration :** elle le même pouvoir sur notre perception du temps que l'amusement, bien qu'une réciprocité d'effet existe entre la concentration et l'amusement, il est tout à fait possible d'être fortement concentré sans pour autant s'amuser. Comme explicité par la réappropriation du modèle de « *Flow* » dédié au jeu vidéo de Sweetser et Wyeth (2005), pour susciter de la concentration, les actions et tâches demandées lors de l'immersion dans un 3D-IVE ne doivent être ni trop simples ni trop compliquées, il est important que l'utilisateur soit convaincu qu'il est capable de relever le défi sans quoi il en résultera un désengagement,

un désintérêt et en conséquence une chute de la présence virtuelle.

- **L'engagement** : décrit par Witmer et Singer (1998, p.227) comme « *L'état psychologique vécu comme une conséquence de la concentration de son énergie et de son attention sur un ensemble cohérent de stimuli ou sur des activités et événements significativement corrélés à ces derniers.* » [Notre traduction]. L'engagement est donc lui aussi facteur du défi proposé mais aussi de la récompense à obtenir ainsi que de la narration.
- **L'émotion** : les émotions, comme par exemple la peur, ont démontré leur impact positif sur la présence virtuelle (Claudio *et al.*, 2015 ; Meehan *et al.*, 2002). Il est intéressant de constater qu'outre le fait qu'il est possible d'induire des émotions au travers des 3D-IVEs (Rodríguez *et al.*, 2013 ; Riva *et al.*, 2007), les études citées ici utilisent la physiologie humaine (relevés ECG⁷, GSR⁸ ou EEG⁹) pour attester de la réalité de l'occurrence des émotions.

2. **L'immersion perceptuelle** : ou simplement immersion, est le pendant technique de la présence. Nous retrouverons dans la partie de ce document réservée à l'expérience utilisateur (section 2.3) cette dichotomie entre psychologie et technique. Elle définit la propension des matériels mis en place pour l'expérience en 3D-IVE à supplanter la réalité primaire par des stimuli générés artificiellement (Slater et Wilbur, 1997). Tout comme pour la présence virtuelle, l'immersion se découpe en plusieurs facteurs.

- **La conscience des interfaces** : liée aux matériels utilisés plus qu'au 3D-IVE lui-même, la conscience de l'interface pose la question des interférences possibles causées par l'équipement de l'utilisateur (Durlach et Held, 1992). Par exemple, le poids d'un casque immersif est-il gênant, le taux de rafraîchissement entre l'image et les capteurs de mouvement est-il suffisant (gyroscopes) ou cause-t-il un malaise ? « *Avoir trop de latence dans nos systèmes est la meilleure façon de rendre les gens malades* » [notre traduction] (Ludwig, 2013).
- **La vivacité sensorielle** : traduit de l'anglais « *Vividness* », elle fait référence à la capacité des technologies à proposer une expérience sensorielle riche, sollicitant plusieurs sens. Comme expliqué par Witmer et Singer (1998) ou Durlach et Held (1992), dans une expérience en 3D-IVE, le visuel est très sollicité, car c'est par ce sens que passe la plupart des informations. Néanmoins, l'utilisation des autres sens, l'ouïe, le toucher, a un apport très positif sur la présence virtuelle : plus le nombre de sens impliqués est important, plus la présence virtuelle tend à augmenter. Notons ici que nous ne faisons pas cas de la ressemblance entre le monde réel et virtuel mais uniquement de l'implication des sens.
- **L'interactivité** : plus l'environnement virtuel sera interactif, c'est-à-dire plus il sera possible de le modifier (p.ex. en ouvrant une porte, une fenêtre, en déplaçant des objets) ou en étant en mesure de détruire ou de construire quelque chose, plus le sentiment de présence virtuelle sera accru (Sheridan, 1992), d'autant plus si ces actions ont un impact non trivial sur l'expérience vécue. Cette notion fait écho à la possibilité du choix : plus l'utilisateur a

7. ECG : abréviation d' Électrocardiogramme, mesure de l'activité cardiaque.

8. GSR : abréviation de l'anglais *Galvanic Skin Response*, mesure des micro variations de la conductance de la peau.

9. EEG : abréviation d' Électroencéphalogramme, mesure de l'activité électrique du cerveau.

la conviction qu'il détermine ce qu'il se passe, plus il se sentira engagé et par corrélation présent (McMahan, 2003). L'interactivité comprend l'interaction sociale, particulièrement dans le cas des agents : un agent qui ne réagit pas lorsque l'utilisateur tente d'entrer en contact avec lui n'aura que très peu d'impact positif sur la présence virtuelle. Il se pourrait même ce que cela amène à une réduction du sentiment de présence virtuelle comme nous le verrons au travers du phénomène de la « *Uncanny Valley* » (voir section 2.2.2.3).

- **Le degré de contrôle et son immédiateté :** dans quelle mesure l'utilisateur a-t-il la sensation d'interagir avec l'environnement virtuel directement plutôt qu'avec un appareil (clavier, manette, bras haptique, Leap Motion¹⁰) qui va agir sur l'environnement virtuel ? Comme nous l'avons vu, plus le lien avec la réalité primaire est ténu, plus la présence virtuelle est bonne. Être conscient d'utiliser un objet extérieur au 3D-IVE entretient ce lien, il est donc néfaste à l'obtention d'un haut niveau de présence virtuelle. Il en va de même pour l'immédiateté de l'interaction : si l'action entreprise dans la réalité primaire n'est pas répercutée immédiatement dans le monde virtuel, l'utilisateur sera confronté à un phénomène qui ne lui semblera pas naturel et tendra à réduire son sentiment de présence virtuelle (Durlach et Held, 1992).
- **Richesse du contenu :** condition *sine qua non* à l'interactivité, le 3D-IVE doit contenir des objets, des personnages avec lesquels il sera possible de proposer des interactions par la suite. Plus importante encore est la richesse du contenu narratif, le 3D-IVE doit raconter une histoire pour porter l'utilisateur, proposer des objectifs explicites favorisant l'engagement. Une ligne narrative non linéaire, influencée par les actes de l'utilisateur est aussi très favorable à l'engagement et donc au sentiment de présence virtuelle (Schubert et Friedmann, 1999).
- **La qualité visuelle :** nous préférons le terme de qualité visuelle à celui de réalisme visuel, en effet le terme réalisme visuel est utilisé par la quasi-totalité des chercheurs cités dans cette section et pourtant la plupart s'accordent à dire qu'un réalisme élevé n'est pas forcément nécessaire pour atteindre un bon niveau de présence virtuelle. Des applications présentant des univers totalement oniriques sont loin d'être visuellement réalistes (voir par exemple le court métrage « *Senza Peso* »¹¹) néanmoins, elles proposent une grande qualité visuelle, par l'utilisation de modèles 3D et de textures haute résolution. Qualité visuelle qui, couplée à une narration travaillée, permet d'obtenir un fort sentiment de présence virtuelle. Il en va de même pour les représentations sociales (propos que nous détaillerons dans la section 2.2.2.3) le réalisme visuel n'est pas une nécessité absolue.
- **Le réalisme comportemental :** a contrario du réalisme visuel, le réalisme comportemental est extrêmement important, que ce soit pour l'environnement virtuel ou pour les représentations le peuplant. Un objet lâché dans le vide est en toute logique soumis à la gravité et doit donc tomber même si la gravité n'est ici que virtuelle. Si toutefois ledit objet ne tombait pas l'utilisateur trouverait cela bizarre sauf si par la narration ou la mise

10. Leap Motion : interface d'interaction sans contact. Voir : www.leapmotion.com

11. Voir : senzapeso.com

en contexte, l'objet est identifié comme n'étant pas soumis à la gravité. Plus que tout, l'important est la cohérence entre les éléments cités dans cette liste de facteurs.

- **La cohérence :** afin de générer un sentiment de présence virtuelle, il est nécessaire d'avoir une bonne cohérence entre les stimuli générés par le 3D-IVE et sa narration mais aussi dans une certaine mesure avec ce que nous pouvons voir dans la réalité primaire. Cette cohérence doit permettre à l'utilisateur de se représenter mentalement les conséquences des actions qu'il va entreprendre et ces conséquences ne doivent pas se trouver en opposition avec la représentation mentale que s'est faite l'utilisateur (Slater, 2009 ; McMahan, 2003 ; Schubert *et al.*, 2002 ; Steuer, 1992 ; Durlach et Held, 1992). Reprenons l'exemple de l'objet et de la gravité, nous avons placé nos utilisateurs dans un contexte de gravité zéro (une station spatiale). Lorsque l'utilisateur va lâcher un objet, conscient du contexte, il va s'attendre à ce que l'objet flotte vis-à-vis de la représentation mentale qu'il peut avoir de la physique s'appliquant en gravité zéro. Mais si l'objet tombe lourdement et rebondit en tout sens, il y a fort à parier que non seulement l'utilisateur va trouver cela étrange mais qu'il risque même de « sortir » de l'application (BIP) afin de demander à une personne présente si cela n'est pas un bug dans le logiciel.

Ce découpage en facteurs de la présence virtuelle va nous être d'une grande aide quant à l'évaluation de l'impact des interactions sociales vocales sur l'expérience utilisateur en 3D-IVE. La présence virtuelle est comme nous le verrons par la suite (voir section 2.3) un élément central de l'expérience utilisateur en 3D-IVE. Cette séparation permet une évaluation compartimentée élargissant les possibilités de mesures et permettant par leur synthèse une évaluation plus robuste du sentiment de présence virtuelle.

2.2.2.2 La co-présence, et la présence sociale « Being there with others »

Nous avons à présent une bonne vision de ce qu'est le sentiment de présence et plus particulièrement de présence virtuelle. Dans notre contexte d'utilisation de l'interaction sociale en 3D-IVE, nous devons porter notre attention sur deux autres phénomènes dérivés de la présence : la co-présence et la présence sociale (Biocca et Harms, 2002), sentiments d'être présent avec les autres : « *Being there with others* ».

Tout comme pour la présence virtuelle, nous sommes ici confrontés à plusieurs terminologies vis-à-vis de la co-présence et de la présence sociale. Il est commun de voir les deux termes utilisés pour parler de la même chose. Ainsi, on voit très souvent apparaître dans la littérature la terminologie suivante : « *co-présence (aussi appelé présence sociale)* ». Or, une différence entre co-présence et présence sociale (sur laquelle nous allons revenir) a pourtant été établie par Biocca et Harms (2002), mais il semble que malgré la citation récurrente du document proposant cette différenciation, les deux termes continuent à être assimilés l'un à l'autre selon le contexte ou les auteurs.

Heeter (1992) définit la présence sociale comme l'existence dans un environnement virtuel d'autres êtres intelligents (vivants ou synthétiques, respectivement, avatars ou agents) que l'utilisateur, ceux-ci

semblant en mesure de réagir aux interactions de l'utilisateur. Pour Heeter (1992), la présence sociale est un élément clé du sentiment de présence virtuelle : « *Le principe de la présence sociale consiste simplement à dire que si d'autres personnes sont dans le monde virtuel, ceci est une preuve de plus que ce monde existe* » [notre traduction]. Schubert *et al.* (2002) ont éprouvé l'importance de ce sentiment de présence de l'autre par la conduite d'une expérimentation exposant les participants à la présence illusoire d'un autre être doté de sentiments. L'illusion produite par l'ajout de bruit de pas dans un environnement virtuel a montré un impact positif sur le sentiment de présence virtuelle des participants. Lee (2004) étend le phénomène au-delà des agents et avatars en y incluant les objets virtuels faisant montre d'un caractère social (p.ex., un arbre, une voiture,...) Ces deux définitions ne sont porteuses que de la notion de présence spatiale de l'autre. On distingue alors une définition proposée par Nowak et Biocca (2003) expliquant la co-présence comme étant le sentiment d'une connexion avec l'esprit de l'autre, connexion intervenant dès lors que nous communiquons. Par nos tentatives d'influencer notre interlocuteur, nous nous représentons ses réactions possibles afin d'orienter l'interaction sociale dans le sens que nous désirons. Le terme de co-présence est ici utilisé pour parler sans distinction du fait d'être spatialement dans le même environnement (virtuel) que d'autres êtres intelligents et du fait d'être connecté et de communiquer avec ces derniers. Conscients de ces confusions communes et afin de les écarter, revenons sur la taxonomie de la présence sociale proposée par Biocca et Harms (2002). Elle définit les niveaux et dimensions de la présence sociale, autrement dit les facteurs y menant (voir figure 2.7).

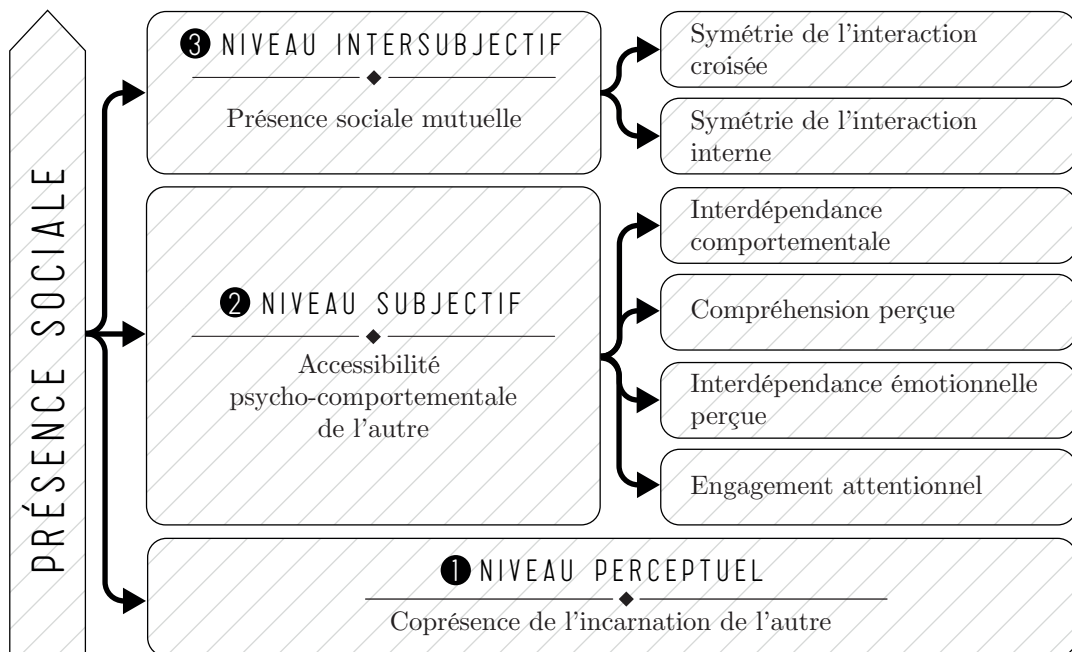


FIGURE 2.7 – Niveaux et dimensions de la présence sociale de Biocca et Harms (2002).

La figure 2.7 décrit bien la différence entre présence sociale et co-présence. La co-présence étant un des facteurs permettant d'atteindre le sentiment de présence sociale. C'est en nous basant sur ce schéma que nous allons à présent lister les facteurs de la présence sociale comme nous l'avons précédemment fait pour la présence virtuelle.

- **Niveau perceptuel, le sentiment de co-présence :** Goffman (1963) dans son ouvrage « *Behavior in Public Places* » définit la co-présence comme suit : « *Les personnes doivent sentir qu'elles sont suffisamment proches pour être perçues quoi qu'elles soient en train de faire, y compris le fait de sentir la présence de l'autre et suffisamment proche pour percevoir une réciprocité dans ce sentiment de perception de l'autre. Dans notre société occidentale, on s'attend à ce que ces conditions soient remplies dès lors que nous nous trouvons dans une pièce, et ce pour toutes les personnes présentes dans cette même pièce* » [notre traduction] (Goffman, 1963, p. 17). Comme le soulignent Biocca et Harms (2002), Goffman fait ici référence à une pièce physique, dans le monde réel. Néanmoins, ce constat s'applique parfaitement aux 3D-IVEs puisque ces derniers sont une sorte de pièce virtuelle dans laquelle plusieurs personnes peuvent se trouver. Afin de stimuler ce sentiment de co-présence, il est nécessaire que les acteurs sociaux en présence dans le 3D-IVE soient définis selon certaines caractéristiques anthropomorphiques visuelles ou comportementales sur lesquelles nous reviendrons par la suite.
- **Niveau subjectif :** les facteurs subjectifs de la présence sociale sont basés sur un concept issu du domaine des sciences cognitives et très largement partagé par la communauté scientifique : la théorie de l'esprit (pour une revue succincte, voir Frith et Frith (2005)). Cette dernière figure du processus cognitif permettant la reconnaissance d'un état mental pour-soi même ou pour un autre individu. Elle correspond aux processus conscients ou inconscients qu'entreprend notre esprit chaque fois que nous interagissons ou tentons d'interagir avec un être intelligent afin de nous représenter son état mental, émotionnel ou ses intentions afin de nous y adapter, ceci nous aidant à atteindre l'objectif de l'interaction sociale entreprise (voir section 2.1).
 - **Engagement attentionnel :** représente l'attention allouée à un autre être sentent lors de l'immersion dans un 3D-IVEs. Cette attention se porte sur les interactions vocales, mais aussi sur la communication non verbale, sur les actions, le comportement. C'est à travers cet effort mental que nous tentons de percevoir l'état mental, les intentions de la personne en co-présence. Glas et Pelachaud (2015) ont récemment produit une revue du concept d'engagement vis-à-vis des agents virtuels. Outre la définition du concept, elles définissent les comportements associés ou suscitant de l'engagement comme le fait de collaborer sur une tâche. Cela fait écho à l'engagement dans le contexte de la présence virtuelle, le besoin d'une narration ou d'objectifs explicites.
 - **Interdépendance émotionnelle perçue :** être en mesure de percevoir et de décrypter les émotions ou simplement l'humeur d'une personne nous permet, comme nous en avons déjà parlé, d'adapter l'interaction sociale en cours ou de décider si nous allons interagir et quel rituel sera le plus adapté. Cette compréhension des émotions mène à l'empathie qui, par la création d'un lien, d'une connexion d'un individu à un autre, va accroître dans le contexte des 3D-IVEs les sentiments de présence sociale et virtuelle des protagonistes. Bien sûr, nous ne sommes pas tous égaux devant la perception des émotions des autres et du sentiment emphatique, certaines personnes ne sont que très peu douées d'empathie, d'autres en revanche le sont énormément et sont très facilement « contaminées » par les émotions des autres (Hatfield *et al.*, 1993). La présence de contenu émotionnel a, comme

nous l'avons vu, un fort impact sur le sentiment de présence virtuelle, il est donc cohérent que cet impact soit encore plus important dans le cadre de la présence sociale puisque, très souvent, une interaction sociale avec un être intelligent implique les émotions.

- **Compréhension perçue :** nous ne parlons pas ici de la compréhension au niveau de l'utilisation d'un même langage ou de mêmes codes sociaux. Ici, la compréhension perçue réside dans le sentiment qu'au travers des interactions sociales (vocales ou non verbales), les protagonistes pensent se comprendre, du moins comprendre les intentions, les motivations et les pensées de leur interlocuteur. Biocca et Harms (2002) insistent sur les termes « *pensent se comprendre* », car il y a peu de moyens d'être certains que l'on a bien compris les motivations ou objectifs de l'autre. C'est pourquoi lors d'une interaction (sociale) vocale, cette compréhension perçue est fréquemment testée, par exemple en répétant ou en reformulant la dernière phrase de l'interlocuteur.
- **Interdépendance comportementale :** elle est l'un des éléments clés de la présence sociale, elle définit la mesure dans laquelle le comportement ou les interactions vocales d'une personne vont influencer le comportement d'une autre personne. Cela peut être intentionnel, par exemple lorsque l'on voit un semblable en difficulté et que l'on vient lui apporter de l'aide, lorsque l'on sent le regard d'une personne sur nous et que l'on décide de répondre à ce regard. La translucence dont nous avons déjà parlé (voir section 2.1.2.3, (Erickson et Kellogg, 2000)), fait partie de l'interdépendance comportementale intentionnelle, elle est sa forme non verbale. Pour ce qui est du non intentionnel ou automatique, on trouve par exemple le bâillement, ou encore le phénomène de mimique qui fait que nous avons tendance à imiter les comportements non verbaux comme les expressions du visage, la posture ou même l'intonation d'une personne avec qui nous parlons, et ce dans le but d'améliorer le lien empathique nous unissant (Hasler *et al.*, 2014 ; Luo *et al.*, 2013 ; Bailenson *et al.*, 2008b ; Bailenson et Yee, 2005 ; Chartrand et Bargh, 1999).
- **Niveau intersubjectif :** afin d'atteindre un sentiment de présence sociale élevé, chaque protagoniste des interactions en cours doit avoir le sentiment, la certitude que son ou ses interlocuteurs sont aussi présents socialement que lui dans l'environnement qu'ils partagent. Dans la réalité, la question ne se pose pas puisque nous sommes certains que nous sommes en face à face dans une même pièce ou dans le même environnement : si nous perdons l'attention de la personne face à nous, nous nous en apercevons rapidement et pouvons y remédier. En revanche, au travers d'un média, la tâche n'est pas aussi aisée. Un exemple simple est celui du téléphone, tout le monde à un jour éprouvé le désagréable sentiment que la personne à l'autre bout du fil est en train de faire quelque chose d'autre au moment où vous arrivez à la partie cruciale de votre récit. Nous sentons alors qu'il ou elle n'est plus aussi présente que nous dans l'interaction, c'est de cela que nous parlons. Dans le contexte d'un 3D-IVE, nous avons non seulement besoin de sentir dans notre esprit la présence de l'autre en plus de la nôtre, mais aussi d'avoir le sentiment que nous sommes aussi présents dans l'esprit de l'autre protagoniste. C'est ce que Biocca et Harms (2002) appellent la symétrie de l'interaction interne et la symétrie de l'interaction croisée (voir figure 2.8).

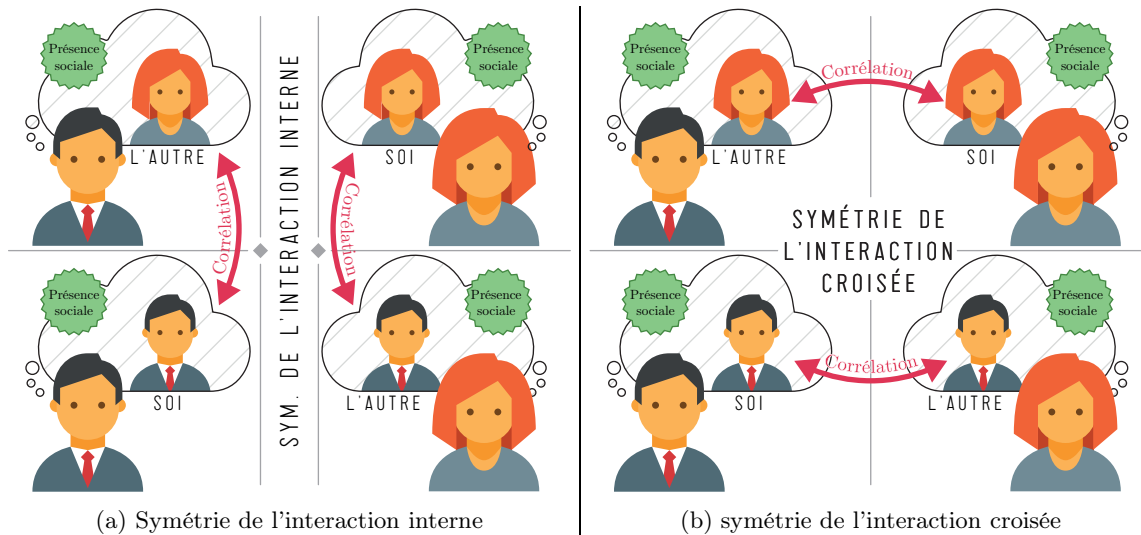


FIGURE 2.8 – Les symétries de l'interaction pour la présence sociale d'après (Biocca et Harms, 2002).

Nous constatons aisément la similarité entre les facteurs tirés de la définition de la présence sociale de Biocca et Harms (2002) listés ci-dessus et ceux que nous avons définis comme « *l'immersion psychologique* » dans le contexte de présence virtuelle. Nous y retrouvons une composante spatiale (présence spatiale / co-présence), des composantes attentionnelles (niveau d'intérêt, concentration, engagement / engagement attentionnel, compréhension perçue, interdépendance comportementale) ainsi qu'une composante émotionnelle (émotion / interdépendance émotionnelle perçue). Il semble donc pertinent de réexaminer les facteurs de la présence virtuelle liée à « *l'immersion perceptuelle* » afin de voir s'ils font écho à des facteurs de la présence sociale dans la littérature.

- **La conscience des interfaces :** ce facteur technique peut jouer un rôle dans le contexte de présence sociale, en effet si la voix de l'utilisateur est mal retransmise, s'il y a des micro-coupures (latence) (Howard *et al.*, 2014) ou de l'écho par exemple, cela renvoie automatiquement l'utilisateur au fait qu'il utilise une technologie pour transporter ses interactions sociales. Ce facteur est donc lié à la qualité de service du modèle de qualité des DIMEs¹² de Wu *et al.* (2009).
- **La vivacité sensorielle :** tout comme pour un environnement virtuel, plus un être virtuel sera riche en termes de sens sollicités, plus ce dernier semblera vivant, présent. Outre l'ouïe que nous avons traitée dans le paragraphe précédent et la vision dont nous parlerons vis-à-vis des niveaux d'anthropomorphisme, de réalisme visuel et de réalisme comportemental, nous savons que le toucher revêt une grande importance dans le cadre des interactions sociales non verbales (voir section 2.1.2.1), ce qui s'applique aussi aux interactions sociales médiées (van Erp et Toet, 2015 ; Haans et Ijsselsteijn, 2005). Rien n'est plus significatif vis-à-vis de la réalité d'une chose ou d'une personne que le fait de pouvoir la toucher et d'en ressentir une émotion (Gaffary, 2015). Reste alors le goût, bien qu'il ne comporte dans notre contexte que peu d'intérêt, il est néanmoins probable malgré le manque d'étude sur le sujet que son impact sur le sentiment de présence sociale soit important. Il en va de même pour l'odorat, que nous savons très difficile à mettre en place dans le contexte des 3D-IVEs.

12. DIME : de l'anglais « *Distributed Interactive Multimedia Environment* »

- **L’interactivité :** Garau *et al.* (2005) ont montré qu’un agent perçu comme potentiellement interactif obtenait une meilleure évaluation du sentiment de co-présence générée mais aussi de l’envie de l’utilisateur d’interagir avec lui. En effet, un agent interactif sera perçu comme plus vivant et donc moins considéré comme un objet de plus dans l’environnement virtuel.
- **Richesse du comportement intrinsèque :** autrement dit, la richesse du comportement de la représentation virtuelle agent ou avatar hors interactions sociales avec l’utilisateur. Toujours en nous référant au travail de Garau *et al.* (2005), nous pouvons dire que l’immobilité totale d’un agent ou avatar va le faire basculer dans la catégorie non intelligent, non doté de sentiments et non interactif. En augmentant ce comportement, par le mouvement et plus encore par des mouvements interdépendants du comportement de l’utilisateur, la représentation virtuelle va gagner en vie et en réalité et donc augmenter le sentiment de présence sociale ou a minima de co-présence de l’utilisateur.
- **Familiarité entre individus :** le fait de connaître ou non la personne avec laquelle vous êtes immergés dans un 3D-IVE a un impact sur le sentiment de présence sociale, tout comme il a un impact sur notre manière de nous comporter dans le monde réel. Hudson *et al.* (2015) ont récemment mené une étude sur ce point dans le contexte des jeux vidéo en ligne de type MOBA (*League of Legend & Dota 2*). Cette étude a révélé que la familiarité entre individus tend à augmenter la présence sociale, mais cette évaluation varie en fonction du jeu vidéo utilisé par les participants.

Les trois facteurs que sont la qualité visuelle, le réalisme comportemental et la cohérence seront traités dans la section suivante. Étant très fortement impliqués dans le processus de présence sociale et afin de mieux comprendre leur entremêlement, nous nous baserons, pour les décrire, sur une théorie bien connue dans le domaine de la réalité virtuelle, la « *Uncanny Valley* ».

2.2.2.3 La théorie de la « *Uncanny Valley* » et l’importance de la cohérence

La théorie de la « *Uncanny Valley* » proposée par Mori *et al.* en 1970 définit la courbe d’affinité et par extension d’acceptabilité par l’humain de la représentation synthétique (physique ou virtuelle) d’une partie ou de la totalité d’un être intelligent. Elle désigne le sentiment de malaise que nous pouvons ressentir au contact de ces représentations synthétiques très proches de l’humain et à la fois perturbantes. Le niveau d’anthropomorphisme¹³ et le réalisme visuel¹⁴ ne sont pas les seuls facteurs pouvant faire tomber une représentation dans la « *Uncanny Valley* », le mouvement est aussi très impliqué dans ce processus. Prenons l’exemple d’un mannequin de magasin de vêtement, s’il ne bouge pas, il vous semble normal, il représente l’humain de manière assez réaliste dans la forme bien que l’on sache qu’il s’agisse d’un mannequin à la vue de ses yeux vides et de la couche de plastique lisse qui lui sert de peau. Mais si ce dernier se met à bouger, il tombe aussitôt dans la « *Uncanny Valley* », puisque le mouvement pour une chose censée être inerte n’est pas normal (voir figure 2.9).

13. Anthropomorphisme : ressemblance anatomique.

14. Réalisme visuel : niveau de photo-réalisme.

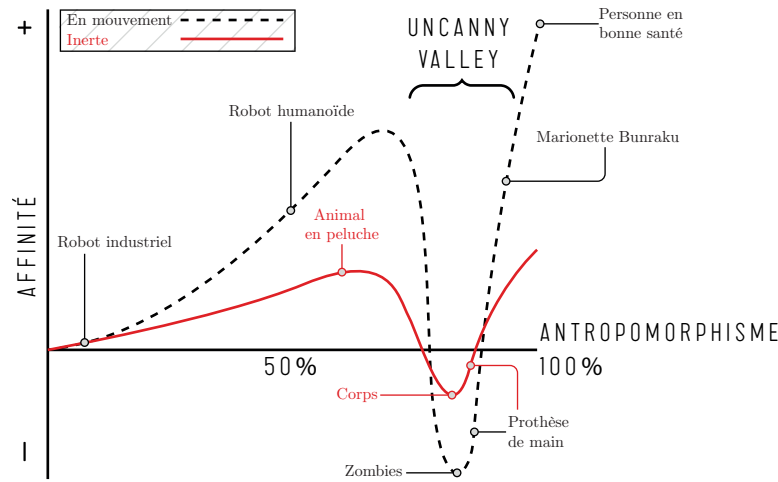


FIGURE 2.9 – Courbe de la « Uncanny Valley » avec effet du mouvement.

Le phénomène décrit par Mori *et al.* soulève depuis sa publication un grand intérêt dans la communauté scientifique. Les résultats obtenus par l'étude empirique sont souvent en désaccord avec la théorie de la « *Uncanny Valley* ». Prenons l'étude de Burleigh *et al.* (2013) qui consiste à l'exposition de participants à une grande variété de visages présentant des variations anthropomorphiques et de réalisme visuel fines. Leurs résultats tendent à prouver que la relation entre l'affinité et le niveau de réalisme est linéaire. Néanmoins, ces résultats sont à pondérer, car aucun des visages présentés ne propose un niveau de réalisme visuel suffisant pour créer le basculement dans la « *Uncanny Valley* ». Les variations n'étant que des modifications du modèle 3D proposé (par exemple allongement des oreilles, agrandissement des yeux, coloration de la peau), les résultats pourraient être différents si les textures utilisées pour les visages provenaient de véritables photographies comme on peut le voir sur la figure 2.10c. Une étude très similaire réalisée par Seyama et Nagayama (2007) va elle dans le sens de la théorie de Mori en affirmant que le réalisme visuel d'une représentation anthropomorphique seul n'est pas suffisant mais demande qu'au moins un élément de la représentation soit anormal. Leur expérimentation consiste à exposer des participants à des visages produits en isolant les étapes d'un « *morphing* » entre par exemple une photographie et un visage issu d'un manga (dessin), un parti pris qui crée un clivage entre la théorie de la « *Uncanny Valley* » et ses applications dans le sens où ils exposent leurs participants à des visages qui ne sauraient être utilisés sciemment sauf dans le cas où l'on voudrait faire ressentir la « *Uncanny Valley* ». Qui plus est, dans ces deux études, les visages sont statiques et ne prennent donc pas en compte la relation du phénomène avec le mouvement. Ce lien est éprouvé par les travaux de Piwek *et al.* (2014). Leur étude consistant à comparer les mêmes représentations anthropomorphiques virtuelles, statiques et en mouvement (allant du robot à l'humain en passant par le zombie), n'a là encore pas confirmé les hypothèses de Mori *et al.*, et au contraire, a montré que même dans le cas du zombie considéré comme étant la plus « *uncanny* » des représentations (voir figure 2.9), le fait de bouger de manière naturelle ou distordue accroissait le niveau d'affinité. Là encore, aucune des représentations n'atteint un niveau de réalisme visuel suffisant pour le basculement et les représentations virtuelles demanderaient à être comparées à des représentations physiques (réelles), car rappelons que la théorie de Mori *et al.* était à l'origine appliquée au domaine de la robotique et non de la réalité virtuelle. Néanmoins, on constate au travers de ces travaux que le

facteur qui semble avoir le plus d'impact sur l'utilisateur est le réalisme comportemental : le fait de se mouvoir comme un humain qui bien sûr demande de l'anthropomorphisme, avoir des jambes, des bras, un corps une tête.

Johansson (1973) a démontré qu'une personne pouvait, en voyant un réseau de treize points en mouvement, identifier quasi-instantanément le mouvement humain et qu'en revanche si ce réseau de points est immobile il est quasi impossible d'identifier à une forme humaine. Plus récemment et cette fois au sein de 3D-IVEs, de nombreuses études ont démontré l'importance du réalisme comportemental. Ainsi, Bailenson *et al.* (2002) ont empiriquement testé les effets du réalisme comportemental visuel (regard mutuel) d'un agent. Les résultats montrent une corrélation entre le niveau de réalisme comportemental et le sentiment de présence sociale, mais aussi un effet de ce même niveau de réalisme sur les performances de l'utilisateur (ici un jeu de questions - réponses). Non seulement le réalisme comportemental est important mais il semble être plus important que le réalisme visuel comme le démontrent les travaux de Bailenson *et al.* (2001a). L'application de textures photo réalistes sur un agent n'augmente pas le sentiment de présence sociale en comparaison d'un agent possédant des textures plus « cartoon » si leur regard a le même niveau de réalisme comportemental. Ce constat est confirmé par les études de Garau (2003) sur différents scénarios (discussion, navigation) qui a montré que l'augmentation du niveau de réalisme visuel d'un avatar n'augmentait pas le sentiment de présence sociale mais qu'au contraire passé un certain niveau, il avait tendance à le faire décroître. Cela est en accord avec la théorie de la « *Uncanny Valley* » de Mori *et al.*. En revanche, un agent ou avatar ne semble pas pouvoir être trop réaliste du point de vue de son comportement. En effet, Bailenson *et al.* ont testé un comportement non verbal très naturel, le phénomène de mimique ou effet caméléon. Ils ont démontré que l'application de l'effet caméléon à un agent ou un avatar, par reproduction décalée de quelques secondes des mouvements de l'utilisateur (technique offrant un très haut niveau de réalisme comportemental puisque reproduisant la réalité), avait un effet bénéfique sur le sentiment de présence sociale (Bailenson *et al.*, 2008b ; Bailenson et Yee, 2005).

Le phénomène de « *Uncanny Valley* » est, tout comme l'expérience utilisateur, variable et différent chez chaque utilisateur, ce qui le rend d'autant plus complexe à analyser, à comprendre. Mori *et al.* émettent l'hypothèse que ce sentiment d'étrangeté et les réactions qu'il suscite viendraient de l'instinct de conservation humain : ce qui n'est pas compris est potentiellement très dangereux ! D'un autre côté, les résultats de Saygin *et al.* (2011) penchent du côté de la « violation des attentes » que l'on peut décrire comme la discordance entre la forme et le mouvement. Dans cette étude, les résultats montrent des réponses plus fortes (plus « *uncanny* ») en face d'un androïde humanoïde (voir Figure 2.10a) mais ne bougeant pas de manière biologique que dans le cas d'un humain bougeant naturellement ou d'un robot bougeant de manière robotique. Résultat portant sur une étude dans le monde réel mais cohérent avec les travaux de Schubert *et al.* (2002) précédemment cités sur la nécessité pour l'utilisateur de se faire une représentation mentale cohérente avec la réalité primaire de ce qu'il se passe et dans le 3D-IVE, et cette dernière de ne pas être en contradiction avec cette représentation mentale.

La « *Uncanny Valley* » traite particulièrement de la perception que nous avons des représentations

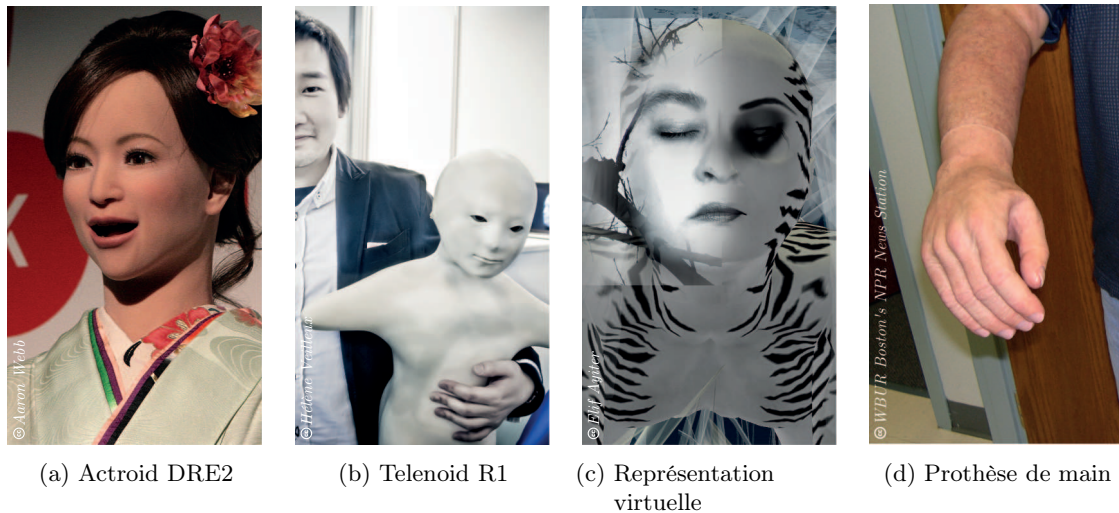
des autres êtres intelligents, aujourd'hui, grâce par exemple aux casques immersifs il est fréquent et conseillé d'offrir à l'utilisateur la possibilité de voir son propre corps, son incarnation virtuelle. Certaines études tendent à démontrer que le phénomène de « *Uncanny Valley* » s'y applique aussi. En effet, une étude menée par Lugin *et al.* (2015) montre qu'au même niveau de réalisme comportemental, les corps virtuels offrant un faible niveau d'anthropomorphisme sont mieux acceptés par les utilisateurs. Les corps virtuels imitant l'humain génèrent une plus forte sensation d'avoir deux corps (un réel / un virtuel) et engendre chez l'utilisateur le sentiment que telle ou telle partie du corps ne correspond pas à son corps réel. Cela tend à rendre l'utilisateur mal à l'aise, autrement dit à plonger dans la « *Uncanny Valley* » et donc faire décroître son sentiment de présence virtuelle. Il semblerait que plus une représentation virtuelle imite l'humain, plus les utilisateurs la comparent inconsciemment à la réalité qu'ils connaissent. La solution serait alors de proposer une reconstruction complète du corps de l'utilisateur avec ses propres proportions et textures. Les projets menés par Saraiji *et al.* (2015; 2014; 2013) vont dans ce sens mais malheureusement ne présentent aucune évaluation du ressenti des utilisateurs vis-à-vis de leur utilisation. Afin que les utilisateurs se sentent à l'aise avec leurs représentations virtuelles et celles des autres, la recherche d'un anthropomorphisme et d'un réalisme visuel poussé semble vaine. Il est préférable de considérer que quoi qu'il en soit l'utilisateur est toujours conscient de son immersion dans un monde qui n'est pas réel, quels que soient les matériels utilisés ou le niveau d'anthropomorphisme et de réalisme visuel des représentations auxquelles il est exposé.

Comme nous l'avons vu, le phénomène de « *Uncanny Valley* » engendre un sentiment de rejet qui est préjudiciable pour l'occurrence d'interactions sociales entre les participants ainsi qu'aux sentiments de présence virtuelle et sociale. Plus important encore, au travers de ce phénomène nous comprenons que l'essentiel ne se situe pas dans les niveaux de réalisme visuel, d'anthropomorphisme et de réalisme comportemental, mais dans la **cohérence entre ces trois facteurs**.

De même, pour la présence virtuelle, ce découpage en facteurs de la présence sociale nous permettra de mieux comprendre son articulation dans les modèles d'expérience utilisateur en 3D-IVEs que nous verrons par la suite. Tout aussi important, ce découpage et les méthodes d'évaluation proposées par Biocca et Harms (2002) permettront une évaluation de la présence sociale robuste, car déjà éprouvée à de nombreuses reprises par des études empiriques.

La figure 2.10 donne quatre exemples de représentations jugées comme tombant dans la « *Uncanny Valley* » :

- **Figure 2.10a :** les « *Actroid* », contraction d'actrice et d'androïde, conçus par l'université d'Osaka (Japon) sont parmi les androïdes les plus réalistes créés jusqu'alors. Néanmoins, de petits détails les font irrémédiablement tomber dans la « *Uncanny Valley* ». Leurs mouvements quelque peu saccadés, l'impression que leur voix ne provient pas de leur bouche et le manque de synchronisation entre le discours et le mouvement des lèvres, la sécheresse apparente de ces dernières ainsi que des yeux et le grain de la peau sont autant d'éléments perturbants. (Pour un aperçu, voir : <https://www.youtube.com/watch?v=V7TKBFJMb0>).
- **Figure 2.10b :** le « *Telenoid* », lui aussi issu des laboratoires de l'université d'Osaka, est, dans

FIGURE 2.10 – Exemples de représentations tombant dans la « *Uncanny Valley* ».

la forme, plus éloigné de l'humain que d'un mannequin de magasin. Mais son usage le rend vraiment étrange. En effet, le « *Telenoid* » est un outil de téléprésence censé reproduire la voix, les expressions faciales et les mouvements d'une personne distante. Il est difficile de s'imaginer avec cet androïde tronc sur les genoux alors que celui-ci narre sa journée avec la voix d'un proche ou d'un enfant. (Pour un aperçu, voir : <https://www.youtube.com/watch?v=r7t4uczfbZY>, on perçoit sur la vidéo que les personnes sont impressionnées par la technologie et à la fois mal à l'aise vis-à-vis de l'interaction avec le « *Telenoid* »).

- **Figure 2.10c :** expérimentation sur le thème de la « *Uncanny Valley* » dans *Second Life* par Elif Ayiter. Ici la perturbation vient de l'utilisation d'un visage photo réaliste sur un corps virtuel dont les couleurs et le motif ne sont pas cohérents avec le visage.
- **Figure 2.10d :** les prothèses sont l'un des exemples donnés par Mori *et al.* dans leur première publication sur la « *Uncanny Valley* ». Malgré un réalisme très poussé, le fait de serrer une main froide ou d'être touché par une main en plastique est perturbant surtout lorsque l'on ne s'y attend pas.

2.2.3 Une interaction sociale convaincante en environnement virtuel, utopie ou réalité ?

Nous avons à présent une vue globale de ce qu'est l'interaction sociale, des phénomènes de présence virtuelle et de présence sociale ainsi que de l'impact des variables liées aux agents ou avatars sur ces phénomènes. Néanmoins, une question demeure, les interactions sociales que nous pouvons créer au sein d'un 3D-IVE sont-elles analogues aux interactions sociales dans le monde réel ? Un avatar est-il capable de nous transmettre une émotion ? De nous faire sentir sa désapprobation, et nous, de la comprendre ?

Il est avéré que la réalité virtuelle peut générer des émotions et que de la génération de ces émotions découle des phénomènes physiologiques (augmentation du rythme cardiaque, variation de la sudation)

qui permettent d'attester de la réalité de l'émotion ressentie comme l'ont par exemple démontré Meehan *et al.* (2002). Il semble en être de même vis-à-vis des interactions sociales, le comportement que nous adoptons dans un environnement virtuel est proche de celui que nous adopterions dans une situation similaire dans le monde réel. Prenons en premier lieu une expérimentation menée par Pertaub *et al.* (2002) sur l'anxiété ressentie lors de la promulgation d'un discours dans un 3D-IVE devant une audience constituée d'agents. Ces agents se comportent de trois manières différentes, soit ils sont neutres et donc ne bougent presque pas, soit ils sont amicaux, souriants et attentifs, soit ils sont hostiles à la présentation et semblent en colère ou inattentifs. Les résultats obtenus par un questionnaire dédié à ce type d'exercices (« *Personal Report of Confidence as a Public Speaker* » ou PRCPS) montrent que dans le cas de l'audience hostile, le niveau d'anxiété est significativement plus haut que dans les deux autres conditions ce qui démontre que le comportement de l'audience même virtuelle a un impact sur le participant. Par la suite, une expérience très similaire a donné des résultats proches, de plus cette dernière utilisait des agents dont le niveau de réalisme visuel était volontairement bas. Là encore, les résultats de l'évaluation de l'anxiété par les questionnaires montrent une plus grande anxiété dans le cadre de l'audience hostile. Ces résultats sont par ailleurs cohérents avec les mesures physiologiques par ECG mises en place. Devant l'audience hostile, le rythme cardiaque a tendance à s'accélérer, signe d'une augmentation de l'anxiété (Slater *et al.*, 2006). Dans un contexte différent, celui des personnes atteintes de paranoïa (syndrome de persécution), Freeman *et al.* (2008) ont montré que les participants immergés dans un 3D-IVE avec des agents rapportent avoir le même sentiment de persécution que dans une situation similaire réelle. Dans cet exemple, les environnements choisis sont des lieux publics, les personnages virtuels affichent une expression neutre mais de manière occasionnelle jettent un regard sur le participant. L'expérimentation menée sur un échantillon large (200 participants) atteste d'une bonne robustesse des résultats et nous donne là encore une preuve de l'impact du comportement social dans le virtuel. Les études décrites ici traitent du comportement non verbal, il est plus difficile de trouver des études où par exemple les participants se feraient invectiver par des agents ou avatars afin de voir si cela génère chez eux de la crainte ou de la colère. Néanmoins, le projet *MACH* (« *My Automated Conversation coach* ») développé par Hoque *et al.* (2013) afin d'aider les personnes à améliorer leurs compétences sociales nous donne un bon indice. Dans la mesure où les utilisateurs de *MACH* ont été en mesure d'améliorer les compétences sociales grâce aux interactions qu'ils ont eues avec l'agent virtuel du logiciel, il paraît cohérent de dire que ces interactions sociales étaient convaincantes et donc proches de la réalité.

Une autre indication permettant d'attester la véracité des interactions sociales en environnement virtuel est la conservation des normes sociales. Une étude menée par Yee *et al.* (2007) sur la plateforme en ligne *Second Life* et consistant en la capture automatique d'image (835 captures exploitables) déclenchée à la détection de l'entrée d'une personne dans la sphère proxémique sociale d'une autre (3,66 m) (Hall, 1966) a montré une conservation des normes sociales durant ces échanges. Les résultats montrent une conservation de « *l'Equilibrium Theory* », décrivant un rapport inverse entre distance proxémique et regard mutuel (plus les acteurs de l'interaction sont proches, moins ils se regardent directement) (Bailenson *et al.*, 2001b ; Argyle et Dean, 1965) ainsi que l'utilisation du regard mu-

tuel comme régulateur de l'interaction sociale (plus deux avatars échangent, plus ils sont enclins à se regarder). Il a aussi été constaté une corrélation entre les comportements dans le réel et dans le virtuel vis-à-vis du genre des utilisateurs. Tout comme dans la réalité, les hommes conservaient une plus grande distance proxémique que les femmes, et maintenaient moins le regard mutuel. Ces deux variables sont également affectées par l'environnement : en intérieur (dans une petite pièce), les distances proxémique sont accrues et le regard mutuel moins maintenu. Bien sûr, cette étude porte sur la plateforme *Second Life* et les résultats pourraient être différents dans un autre environnement virtuel, mais il faut néanmoins les prendre en considération au regard du nombre d'interactions sociales étudiées. Il faut toutefois signaler que cette étude ne tient pas compte de la familiarité potentielle entre individus, un facteur qui a une grande influence sur la proxémique et le regard mutuel, comme nous l'avons déjà évoqué. D'un autre côté, Bailenson *et al.* dans leurs études de l'effet caméléon, ont eux aussi constaté une conservation des lois régissant ce phénomène. L'effet caméléon a un effet bénéfique sur l'interaction sociale en accroissant le sentiment d'empathie, en créant du lien, mais cet effet s'inverse de manière probante dès que la personne qui est imitée détecte cette imitation. En résulte alors une forme de rejet de l'autre, car la personne se sent « singé » ce qui est pris comme une forme de moquerie. Les résultats de leurs études empiriques mettent en exergue ces variations de l'empathie dans les 3D-IVEs (Bailenson *et al.*, 2008b ; Bailenson et Yee, 2005).

Émotions, stimulation physiologique, normes sociales, les études citées ici montrent que tous ces éléments sont conservés lorsque nous interagissons socialement dans un environnement virtuel. Si les personnes impliquées dans une interaction sociale en environnement virtuel agissent et réagissent de la même manière que dans la réalité, et ce malgré le fait que les modalités d'interaction et de navigation soient différentes (utiliser une souris, un clavier, une manette au lieu de nos bras et jambes, parler dans un microphone devant un écran plutôt qu'en face à face), alors nous pouvons affirmer qu'il est possible d'étudier les interactions sociales en environnements virtuels et donc potentiellement dans notre contexte leur impact sur l'expérience utilisateur.



Nous avons à présent une vue d'ensemble sur le fonctionnement de l'interaction sociale, de la raison de son utilisation et des modalités de ces interactions sociales en environnements virtuels. Nous avons souligné le fait que les facteurs les plus importants de l'interaction sociale en 3D-IVE sont le réalisme comportemental et la cohérence de ce dernier avec le niveau d'anthropomorphisme ainsi qu'avec le réalisme visuel de l'avatar ou de l'agent mais aussi de l'environnement virtuel lui-même. Nous allons à présent pouvoir nous intéresser à l'expérience utilisateur afin d'en définir les facteurs en 3D-IVEs mais aussi afin de voir comment elle s'articule avec les interactions sociales vocales et dans quelle mesure il est possible d'en évaluer l'impact.

2.3 L'expérience utilisateur

L'expérience utilisateur, abrégée UX est un concept décrivant l'expérience d'une ou de plusieurs personnes avec un produit ou un service et avec la prestation qu'il propose. Avant de nous attarder sur sa définition, intéressons-nous rapidement à son histoire, son émergence. Nous pourrions ensuite nous focaliser sur les deux modèles d'expérience utilisateur dédiés spécifiquement aux 3D-IVEs et sur lesquels reposera notre étude empirique, le modèle de Wu *et al.* (2009), et celui en découlant de Pallot *et al.* (2013). Au travers de ces explications, nous mettrons ces deux modèles en parallèle avec les concepts de présence virtuelle et sociale afin de comprendre l'articulation de l'expérience utilisateur en 3D-IVE. L'objectif est ici de démontrer la nature holistique de ces modèles, et donc de la pertinence de ceux-ci dans le cadre de notre travail de recherche.

2.3.1 De l'utilisabilité à l'expérience utilisateur, évolution d'un concept multimodal

À l'origine du concept d'expérience utilisateur est le concept d'utilisabilité, concept ayant pour objectif de comprendre et d'évaluer la propension d'un système ou d'un service à permettre à un utilisateur d'atteindre un objectif précis. De nombreuses définitions de l'utilisabilité ont bien sûr vu le jour au fil du temps, mais comme là n'est pas le propos premier de cette revue de littérature, nous nous baserons pour l'expliquer sur la définition la plus communément admise, celle donnée par le standard ISO 9241-11 (1998) : « *le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié* ». Cette définition nous permet de pointer les termes importants la composant (Rusu *et al.*, 2015) :

- **Utilisateur** : personne(s) interagissant avec le produit ;
- **buts définis** : résultats escomptés ;
- **efficacité** : exactitude et exhaustivité avec lesquelles l'utilisateur atteint les objectifs fixés ;
- **efficience** : ressources utilisées en relation avec l'exactitude et l'exhaustivité avec lesquelles l'utilisateur atteint les objectifs fixés ;
- **satisfaction** : absence d'inconfort et attitude positive vis-à-vis de l'utilisation du produit ou service ;
- **contexte** : l'utilisateur lui-même, les tâches mises en place en vue de l'atteinte des objectifs fixés, l'équipement (matériel, logiciel) ainsi que les contextes physiques et sociaux dans lesquels le produit ou service est utilisé.

L'année suivant la définition de l'utilisabilité par le standard ISO, un autre standard a été rédigé vis-à-vis de l'évaluation de la qualité d'un logiciel. Le standard ISO/IEC 25010 (2001) (remplacé depuis par ISO/IEC 25010 (2011)) qui reprend en grande partie la définition de l'utilisabilité et définit « *la qualité d'utilisation* » comme suit : « *La capacité d'un produit logiciel de permettre à des utilisateurs spécifiques d'atteindre des buts spécifiques avec efficacité, productivité, sécurité et satisfaction, dans*

un contexte d'utilisation spécifié ». De cette définition, découlent les quatre attributs de « *la qualité d'utilisation* » qui sont définis comme la capacité du produit logiciel de :

- Permettre à l'utilisateur d'atteindre les buts définis avec précision et de manière exhaustive dans le contexte d'utilisation spécifié (**efficacité**) ;
- permettre à l'utilisateur une utilisation des ressources appropriée en relation avec l'efficacité atteinte dans le contexte d'utilisation spécifié (**productivité**) ;
- ne pas dépasser un niveau acceptable de risque de préjudices, pour les personnes, les entreprises, les logiciels, les propriétés ou l'environnement dans le contexte d'utilisation spécifié (**sécurité**) ;
- satisfaire les utilisateurs dans le contexte d'utilisation spécifié (**satisfaction**).

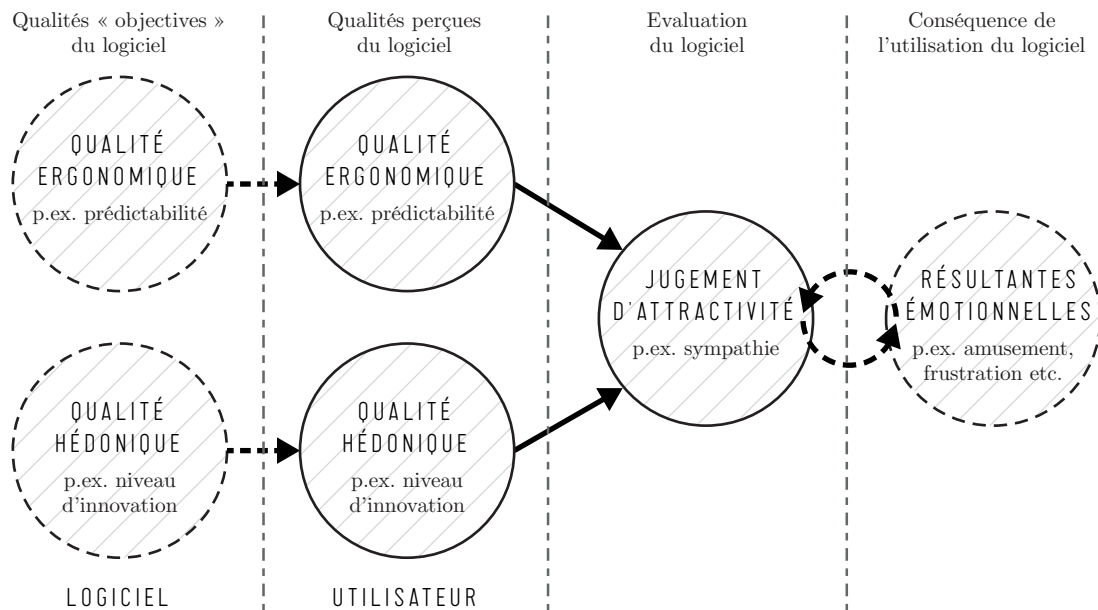
De ces deux définitions, nous pouvons retenir plusieurs choses. La première est qu'elles proposent toutes deux une définition multimodale au travers de facteurs indépendants mais dont l'interdépendance conditionne l'utilisabilité ou la qualité d'utilisation du logiciel. La seconde est qu'elles sont exclusivement orientées sur des objectifs spécifiques au travers de tâches précises, elles sont donc adaptées à l'évaluation d'un outil de travail, mais pas à celle d'un jeu, d'un voyage, ou tout autre produit ou service faisant appel au ressenti hédonique¹⁵ de l'utilisateur. C'est ici qu'intervient l'expérience utilisateur, le terme est apparu avec l'inclusion du concept de plaisir, ce qui a donc poussé à une vision plus holistique de l'expérience générée par un produit ou un service. L'une des premières équipes de chercheurs qui s'est intéressées aux sentiments hédoniques de l'utilisateur lors de l'utilisation d'un système interactif (logiciel) est Hassenzahl *et al.* (2000). Ils proposent alors un modèle (voir figure 2.11) comprenant deux dimensions vis-à-vis de l'évaluation d'un système interactif :

- **Les qualités ergonomiques** : constituées des facteurs de l'utilisabilité traditionnelle, entre autres, l'efficacité et l'efficacé (voir ISO 9241-11 (1998)), donc focalisées sur la tâche et sur la conception du système interactif ;
- **les qualités hédoniques** : constituées de facteurs n'ayant pas de relations évidentes avec la tâche que l'utilisateur doit accomplir mais avec le système interactif comme son originalité, son niveau d'innovation, sa beauté, etc. Ces qualités hédoniques, bien que non directement corrélées à la tâche ou aux objectifs spécifiés, revêtent une grande importance pour l'utilisateur au regard de l'intérêt qu'il trouve dans l'utilisation du système interactif et qui le poussera ou non à sa réutilisation.

Ce modèle étendu de l'utilisabilité par Marc Hassenzahl *et al.*, qu'ils ont eux-même qualifié comme étant **prometteur** et **stimulant** pour les futures recherches (Hassenzahl *et al.*, 2000, p.207), a par la suite, et conjointement à l'intérêt des industriels et aux travaux de multiples chercheurs, donné lieu à ce que nous appelons aujourd'hui l'expérience utilisateur.

Le standard ISO 9241-11 (2010) définit l'expérience utilisateur comme, « *l'ensemble des réponses et des perceptions d'une personne résultant de l'usage ou de l'anticipation de l'usage d'un produit, d'un service ou d'un système* ». À cette définition s'ajoutent les notes explicatives suivantes : « *l'expérience*

15. Hédonique : d'hédonisme, système philosophique qui fait du plaisir le but de la vie (Larousse 2015). Le ressenti hédonique exprime selon s'il est positif ou négatif le fait d'éprouver ou non du plaisir.

FIGURE 2.11 – Modèle théorique d'évaluation d'un système interactif de Hassenzahl *et al.* (2000).

de l'utilisateur inclut toutes les émotions, croyances, préférences, perceptions, réponses physiologiques et psychologiques, comportements et réussites de l'utilisateur se produisant avant, pendant et après l'utilisation » et « l'expérience utilisateur est une conséquence de l'image de marque, de la présentation, des fonctionnalités, de la performance, du comportement interactif et des capacités d'assistance du système interactif, de l'état spirituel et physique de l'utilisateur résultant d'expériences antérieures, de ses attitudes, de ses compétences et de sa personnalité ainsi que du contexte d'utilisation ». La définition standard inclut aussi une note sur les aspects de l'utilisabilité dans l'expérience utilisateur : « l'utilisabilité, lorsqu'elle est interprétée depuis la perspective des objectifs personnels des utilisateurs, peut inclure le type d'aspect perceptuel et émotionnel généralement associé à l'expérience utilisateur. Les critères issus de l'utilisabilité peuvent être utilisés pour évaluer certains aspects de l'expérience de l'utilisateur » (résumé du standard ISO 9241-11 (2010) issu de Scapin *et al.* (2012)). On s'aperçoit ici que l'expérience utilisateur, en plus de prendre en compte les aspects perceptuels et émotionnels de l'utilisateur, inclut comme facteur influent l'image de marque de la compagnie proposant le produit ou le service. Ces trois facteurs que sont la perception, l'émotion et l'image de marque, en étant placés au centre du processus de conception de produit, ont permis à des entreprises comme *Apple* d'acquérir une grande renommée, et ce malgré le fait que leurs produits ne soient pas les plus aboutis techniquement ou d'un point de vue matériel, mais, simplement, ils génèrent de l'émotion : « *On ne fera jamais une bonne publicité en mettant en avant des caractéristiques techniques, de Giga-octet de RAM, des tableaux et des comparatifs. Il faut transmettre une émotion* » [notre traduction] (Steve Jobs — *The Apple Way* — 2006). Nous trouvons ici un indice sur l'importance de l'émotion dans l'expérience utilisateur, autrement dit d'un processus mental très intuitif et ne faisant que peu appel à notre cognition comme nous le verrons par la suite.

Les définitions du standard ISO 9241-11 (2010) mettent en lumière l'aspect multimodal de l'expérience utilisateur mais il faut aussi être conscient qu'une expérience globale n'est pas uniquement mul-

timodale, elle est aussi composée des multiples interactions entre les expériences elles-mêmes comme l'illustre l'exemple suivant issu des travaux de Forlizzi et Battarbee (2004) : « *Les produits à emporter lors d'un voyage en camping peuvent inclure des téléphones, des cartes, des outils pour les vélos ainsi que du matériel de cuisine. Le voyage lui aussi est une expérience : les préparatifs sont bouclés, le voyage commence. Au fil du temps, l'expérience du voyage se cumule à de nombreuses autres petites expériences. Dans ce processus, l'expérience d'un produit peut tout changer, une tente peut commencer à fuir, un téléphone peut sauver la journée* » [notre traduction]. Si nous ramenons cet exemple à notre contexte, l'étude empirique en 3D-IVE, il est primordial de comprendre que l'expérience des utilisateurs ne se cantonne pas au seul temps d'immersion, elle s'étend aussi à l'environnement réel dans lequel ils seront, au comportement des personnes les accueillant, à la manière dont l'expérimentation leur sera présentée et toute autre chose qu'ils seraient en mesure de percevoir avant, pendant ou après leur immersion en environnement virtuel. Nous pointons ici un dernier élément extrêmement important de l'expérience utilisateur, et qui lui confère sa complexité quant à son évaluation : sa nature dynamique, versatile en fonction du temps et du contexte (Law *et al.*, 2009 ; Komulainen *et al.*, 2008). Outre les événements pouvant arriver au cours d'une expérience et la faisant varier, Kankainen (2002) argue que cette versatilité peut être en partie expliquée par le fait qu'une personne porte en elle ses expériences antérieures, ainsi ses souvenirs des expériences passées ont un impact important sur l'expérience présente. Pour illustrer ce propos, imaginons que nous nous trouvons face à une porte, cette dernière présente une poignée d'un type que nous n'avons jamais rencontré (voir figure 2.12), nos expériences préalables nous dictent que son fonctionnement doit-être proche de celui des poignées que nous connaissons déjà, une rotation vers le bas devrait actionner le pêne et permettre l'ouverture de la porte. Cependant, sa forme ne nous parle pas immédiatement, mais au regard du contexte dans lequel se trouve cette poignée, ici, un lieu médicalisé, notre esprit va faire le lien entre la forme de la poignée et le fait que ce lieu est soumis à des risques de contamination. Nous comprenons alors (sans doute) qu'il faut actionner cette poignée à l'aide de notre avant-bras. L'expérience que nous avons eue avec cette poignée de porte restera dans notre esprit, que nous ayons trouvé cette poignée utile, inutile, voire même pénible, va conditionner la réaction que nous aurons face à une poignée de porte ordinaire. Peut-être, allons-nous être soulagés par sa « normalité » ou repoussé à la pensée de toutes les mains l'ayant actionnée et y ayant laissé de petites particules organiques. Les réactions futures sont ici bien évidemment variables selon la personne mais il est certain que cette expérience laissera sa marque dans l'esprit de l'utilisateur à plus ou moins long terme.

Il nous faut retenir de cette introduction au concept d'expérience utilisateur qu'elle est multimodale, impliquant des facteurs internes à l'utilisateur (état d'esprit, expérience antérieure, sensibilité, connaissances, etc.), des facteurs inhérents à la conception du produit ou du service ainsi que des facteurs environnementaux. L'aspect le plus important de l'expérience utilisateur et sans doute sa variabilité en fonction du temps (avant, pendant, après) mais aussi du contexte d'utilisation. Avoir une vision holistique de l'expérience utilisateur n'est donc pas chose aisée, il est primordial lors de la conception centrée utilisateur d'un système de réalité virtuelle ou lors de recherche tendant à l'évaluation de l'expérience utilisateur, d'être vigilant vis-à-vis des facteurs pouvant varier au cours de



(Voir le site *Ulna System* : www.ulna-system.com/fr/liste-des-produits-ulna/ulna-initial.html)

FIGURE 2.12 – Ulna, poignée de porte anti-contamination.

l'expérience proposée et risquant dans l'orienter dans une direction non désirée. Dans le cadre d'une évaluation, il est aussi très important de bien choisir les facteurs sur lesquels travailler afin d'obtenir une vision aussi holistique que possible de l'expérience utilisateur tout en proposant des moyens robustes de mesurer ces facteurs.

2.3.2 Modèles d'expérience utilisateur retenus

Plutôt que de décrire un par un les modèles existants de l'expérience utilisateur, et de submerger le lecteur sous les facteurs les composant, nous allons directement nous intéresser aux deux modèles particuliers retenus pour étayer nos travaux : celui de Wu *et al.* (2009) et celui de Pallot *et al.* (2013). Ce choix sur base sur deux points principaux, les modèles décrits ici, sont dédiés spécifiquement aux systèmes interactifs tels que la réalité virtuelle. Ils présentent une taxonomie de facteurs cohérente avec les phénomènes de présence virtuelle et de présence sociale centraux dans notre travail. Nous les mettrons en parallèle au fil des explications avec les autres modèles existants afin de montrer leurs points communs, et ainsi la vision holistique qu'ils proposent.

2.3.2.1 Le modèle de qualité des DIMEs de Wu et al. (2009), dichotomie entre Qualité de Service et Qualité de l'Expérience

Le développement de ce modèle de qualité des environnements multimédias interactifs connectés (ou distribués) (DIMEs¹⁶) amène une nouvelle vision dans laquelle l'expérience de l'utilisateur et décentrée du projet, de l'application. En effet, contrairement aux modèles proposés par exemple par Morville (2004) (voir figure 2.13) ou Revang (2007) définissant l'expérience utilisateur selon des facteurs auxquels doit répondre l'application (crédibilité, accessibilité, utilité, etc.). Wu *et al.* proposent une dichotomie des facteurs technologiques et des facteurs humains en deux entités distinctes, respectivement la Qualité de Service (*Quality of Service* : QoS) et la Qualité de l'Expérience (*Quality of Experience* : QoE) ainsi que la modélisation de l'impact de la QoS sur la QoE. Faisant ainsi passer son modèle d'une logique centrée sur le système et prenant en compte l'utilisateur comme une partie de ce

16. DIME : de l'anglais « *Distributed Interactive Multimedia Environment* »

dernier, à une logique centrée sur l'expérience de l'utilisateur sur laquelle le système va venir influencer. La chaîne de causalité QoS-QoE (voir figure 2.14) illustre l'impact des composantes de la QoS en tant que facteurs environnementaux sur la QoE de l'utilisateur ainsi que les retours des utilisateurs de la QoE vers la QoS pouvant amener à une modification de la configuration liée à la QoS (retour illustré par une flèche en pointillé, car ne se faisant pas en temps en réel mais a posteriori de l'utilisation).

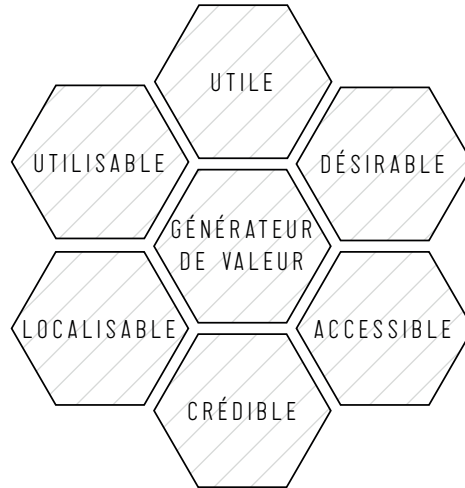


FIGURE 2.13 – Modèle d'expérience utilisateur en nid d'abeille (Morville, 2004).

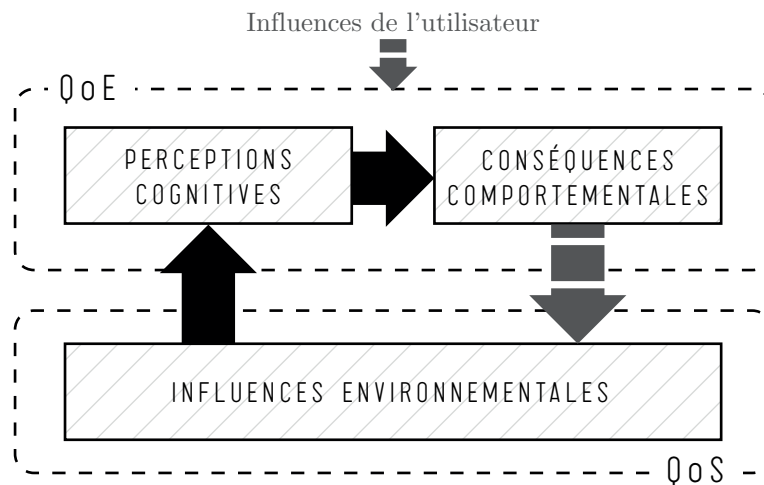


FIGURE 2.14 – Chaîne de causalité entre QoS & QoE (Wu *et al.*, 2009).

La QoS, comme définie par Nahrstedt *et al.* (2011), correspond aux ressources technologiques d'un système interactif et à leur propension à satisfaire les besoins et objectifs de ce système. Elle définit deux catégories de ressources :

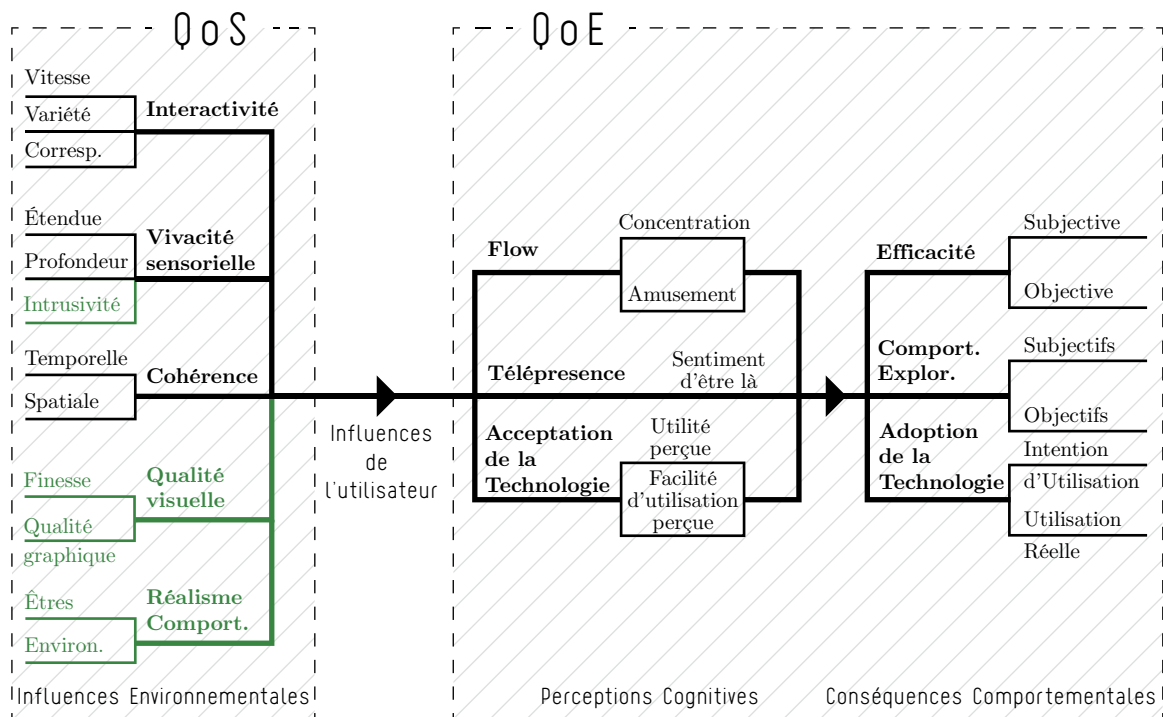
- **Les ressources informatiques :** la puissance de traitement de l'ordinateur utilisé. Le matériel doit être en mesure de fournir suffisamment de puissance pour que l'application fonctionne de manière fluide ;
- **les ressources réseau :** la bande passante du réseau, doit être suffisante pour transmettre en temps réel les informations nécessaires au bon fonctionnement du système. Une ressource extrêmement importante dans le cadre d'une vidéoconférence ou du partage d'un environnement hébergé sur des serveurs distants (MMORPG ou *Second Life* par exemple), mais qui n'est

pas forcément importante dans le cas où les informations passeraient par un réseau local, les ralentissements ou pertes de paquets étant négligeables dans ce contexte.

Nous pouvons considérer cette définition comme un premier niveau de QoS. En effet, sans une capacité technique suffisante autrement dit si l'application, quelle qu'elle soit, ne fonctionne pas comme elle le devrait, la QoE va se voir influencée de manière négative. Dans le cadre de leur modèle de qualité des DIMEs Wu *et al.*, proposent une définition de la QoS selon trois facteurs. Ces facteurs sont issus directement des travaux de Steuer (1992) déjà cités dans ce manuscrit pour définir la présence virtuelle dans les 3D-IVEs. Lors de cette définition de la présence virtuelle, nous avons mis en exergue la similarité entre les facteurs de la présence virtuelle proposés par Steuer (1992) et ceux proposés par Biocca et Delaney (1995) sous l'appellation d'**immersion perceptuelle** que nous avons définis comme le pendant technique du sentiment de présence virtuelle (voir section 2.2.2.1). Il apparaît alors clairement que l'immersion perceptuelle et la QoS sont deux dénominations différentes des facteurs impactant la QoE dont fait partie intégrante la présence virtuelle (constat aussi en accord avec Slater (2009)). Étant donné ce lien, nous ne détaillerons pas de manière aussi précise que précédemment les facteurs de la QoS / immersion perceptuelle, nous allons en revanche nous intéresser aux variables de ces facteurs proposées par Wu *et al.* :

- **L'interactivité** : « mesure dans laquelle les utilisateurs peuvent participer à la modification de la forme et du contenu d'un environnement médiat en temps réel » [notre traduction] (Steuer, 1992) est découpée par Wu *et al.* selon :
 - **La vitesse**, temps de réaction de l'application entre le moment où l'utilisateur entreprend l'action dans le réel et où celle-ci est répercutée dans le virtuel ;
 - **la variété**, les possibilités offertes à l'utilisateur en ce qui concerne l'interaction avec l'environnement virtuel, par exemple le fait de pouvoir changer de point de vue (passer à la troisième personne), le fait de pouvoir interagir avec les objets, de modifier l'environnement lui-même ;
 - **la correspondance**, capacité de l'environnement virtuel à restituer les mouvements de l'utilisateur (à travers un clavier, une manette ou un outil de capture).
- **La vivacité sensorielle** : « richesse de représentation d'un environnement médiat » [notre traduction] (Steuer, 1992) découpée selon :
 - **Son étendue**, nombre de canaux sensoriels sollicités ;
 - **sa profondeur**, résolution ou acuité avec lesquelles ces canaux sensoriels sont sollicités (par exemple, la granularité d'un système haptique, la résolution des écrans utilisés, ou le taux de rafraîchissement de l'image (Barfield et Hendrix, 1995)).
- **La cohérence** : entre l'environnement médiat, les attentes et les représentations mentales de l'utilisateur. Découpée selon deux dimensions :
 - **Spatiale**, voir l'exemple de la gravité et des objets dans la section 2.2.2.1 ;
 - **temporelle**, le fait que le passage du temps semble cohérent entre les différents éléments de l'environnement médiat (sauf dans le cas où cela est explicitement dit à l'utilisateur) et que les différents éléments et utilisateurs semblent évoluer dans une même temporalité.

Si nous comparons les facteurs de l'immersion perceptuelle de Biocca et Delaney (1995) à ceux présentés ici dans la définition de la QoS de Wu *et al.* (2009), nous constatons que certains facteurs de l'immersion perceptuelle ne sont pas présents dans la modélisation de la QoS. En premier lieu la **conscience des interfaces** que nous pourrions néanmoins définir comme une variable **d'intrusivité** de la **vivacité sensorielle**. En effet, plus l'utilisateur est conscient d'être équipé de matériel, plus ce dernier y allouera d'attention sensorielle parasitant ainsi les stimuli proposés par l'application, l'environnement virtuel. Reste alors la **qualité visuelle** et le **réalisme comportemental** que nous ne pouvons inclure dans aucun des trois facteurs de la QoS, leur intégration dans cette modélisation demande leur ajout en tant que facteurs. Ainsi, la **qualité visuelle** pourrait être découpée selon deux variables, sa **finesse** représentant la qualité (résolution) des textures utilisées. Et selon sa **qualité graphique** ou beauté qui bien que faisant appel à la subjectivité de l'utilisateur reste néanmoins un facteur technique de l'environnement. Le **réalisme comportemental** quant à lui peut-être découpé selon le réalisme comportemental des **êtres présents** dans l'environnement (agents, avatars) et selon celui de **l'environnement** lui-même caractérisé par exemple par la présence de *shaders*¹⁷ permettant de restituer les effets de la profondeur de champ en fonction des fixations du regard de l'utilisateur ou de faire varier l'aspect des textures en fonction des aléas climatiques (aspect mouillé). Ces trois facteurs ont été intégrés dans la partie QoS du modèle de qualité des DIMEs de Wu *et al.* (2009) en **vert** sur la figure 2.15 afin de les différencier du modèle original.

FIGURE 2.15 – Modèle étendu de qualité des DIMEs de Wu *et al.* (2009).

Nous avons compris l'importance de la QoS dans le processus d'expérience utilisateur et le lien que celle-ci avait avec les facteurs de l'immersion perceptuelle précédemment définie. La dichotomie

17. Un *shader*, en français, nuanceur est un programme ou un script permettant de définir certains paramètres du rendu graphique d'une scène 3D. Par exemple l'absorption et la diffusion de la lumière, les réflexions et réfractions, le type d'ombrage. Ils permettent entre autres de créer des effets tels que du brouillard, du feu ou une étendue d'eau mouvante.

de Wu *et al.*, en plaçant la technologie comme une influence sur l'expérience utilisateur et non pas comme une composante de cette dernière, et les expérimentations menées dans le cadre de la validation de ce modèle de qualité des DIMEs (voir Wu *et al.* (2009, p.485)), nous permettent de dire qu'il est possible de mesurer l'impact d'un facteur particulier de la QoS sur la QoE. Alors que dire des interactions sociales vocales dans ce contexte ? Le fait de pouvoir ou de ne pas pouvoir se parler dans un environnement virtuel n'est-il pas un élément du **réalisme comportemental** des avatars dans le 3D-IVE ? En conservant une QoS constante et en ne faisant varier que la possibilité ou non de se parler entre utilisateurs, nous serions en mesure de comprendre son impact sur la QoE autrement dit sur l'expérience vécue par l'utilisateur.

Vis-à-vis de la modélisation de la QoE de Wu *et al.* (2009) nous constatons que, malgré la proposition d'un modèle de communication en DIME (voir figure 2.16), les facteurs sociaux de l'expérience utilisateur ne sont pas intégrés. Pour cette raison, nous allons détailler la QoE non pas au travers du modèle de Wu *et al.* mais de son extension par Pallot *et al.* (2013) les prenant en compte.

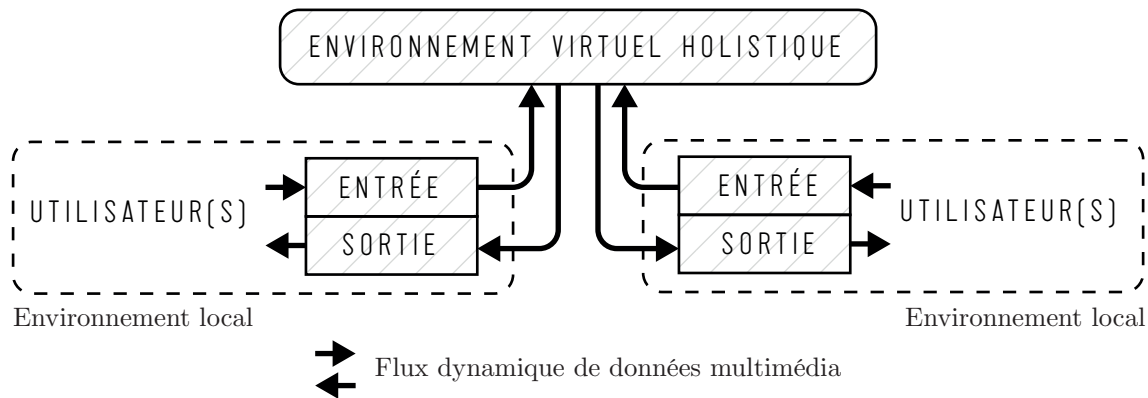


FIGURE 2.16 – Modèle de communication en DIMEs (Wu *et al.*, 2009).

2.3.2.2 Le modèle d'expérience utilisateur de Pallot et al. (2013), Expérience (Utilisateur) Immersive & Qualité de l'Expérience

Le modèle d'expérience utilisateur de Pallot *et al.* (2013) a été développé dans le contexte du projet européen 3D Live (The 3DLive Consortium, 2015). Ce projet avait pour objectif la création d'une expérience utilisateur commune à au moins deux personnes dans le contexte de la pratique d'un sport à distance (ski, golf) entre réel et virtuel, c'est-à-dire avec des utilisateurs dans le monde réel et des utilisateurs immergés dans un environnement virtuel. Le projet a donc vu entre autres la création d'une piste de ski virtuelle calquée sur une piste de ski réelle en Autriche. Les utilisateurs réellement sur la piste, équipés de lunettes de réalité augmentée étaient en mesure de voir les utilisateurs distants immergés dans l'environnement virtuel par une reconstitution 3D temps réel de la personne (scan 3D via trois *Kinect*). Les utilisateurs immergés dans le monde virtuel voyaient les utilisateurs du monde réel au travers d'un avatar. Ainsi, connectés au même environnement, les utilisateurs étaient en mesure de partager l'expérience du ski.

Traitant du partage d'une expérience commune, le projet nécessitait le développement d'un mo-

dèle d'expérience utilisateur prenant en compte les aspects sociaux de l'expérience utilisateur. C'est pourquoi, à la différence du modèle de qualité des DIMEs (Wu *et al.*, 2009), le modèle de Pallot *et al.* propose une division de l'**Expérience (Utilisateur) Immersive** suivant les processus mentaux **rationnel** et **expérientiel** qui la régissent. Le processus rationnel reprend la partie perceptions cognitives du modèle de Wu *et al.* (2009) tandis que le processus expérientiel repose sur les émotions et les intuitions. Cette dichotomie se base sur la théorie de la cognition-expérientielle autonome (*Cognitive-Experiential Self-Theory* ou CEST) de Epstein (1998), née de l'idée selon laquelle les humains utilisent deux processus mentaux séparés pour analyser les informations, un processus analytique-rationnel et un processus intuitif-expérientiel. Le processus rationnel est activé de manière délibérée et fait appel à la logique tandis que le processus expérientiel est autonome (comme l'est un réflexe) et fait appel à nos émotions, à nos intuitions. Ces deux processus sont indépendants mais fonctionnent de concert, en résultent nos comportements et nos pensées conscientes. Selon Norris et Epstein (2011), les deux systèmes présentent des avantages et des inconvénients uniques. Ainsi, le processus rationnel est bien supérieur au processus expérientiel en ce qui concerne la pensée abstraite, la compréhension des relations de cause à effet, quand il s'agit de faire preuve de patience, ou de planifier un avenir lointain. En revanche, le processus expérientiel lui est largement supérieur vis-à-vis des comportements quotidiens qui se produisent sans effort et donc en ne faisant appel qu'à très peu de nos ressources cognitives. D'après ce constat, il est impossible d'établir un ordre d'importance entre ces deux processus mentaux, ils ne sont pas utilisés aux mêmes fins et font montre d'un apport mutuel lorsqu'ils sont utilisés en synergie, en entraînant un impact positif sur les performances, la réflexion ou la créativité des utilisateurs (Novak et Hoffman, 2009 ; Hoffman et Novak, 2008 ; Norris et Epstein, 2002).

Nous allons à présent, suivant le modèle établi par Pallot *et al.* (2013) sur la base du travail de Wu *et al.* (2009), décrire chaque facteur de l'expérience (utilisateur) immersive ainsi que les facteurs de la QoE sur lesquels ces derniers ont un impact. Le modèle de Pallot *et al.* (2013) est présenté sur la figure 2.17.

***NOTA :** A début des paragraphes à venir les textes en **gras** font référence à la figure 2.17. A la manière d'un fil d'Ariane ils permettent de situer le facteur décrit dans le modèle de Pallot et al. (2013).*

Expérience (Utilisateur) Immersive ► Processus Expérientiel :

Le processus expérientiel est, comme nous l'avons dit, un processus intuitif, il comprend nos émotions, notre état d'esprit mais aussi les aspects sociaux de ces mêmes facteurs. En effet, une émotion partagée est différente d'une émotion ressentie seul, par exemple la peur est moins poignante lorsque l'on est pas seul face à cette dernière, la joie est un sentiment encore plus grisant lorsque l'on peut la partager, l'exprimer verbalement à une personne, expression qui aura une influence sur notre interlocuteur et sur nous-même (Hatfield *et al.*, 1994). Ces facteurs émotionnels sociaux occupent une grande place dans l'expérience utilisateur comme nous l'avons déjà évoqué au travers de la popularité des jeux vidéo en ligne ou de la place importante de l'image de marque d'un produit. La communication, l'expression de nos sentiments et états mentaux est un besoin bien plus qu'une envie (voir section 2.1).

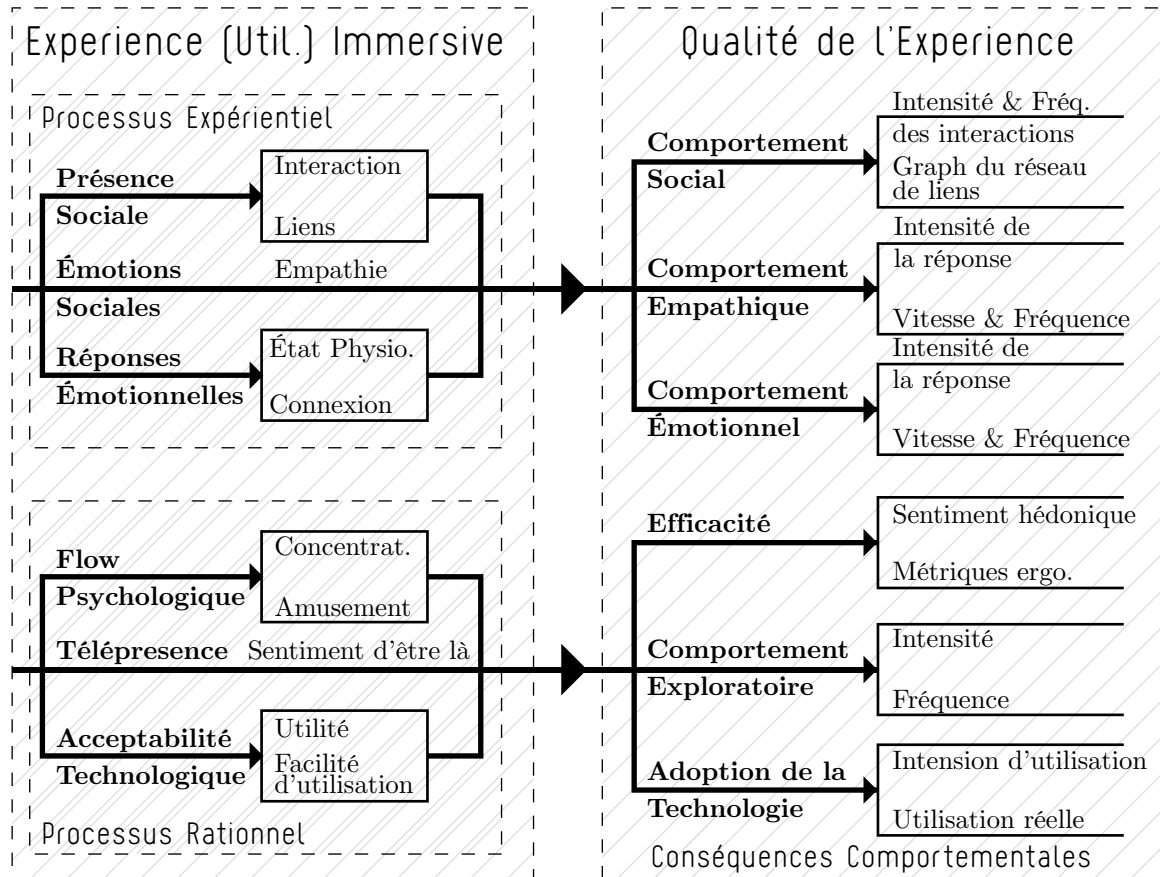


FIGURE 2.17 – Modèle d'expérience utilisateur théorique de Pallot *et al.* (2013) d'après Wu *et al.* (2009).

Dans le modèle présenté ici, le processus expérientiel est découpé selon trois facteurs :

► **Présence Sociale** : nous avons longuement disserté de la présence sociale dans les sections 2.2.2.2 & 2.2.2.3, sur sa nature multimodale et donc sur les facteurs qui, dans un 3D-IVE, conditionnent ce sentiment d'être présent avec d'autres êtres intelligents. Nous citerons simplement à nouveau Heeter (1992) afin de souligner l'importance de ce facteur de l'expérience (utilisateur) immersive : « *le principe de la présence sociale consiste simplement à dire que si d'autres personnes sont dans le monde virtuel, ceci est une preuve de plus que ce monde existe* » [notre traduction].

► **Présence Sociale** ► *Interactions & Liens* : les variables interactions et liens font respectivement référence aux interactions sociales produites entre les protagonistes immergés dans le même environnement et la familiarité entre ces derniers (voir le facteur **familiarité** de la présence sociale en page 33).

► **Émotions Sociales** : les émotions sociales sont des émotions nécessitant la représentation de l'état mental des autres personnes. Par exemple, l'embarras, la culpabilité, la honte, la jalousie ou encore l'amour. À ceci s'opposent les émotions basiques, comme la joie ou la colère qui ne nécessitent, quand elles sont dirigées vers une personne, que sa présence physique (réelle ou mentalement représentée). Cette connexion mentale mène comme nous l'avons précédemment expliqué à l'empathie qui tend à renforcer la connexion entre les individus (voir le facteur interdépendance émotionnelle de la

présence sociale en page 30).

► **Réponses Émotionnelles** : dans le contexte de l'expérience utilisateur en 3D-IVEs, les émotions peuvent être générées de plusieurs manières, par l'environnement lui-même, au travers de sa forme, de ses couleurs (textures) ou de son environnement sonore. Tout comme dans le monde réel, une ville abandonnée générera chez certaines personnes un sentiment de crainte, voire de peur, une ruelle sombre et mal famée, un sentiment de dégoût, de répulsion. Les environnements virtuels sont capables d'induire des émotions chez l'utilisateur allant de la joie à la tristesse comme en témoignent les travaux de Felnhofer *et al.* (2015), de Rodríguez *et al.* (2013) et de Baños *et al.* (2005). Ensuite, il est possible d'agir sur l'état émotionnel de l'utilisateur par la narration, évoluer dans les couloirs d'un hôpital psychiatrique virtuel désert est une chose, le faire alors qu'une voix conte les atrocités qui se sont produites entre ces murs va renforcer le sentiment de malaise et mettre l'utilisateur dans un état émotionnel favorisant l'occurrence de la peur (Geslin, 2013). Vis-à-vis des émotions générées par l'interaction sociale, nous avons déjà exprimé les difficultés inhérentes à leur restitution, en effet la capture des émotions par le biais des expressions faciales est possible mais sa restitution en environnement virtuel n'est pas aisée et est influencée par la QoS, en particulier par les facteurs que sont l'**interactivité**, la **vivacité sensorielle** et le **réalisme comportemental**. Dans le modèle d'expérience utilisateur que nous décrivons ici, les émotions ne sont pas directement connectées au sentiment de présence virtuelle, mais nous savons néanmoins que ces dernières, par leur présence et leur partage, ont un fort impact sur ce phénomène (Claudio *et al.*, 2015 ; Meehan *et al.*, 2002), ce qui implique par corrélation un impact sur les facteurs de la QoE corrélés au sentiment de présence virtuelle comme l'**efficacité** ou le **comportement exploratoire**. Cette implication à plusieurs niveaux des émotions dans l'expérience utilisateur fait qu'il est crucial dans le cadre d'une expérience partagée de pouvoir transmettre ces émotions que ce soit par la voix ou la restitution des expressions faciales, les utilisateurs doivent se sentir en présence d'un être non seulement intelligent mais aussi doté de sentiments.

► **Réponses Émotionnelles** ► *État Physiologique* : variable faisant référence aux réponses physiologiques de l'utilisateur à l'occurrence des émotions. Les méthodes de détection des émotions par la physiologie sont largement utilisées et dans l'ensemble sont fiables (Wen *et al.*, 2014 ; Monkaresi et Calvo, 2014 ; Powar, 2013 ; Chang *et al.*, 2013 ; Chittaro et Sioni, 2013 ; Rainville *et al.*, 2006) mais les difficultés principales résident encore une fois plutôt dans la restitution de ces émotions que dans leur captation.

► **Réponses Émotionnelles** ► *Connexion* : variable faisant référence, comme pour les émotions sociales, au niveau de connexion entre les utilisateurs (voir section 2.8) mais aussi à l'immédiateté de la transmission des émotions nécessaire à un retour spontané d'un utilisateur envers l'autre.

Expérience (Utilisateur) Immersive ► Processus Rationnel

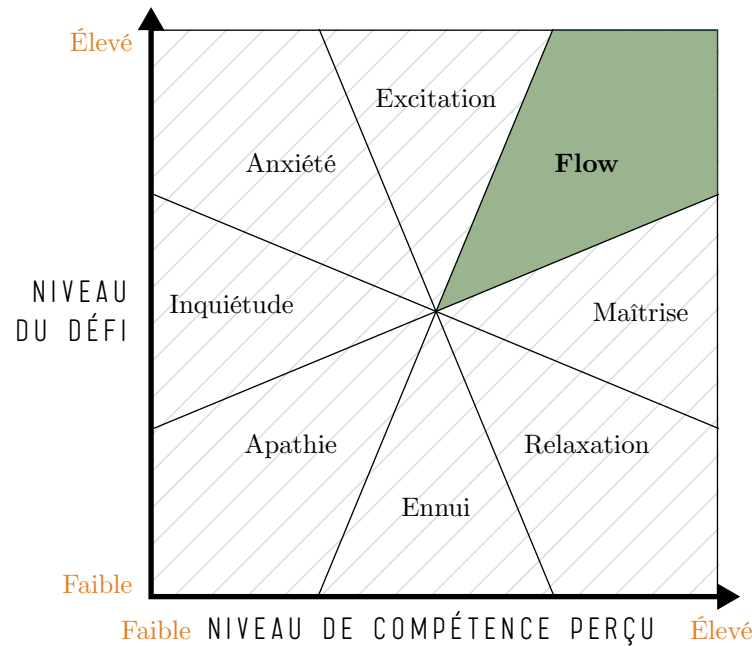
La partie rationnelle du modèle de Pallot *et al.* (2013) reprend l'intégralité de la partie **perceptions cognitive** du modèle de qualité des DIMEs de Wu *et al.* (2009).

► **Flow Psychologique :** le *flow* (ou flux) psychologique représente un état psychologique dans lequel une personne est totalement investie dans ce qu'elle fait, lui procurant ainsi un fort sentiment d'amusement et de maîtrise. Ce concept fut défini en premier lieu par Mihaly Csikszentmihalyi en 1990¹⁸ puis modélisé en 1997 sous une forme très largement utilisée aujourd'hui (Csikszentmihalyi, 1997), présentée en figure 2.18. Csikszentmihalyi définit le *flow* comme étant l'équilibre entre le niveau de défi proposé par l'activité, autrement dit sa difficulté, et le niveau de compétence perçu (ou estimé) par l'utilisateur. Le modèle propose aussi, outre l'état de *flow*, sept états mentaux résultants d'un déséquilibre entre les deux facteurs précédemment cités comme l'apathie, l'anxiété ou l'excitation. Placer l'utilisateur dans un état de *flow* fait aussi appel à un autre élément que nous avons identifié précédemment comme étant impliqué dans la partie **immersion psychologique** du sentiment de présence virtuelle (voir Lombard et Ditton (1997) ; Biocca et Delaney (1995) dans la section 2.2.2.1), **la concentration**. Csikszentmihalyi définit alors huit éléments nécessaires à l'accès à l'état de *flow* : (1) la tâche proposée doit pouvoir être réalisée ; (2) l'utilisateur doit être en mesure de se concentrer sur la tâche ; (3) la concentration est possible, car la tâche a un objectif précis ; (4) la concentration est possible, car la tâche offre un retour immédiat (visibilité de la progression en temps réel) ; (5) l'utilisateur doit avoir la capacité d'exercer un certain contrôle sur les actions (notion d'interactivité) ; (6) l'utilisateur doit ressentir une profonde implication occultant la conscience des frustrations de la vie quotidienne, et ce sans efforts ; (7) l'utilisateur doit perdre ses considérations égocentrées, mais son ego doit se voir renforcé par la suite (notion de satisfaction) ; et (8) la perception du temps est altérée. La logique du *flow* peut s'appliquer à de nombreuses activités, la lecture, le sport, l'apprentissage. Des modèles spécifiques aux contextes étudiés ont donc vu le jour, par exemple dans le cadre de la pratique du jeu-vidéo (Sweetser et Wyeth, 2005) ou encore dans le cadre professionnel (où le *flow* est alors appelé « *workflow* »)(Vass *et al.*, 2002). Il nous faut retenir du *flow* qu'un bon équilibre entre les compétences de l'utilisateur et le défi proposé favorise la **concentration** ainsi que l'**amusement** qui à leur tour auront un impact sur la perception du temps de l'utilisateur et donc sur la qualité de l'expérience globale. Il est aussi important de noter l'interpénétration de concept entre le *flow* et la présence virtuelle, puisque nous avons vu précédemment que la **concentration** et l'**amusement** avaient tous deux un fort impact sur la présence virtuelle. Nous pouvons donc dire qu'un bon équilibre vis-à-vis du *flow* favorisera le sentiment de présence virtuelle dans le contexte des 3D-IVEs.

► **Téléprésence :** tout comme pour la présence sociale nous avons largement abordé le sujet de la présence virtuelle appelée ici téléprésence (voir section 2.2.2.1). Nous rappellerons simplement qu'elle représente le sentiment « *d'être là* » et dans le contexte des 3D-IVEs, de ressentir que nous sommes plus présents physiquement dans l'environnement virtuel que dans l'environnement réel.

► **Acceptabilité Technologique :** dans la partie « **perceptions cognitives** » du modèle de qualité des DIMEs de Wu *et al.*, le facteur **acceptation technologique** est basé sur les travaux de Davis (1985). Ce facteur a pour objectif d'examiner les perceptions et les attitudes de l'utilisateur générées par l'utilisation d'une technologie, ce au travers de deux variables, l'**utilité perçue** et la **facilité d'utilisation perçue**. Dans le modèle de Davis, l'**utilité perçue** représente la mesure dans

18. Voir l'ouvrage : « *Flow: The Psychology of Optimal Experience* » — Mihaly Csikszentmihalyi, 1990.

FIGURE 2.18 – Modèle du *flow* psychologique de Csikszentmihalyi (1997).

laquelle l'utilisateur estime que l'utilisation d'une technologie lui permet d'accroître ses performances et est donc corrélée aux objectifs de l'utilisateur. Mais dans notre contexte des 3D-IVEs ou dans celui des jeux vidéo, l'objectif de l'utilisateur peut simplement être de prendre du plaisir, de s'amuser, il est alors plus pertinent de dire que **l'utilité perçue** correspond à la propension de la technologie utilisée de permettre à l'utilisateur d'atteindre l'objectif fixé lors de la décision de l'utilisation. **La facilité d'utilisation** perçue, quant à elle, représente le degré d'intuitivité de la technologie, le niveau d'effort à déployer pour son utilisation. Elle sera fortement influencée par **la correspondance** du facteur **interactivité** de la QoS puisque c'est ce facteur technique qui définira **le degré de contrôle** possible ainsi que **l'immédiateté du contrôle** de l'utilisateur sur l'environnement et ses artefacts (ici encore, on constate une interpénétration avec le concept de présence virtuelle comme défini par Biocca et Delaney (1995), voir la section 2.2.2.1).

Qualité de l'Expérience

Nous avons vu les deux processus composant l'Expérience (Utilisateur) Immersive représentant l'expérience telle quelle se dessine de manière instantanée dans l'esprit de l'utilisateur. La QoE représente les conséquences comportementales, émotionnelles que va avoir l'expérience immersive sur l'utilisateur.

► **Comportement Social** : le comportement social représente les réactions des utilisateurs vis-à-vis des contenus sociaux du 3D-IVE, que ces derniers soient générés par un autre humain par le biais d'un avatar ou directement par l'environnement virtuel. Il s'illustre donc par les réactions générées non seulement par les interactions sociales vocales et non verbales, mais aussi par le comportement des êtres sociaux en présence de l'utilisateur.

► **Comportement Empathique :** le comportement empathique représente la réaction de l'utilisateur résultant de l'expression d'une émotion sociale ou primaire par un autre individu lors d'une interaction sociale. Ce comportement s'illustre par exemple par l'émission d'un message d'encouragement ou la proposition d'une aide. Ces comportements empathiques peuvent être capturés (par la vidéo ou le son) de la même manière que les autres comportements sociaux. Notons ici qu'à la confluence du comportement social et empathique se trouve la translucence que nous avons décrite dans la section 2.1.2.3. Ce phénomène comportemental est bien dû à l'empathie, puisque c'est en nous plaçant dans la position de l'autre que nous sommes capables d'adapter notre comportement. Il fait aussi appel à notre comportement social, puisque largement régi par les normes et codes de la société dans laquelle nous évoluons.

► **Comportement Émotionnel :** le modèle de qualité des DIMEs de Wu *et al.* (2009) décrit le comportement émotionnel comme l'ensemble des réponses émotionnelles de l'utilisateur soit dans un contexte individuel, lors de son immersion et de ces interactions avec l'environnement virtuel, soit dans un contexte social durant une immersion ou une activité partagée.

► **Efficacité :** le modèle de qualité des DIMEs de Wu *et al.* (2009) décrit l'efficacité comme étant l'accroissement des performances d'un individu vis-à-vis d'une tâche et propose sa mesure de manière objective et subjective (voir le chapitre Méthodologie en page 61). Dans le modèle dérivé de Pallot *et al.* (2013), on constate que ces variables objectives et subjectives cèdent la place à l'aspect hédonique (plaisir ressenti) et l'aspect ergonomique (effort mis en jeu). En effet, le plaisir éprouvé lors d'une réussite dans la complétion d'une tâche, par un accroissement de la confiance en soi aura un impact sur les futures performances de l'utilisateur sur la même tâche ou sur le même type de travaux. L'aspect ergonomique est lui tout à fait similaire avec la variable objective proposée par Wu *et al.*.

► **Comportement Exploratoire :** le modèle de qualité des DIMEs de Wu *et al.* (2009) décrit le comportement exploratoire comme les comportements motivés par la curiosité de l'utilisateur à explorer spontanément la technologie en présence (dans notre contexte, l'environnement virtuel) sans plans ou objectifs précis. Ces comportements exploratoires peuvent être évalués de manière objective au travers des actions effectuées hors du cadre défini dans le contexte d'une tâche, ou vis-à-vis des déplacements de l'utilisateur dans des zones d'un environnement virtuel non directement corrélées aux objectifs ou à la narration de l'environnement. À cela, Wu *et al.* ajoutent des mesures quantitatives subjectives en donnant l'exemple suivant tiré d'un questionnaire d'évaluation des comportements exploratoires sur internet : « *je clique souvent sur un lien juste par curiosité* » et « *surfer sur internet pour voir ce qu'il y a de nouveau est une perte de temps* » (échelle-inverse). Là encore, Pallot *et al.* proposent une construction légèrement différente par une évaluation au travers de l'**intensité** que l'on pourrait traduire par la durée de la « divagation » et la **fréquence** de l'occurrence de ces parenthèses exploratoires vis-à-vis des objectifs ou du contexte.

► **Adoption Technologique :** le modèle de qualité des DIMEs de Wu *et al.* (2009) décrit l'adoption technologique selon deux variables, l'une subjective, l'**intention d'utilisation** et l'autre objective, l'**utilisation réelle**. Ces deux variables sont directement influencées par le facteur **acceptabilité**

technologique de l'expérience (utilisateur) immersive (processus rationnel). L'**intention d'utilisation** est une variable purement subjective mais qui est néanmoins considérée comme une mesure importante de l'expérience utilisateur (Wu *et al.*, 2009, p.484) et comme l'indiquent Wu *et al.* en se basant sur la « *théorie du comportement planifié* » (Ajzen, 1991), elle permet aussi de prédire dans une certaine mesure l'**utilisation réelle** si bien qu'elle lui est parfois substituée dans certaines études. L'**utilisation réelle**, objective, pose quant à elle un problème d'ordre logistique, son évaluation demande le suivi des utilisateurs sur une durée de plusieurs mois pour être utile, ce qui n'est pas forcément aisé à mettre en place tant du point de vue des chercheurs que de celui des utilisateurs.



L'expérience utilisateur, comme nous avons pu le voir au travers de la description des modèles de Wu *et al.* (2009) et de Pallot *et al.* (2013), est composée de multiples facteurs qui, malgré la taxonomie proposée, présentent des interpénétrations et des impacts entre facteurs, comme par exemple le fait que la concentration incluse dans le *flow* psychologique a également un grand impact sur le sentiment de présence virtuelle comme nous l'avons vu dans la section 2.2.2.1. De plus, sa nature dynamique et dépendante du contexte rend difficile la détermination d'une hiérarchie des facteurs plus ou moins influents comme en témoigne l'évaluation de l'impact de la QoS sur la QoE proposée par Wu *et al.* (2009, p.489). Dans cette évaluation, les résultats sont décrits comme issus pour la plupart de facteurs liés au contexte plus qu'à une réelle constante de l'impact d'un facteur sur l'autre. Néanmoins, les taxonomies proposées nous donnent une base de travail vis-à-vis des mesures possibles des facteurs de l'expérience utilisateur comme nous le verrons par la suite dans le chapitre dédié à la méthodologie appliquée pour nos travaux. Le choix de ces deux modèles est aussi basé sur le fait qu'ils restent, tout en définissant un cadre d'utilisation (DIMEs, 3D-IVEs), très généralistes dans le sens où ils ne sont pas liés à un contexte précis comme l'éducation (Shin *et al.*, 2013) ou le divertissement (Cheng *et al.*, 2014), comme cela est souvent le cas dans la littérature. Cette nature généraliste offre une plus grande liberté de choix dans la définition des tâches et interactions proposées lors de la conception d'un système de réalité virtuelle.

2.4 Résumé de l'état de l'art

Afin de résumer la revue de littérature menée dans le cadre de cette thèse de doctorat, nous allons reprendre les questions de recherche soulevées au début de ce manuscrit (voir section 1.2) afin de nous assurer que nous avons apporté les réponses nécessaires à la continuation de ces travaux.

Q.1.3: *Quelles sont les composantes de l'interaction sociale verbale, les stimulations induites et les réponses produites vis-à-vis de l'interlocuteur ?*

Nous avons tout d'abord cherché à comprendre les interactions sociales dans leur globalité et avons, au travers des travaux de Maslow (1943), énoncé le fait que nous interagissons socialement pour répondre à une variété de besoins inhérents à notre condition d'humain. Ces besoins nécessitent

pour être comblés que nous interagissions avec nos semblables et plus encore que nous les influençons par nos interactions dans leurs actions, leurs pensées, leurs sentiments. Nous nous sommes par la suite intéressés au fonctionnement global de l'interaction sociale au travers des boucles d'encodage / décodage de Schramm (1954) mais aussi des protocoles d'ouverture, de transition et de rupture de l'interaction sociale (Picard, 2001).

Nous avons alors fait la dichotomie des interactions sociales vis-à-vis de leurs composantes verbales, vocales et non verbales, en portant plus particulièrement notre intérêt sur l'interaction sociale vocale en raison des contraintes inhérentes à l'exploitation du non verbal dans le contexte des 3D-IVEs (voir 2.5). Néanmoins, les interactions sociales non verbales ne sont pas totalement exclues des 3D-IVEs comme nous avons pu le constater vis-à-vis de la proxémie et du phénomène de translucence que nous pouvons considérer comme une forme de d'interaction sociale par le comportement. Par la suite, l'étude des phénomènes physiologiques permettant la création des interactions sociales vocales (hauteur, intensité, timbre et intonation) nous a permis une investigation des technologies permettant de communiquer verbalement à distance avec plus ou moins d'acuité (SMS, messageries instantanées, vidéoconférence, etc...), nous menant au centre de notre contexte dans les environnements virtuels.

Aux regards des travaux menés par Jeremy Bailenson, Rosanna Guadagno, Maia Garau ou Masahiro Mori, nous avons souligné les différences entre agents et avatars et surtout les prérequis d'une interaction sociale convaincante en environnement virtuel. Nous avons identifié ceux-ci comme étant une bonne cohérence entre anthropomorphisme, réalisme visuel et surtout, réalisme comportemental. Par l'étude de ces facteurs, nous avons aussi mis en lumière les réactions émotionnelles et comportementales que peuvent induire les interactions sociales sur nos interlocuteurs et nous avons découvert la possibilité et la véridicité de ces réactions lors d'une immersion en 3D-IVE.

Il reste difficile à affirmer qu'il existe un rapport de force entre les composantes de l'interaction sociale (**Q.1.3.1**), car nous avons vu que la structure de notre cerveau fait que nous analysons en premier le langage verbal pour ensuite nous intéresser au non verbal afin de récolter des informations supplémentaires. En revanche, le pouvoir des émotions est clairement établi, et nous pouvons donc dire qu'un message chargé d'une émotion, quelle qu'elle soit, aura plus de portée qu'un message neutre que ce dernier soit de l'ordre du verbal ou du non verbal.

Q.1.1 : *Quels sont les facteurs mesurables des modèles existants de l'expérience utilisateur ?*

Après un bref historique de l'expérience utilisateur décrivant sa naissance par la rencontre du concept d'utilisabilité et des sentiments hédoniques de l'utilisateur, nous nous sommes penchés sur sa nature. Celle-ci est soumise à des variations dues à l'utilisateur lui-même, à son état d'esprit, à ses expériences passées (Kankainen, 2002), au contexte mais aussi à l'accumulation de plusieurs expériences afin de constituer une expérience globale d'un moment de vie (Forlizzi et Battarbee, 2004) dans lequel l'utilisateur aura pu, par exemple, être immergé dans un 3D-IVE.

Par la suite, nous avons focalisé notre étude sur deux modèles en particulier, le modèle de qualité des DIMEs de Wanmin Wu et le modèle d'expérience utilisateur de Marc Pallot en étant issu, que

nous avons choisis pour leur application au contexte des système interactifs et pour leur cohérence avec les phénomènes de présence virtuelle et de présence sociale. Ces modèles proposent tous deux une vision centrée sur l'expérience vécue plutôt que sur l'application ou le système amenant à vivre l'expérience. Et, par leur dichotomie entre facteurs techniques (QoS) et facteurs humains (QoE), ils offrent une vision théorique des conséquences des choix de conception sur l'expérience utilisateur (chaîne de causalité QoS-QoE). La description des facteurs de ces modèles nous a permis de mettre en exergue les similarités entre les facteurs décrits comme influant sur les phénomènes de présence virtuelle, de présence sociale et d'expérience utilisateur ainsi que la découverte des interpénétrations de certains facteurs entre eux (le fait par exemple que la concentration impact non seulement le *flow* psychologique mais aussi la présence virtuelle). Nous avons aussi exprimé le fort potentiel des émotions personnelles ou sociales dans le cadre de la création d'une expérience utilisateur de qualité.

Nous avons établi, au travers de la description du modèle de Wu *et al.* étendu par Pallot *et al.*, que certains des facteurs sont de nature objectives et donc mesurables de manière objective et subjective (**Q.1.2**). Ce constat a été étayé par la revue des phénomènes de présence virtuelle et sociale pour lesquels des méthodes de mesures sont proposées, propos que nous détaillerons dans le chapitre consacré à la méthodologie de ce manuscrit.

Reste la question du rapport de force entre facteurs (**Q.1.1.1**) : le phénomène d'expérience utilisateur étant variable et fortement soumis au contexte, un rapport de force est sans doute présent mais il est difficile de définir une constante car, selon le contexte d'utilisation, le rapport pourrait être très modifié voire inversé. Néanmoins, là encore, nous sommes conscients de la puissance des émotions et des sentiments hédoniques de l'utilisateur. Ainsi, avec un contenu proposant la transmission à l'utilisateur d'émotions fortes, ces dernières amèneront à une forme de plaisir (se faire peur dans un certain contexte est en soi très amusant), ce qui menera à une expérience utilisateur très satisfaisante.

La revue de littérature conduite est encourageante quant à l'existence d'un impact de l'interaction sociale verbale sur l'expérience utilisateur en 3D-IVE. Malgré le manque de telles démarches dans la littérature, le travail mené jusqu'alors nous a permis de nous assurer qu'au travers de la dichotomie de facteurs proposée par les modèles de Wu *et al.* (2009) et de Pallot *et al.* (2013), nous serions en mesure d'évaluer l'expérience utilisateur. Le point crucial réside à présent dans les moyens à mettre en place pour permettre cette évaluation des facteurs clés de l'expérience utilisateur (en accord avec le contexte proposé) et de comparer ses facteurs vis-à-vis de la présence ou non d'interactions sociales vocales.

2.5 Hypothèses et poursuite des travaux

Avant d'exposer les hypothèses émissent à la suite de notre revue de littérature en vue de la continuité de ce travail, il nous est apparu comme nécessaire d'apporter des éclaircissements vis-à-vis de l'orientation de nos travaux sur l'interaction sociale vocale. Outre notre constat par les travaux de Delamarre (2011) du fait que notre structure neuronale priorise la lecture du langage verbal,

cette orientation est très largement liée aux contraintes inhérentes à notre domaine d'investigation, les environnements virtuels immersifs. En effet, les éléments non verbaux de l'interaction sociale, telle que la reconnaissance faciale des émotions, sont, dans le cadre de nos travaux, très difficiles à utiliser. Dans notre contexte, l'immersion en 3D-IVE requiert l'utilisation de matériels tels que les casques immersifs ou les lunettes de stéréoscopie. Ces derniers occultent tout ou partie du visage de l'utilisateur, rendant la détection du mouvement des sourcils ou de l'ouverture des yeux difficile à mettre en place. Cependant, ces limitations tendent à disparaître : nous avons vu apparaître en 2014 et 2015 plusieurs solutions de casques immersifs proposant de l'oculométrie¹⁹ comme le *FOVE*²⁰ dont les premières versions dédiées aux développeurs devraient être distribuées au printemps 2016. Une proposition à faible coût de Stengel *et al.* (2015) a été présentée lors de la conférence ACM Multimedia 2015.

Vis-à-vis des gestes, ou autres indications non verbales produites par le corps, nous étions là aussi confrontés à certaines contraintes très pragmatiques. Travailler sur les 3D-IVEs implique de l'interaction avec le monde virtuel, ce qui nécessite l'utilisation d'interactions digitales (avec les mains : souris, clavier, manette, bras haptique, etc...) sans quoi les possibilités sont relativement réduites. L'utilisation de ces systèmes d'interaction réduit de manière drastique les gestes de l'utilisateur et conditionne aussi sa posture, ses attitudes. De leur côté, les systèmes d'interaction sans les mains comme par exemple les casques EEG⁹ demandent de l'entraînement et ne sont pas aussi flexibles qu'une simple manette quant aux interactions possibles avec l'environnement. Qu'en est-il alors de la posture du corps, de la manière de se déplacer ou de la tenue vestimentaire ? Ces trois éléments sont tout à fait enregistrables via une *Kinect*[®], la complexité se situe plutôt dans la restitution. La recomposition fidèle d'un humain sur un avatar virtuel en temps réel a été accomplie dans la cadre du projet *3D Live* (The 3DLive Consortium, 2015) mais demande une grande puissance de calcul et risque donc de se faire au détriment du taux de rafraîchissement, ce qui impacte fortement la QoS. Ou dans le cadre du projet *Real Virtuality* (Chague et Charbonnier, 2015) qui implique l'utilisation d'un scanner 3D très encombrant et très coûteux. De plus, ici encore, dans le cas de casques immersifs, leur utilisation en position debout peut s'avérer dangereuse, car le conflit entre les stimuli visuels et ceux reçus par l'oreille interne peuvent provoquer un déséquilibre puis une chute. Des produits comme le *Virtuix Omni*²¹ permettent l'utilisation d'un casque immersif en mouvement mais la structure entrerait alors en conflit avec le système de détection via la *Kinect*[®]. Malgré ces contraintes fortes, l'ensemble de notre revue de littérature des interactions sociales, de l'expérience utilisateur et des liens l'unissant aux phénomènes de présence virtuelle et de présence sociale saura être pleinement mise à profit lorsque les verrous techniques cités n'en seront plus. Au regard de ces constats, nous nous évertuerons alors à proposer une démarche cohérente avec l'étude des interactions sociales vocales uniquement.

Au cours de notre revue de littérature, nous avons pu constater les similarités entre présence virtuelle et présence sociale. Ces deux concepts présentant des gammes de facteurs proches, à la diffé-

19. Oculométrie : technique de suivi de l'œil

20. Voir : <http://www.getfove.com>

21. Voir : <http://www.virtuix.com>

rence que la présence sociale est dépendante des autres êtres partageant l'environnement virtuel et de leurs représentations virtuelles, tandis que la présence virtuelle est plus liée à l'environnement virtuel. Néanmoins, les travaux de Garau *et al.* (2005), Sweetser et Wyeth (2005) et de Cole et Griffiths (2007) décrivent les effets positifs de la présence sociale sur la présence virtuelle (dans le cas de l'étude menée par Cole et Griffiths (2007), l'impact se traduit par une augmentation de l'amusement). La présence d'interactions sociales (verbales ou non verbales) est un facteur déterminant de la présence sociale permettant son accroissement. Par corrélation, une augmentation de la présence sociale doit mener à une augmentation de la présence virtuelle et donc à une amélioration de l'expérience utilisateur. Nous pouvons alors poser l'hypothèse suivante :

H.1 : *Les interactions sociales vocales ont un impact bénéfique sur l'expérience utilisateur en 3D-IVE.*

Afin de valider ou d'infirmer cette hypothèse, nous avons mis en place une étude empirique nous permettant d'évaluer l'expérience utilisateur au sein d'un 3D-IVE dans deux contextes. L'un supportant les interactions sociales vocales, l'autre les interdisant. L'objectif de cette configuration à deux contextes est d'être en mesure d'opérer une comparaison entre les évaluations de l'expérience utilisateur afin d'observer les possibles différences. Si nous observons des différences significatives, nous serons alors en mesure de les imputer à la possibilité de communiquer vocalement.

Notre étude empirique devra aussi tester une seconde hypothèse liée à la dépendance à la nature de la tâche (coolaborative / coopérative) de l'impact de l'interaction sociale vocale sur l'expérience utilisateur. L'objectif ici est d'observer s'il existe des similitudes de résultats vis-à-vis de l'impact de l'interaction sociale vocale sur l'expérience utilisateur dans ces deux modes d'activité. Il serait alors possible d'admettre que l'impact de l'interaction sociale vocale sur tel ou tel facteur peut-être généralisé. Sachant que l'expérience utilisateur est très fortement soumise au contexte, que la moindre variation dans l'activité proposée ou dans la manière d'interagir avec le 3D-IVE peut fortement la modifier, nous formulons l'hypothèse suivante, que nous éprouverons grâce aux différentes expérimentations de notre étude empirique :

H.2 : *L'impact de l'interaction sociale vocale sur l'expérience utilisateur en 3D-IVE est dépendant de la nature de l'activité proposée.*

La validation ou l'infirmer de deux hypothèses présentées nous permettra de répondre à l'objectif de ce travail de thèse, offrir une meilleure compréhension de l'impact de l'interaction sociale vocale sur l'expérience utilisateur lors de l'immersion dans un 3D-IVE partagé. Ainsi que par corrélation un apport à la connaissance générale de l'expérience utilisateur en 3D-IVE.

Chapitre 3

Méthodologie

*« La science n'est rien que le bon sens et
un raisonnement sain. »*

Stanislas Leszczynski
Roi de Pologne - 1677-1766

Sommaire

3.1	Typologies des méthodes	62
3.1.1	Méthodes quantitatives	62
3.1.2	Méthodes qualitatives	64
3.2	Croiser les données par triangulation	66
3.3	Facteurs mesurés et moyens de mesure	67
3.4	Méthodologie d'étude empirique, comparaisons et itérations	75
3.5	Résumé du chapitre Méthodologie	76

Ce chapitre présente la méthodologie appliquée lors de notre étude empirique visant à déterminer l'impact des interactions sociales vocales sur l'expérience utilisateur en 3D-IVE. Comme nous l'avons décrit dans le chapitre précédent, l'expérience utilisateur varie en fonction des individus et du contexte. Dégager des résultats représentatifs est alors un travail demandant de l'exhaustivité dans le type de mesures mises en place. C'est pourquoi nous avons opté pour une méthode de triangulation entre des mesures quantitatives et qualitatives, objectives et subjectives. Bien sûr, comme nous l'avons précédemment évoqué, les facteurs de l'expérience utilisateur ne sont pas tous mesurables par tous ces types de mesure, néanmoins, conformément à l'interpénétration fréquemment observée entre les facteurs et selon les méthodes utilisées, il est possible d'obtenir des résultats permettant une interprétation, vis-à-vis de l'impact des interactions sociales vocales sur l'expérience utilisateur en 3D-IVE. Ensuite, nous détaillerons les mesures possibles des facteurs de l'expérience utilisateur proposées par le modèle de Wu *et al.* (2009) et son extension par Pallot *et al.* (2013) et l'approche comparative mise en place afin de mettre en exergue les différences entre une expérience en 3D-IVE vécue avec ou sans la possibilité de communiquer vocalement. Nous clôturerons ce chapitre par une explication de la méthode itérative qui nous a permis, en fonction des résultats de la première expérimentation, de mettre en place et de mener les deux suivantes.

3.1 Typologies des méthodes

3.1.1 Méthodes quantitatives

Afin de définir la méthodologie mise en place pour notre étude, nous nous sommes en premier lieu basés sur notre revue de littérature : il en ressort qu'une grande partie des études empiriques que nous avons analysées s'appuient sur des méthodes quantitatives, autrement dit des méthodes générant des données numériques ou des informations qui peuvent être converties en chiffres, permettant par la suite une analyse mathématique (statistique) (Creswell, 2009). Les méthodes quantitatives sont considérées comme fondamentalement objectives et utilisées dans le cadre de deux principaux types d'étude. En premier lieu, les études descriptives, lorsque l'on cherche à décrire une situation. Par exemple, lors d'une élection, la question : « *combien de personnes ont voté pour le candidat X ?* » ne peut trouver de réponse que par le quantitatif, compter le nombre de voix. Il en va de même pour la mesure de l'évolution de la natalité d'un pays : on décrit simplement une évolution numérique au travers d'un laps de temps fini.

Sonder une population sur leurs opinions, leurs comportements ou leurs attitudes permet d'obtenir des résultats quantitatifs à partir de données subjectives codées numériquement (e.g. : « *Les transports dans l'agglomération lavalloise sont : (A) Bons - (B) Passables - (C) Mauvais* ». C'est dans ce cadre que s'inscrit aussi la définition de questionnaires tendant à évaluer le niveau de présence virtuelle ressentie par les utilisateurs lors d'une immersion en 3D-IVE (Witmer et Singer, 1998). Les données alors recueillies sont analysées par le ou les protocoles statistiques adéquats vis-à-vis du nombre de participants et du type d'échantillons (Sukamolson, 2007).

Le second type d'étude auquel s'applique les méthodes quantitatives sont les études explicatives, lorsque l'on désire expliquer un phénomène ou valider / infirmer une hypothèse. Nous pouvons par exemple citer le travail de Jeremy Bailenson *et al.* (2006) visant à comprendre l'impact des niveaux de réalisme comportemental et visuel (facteurs variables) d'agents sur les sentiments de présence virtuelle, de présence sociale ou encore sur les comportements des utilisateurs. Dans l'étude citée, les utilisateurs sont divisés en plusieurs échantillons qui seront exposés à des agents présentant des niveaux de réalisme visuel et comportemental différents. Ces utilisateurs rempliront par la suite des questionnaires identiques qui, une fois analysés par l'étude statistique, permettront de mettre en lumière les différences entre les échantillons et donc les éventuels effets des facteurs variables (Sukamolson, 2007).

De part la définition ISO de l'expérience utilisateur, celle-ci comporte en majorité des facteurs de nature subjective qui la rendent changeante d'un utilisateur à un autre. Dans le cas de l'utilisation de questionnaires à échelle de Likert¹ (Likert, 1932) ou bipolaire sémantique², le même questionnaire, soumis à deux utilisateurs ayant vécu exactement la même expérience ne donnera pas forcément les mêmes résultats du fait de la subjectivité de certains facteurs. Selon l'état d'esprit, les convictions ou les expériences passées, les réponses données pourront donc largement différer. Ce type de mesures est soumis à l'appréciation spécifique de l'utilisateur. Il est alors communément admis que plus l'échantillon est important, plus les résultats obtenus après analyse statistique seront robustes. Il est assez difficile de définir le nombre minimal de participants nécessaire : dans la littérature étudiée, ce nombre varie entre une vingtaine et plusieurs centaines de participants. Il est aussi pertinent de penser que plus le phénomène étudié est subjectif, complexe et variable (comme l'est l'expérience utilisateur), plus le nombre de participants doit être important. Clôturons cette digression en identifiant les deux facteurs déterminant le nombre de participants à une étude empirique : le phénomène étudié ou l'hypothèse à valider / infirmer et la capacité des chercheurs à atteindre et réunir une large population, nous reviendrons sur ce point dans la suite de ce chapitre.

Nous souhaitons aussi parler d'un autre type de méthode quantitative, les mesures physiologiques prises sur les utilisateurs. Ce type de mesure n'est pas là encore totalement objectif. En effet, si l'on enregistre l'activité cardiaque et respiratoire, la conductance de la peau (Rainville *et al.*, 2006 ; Collet *et al.*, 1997) ou encore les clignements oculaires (Chittaro et Sioni, 2013) afin de déterminer les émotions ressenties par un utilisateur, les données recueillies sont objectives, en revanche l'interprétation qui en est faite reste subjective comme l'indique d'ailleurs l'utilisation du terme *interprétation*³. Cela, même si, comme dans l'exemple donné, les organes engagés dans le processus ont été le sujet de très nombreuses études, ce qui apporte ainsi une certaine robustesse aux analyses. Il est néanmoins nécessaire de rester vigilants quant aux conclusions à en tirer. Ceci est encore plus vrai dans le cadre de l'étude du cerveau dont — malgré de longues années d'observation — nous connaissons encore assez mal le fonctionnement. Il est alors préférable de rester dans l'observation et dans la corrélation

1. Échelle de Likert : échelle de jugement très répandue dans les questionnaires psychométriques par laquelle la personne interrogée exprime son degré d'accord ou de désaccord vis-à-vis d'un énoncé sous la forme d'une affirmation.

2. Échelle bipolaire : contrairement à l'échelle de Likert, les pôles d'une échelle bipolaire peuvent représenter autre chose que le niveau d'approbation, par exemple pour l'évaluation d'un état émotionnel, 0 = Tristesse et 5 = Joie.

3. Interpréter : donner à des propos, à un événement, à un acte telle signification, les comprendre en fonction de sa vision personnelle : « *J'interprétais son silence comme une acceptation* ». (Larousse 2015)

entre observations passées que de déduire un quelconque résultat arrêté.

3.1.2 Méthodes qualitatives

En marge des méthodes quantitatives génératrices de chiffres et de statistiques, nous trouvons les méthodes qualitatives qui mettent l'accent sur la collecte de données principalement verbales. La définition suivante donne une vision assez globale de ce qu'est la recherche qualitative et de ses objectifs : *« la recherche qualitative est une activité située qui localise l'observateur dans le monde. Elle se compose d'un ensemble de pratiques et matériels interprétatifs qui rendent ce monde visible. Ces pratiques [...] transforment le monde en une série de représentations comprenant des notes de terrain, des interviews, des conversations, des photographies, des enregistrements et des notes personnels. À ce stade, la recherche qualitative implique une approche interprétative, naturaliste du monde. Cela signifie que les chercheurs qualitatifs étudient les choses dans leurs milieux naturels, en essayant de donner un sens, ou d'interpréter, des phénomènes par la signification que les gens eux-mêmes y apportent. »* [notre traduction] (Denzin et Lincoln, 1994). Les principes de l'étude qualitative mis en exergue par Denzin et Lincoln sont aussi présents dans d'autres définitions, confirmant ainsi le consensus établi sur la nature de l'approche qualitative comme étant subjective, interprétative et visant à comprendre des phénomènes au travers de la vision des personnes et non pas de celle des données : *« par le terme « recherche qualitative », nous entendons tout type de recherche qui produit des résultats par d'autres moyens que les procédures statistiques ou autres méthodes quantitatives »* [notre traduction] (Corbin et Strauss, 1990).

Nous devons à présent définir les outils de mesures à mettre en place dans notre contexte d'étude. L'utilisation de questionnaires, tout comme dans un cadre quantitatif est possible, néanmoins les questions posées doivent être ouvertes afin de laisser place à l'interprétation du phénomène étudié par l'utilisateur. Si nous reprenons l'exemple précédent des transports dans l'agglomération lavalloise, la question posée devrait alors être : *« Quel est votre sentiment vis-à-vis des transports dans l'agglomération lavalloise ? »*. Laisser une place à la fin d'un questionnaire quantitatif proposant au participant de l'étude d'y inscrire « ses remarques » permet de récolter des informations qualitatives bien que dans ce cas la question n'étant pas suffisamment ciblée, la plupart des réponses risquent de ne pas concerner les phénomènes importants pour l'étude. L'analyse et la retranscription des échanges entre utilisateurs sur la base d'enregistrements audio ou vidéo capturés durant une expérimentation peuvent aussi donner des résultats pertinents. Mais il faut être conscient que le temps demandé pour de telles analyses est important (même en omettant le temps dédié à la retranscription verbatim de l'audio en texte) et bien évidemment ce temps croît de manière proportionnelle au nombre de participants à l'étude.

Autre outil, utilisé dans le cadre des méthodes qualitatives, les groupes de discussion (*focus groups* en anglais). D'après Lederman (voir Rabiee (2007)), les groupes de discussions sont : *« Une technique impliquant l'utilisation d'interviews en profondeur de groupes dans lesquels les participants sont choisis intentionnellement, l'échantillon ne doit pas nécessairement être représentatif d'une population spéci-*

fique, le groupe est « focalisé » sur un sujet spécifié » [notre traduction]. Un élément important n'est pas exposé dans cette définition, la présence d'un « guide » (bien souvent en la personne du chercheur) pour poser les questions et recentrer la conversation sur les points d'intérêt lorsque le groupe part en digression. Stewart et Shamdasani (2015) nous donnent la liste suivante quant aux objectifs des groupes de discussions :

- Obtenir des renseignements d'ordre général sur un sujet d'intérêt ;
- générer des hypothèses de recherche pouvant par la suite être soumises à une nouvelle étude utilisant une méthode quantitative ;
- stimuler la créativité, la génération de nouvelles idées et concepts créatifs ;
- diagnostiquer les problèmes potentiels vis-à-vis d'un nouveau logiciel, produit ou service ;
- mettre en lumière le ressenti global vis-à-vis d'un produit, d'un logiciel, d'un service, d'une institution ou de tout autre sujet d'intérêt ;
- en apprendre plus sur la manière dont les participants parlent, échange vis-à-vis du sujet d'intérêt, cela pouvant faciliter la mise en place future de mesures quantitatives ;
- interpréter des résultats qualitatifs précédemment obtenus.

Dans le cadre de notre étude, l'objectif principal de l'utilisation des groupes de discussion est double : confirmer et comprendre les résultats quantitatifs précédemment obtenus. De plus les groupes de discussion permettent d'avoir une vision plus humaine de l'étude menée, de soulever des points d'intérêt auxquels les chercheurs n'auraient pas pensé sans la conduite de ces groupes de discussion. Lors de la conduite de groupe de discussion, il faut être vigilant au laps de temps laissé entre l'expérience et le groupe de discussion. En effet, en fonction des objectifs fixés, les groupes de discussion peuvent être conduits avant ou après une étude empirique (ou l'utilisation d'un produit ou service) mais aussi plusieurs mois après celle-ci, s'il est intéressant de faire appel aux souvenirs bien ancrés dans la mémoire des participants. Stewart et Shamdasani donnent également certaines directives quant à la conduite d'un groupe de discussion : le minimum de participants est de quatre, sans quoi il est difficile de créer une vraie dynamique de conversation. Le maximum conseillé est de douze participants afin de permettre à la personne conduisant le groupe de discussion de garder le « contrôle » de la situation. Ces tailles de groupes relativement précises impliquent que le nombre de groupes de discussions à conduire est facteur de la taille de l'échantillon de population ayant participé à l'étude. Vis-à-vis du contenu, il est conseillé de ne pas avoir plus de dix questions à traiter. En effet, si la durée excède quarante-cinq minutes, l'attention des participants chute et ils ne sont plus aussi focalisés sur la discussion qu'ils le devraient. Pour ces mêmes raisons, les questions les plus importantes doivent être traitées le plus tôt possible, et si plusieurs questions traitent exactement du même point, il est conseillé de les classer de la plus générale à la plus spécifique.

Afin de mieux comprendre les différences entre quantitatif et qualitatif et donc l'intérêt d'exploiter ces deux types d'évaluation sur une même problématique, le tableau ci-dessous (tableau 3.1), extrait des travaux de (Creswell, 2009), propose une classification selon différentes hypothèses philosophiques. Cet aspect philosophique n'a pas grande importance dans notre contexte d'étude, mais Creswell pose

pour chaque segment une question permettant de mieux se représenter le propos décrit, ce qui a motivé le choix de ce tableau plutôt que d'un autre.

Hypothèse	Question	Quantitatif	Qualitatif
<i>Hypothèse ontologique</i>	Quelle est la nature de la réalité ?	La réalité est objective et unique exception faite du point de vue du chercheur.	La réalité est subjective et multiple au travers du regard des participants à l'étude.
<i>Hypothèse épistémologique</i>	Quelle est la relation du chercheur avec l'objet de l'étude ?	Le chercheur est indépendant de l'objet de l'étude.	Le chercheur interagit avec l'objet de l'étude.
<i>Hypothèse axiologique</i>	Quel est le rôle des valeurs morales ?	Sans valeur morale et impartial.	Comprend des jugements de valeur et est biaisée.
<i>Hypothèse rhétorique</i>	Quel est le langage de l'étude ?	Formel, impersonnel, basé sur les définitions de termes quantitatifs acceptés.	Informel, évolutif, personnel, utilisation de termes qualitatifs admis.
<i>Hypothèse méthodologique</i>	Quel est le protocole de l'étude ?	Protocole déductif, causalité, catégories statiques isolées avant l'étude, généralisations conduisant à une prédiction, explication et compréhension, exactitude (validité) et fiabilité.	Processus inductif, catégories identifiables au cours de l'étude, lié au contexte, modèles et théories développés dans un but de compréhension, précis et fiable grâce aux possibilités de vérification.

TABLEAU 3.1 – Différence de paradigme entre mesures quantitatives et qualitatives (Creswell, 2009).

3.2 Croiser les données par triangulation

La méthode de triangulation des données émerge du fait qu'il est plus pertinent de voir les données quantitatives et qualitatives comme deux éléments complémentaires plutôt qu'opposés par leurs différences. Denzin et Lincoln (1994) définissent la méthode de triangulation comme étant : « *la combinaison de méthodes différentes pour l'étude d'un même phénomène* » [notre traduction]. L'intérêt d'une telle méthodologie réside dans cette combinaison permettant d'avoir un point de vue plus holistique du phénomène étudié et ainsi d'obtenir une plus grande précision dans les mesures et interprétations découlant des données récoltées. Cela est d'autant plus intéressant lorsque l'on étudie un phénomène complexe aux multiples facettes comme l'expérience utilisateur.

Dans notre contexte, nous souhaiterions appliquer la méthode suivante : trianguler les données quantitatives issues de questionnaires post-expérimentation avec les données issues de mesures faites durant l'expérimentation (e.g. données de plateforme, données physiologiques) et les données issues de groupes de discussions. Le but est bien sûr d'obtenir une vision plus globale et précise des variations possibles de l'expérience utilisateur en 3D-IVE selon s'il est possible de communiquer verbalement ou non, mais aussi de confronter les résultats obtenus grâce à chaque méthode avec les autres. L'inclusion

de la méthode qualitative peut aussi permettre de mettre en exergue des phénomènes n'ayant pas été décelés par le chercheur, mais par les utilisateurs, permettant ainsi d'avoir un regard différent sur les résultats quantitatifs. La figure 3.1 présente la méthode que nous souhaitons utiliser pour cette triangulation des données. Les mesures à mettre en place pour chaque type de données seront plus longuement décrites en rapport avec les facteurs mesurés dans la section suivante (section 3.3).

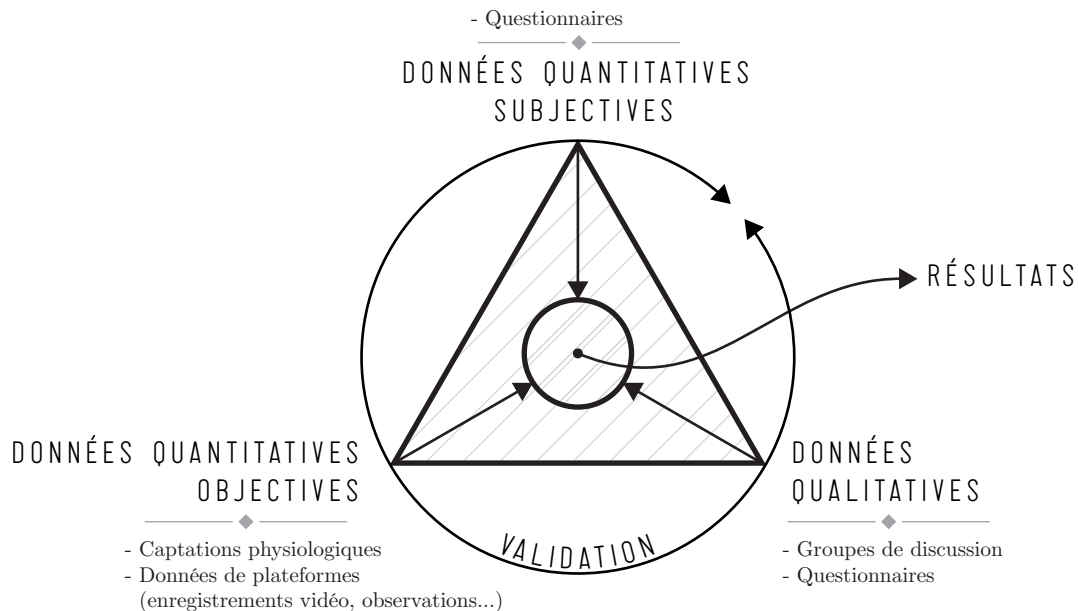


FIGURE 3.1 – Méthodologie par triangulation pour l'analyse des données.

3.3 Facteurs mesurés et moyens de mesure

Nous allons dans cette section reprendre les facteurs du modèle d'expérience utilisateur de Wu *et al.* (2009) étendu par Pallot *et al.* (2013) afin d'en présenter les différentes mesures proposées et expérimentées dans la littérature. Ainsi nous verrons les différents moyens de mesures de la présence sociale, de la présence virtuelle mais aussi du *flow* psychologique, etc...

Expérience (Utilisateur) Immersive ► Processus Expérientiel

► **Présence Sociale** : la définition que nous avons donnée de la présence sociale se base principalement sur les travaux de Biocca et Harms (2002) confirmés dans leur fiabilité par les travaux de Garau *et al.* (2005). C'est donc naturellement que nous avons étudié les mesures de la présence sociale proposées par ces mêmes chercheurs.

Biocca et Harms proposent en premier lieu des mesures basées sur l'observation. Il est en effet possible de quantifier les interactions sociales, et par corrélation la présence sociale, grâce au temps de parole de chaque participant ou par leurs fixations oculaires sur la représentation virtuelle de leur partenaire, qui sont deux bons indicateurs de l'**engagement attentionnel**. Les comportements proxémiques et plus particulièrement leur respect comme nous le ferions dans le réel (voir Bailenson *et al.*

(2001b)) sont aussi un bon indicateur du niveau de présence sociale. Si nous nous reportons au modèle d'expérience utilisateur de Pallot *et al.* (2013), ces mesures sont à classer sous la variable *interaction* du facteur présence sociale. Dans certains cas, il est aussi possible, par exemple lors de la présentation d'un discours devant une audience (voir Pertaub *et al.* (2002)), de mesurer les effets de la présence sociale sur l'utilisateur par ses constantes physiologiques : augmentation du stress ou de l'excitation engendrant une accélération du rythme cardiaque ou une variation de la conductance de la peau. Ces mesures, bien que très pertinentes par leur objectivité, présentent certains obstacles. Biocca et Harms évoquent des difficultés de mesure mais, depuis la publication de ce papier, la technologie a évolué et ces difficultés sont aujourd'hui surmontables. Le problème réside plutôt dans la nature de la présence sociale. En effet, il n'est en aucun cas nécessaire de regarder une personne (réelle ou virtuelle) pour la sentir présente à nos côtés, et les règles proxémiques peuvent quant à elle être respectées à un instant T, mais, selon l'activité en cours, être totalement occultées la minute suivante. Pour ce qui est des mesures par la physiologie, il est difficile de dire si les variations physiologiques observées sont à attribuer à la seule présence de l'autre ou à d'autres phénomènes comme une difficulté dans l'exécution de l'activité en cours.

Avant la rédaction de leur article le plus conséquent sur la définition de la présence sociale, Biocca *et al.* (2001) proposaient déjà son évaluation au travers d'un questionnaire, le « *Networked Minds social presence questionnaire* ». Ce questionnaire, bien qu'antérieur à la définition de la présence sociale proposée, a été empiriquement testé par une comparaison entre les mêmes situations en face à face réel et virtuel (Biocca *et al.*, 2001). Cette comparaison a montré la cohérence des résultats entre les deux états. C'est pourquoi, en dépit des différences mineures qu'il présente avec la définition de la présence sociale précédemment exposée, nous considérons ce questionnaire comme une source fiable pour la mesure de celle-ci. Composé de trente-huit questions à échelle de Likert (sept niveaux), ce questionnaire évalue les facteurs suivants de la présence sociale :

- **La co-présence** (niveau perceptuel voir figure 2.7) qui représente le sentiment d'être dans un même espace avec un autre être doté de sentiment, est évaluée au travers de trois variables, **l'isolation / solitude**, « *ai-je le sentiment d'être seul dans l'environnement virtuel ?* » (2 questions), **la conscience mutuelle**, « *dans quelle mesure suis-je conscient de la présence d'une autre personne ?* » (6 questions) et **l'attention allouée**, « *fais-je particulièrement attention à l'autre ?* » (8 questions) ;
- **l'engagement psychologique** (niveau subjectif, voir figure 2.7) évalué selon deux variables, **l'empathie** ou **contagion émotionnelle** (6 questions) et la **compréhension mutuelle**, « *est-ce que je comprends l'autre et inversement ?* » (6 questions) ;
- **l'engagement comportemental** (niveau subjectif, voir figure 2.7) évalué selon trois variables, **l'interdépendance comportementale**, « *mon comportement est-il induit par celui de l'autre ?* » (6 questions), **l'assistance mutuelle**, « *nous sommes-nous assistés dans ce que nous avons à faire ?* » (4 questions) et la **dépendance des actions**, « *mes actions sont elles dépendantes de celles de l'autre ?* » (2 questions).

Comme nous l'avons souligné, ce questionnaire ne correspond pas exactement à la définition de

la présence sociale proposée ayant été produit antérieurement à la définition de Biocca *et al.* (2001), néanmoins nous y retrouvons les facteurs importants des niveaux perceptuel et subjectif présentés en figure 2.7. Le niveau intersubjectif est quant à lui difficile à évaluer par un questionnaire, le phénomène décrit étant à la fois automatique et abstrait, définir des questions permettant son évaluation sans causer d'incompréhension chez l'utilisateur paraît complexe.

Nous avons précédemment cité le travail de Garau *et al.* (2005) visant à évaluer les variations de la présence virtuelle induites par différents niveaux de réalisme comportemental chez des agents virtuels. Le contexte est différent, car l'objectif de l'étude n'était pas de mesurer la présence sociale. Néanmoins, le questionnaire proposé, basé sur des échelles bipolaires sémantiques (sept niveaux), évalue certains aspects de la présence sociale. Les deux questions suivantes (extraites d'une série de cinq) évaluent le sentiment de co-présence :

- « *Au cours de l'expérience, aviez-vous le sentiment d'être dans la salle avec d'autres personnes ou d'être seul ?* » (Avec d'autres personnes = 1, Seul = 7). Cela correspond au facteur **isolation** / **solitude** de Biocca *et al.* (2001).
- « *Maintenant, considérez votre réponse vis-à-vis de toute l'expérience. Dans quelle mesure avez-vous réagi aux autres personnes comme s'ils étaient des gens ?* » (Pas du tout = 1, Très = 7). Cette question renvoie au facteur **interdépendance comportementale** de Biocca *et al.* (2001).

L'intérêt de ce questionnaire réside dans le fait qu'à la différence de Biocca *et al.*, Garau *et al.* demandent aux participants de prendre du recul et de considérer la globalité de l'expérience, ce qui est un aspect intéressant quand on prend en compte la nature fugace du sentiment de présence sociale, comme nous l'avons évoqué en début de cette section.

Le questionnaire est donc le moyen le plus simple et le plus usité de mesurer la présence sociale dans un environnement virtuel, mais il est aussi possible d'obtenir plus d'informations par la conduite de groupes de discussion. Ne pouvant être menés juste après la phase d'immersion, ces derniers présentent l'avantage de laisser le temps aux utilisateurs de prendre du recul sur l'expérience vécue, les amenant à parler de leur expérience globale et non pas des dernières minutes vécues comme il est possible qu'ils le fassent lors de la complétion d'un questionnaire post-expérimentation. Cela reste par ailleurs valable pour la majorité des facteurs de l'expérience utilisateur. Nous considérons donc comme acquis pour la suite de ce chapitre que la conduite de groupes de discussion offre un apport important aux mesures quantitatives proposées.

► **Liens** : un court questionnaire précédant l'expérimentation est la solution la plus simple pour être informé des liens possibles unissant les participants d'une expérimentation. Il est important de connaître ces hypothétiques liens entre individus, ils permettent lors de l'analyse des mesures de la présence sociale soit d'exclure certains participants, soit d'affiner l'analyse des résultats en testant l'effet de la familiarité sur tel ou tel facteur.

► **Émotions Sociales** : nous venons de voir que Biocca *et al.* (2001) intègre l'**empathie** dans les mesures de la présence sociale, il est néanmoins une forme d'empathie comportementale que son

questionnaire ne prend pas directement en compte : la translucence. Nous pouvons tout de même dire que la translucence étant une forme d'empathie comportementale, le croisement des questions pourtant sur l'**empathie** et sur l'**engagement comportemental** peut nous donner de bonnes indications sur l'occurrence de ce phénomène, surtout dans un contexte où les utilisateurs ne seraient pas en mesure de communiquer vocalement. Néanmoins, le concept même de translucence n'est pas simple à comprendre pour l'utilisateur et nous pensons qu'il est encore moins facile à évaluer dans un questionnaire à échelle de Likert ou bipolaire. Dans ce cas, la translucence demande une évaluation qualitative par le groupe de discussion. Pour ce qui est des autres émotions sociales (embarras, culpabilité, honte, etc. . .), leur évaluation objective est là aussi d'une grande complexité. Comme nous allons le voir par la suite, de nombreuses études ont mesuré la physiologie des émotions, mais ces études se concentrent sur les six émotions fondamentales (voir figure 2.5) présentant des schémas physiologiques distincts. En revanche, dans le cadre des émotions sociales, nous n'avons que peu d'exemples de mesures physiologiques : le travail de Blau (2010) présente une mesure de la jalousie chez de très jeunes enfants (6 à 12 mois) via EEG, mais étant donné la population sondée nous ne pouvons nous baser sur ces travaux. Une autre possibilité réside dans l'analyse des interactions sociales vocales, au travers du ton, de l'intonation et de l'intensité afin de détecter les émotions (fondamentales) (Chang *et al.*, 2013). Le contenu verbal de ces interactions apportant d'autres éléments, une analyse qualitative poussée pourrait permettre de détecter les émotions sociales. Néanmoins, une telle analyse, outre le temps requis, reste totalement soumise à la subjectivité de la personne menant les analyses.

► **Réponses Émotionnelles :** l'état émotionnel, au regard de son implication dans l'attirance, l'expérience et par corrélation l'envie de réutilisation, est un facteur qui a été très largement étudié dans la littérature, que ce soit dans le domaine du jeu vidéo ou dans les 3D-IVEs. Dans cette recherche d'objectivité, les études visant à mesurer les émotions sont nombreuses. Les constantes physiologiques (**état physiologique** dans le modèle de Pallot *et al.* (2013)) les plus souvent utilisées sont le rythme cardia respiratoire et la conductance de la peau (Monkaresi et Calvo, 2014 ; Wen *et al.*, 2014 ; Lee *et al.*, 2006 ; Rainville *et al.*, 2006 ; Nakasone *et al.*, 2005 ; Collet *et al.*, 1997 ; Sinha *et al.*, 1992 ; Schwartz et Weinberger, 1981). Autre aspect physiologique très utilisé, la reconnaissance des expressions faciales, soit grâce à des capteurs EMG⁴, soit par reconnaissance des schémas émotionnels faciaux dans un flux vidéo (Soleymani *et al.*, 2014). Plus récemment l'utilisation des casques EEG est devenue très populaire au sein de la communauté scientifique : depuis 2005 le nombre d'études menées sur la détection des émotions utilisant ces outils présente une croissance importante (Rodríguez *et al.*, 2015 ; 2013 ; Valenzi *et al.*, 2014 ; Liu *et al.*, 2010 ; Petrantonakis et Hadjileontiadis, 2009). D'autres mesures physiologiques ont aussi montré leur intérêt dans cette quête de mesure objective de l'émotion : le clignement oculaire (Chittaro et Sioni, 2013), la détection des variations thermiques du visage (Powar, 2013) ou encore l'analyse du ton, de l'intonation et de l'intensité de la voix (Chang *et al.*, 2013 ; Yang *et al.*, 2012). Enfin, certains chercheurs ont tenté de combiner le plus possible de ces mesures physiologiques dans la même étude afin d'établir des corrélations entre les différents changements physiologiques observés : l'étude de Liu *et al.* (2014), par exemple, combine EEG, EMG,

4. EMG : électromyogramme, enregistrement de l'activité électrique spontanée (contraction, relâchement) d'un muscle ou d'un nerf.

ECG et GSR. Un questionnaire spécifique à l'évaluation des émotions, le « *Self-assessment Manikin questionnaire* » (SAM) de Bradley et Lang (1994), vient compléter cette gamme de mesures objectives très complète en proposant une évaluation par l'utilisateur de son état émotionnel selon sa valence (positive / négative), la stimulation engendrée (faible / forte) et sa dominance ou notion de contrôle de l'émotion (faible / forte) (voir figure A.4 en annexe A).

Expérience (Utilisateur) Immersive ► Processus Rationnel

► **Flow Psychologique** : le *flow* psychologique implique trois variables, la **compétence perçue** de l'utilisateur, le **niveau de challenge** proposé ainsi que la **concentration**. De l'équilibre entre ces variables résulte un fort engagement vis-à-vis de l'activité proposée dans le 3D-IVE. Dans notre contexte et lors de la création d'une première expérimentation, le **niveau de challenge** peut être estimé de deux manières, par le chercheur-concepteur de l'expérimentation ou en faisant tester la tâche proposée par des participants hors contexte expérimental. Après analyse d'une première occurrence expérimentale, il est possible de juger plus objectivement du **niveau de challenge** proposé en observant les performances des utilisateurs. Si aucun d'entre eux (ou une faible proportion) n'a été en mesure d'accomplir l'activité proposée, il est évident que celle-ci est trop difficile. Bien sûr, la difficulté est facteur de la **compétence perçue**, il est donc nécessaire de connaître cette dernière. Le seul moyen est de demander aux utilisateurs, par un questionnaire pré expérimental, quel est le niveau de compétence qu'ils estiment posséder vis-à-vis de la navigation en environnement virtuel, de l'utilisation de tel ou tel système de visualisation (HMD, lunettes 3D) ou de contrôle (clavier, souris, manette). Le niveau de compétence estimé est fortement dépendant du contexte, du logiciel et du système de contrôle que le chercheur décide de mettre en place. Qui plus est, le **niveau de compétence perçue**, que nous pouvons aussi qualifier d'expérience préalable, aura un impact sur les performances et sur l'engagement (une tâche trop facile risque d'entraîner l'utilisateur vers l'ennui) lors d'une expérimentation. Sa connaissance est donc primordiale afin de tester ses effets sur la performance. Sans cette connaissance, il est plus difficile d'imputer les variations de la performance à d'autres facteurs tels que le fait ou non de pouvoir communiquer vocalement. La **concentration** et l'**engagement** sont mesurables ici encore grâce au questionnaire, en demandant aux utilisateurs s'ils étaient focalisés et s'ils se sont sentis investis (volonté de terminer ou de faire bien) sur la tâche ou sur les événements tenant place lors de l'immersion.

► **Téléprésence** : la téléprésence ou présence virtuelle a elle aussi été très largement étudiée et mesurée. Nous retrouvons ici les moyens de mesures quantitatifs que sont le questionnaire, le comportement et la physiologie. Avant de nous intéresser à ces mesures, nous devons faire un court récapitulatif. En effet, certains facteurs de la définition de la présence virtuelle que nous avons présentée (voir Lombard et Ditton (1997) ; Biocca et Delaney (1995) en section 2.2.2.1) sont, dans le modèle d'expérience utilisateur de (Wu *et al.*, 2009) étendu par Pallot *et al.* (2013), présentés en parallèle du facteur présence virtuelle et nous avons donc déjà proposé des mesures pour ces derniers. Si nous prenons la partie **immersion psychologique**, la **concentration** et l'**amusement**, ou le **plaisir** sont les variables du *flow* psychologique du processus rationnel et l'**engagement** en est la résultante

(voir paragraphe ci-dessus). Les **émotions** sont elles aussi un facteur distinct dans le **processus expérientiel**. Le **niveau d'intérêt** (par anticipation de l'usage) est lui une variable du facteur **acceptabilité technologique** que nous traiterons par la suite. Vis-à-vis de la partie **immersion perceptuelle**, nous avons déjà établi qu'elle correspondait à la QoS du modèle de qualité des DIMEs de Wu *et al.* (2009) (voir section 2.3), et nous considérons qu'au cours de nos expérimentations tout sera mis en œuvre afin de maintenir un niveau d'**immersion perceptuelle** constant. Reste alors un seul facteur, le sentiment de **présence spatiale**, qui est le rapport entre le corps de l'utilisateur et l'environnement dans lequel il est immergé. Ce court récapitulatif nous permet encore de fois de saisir l'intrication profonde entre les divers phénomènes que nous étudions et faisant partie de l'expérience utilisateur. D'un point de vue méthodologique, son objectif est surtout de nous permettre de faire le « tri » au sein des questionnaires proposés afin de ne pas inclure de redondance dans nos questions lors de la création de nos propres questionnaires.

Intéressons-nous en premier lieu aux mesures de la présence virtuelle par questionnaire. Là encore, l'utilisation de questionnaire à échelles de Likert ou bipolaires sémantiques est prédominante. Parmi les plus couramment utilisés, nous trouvons, le « *ITC-Sense of Presence Inventory* » (ITC-SOPI) de Freeman *et al.* (2001), le « *Igroup Presence Questionnaire* » (IPQ) de Schubert *et al.* (2001) et le « *Presence Questionnaire* » (PQ) de Witmer et Singer (1998). Ces questionnaires sont basés sur une analyse des facteurs de la présence virtuelle en corrélation avec la littérature et proposent une gamme de questions pour chaque facteur afin de le cerner au mieux. Ces questionnaires ont par la suite été validés empiriquement afin d'attester de la fiabilité des résultats obtenus par leur utilisation. Nous ne détaillerons que brièvement leur contenu, l'objectif étant plutôt de lister les facteurs mesurés afin d'établir une corrélation avec la définition de la présence virtuelle que nous avons exposée ainsi qu'avec les facteurs du modèle d'expérience utilisateur de Pallot *et al.* (2013).

Le questionnaire ITC-SOPI (Freeman *et al.*, 2001), propose la mesure de la **présence spatiale** au travers de questions comme « *J'ai eu le sentiment d'être dans la scène présentée* » ou « *J'ai eu l'impression que les objets ou les personnages pouvaient presque me toucher* ». De l'**engagement** avec des questions telles que « *J'ai été captivé* » ou « *J'ai perdu la notion du temps* ». De la **validité écologique** qui est une autre dénomination de la **cohérence** au travers de question comme « *Le contenu proposé m'est apparu crédible / plausible* » ou « *L'environnement affiché m'a paru naturel* ». Et des **effets négatifs** que nous identifions comme les gênes occasionnées par le matériel utilisé et que nous pouvons donc intégrer dans le facteur **vivacité sensorielle** de la QoS au travers de la variable **intrusivité** que nous lui avons ajoutée. Mesure effectuée par des affirmations telles que « *Je me suis senti nauséux* » ou « *J'ai ressenti une fatigue oculaire* ».

L'IPQ (Schubert *et al.*, 2001) mesure dans son cas trois facteurs, évidemment, la **présence spatiale** avec des questions telles que « *D'une certaine façon, j'ai eu l'impression que le monde virtuel m'entourait* » ou « *J'avais l'impression que j'étais juste en train de percevoir des images* ». L'**engagement** avec des questions telles que « *Je n'étais pas conscient de mon environnement réel* » ou « *J'étais complètement captivé par le monde virtuel* ». Et le **réalisme** qui, comme le montrent les questions à suivre,

mesure aussi la **cohérence**, « *A quel point votre expérience dans l'environnement virtuel vous a-t-elle semblée cohérente avec votre expérience dans le monde réel ?* » ou « *A quel point le monde virtuel vous a-t-il semblé réel ?* ».

Le PQ (version 3) (Witmer et Singer, 1998) mesure un plus grand nombre de facteurs. Ainsi, nous y retrouvons le **réalisme**, qui amalgame **réalisme**, **degré de contrôle** (voir section 2.2.2.1), et **cohérence**. La **possibilité d'interaction** qui amalgame **interactivité**, **degré de contrôle** et **cohérence**. La **qualité des interfaces**, qui amalgame le **degré de contrôle** et la variable **intrusivité** du facteur **vivacité sensorielle** de la QoS. La **possibilité d'examiner** qui mesure la possibilité de s'approcher d'un objet et de l'examiner sous toutes les coutures ainsi que l'**engagement**. Pour finir le PQ propose une **auto-évaluation** de la **performance** qui mesure plus l'adaptabilité de l'utilisateur au 3D-IVE que ses performances vis-à-vis d'une tâche ou d'une activité. Les nombreux amalgames dans ce questionnaire lui ont valu les critiques de Slater (1999) (l'un des auteurs sur lesquels nous basons notre définition de la présence virtuelle) lui reprochant de ne pas permettre une mesure fiable de la présence, point sur lequel nous ne pouvons qu'être d'accord. Néanmoins, ce questionnaire fut et est largement utilisé.

Nous avons pu constater la cohérence des facteurs mesurés par les questionnaires présentés et les facteurs du modèle d'expérience utilisateur de Pallot *et al.* (2013) et du modèle de qualité des DIMEs de (Wu *et al.*, 2009). Ce constat, ainsi que le fait que la validité des questionnaires présentés soit attestée, implique nous pourrions nous servir de ces derniers comme base de travail en vue de l'établissement de nos propres questionnaires pour notre étude.

Une seconde méthode permettant d'évaluer la présence spatiale est l'étude du comportement de l'utilisateur en partant du postulat que si un utilisateur agit de manière similaire dans le monde réel et dans un environnement virtuel cela indique qu'il se sent présent dans l'environnement (Slater, 1999). L'observation de cette similitude, pour être mesurée, demande l'intégration d'objets ou d'interactions dans le 3D-IVE générant une réponse corporelle chez l'utilisateur. Par exemple, baisser la tête pour esquiver un objet en mouvement, ou le fait de se pencher pour suivre une courbe prise à grande vitesse dans un véhicule (Freeman *et al.*, 2000). Cette réaction comportementale trouve sa symétrie vis-à-vis de la présence sociale par le respect des règles proxémiques ou par le fait de regarder l'avatar d'une personne lors qu'on lui adresse la parole.

Une approche particulière de cette étude du comportement est l'utilisation de mesures physiologiques. Comme nous l'avons déjà vu au travers des travaux de Pertaub *et al.* (2002), une audience virtuelle est capable de générer du stress chez un utilisateur comme cela serait le cas dans le monde réel. Le même type d'étude a été mené par Meehan *et al.* (2002) afin d'observer l'apparition du stress lors de l'exposition d'un utilisateur à un grand trou dans une pièce virtuelle : l'occurrence du stress généré par la peur de la chute étant une indication de la similitude entre réactions dans le monde réel et virtuel attestant ainsi du sentiment de présence de l'utilisateur. Une autre méthode plus récente est la mesure EEG durant une immersion en 3D-IVE, puis la mise en parallèle de ces mesures avec d'autres mesures physiologiques (ECG, GSR) ainsi qu'avec des questionnaires afin de définir les

schémas neuraux de la présence virtuelle (Clemente *et al.*, 2013 ; Baumgartner *et al.*, 2008; 2006).

► **Acceptation technologique** : en accord avec la *Technology Acceptance Model* de Davis (1989) sur lequel se basent Wu *et al.* (2009) dans la définition de leur modèle de qualité des DIMEs, il est possible de mesurer l'acceptation technologique par un questionnaire basé sur des échelles bipolaires sémantiques à sept niveaux. Pour rappel, dans le modèle d'expérience utilisateur de Pallot *et al.* (2013), l'**acceptation technologique** est définie par deux variables, l'**utilité** (perçue) et la **facilité d'utilisation** (perçue) menant à l'**adoption de la technologie** (QoE) elle-même caractérisée par l'intention d'**utilisation** (future) et l'**utilisation réelle**. Dans le contexte d'une étude empirique comme la nôtre, il n'y aura pas de réutilisation donc pas d'**adoption de la technologie**. Toutefois, nous pourrions questionner nos participants vis-à-vis de l'**acceptation technologique** ainsi que d'une potentielle réutilisation, en d'autres termes, leur envie de reproduire l'expérience à laquelle ils ont participé et sous quelles conditions. Outre des questions que nous pourrions définir sur la base des travaux de Davis (1989), il nous paraît important vis-à-vis de ce facteur d'obtenir des informations qualitatives par les groupes de discussion, étant donné que nos expérimentations n'auront en soit pas d'utilité pour l'utilisateur et feront donc plus appel à un sentiment hédonique de plaisir ou d'amusement plutôt qu'à la perception d'un réel apport dans l'accomplissement d'une tâche ou d'un travail récurrent.

Qualité de l'expérience

► **Comportement social** : le comportement social peut-être mesuré de manière objective au travers de plusieurs variables, comme exposé dans le modèle de (Wu *et al.*, 2009) étendu par Pallot *et al.* (2013) : il est possible d'enregistrer la fréquence et l'intensité (notion d'engagement) des interactions, mais aussi de manière plus qualitative par une analyse du discours ou du comportement des utilisateurs. Si une certaine familiarité entre les utilisateurs existe, il est alors pertinent de la modéliser par un graphique par exemple afin d'obtenir des informations supplémentaires et de voir quels comportements ont tendance à apparaître vis-à-vis du niveau de familiarité. Cela est possible dans le cadre des 3D-IVEs par l'interrogation des participants, mais peut aussi être produit par l'analyse des réseaux sociaux auxquels adhèrent les utilisateurs.

► **Comportement Empathique** : comme pour le comportement social, le comportement empathique peut être évalué de manière objective par la vitesse et l'intensité de la réaction ainsi que par la fréquence de leur occurrence. A contrario, l'identification de l'occurrence et le décryptage de comportement en réaction aux émotions dites sociales (embarras, culpabilité, honte, etc.) sont plus complexes puisqu'intimement lié à l'état mental des utilisateurs. Il est alors nécessaire de faire appel à des mesures subjectives par le biais de questionnaires à échelles de Likert ou bipolaires à échelles sémantiques.

► **Comportement Émotionnel** : comme précédemment, il est possible d'évaluer le comportement émotionnel d'un individu au regard de la vitesse et l'intensité de ses réactions (par exemple, sourire ou faire une moue de dégoût lors de l'apparition d'un élément dans le 3D-IVE ou lors d'une interaction) ainsi que par la fréquence de l'occurrence des émotions. Il est là encore possible d'utiliser deux types de mesures. Quantitatives au travers de l'observation des constantes physiologiques (Wen *et al.*, 2014 ;

Rainville *et al.*, 2006), de l'activité cérébrale de l'utilisateur (Ping *et al.*, 2013) ou dans un contexte plus social, nécessitant des échanges vocaux par l'analyse de la voix (Chang *et al.*, 2013) et par la reconnaissance faciale des émotions (que nous avons exclus de nos travaux pour les raisons citées en section 2.5). Ainsi que des mesures par questionnaire, le « *Self-assessment Manikin questionnaire* » (SAM) (Bradley et Lang, 1994) que nous avons déjà évoqué et qui est particulièrement indiqué dans ce contexte.

► **Efficacité** : l'efficacité est mesurable pour son aspect **hédonique** uniquement de manière subjective puisqu'elle fait référence au plaisir éprouvé par l'utilisateur, et le seul moyen de connaître le niveau de plaisir ressenti et de questionner l'utilisateur (questionnaire). Bien sûr, le plaisir hédonique est corrélé à l'occurrence de la joie, néanmoins les émotions sont, comme nous l'avons déjà évoqué, fugaces. Un utilisateur peut donc éprouver beaucoup de plaisir lors de l'utilisation d'une application, d'un produit ou lors d'une immersion dans un 3D-IVE sans pour autant être en permanence empreint d'une joie ou d'une euphorie profonde. En revanche, l'aspect **ergonomique** de l'efficacité est aisément mesurable de manière objective, le temps de complétion d'une tâche, le nombre d'erreurs commises en sont les variables les plus parlantes.

► **Comportement Exploratoire** : afin de ne pas être redondants dans nos explications nous ferons ici référence au paragraphe concernant le comportement exploratoire dans la section dédiée à l'expérience utilisateur (section 2.3) de notre revue de littérature en page 54. En effet, pour être claire la définition donnée alors comprenait un exposé des mesures de ce facteur.

► **Adoption Technologique** : voir le paragraphe sur le facteur **acceptation technologique** en page 74.

3.4 Méthodologie d'étude empirique, comparaisons et itérations

Comment mesurer l'impact d'une chose sur une autre ? Que ce soit en biologie, en médecine ou en réalité virtuelle, la solution la plus logique est la comparaison de deux échantillons. Un échantillon n'étant pas exposé au phénomène testé servira de témoin, d'état neutre, et le second échantillon, lui exposé au phénomène, permettra de mettre en exergue les éventuelles différences entre les mesures prises sur les deux échantillons. Il est difficile de fournir une littérature exposant le bien-fondé de cette méthode tant cette dernière est commune. Dans notre contexte, avec pour objectif l'évaluation de l'impact de communication vocale sur l'expérience utilisateur lors d'une immersion en 3D-IVE, cette méthode nous est apparue comme la plus logique. Notre comparaison sera donc la suivante : durant son immersion notre, échantillon témoin ne sera pas en mesure de communiquer vocalement, a contrario l'autre échantillon aura cette possibilité. Nous devons nous efforcer de faire en sorte que cette différence soit la seule entre les deux échantillons au sein d'une expérimentation, que ce soit du point de vue du 3D-IVE proposé, ou de celui de l'homogénéité des échantillons vis-à-vis de leurs expériences préalables ou de leur familiarité. Nous effectuerons sur nos deux échantillons les mêmes mesures, quantitatives et qualitatives pour, grâce aux protocoles statistiques adaptés, isoler

les différences significatives entre les deux échantillons. Afin de tester la dépendance de l'impact des interactions sociales vocales à la nature de l'activité proposée (collaboration / coopération) notre étude devra comporter aux moins deux expérimentations.

Outre le fait de comparer ces deux méthodes de travail commun quant à l'impact des interactions sociales vocales, nous serons aussi en mesure de retravailler notre protocole afin de mieux répondre aux exigences de notre première évaluation (e.g. modifier les questionnaires si nous jugeons que les questions utilisées pour évaluer un facteur auraient pu être plus pertinentes, ou modifier un aspect technique qui aurait fortement perturbé l'expérience). Nous devons néanmoins être extrêmement vigilant afin que ces changements s'ils ont lieu restent suffisamment fins pour ne pas modifier l'expérience dans une mesure qui la rendrait incompatible à la comparaison. Enfin, cette méthode itérative nous aidera à atteindre une plus grande exhaustivité quant à la validation ou non des hypothèses proposées.

3.5 Résumé du chapitre Méthodologie

Ce chapitre nous a permis d'exposer la méthodologie générale de mesure et d'analyse que nous allons employer au cours de notre étude empirique. Nous avons vu les types de mesures quantitatives et qualitatives proposées dans la littérature afin de mesurer les différents facteurs de l'expérience utilisateur du modèle de (Wu *et al.*, 2009) étendu par Pallot *et al.* (2013). Un tableau récapitulatif de ces moyens de mesure est présenté dans les pages suivantes. Grâce à ces différents types de mesures, nous serons en position de trianguler les résultats obtenus par la comparaison de deux échantillons, l'un étant capable de communiquer vocalement, l'autre non. Avec pour objectif d'avoir une vision holistique de l'impact de l'interaction sociale vocale sur l'expérience utilisateur en 3D-IVE.

Nous reviendrons au cours de la description de nos expérimentations sur les méthodologies utilisées, vis-à-vis de la constitution des échantillons, des procédures expérimentales ainsi que des protocoles statistiques utilisés. Par souci de compréhension, nous avons préféré décrire ces méthodes en corrélation avec le contexte d'utilisation et non pas dans ce chapitre plus général.

Facteur	Quanti. Subj.	Quanti. Obj.	Qualitatif
<i>Présence Sociale (Interactions)</i>	► Questionnaire post expérience (Garau <i>et al.</i> , 2005 ; Biocca <i>et al.</i> , 2001)	► Observation du temps de parole, de la fixation du regard sur la représentation virtuelle de l'autre ; ► observation du comportement proxémique ; ► mesures physiologiques de l'occurrence du stress.	► Groupes de discussion.
<i>Présence Sociale (Liens)</i>	► Questionnaire pré expérience.	► N/A	► N/A
<i>Émotions sociales</i>	► Questionnaire post expérience (Biocca <i>et al.</i> , 2001).	Physiologie : ► EEG (Blau, 2010) ; ► analyse de la voix (Chang <i>et al.</i> , 2013).	► Groupes de discussion (translucence) ; ► analyse du contenu verbale des interactions.
<i>Réponses Émotionnelles</i>	► Self-assessment Manikin questionnaire (Bradley et Lang, 1994).	► Enregistrement cardia respiratoire (Monkaresi et Calvo, 2014) ; ► enreg. des variations thermiques du visage (Powar, 2013) ; ► enreg. des clignements oculaire (Chittaro et Sioni, 2013) ; ► enreg. EEG (Rodríguez <i>et al.</i> , 2015; 2013) ; ► reconnaissance faciale des émotions par EMG ; ► reconnaissance faciale des émotions par analyse de flux vidéo (Soleymani <i>et al.</i> , 2014) ; ► analyse de la voix (Chang <i>et al.</i> , 2013).	► Analyse du contenu verbal des interactions ; ► groupes de discussion.
<i>Flow Psychologique (Niv. de défi perçu)</i>	► Questionnaire post expérience.	► Tests utilisateur a priori de l'expérience.	► Groupes de discussion.
<i>Flow Psychologique (Niv. de compétence perçu)</i>	► Questionnaire pré expérience.	► N/A	► Groupes de discussion.
<i>Flow Psychologique (Concentration & Engagement)</i>	► Questionnaire post expérience.	► N/A	► Groupes de discussion.

TABLEAU 3.2 – Récapitulatif des moyens de mesures des facteurs de l'expérience utilisateur en 3D-IVE - partie 1.

Facteur	Quanti. Subj.	Quanti. Obj.	Qualitatif
<i>Présence Virtuelle (Téléprésence)</i>	Questionnaire post expérience : ► ITC-SOPI (Freeman <i>et al.</i> , 2001) ; ► IPQ (Schubert <i>et al.</i> , 2001) ; ► PQ (Witmer et Singer, 1998).	► Analyse du contenu verbal des interactions ; ► étude du comportement, similarité entre réel et virtuel (Freeman <i>et al.</i> , 2000 ; Slater, 1999) ; ► étude du comportement physiologique (stress) (Pertaud <i>et al.</i> , 2002) ; ► enreg. EEG (Clemente <i>et al.</i> , 2013).	► Groupes de discussion.
<i>Acceptation technologique</i>	► Questionnaire post expérience (Davis, 1989).	► N/A	► Groupes de discussion.
<i>Comportement social</i>	► Questionnaire post expérience.	► Enregistrement de la fréquence des interactions sociales.	► Groupes de discussion.
<i>Comportement Empathique</i>	► Questionnaire post expérience.	► Enregistrement de la vitesse et de l'intensité de la réaction à un stimuli social.	► Groupes de discussion.
<i>Comportement Émotionnel</i>	► Self-assessment Manikin questionnaire (Bradley et Lang, 1994).	► Identification de la vitesse et de l'intensité des réactions émotionnelles (reconnaissance physiologique des émotions) (Wen <i>et al.</i> , 2014) ; ► enreg. EEG (Ping <i>et al.</i> , 2013) ; ► analyse de la voix (Chang <i>et al.</i> , 2013).	► Groupes de discussion ; ► analyse du contenu verbal des interactions.
<i>Efficacité</i>	► Questionnaire post expérience (aspect hédonique).	► Observations : temps de complétion d'une tâche, erreur (aspect ergonomique).	► N/A
<i>Comportement Exploratoire</i>	► Questionnaire post expérience.	► Observation : actions, déplacements hors du cadre défini, durée et fréquence des parenthèses exploratoires.	► Groupes de discussion.
<i>Adoption Technologique</i>	► Questionnaire post expérience (Davis, 1989).		► Groupes de discussion.

TABLEAU 3.3 – Récapitulatif des moyens de mesures des facteurs de l'expérience utilisateur en 3D-IVE - partie 2.

Chapitre 4

Expérimentations

« *Le bon sens, quoi qu'il fasse, ne peut
manquer de se laisser surprendre à
l'occasion.* »

Alfred Elton van Vogt
Le Monde des A (1945)

Sommaire

4.1	Choix de l'environnement virtuel	81
4.1.1	Minecraft, bac à sable virtuel	81
4.1.2	Pourquoi Minecraft ?	84
4.2	Première expérimentation	87
4.2.1	Objectif et activité proposée	87
4.2.2	Environnement Virtuel	89
4.2.3	Participants	93
4.2.4	Mesures	94
4.2.5	Équipement des participants	97
4.2.6	Procédure	99
4.2.7	Résultats	100
4.2.8	Analyse et synthèse	111
4.3	Seconde expérimentation	117
4.3.1	Objectif et activité proposée	117
4.3.2	Environnement Virtuel	120
4.3.3	Participants	122
4.3.4	Mesures	122
4.3.5	Équipement des participants	125
4.3.6	Procédure	126

4.3.7	Résultats	127
4.3.8	Analyse et synthèse	145
4.4	Expérimentation complémentaire	151
4.4.1	Objectifs de l'expérimentation	151
4.4.2	Environnement virtuel	151
4.4.3	Participants	151
4.4.4	Mesures	152
4.4.5	Équipement des participants	153
4.4.6	Procédure	153
4.4.7	Résultats	154
4.4.8	Analyse et synthèse	168

Ce chapitre présente le travail d'expérimentation conduit dans le cadre de notre étude empirique visant à étudier l'impact des interactions sociales vocales sur l'expérience utilisateur en 3D-IVE. Nous débuterons par des explications quant au choix du logiciel *Minecraft* comme environnement virtuel. Puis nous poursuivrons en décrivant nos deux expérimentations quantitatives comparant deux contextes, avec et sans communication vocale. Ensuite, nous décrirons l'expérimentation qualitative qui reprend pour un même panel d'utilisateurs les deux expérimentations précédentes en proposant en plus des mesures quantitatives, une série de groupes de discussion. Chaque expérimentation sera décrite, au travers d'un exposé des objectifs de l'expérimentation, d'une description de l'environnement virtuel, d'un exposé des données afférentes aux participants, des mesures mises en place, de la configuration matérielle utilisée, du protocole expérimental, des résultats obtenus puis, pour finir, d'une analyse et synthèse proposant un regard critique sur les résultats vis-à-vis de la littérature ainsi que des propositions pour la suite du travail. Une discussion sur l'ensemble des résultats de cette étude sera proposée par la suite dans le chapitre *discussion* (chapitre 5).

4.1 Choix de l'environnement virtuel

4.1.1 Minecraft, bac à sable virtuel

La première étape permettant la construction de cette expérimentation était le choix d'un environnement virtuel. Deux choix s'offraient alors, contruire un environnement de toutes pièces, ou choisir une application proposant un environnement existant et suffisamment modulable pour répondre à nos exigences. Pour des raisons de temps, de moyens et d'attraction vis-à-vis des futurs participants (que nous détaillerons dans la section 4.1.2), nous avons choisi la seconde solution, ici par l'utilisation du logiciel *Minecraft*.

Minecraft est un jeu vidéo 3D créé par *Markus Persson* et est sorti officiellement le 19 novembre 2011. Précurseur d'un type de jeu vidéo aujourd'hui extrêmement populaire, les « *sand boxes* » ou en français, bacs à sable, *Minecraft* fut tout d'abord disponible uniquement sur ordinateur (sur les trois systèmes d'exploitation principaux, *Windows*, *OSX* et *Linux*, le jeu étant basé sur un langage de programmation multiplateformes, *Java*). Puis, il s'est vu adapté sur tous les supports existants, consoles de salon, téléphones et tablettes et a atteint en 2014 un total de cinquante-quatre millions de licences vendues (et au moins autant de copies pirates). La popularité du jeu est devenue telle que *Microsoft* a racheté en 2014 le studio de développement pour un total de deux milliards et demi de dollars, avec pour objectif de faire de *Minecraft* le logiciel phare de son système de réalité virtuelle l'*Hololens*. Très récemment (19 Janvier 2016) *Microsoft* a annoncé la sortie d'une nouvelle version de *Minecraft* dédiée spécifiquement à l'éducation¹, démontrant les possibilités de l'application au delà du jeu et de l'amusement.

Ce type de jeu propose bien sûr une partie narrative, bien que minimaliste dans le cas de *Minecraft*,

1. Voir : <http://education.minecraft.net/announce011916/>

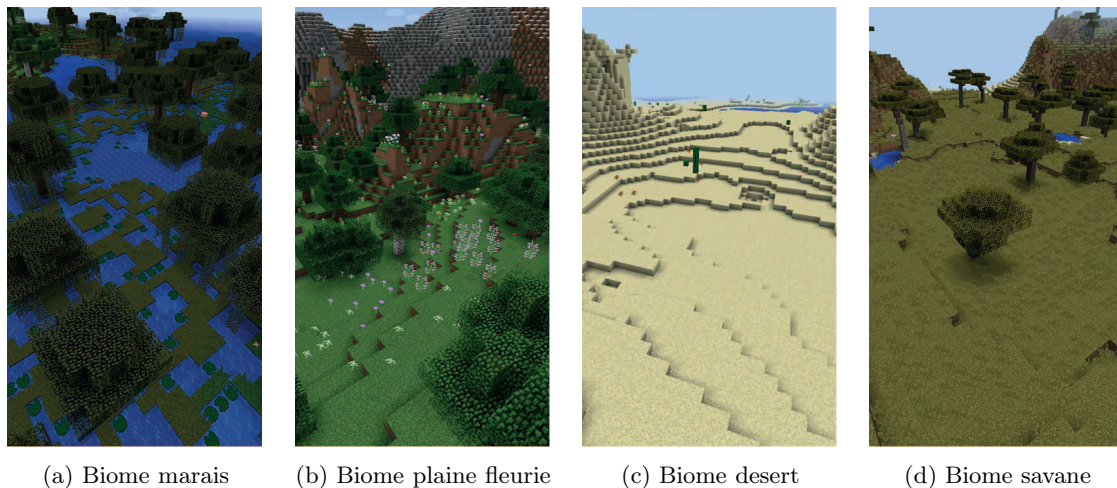
avec des personnages, des objectifs et un semblant de récit, mais offre surtout aux joueurs — et c’est là la grande différence avec la plupart des jeux vidéo — une liberté presque sans limites quant à la transformation des mondes qu’ils ont créés. En effet, lors de la création d’un nouveau monde, *Minecraft* génère un environnement tridimensionnel procédural unique, d’une taille telle qu’il est impossible d’en atteindre les limites sans l’utilisation d’un logiciel externe. La taille standard d’une carte est de 30 000 km de côté proposant donc une surface de 900 millions de km² (pour comparaison la terre a une superficie d’un peu plus de 510 millions de km²) pour une profondeur et une altitude de 128 mètres (soit 256 mètres de la limite haute à la limite basse) (Persson, 2010), représentant un volume totalement éditable par le ou les joueurs. Bien qu’il soit possible de créer des mondes plats et totalement vides, un monde par défaut se compose de différents biomes, forêts, déserts, zones enneigées (taïga), savanes (voir figure 4.2) et diverses variations sur chaque biome ainsi que des biomes spéciaux les enfers et l’*Ender*. Chacun de ces biomes est constitué d’une faune, d’une flore, d’un relief spécifique ainsi que de villages peuplés par des PNJ.

Mais la particularité de *Minecraft* se situe dans la matière constituant ces mondes et les éléments les composant. En effet, tout (à quelques exceptions près) se base sur des blocs cubiques d’un même format qui, par leurs assemblages, composent tout les éléments du jeu vidéo et donnent son aspect visuel — dont la figure 4.1 donne un aperçu — si reconnaissable à *Minecraft*. Un bloc mesure 1 m³ (Persson, 2010), et un personnage mesure deux blocs de haut, ainsi, les proportions du monde réel quant à la taille d’un arbre ou d’un bâtiment sont globalement respectées. Les blocs ont donc une taille unique, en revanche, ils proposent une très grande variété de matériaux aux propriétés différentes. Ainsi on trouvera du bois pour les blocs composant les arbres, de la terre, de la pierre, du sable ou encore de l’argile pour le sol, de l’eau, de la lave, des minerais comme le charbon, le fer, l’argent, l’or et une multitude d’autres encore. Le joueur devra pour obtenir ces différents blocs aller les extraire (miner) dans l’environnement, pour par la suite les combiner sous différentes formes, tel un artisan, afin de produire des outils, des armes, des armures ou des éléments d’architecture. Les termes anglais désignant ces deux actions, miner (*mine*) et fabriquer (artisanalement) (*craft*) ont donné son nom au jeu, *Minecraft*, et sont les pivots de son *gameplay*².



FIGURE 4.1 – Vue d’un monde *Minecraft* à l’aube.

2. Gameplay : anglicisme ayant pour équivalent français le terme jouabilité. Il désigne les éléments d’interaction avec l’utilisateur d’un produit vidéo ludique.

FIGURE 4.2 – Différents types de biomes dans *Minecraft*.

Afin de mieux comprendre ce fameux *gameplay*, parlons à présent des modes de jeu. *Minecraft* propose deux modes, le premier le **mode survie** qui comme son nom l'indique invite le joueur à survivre. Il débute alors en un point aléatoire du monde généré et devra, à partir de cet instant, tout faire pour survivre, c'est-à-dire se nourrir, se construire un abri et se défendre des monstres qui peuplent les nuits *minecraftiennes*. Afin d'illustrer ce **mode survie** sans de longues explications écrites, une vidéo créée par nos soins est disponible à l'adresse suivante : <http://www.remy-eynard.net/minecraft-exp/gameplay.html>. Malgré sa complexité et son contenu très fourni, le **mode survie** n'est pas le mode le plus joué sur *Minecraft*. En effet, le **mode créatif** est beaucoup plus populaire. Ce mode a pour unique objectif de laisser les joueurs construire tout ce qu'ils désirent. Pour faciliter la construction, ils sont en mesure de voler, mais surtout ont un accès illimité à tous les types de blocs disponibles dans le jeu. Depuis la sortie du jeu, les joueurs ont créé nombre de productions plus incroyables les unes que les autres, allant de la construction architecturale monumentale à la création de systèmes interactifs en utilisant la *redstone*. La *redstone* est une catégorie de blocs particuliers qui permettent de créer des circuits semblables aux circuits électroniques simples que nous utilisons dans la vie réelle. Ainsi cette catégorie se compose de sources d'alimentation (bloc ou torche) de fils conducteurs et deux éléments singuliers que sont le **répéteur** pouvant servir de retardateur ou de diode et le **comparateur** agissant comme un transistor. La combinaison de ces quelques éléments avec les leviers et autres actionneurs proposés par le jeu permet la création simple et intuitive de portes logiques ouvrant ainsi le champ des possibilités de création sur un univers nouveau. La liste suivante propose un petit échantillon de création impressionnantes n'utilisant que les fonctionnalités offertes par *Minecraft*, car comme nous allons le voir, le jeu étant développé en *Java* et de manière ouverte, une très grande communauté de joueurs / développeurs s'est employée à étendre les possibilités déjà immenses du logiciel.

- Construction du château de « *Laputa* » extrait de l'œuvre d'animation « *Le château dans le ciel* » de Hayao Miyazaki par *adybuddy_* : <http://goo.gl/hDdkLH>
- Construction de monuments historiques par divers joueurs : <https://youtu.be/lizEvWNEOh8>
- Calculatrice scientifique fonctionnelle, construite uniquement grâce aux circuits de *redstone* par *EvilCraftProductions* : <https://youtu.be/PiKWqpnqY7o>

— Construction du monde de « *Westeros* » issu de la série « *Game of Thrones* » par l'équipe *WesterosCraft* : <https://youtu.be/5wOc-N2NuEM>

Nous parlions d'étendre les possibilités du jeu, c'est ici qu'interviennent les *mods*, qui est un terme désignant des contenus additionnels, des scripts ou des programmes développés par la communauté des joueurs pour un jeu vidéo. Pour *Minecraft*, le nombre de *mods* existant est très important. Ils proposent par exemple l'ajout de lignes de commande permettant de poser une surface complète de bloc en quelques secondes, l'ajout de nouveaux éléments de *gameplay* (nouveaux blocs, nouvelles fonctionnalités), l'ajout de *shaders*, ou encore de nouvelles textures afin d'améliorer le rendu visuel. Nous reviendrons sur les *mods* de manière plus spécifique lors de la description des environnements créés pour nos expérimentations, et nous décrirons alors leurs fonctionnalités.

Pour finir cette section, nous souhaitons ajouter que *Minecraft* comporte bien sûr des fonctionnalités multijoueurs. Le grand intérêt de ces fonctionnalités réside dans le fait qu'il est possible d'héberger soi-même son serveur de jeu. Ainsi, le poste serveur héberge le monde auquel les joueurs vont venir se connecter. Cette possibilité d'auto hébergement permet de jouer à *Minecraft* sur un réseau local supprimant ainsi les effets néfastes des connexions distantes (ralentissements, pertes de paquet).

En définitive, *Minecraft* est un logiciel unique, en raison de sa conception et des nombreux développements ultérieurs : les frontières du jeu sont sans cesse repoussées, et la seule limite admise est la créativité de l'utilisateur.

4.1.2 Pourquoi Minecraft ?

La première raison ayant motivé l'utilisation de *Minecraft* pour nos expérimentations est purement « logistique ». La création d'un 3D-IVE demande beaucoup de temps et de ressources humaines. Modéliser, texturer, animer puis rendre interactif un environnement même limité n'est pas l'affaire de quelques jours, mais plutôt de quelques mois. Dans le cadre d'une thèse de doctorat, le temps est un élément précieux qu'il est préférable pour le doctorant de dépenser en réflexion et en tâches afférentes à la résolution des hypothèses émises plutôt qu'en développement. D'un autre côté, l'équipe dans laquelle cette thèse s'est tenue n'est pas une grande équipe du point de vue des ressources humaines, aucun ingénieur de recherche n'était disponible pour participer à la création d'un tel environnement. Il semble donc pertinent de penser que, pour une personne seule dans un laps de temps limité, se lancer dans une phase de développement longue et complexe présente des risques importants, soit de ne pas produire un environnement virtuel suffisamment abouti, soit de se voir pris par le temps.

Convaincus du fait qu'il nous fallait utiliser l'existant pour mener à bien nos travaux, nous avons alors réfléchi au type d'activité à proposer afin de générer une expérience intéressante et satisfaisante pour nos participants, mais surtout favorisant l'occurrence d'interactions sociales vocales. Contrainte supplémentaire, afin d'être en mesure de comparer nos deux contextes, avec et sans communication vocale, l'activité proposée devait être réalisable par au moins deux participants dans ces deux contextes, excluant ainsi les activités de type jeu de questions / réponses ou jeu de *guide d'aveugle*. L'idée d'une

tâche de construction nous a alors semblé la plus pertinente. En effet, si tant est que les participants possèdent tous un niveau d'information (plan) leur permettant de construire quelque chose, il n'est dans ce cas pas absolument nécessaire de communiquer vocalement pour atteindre l'objectif fixé.

Suite à quelques recherches préliminaires, nous avons eu la confirmation que *Minecraft*, outre le fait de nous permettre très facilement de mettre en place une activité de construction (étant la base même du jeu), nous donnerait un contrôle total de l'environnement. Par l'utilisation de *mods* et de blocs spécifiques — que nous détaillerons dans les sections environnement virtuel de chaque expérimentation — nous serions en mesure par exemple de remplir les inventaires des participants avec la variété et le nombre de blocs voulus. De plus, par l'utilisation d'un serveur local nous aurions la possibilité de conserver tous les environnements utilisés pour analyse ultérieure afin d'observer a posteriori les performances des participants.

Avant de nous décider à l'utilisation de *Minecraft* pour notre étude empirique, nous avons mené un rapide état de l'art afin de déterminer si l'utilisation de ce logiciel dans un contexte scientifique avait des précédents et ainsi valider son utilisation dans un tel contexte. Nous avons été surpris par l'évolution des publications comprenant le terme *Minecraft* entre 2010 (date de sortie de la bêta) et aujourd'hui (voir figure 4.3). Cette évolution atteste d'un intérêt certain de la part de la communauté scientifique pour le logiciel. Nous avons constaté que la plupart des utilisations de *Minecraft* donnant lieu à une publication concernent l'éducation, l'apprentissage d'une langue étrangère (Koichi, 2014), de concepts scientifiques comme la table des éléments périodiques (Short, 2012) ou encore de la programmation (Finley, 2014). Tous ces travaux se basent sur le concept de *l'edutainment*, autrement dit l'inclusion dans un jeu d'activités ou de fonctionnalités permettant d'apprendre en s'amusant. Nous avons aussi pu lire un article concernant l'utilisation de *Minecraft* dans le cadre de soins psychologiques aux soldats victimes de chocs post-traumatiques (Neltz, 2014), ainsi qu'une revue de littérature proposant son utilisation dans la cadre de psychothérapie dédiée aux enfants et aux adolescents (Steadman *et al.*, 2014).

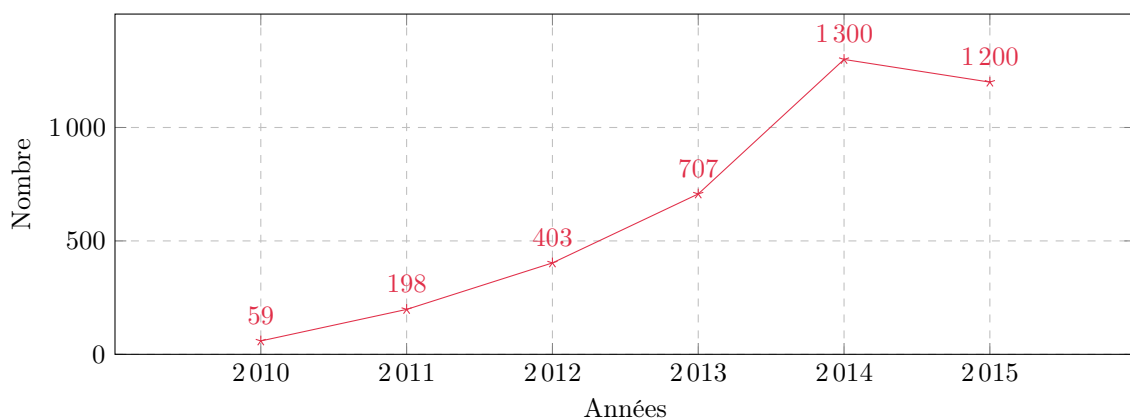


FIGURE 4.3 – Evolution du nombre de publications d'après la recherche "*Minecraft*" -*patent* sur Google Scholar

Dès lors, l'utilisation de *Minecraft* dans un contexte tel que le nôtre s'avérerait pertinent. Nous avons alors entrepris une évaluation du logiciel selon la QoS proposée dans le modèle de Wu *et al.* (2009)

(voir figures 2.15 en section 2.3) afin de nous assurer que *Minecraft* était techniquement en mesure de générer une expérience utilisateur satisfaisante.

Qualité de Service

► **Interactivité :** nous avons défini l’interactivité, selon Steuer (1992), comme étant la « *mesure dans laquelle les utilisateurs peuvent participer à la modification de la forme et du contenu d’un environnement médiat en temps réel* » [notre traduction]. Un environnement en temps-réel dont tous les éléments sont destructibles et modulaires sans limites ou presque pourrait être qualifié de pinacle de l’interactivité. La **variété** des interactions n’est quant à elle pas très élevée, puisque basée sur trois actions principales, miner, récolter, construire. Néanmoins, les possibilités d’applications de ces trois actions pallient largement à ce manque de **variété** et les risques de voir les participants s’ennuyer lors d’une immersion courte sont assez réduits. Vis-à-vis des variables **vitesse** et **correspondance** (voir page 46), elles dépendent uniquement de la puissance des machines utilisées, il est donc de notre ressort de nous servir de machines suffisamment puissantes ou d’optimiser *Minecraft* afin d’atteindre les 60 images par seconde préconisées dans l’utilisation des 3D-IVEs ainsi que la réactivité des contrôles nécessaire pour que le participant ne soit pas perturbé par des ralentissements.

► **Vivacité Sensorielle :** comme la plupart des 3D-IVEs, l’étendue de la **vivacité sensorielle** que propose *Minecraft* se résume au canal visuel et auditif. Quant à sa **profondeur** et à son **intrusivité**, cela dépendra des matériels d’affichage et de transmission vocale que nous utiliserons (HMD, lunettes 3D, écrans, microphone, casque audio).

► **Cohérence :** *Minecraft* respecte les règles basiques de la physique qui régissent notre environnement réel. Un objet lâché tombe, le temps passe, un bloc frappé par l’utilisateur finit par être détruit, un arbre dépouillé de son tronc verra ses feuilles disparaître, le relief malgré son aspect cubique ne présente pas de formes totalement irréelles. Le premier « choc » visuel passé, le logiciel ne présente aucun élément pouvant frapper l’utilisateur par son manque de cohérence avec le monde réel ou les représentations mentales qu’il pourrait s’être fait du fonctionnement de ce monde virtuel par anticipation de son utilisation.

► **Qualité Visuelle :** comme nous venons de le dire, l’aspect cubique de *Minecraft* peut être quelque peu déroutant de prime abord et les textures par défaut, de type pixélisé et de basse résolution (16 x 16px) (voir images et vidéo dans la section 4.1.1), renforcent cet aspect qui peut déranger. Néanmoins, il est possible de changer ces éléments pour des textures de plus haute résolution et présentant un graphisme plus lisse, plus agréable à l’œil. Il faut néanmoins être vigilant sur deux points, conserver des ressources matérielles suffisantes pour assurer un bon niveau d’**interactivité** et rester cohérent avec les modélisations. En effet, l’utilisation de textures photo réalistes sur les modélisations cubiques de *Minecraft* pourrait être plus perturbante que les textures par défaut (voir figure 4.4a). Vis-à-vis de ce facteur, ce que nous venons de dire au sujet de l’environnement s’applique aussi aux avatars dont les textures sont là aussi modifiables. Malgré leur aspect cubique, les avatars possèdent, un tronc, une tête et les quatre membres qui caractérisent l’humain ; ils présentent donc un niveau d’anthropomorphisme élevé.

► **Réalisme Comportemental** : nous avons déjà parlé du **réalisme comportemental** de l'environnement lorsque nous avons traité le facteur **cohérence** de la QoS. Nous ajoutons qu'il est aussi possible d'utiliser des *shaders* afin d'améliorer le rendu graphique de l'eau ou de la lumière, accentuant ainsi la ressemblance entre les mondes virtuels et réels. Par exemple, il est possible d'ajouter aux étendues d'eau de la réfraction ou la réflexion des sources de lumière sur les matériaux (voir figure 4.4b). Vis-à-vis des avatars représentant les utilisateurs, leur comportement est cohérent avec ce que nous pouvons voir dans la réalité, lorsqu'ils se déplacent leurs jambes et leurs bras bougent, il en va de même lorsqu'ils font une action, casser, poser un bloc ou actionner un levier. En revanche, ils ne font par défaut montre d'aucune expression faciale ni ne bougent les lèvres quand l'utilisateur parle. Nous avons tenté de corriger ce défaut majeur grâce à un *mod*, mais malheureusement sans succès car ce dernier n'était pas compatible avec la version de *Minecraft* que nous étions contraint d'utiliser (voir section 4.2.2).



(a) Exemple de textures photoréalistes*.

(b) Exemple de *shader***.

*Capture d'écran issue du *SEK Photo Realism Resource Pack 1.8.8* par *SavageStyle*.

**Capture d'écran issue du *Realistico Resource Pack 1.8.8* par *matteorizzo*.

FIGURE 4.4 – Exemple de modifications de la qualité visuelle de *Minecraft*.

Un dernier point a entériné notre choix, la popularité de *Minecraft*. En effet, il est difficile de trouver des participants pour les expérimentations, et nous avons pensé que l'évocation de *Minecraft* saurait motiver les joueurs invétérés comme les néophytes curieux. Ce postulat s'est confirmé lors des courts échanges que nous avons eu avec les participants de nos expérimentations. Le choix de notre 3D-IVE arrêté, il nous restait alors à définir la tâche à proposer aux participants ainsi que les matériels et mesures à mettre en place. Ces points seront détaillés dans les sections liées à chaque expérimentation.

4.2 Première expérimentation

4.2.1 Objectif et activité proposée

L'objectif de ces travaux de thèse est d'étudier l'impact des interactions sociales vocales sur l'expérience utilisateur dans les 3D-IVEs. Par cette première expérimentation, nous proposons, grâce à la comparaison de deux contextes expérimentaux dont la seule différence réside dans la possibilité ou non de communiquer vocalement, de répondre à l'objectif suivant : observer, évaluer et enregistrer les

éventuelles différences, variations entre ces deux contextes vis-à-vis des facteurs de l'expérience utilisateur que nous avons mis en exergue durant notre étude de la littérature, cela en utilisant les moyens de mesures proposés dans le chapitre méthodologie.

Afin de favoriser l'occurrence d'interactions sociales vocales, tout en permettant aux participants n'étant pas en mesure de communiquer de participer et d'apprécier l'activité proposée, le choix de cette dernière était à la fois fondamental et délicat. Nous l'avons déjà évoqué (voir section 4.1.2), la nécessité d'une activité réalisable à la fois avec et sans la possibilité de communiquer était primordial. C'est pourquoi nous avons mis en place une tâche de construction collaborative où les participants devaient travailler de concert sur la même construction afin de la terminer dans le temps imparti. Pour ce faire, les participants avaient à leur disposition, un plan présentant la construction et les matériaux *Minecraft* dédiés à chaque partie de cette dernière, ainsi qu'un modèle de taille réelle (détails en figure 4.8). Dans le but de proposer une expérience aussi satisfaisante que possible, le choix du sujet de la construction était lui aussi important, se devant d'être à la fois amusant et valorisant pour l'utilisateur une fois la construction terminée. Au regard de la constitution potentielle de notre échantillon de participants, des personnes plutôt jeunes (dont le choix sera expliqué dans la section 4.2.3), nous avons décidé de leur faire construire une statue représentant un personnage issu de l'univers « *Marvel* » : « *Loki* » le frère maléfique de « *Thor* », dieu du tonnerre. Afin que la tâche soit réalisable et cohérente avec l'utilisation des blocs de *Minecraft*, nous ne pouvions proposer un modèle réaliste et avons opté pour une version *chibi*³ *pixel art*⁴ de « *Loki* » créée par *Mattmadeacom*⁵ et inspiré des personnages du jeu vidéo pour mobiles et tablettes « *Tiny Tower* »⁶. Le modèle est présenté en figure 4.5. Cette statue, mesurant 23 blocs de hauteur et 14 de largeur au maximum, est composée de 360 blocs (deux couches de 180 blocs) de 7 matériaux différents. Nous avons par la suite testé la faisabilité de la tâche, d'une part, nous même et d'autre part avec le concours de joueurs néophytes sur *Minecraft*. Ces derniers ont testés la construction avec et sans communication vocale, au regard de leurs performances, nous pouvions affirmer que la construction en binôme, échanges vocaux ou non, était faisable dans le temps imparti, soit 25 minutes.

Nous étions aussi désireux de mesurer la présence virtuelle par la physiologie pour appliquer la méthode de la triangulation de données proposée dans le chapitre méthodologie. Afin de répondre à cet objectif, une courte activité de cinq minutes tenait place à la fin de la session d'immersion. Les participants devaient défendre leur construction contre une horde de monstres déferlant de toutes parts. L'unique but de cette phase était de générer du stress afin d'en mesurer l'occurrence et ainsi de pouvoir, comme l'a fait Meehan *et al.* (2002), faire le lien entre réactions physiologiques dans le réel et dans le virtuel afin d'attester du sentiment de présence virtuelle.

3. Chibi: terme japonais qui définit une petite personne, un bébé ou un enfant. Le terme comparable en français est avorton ou vilain. Le terme est utilisé pour désigner la version « jeune » d'un personnage de manga ou tout autre personnage fictif redessiné pour avoir l'air d'un enfant ou d'un bébé.

4. Pixel art: terme désignant une composition numérique qui utilise une définition d'écran basse et un nombre de couleurs limité et où le positionnement précis de chaque pixel a son importance.

5. Voir : <http://mattmadeacom.com/art/Avenge-me-bitizens-325798812>

6. Voir sur l'Apple AppStore : <https://itunes.apple.com/fr/app/tiny-tower-free-city-building/id422667065?mt=8>



FIGURE 4.5 – Version *chibi pixel art* du personnage *Loki* par *Mattmadea.comic*.

4.2.2 Environnement Virtuel

Il est tout d'abord important, dans un souci de reproductibilité, de citer la version de *Minecraft* utilisée. En effet, certains des *mods* dont nous avons besoin ne sont pas compatibles avec la dernière version du logiciel. Dans le cadre de notre première expérimentation, nous avons donc utilisé la version 1.6.2 de *Minecraft*. Les explications nécessaires à la mise en place de l'environnement utilisé sont listées en annexe de ce document, voir section B.2. Cette annexe comprend un lien de téléchargement de l'environnement, une liste des *mods* utilisés et leurs descriptions ainsi que des directives d'installation et de configuration.

L'environnement virtuel que nous avons utilisé pour cette première expérimentation se découpe en deux parties : l'une est une zone réservée aux participants où se tenait l'activité de construction et l'autre une zone cachée, réservée à l'expérimentateur. Nous allons en premier lieu traiter cet environnement d'un point de vue général pour ensuite donner des explications plus précises de ces deux parties.

Nous l'avons brièvement évoqué, il est possible dans *Minecraft* de créer des environnements vides et plats. Nous pourrions penser que l'utilisation d'un tel environnement favoriserait la concentration des participants sur l'activité proposée. Néanmoins, nous avons préféré conserver un environnement « normal » afin que ce dernier paraisse plus vivant, soit plus agréable à l'œil et ne déçoive pas les attentes des participants quant à l'immersion dans le monde de *Minecraft* et non pas dans un environnement vide utilisant les fonctionnalités du logiciel. Toujours dans un souci esthétique, nous avons changé le pack de textures par défaut par le pack *Sphax PureBDcraft 64x64* (voir lien en annexe, section B.2). La figure 4.6 présente une comparaison entre le pack de textures par défaut et celui utilisé. Ce pack de texture en 64 x 64 pixels améliore grandement la qualité visuelle du logiciel sans pour autant demander trop de ressources matérielles supplémentaires. Il a aussi été choisi pour son

aspect BD peu réaliste qui correspond bien à l'esprit *Minecraft*, afin de ne pas créer d'incohérences entre textures et modélisations. Vis-à-vis des avatars, nous avons conservé le *skin*⁷ par défaut. Pour figer le cycle jour / nuit et les conditions météorologiques de *Minecraft* et donc garantir des conditions d'expérimentation identiques d'une session sur l'autre, un système impliquant des entonnoirs et de la *redstone* permettait l'activation toutes les secondes d'un *command block* (en annexe, voir figure B.5). Les *command blocks* sont un type de bloc spécial dans lequel il est possible d'inscrire une commande ; à l'activation du bloc par un signal *redstone* (bouton, levier ou autre) la commande est exécutée. Dans le cas présent, la commande permettait de bloquer l'heure sur midi et le temps sur ensoleillé.



(a) Pack de textures par défaut.

(b) Pack de textures *Sphax PureBDcraft 64x64*.

FIGURE 4.6 – Différence entre les packs de textures par défaut et *Sphax PureBDcraft 64x64*.

Environnement virtuel → Partie participants : avant de nous intéresser à l'environnement en lui-même, donnons un aperçu de l'interface utilisateur de *Minecraft*. Dans le cadre de nos expérimentations, l'interface utilisateur était réduite à quatre éléments, en effet nous avons interdit l'ouverture de l'inventaire et donc des fonctionnalités de fabrication n'ayant pas lieu d'être utilisées pour la complétion de la tâche demandée. Les points suivants font référence à la figure 4.7 :

1. Bras de l'avatar de l'utilisateur, bouge légèrement dans un mouvement de balancement lorsque l'utilisateur se déplace et affiche le bloc sélectionné dans l'inventaire ;
2. l'inventaire dans lequel sont affichés tous les objets (matériaux ou outils) en possession de l'utilisateur. L'hexagone rouge indique l'emplacement et donc l'objet sélectionné ;
3. viseur permettant à l'utilisateur de voir avec quel bloc ou élément il va entrer en interaction ;
4. liseré noir apportant une aide visuelle supplémentaire en soulignant l'emplacement où sera posé un bloc le cas échéant.

La zone dans laquelle évoluent les participants était constituée de plusieurs éléments afin de per-

7. Skin : en français, *peau*, désigne la texture enrobante d'un objet virtuel, plus particulièrement d'un avatar ou d'un agent.

FIGURE 4.7 – Éléments de l'interface utilisateur de *Minecraft*.

mettre la bonne conduite de la tâche demandée. Les points suivants font référence à la figure 4.8 :

1. Panneau d'information proposant le plan de la statue à construire ainsi que, pour chaque partie, le nom du matériau correspondant (nom visible sur l'inventaire lors de la sélection). Le plan présente aussi un liseré noir délimitant chaque bloc afin de faciliter leur comptage ;
2. modèle taille réelle de la statue à construire. Nous avons jugé qu'en plus du plan, un modèle de taille réelle serait d'une grande aide pour les participants ;
3. blocs brillants de positionnement des premiers blocs de la statue, placés afin que les participants ne perdent pas de temps à décider de l'emplacement de leur construction ;
4. panneau d'information présentant les contrôles à la manette du jeu. Comme l'utilisation de la manette n'est pas usitée dans le cadre de *Minecraft* (nous reviendrons sur ce choix), ce panneau présente la fonction de chaque bouton afin que les participants n'éprouvent pas de difficultés pouvant réduire leur sentiment de présence virtuelle. Un visuel plus détaillé des panneaux 3 & 4 est disponible en annexe de ce document, voir figure B.1.

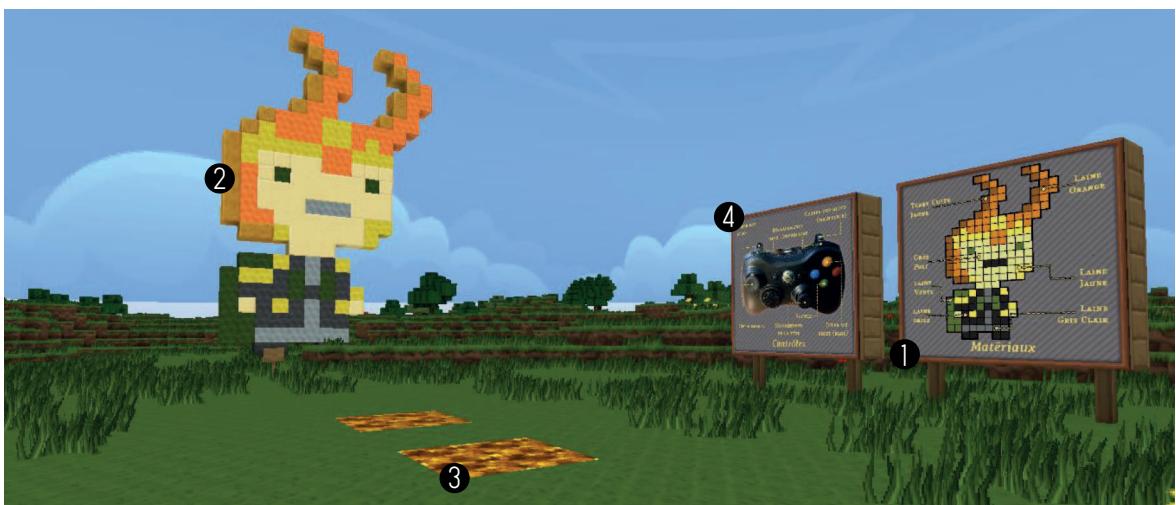


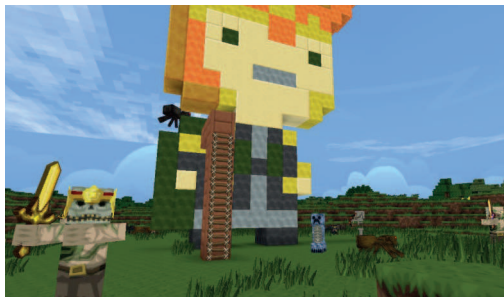
FIGURE 4.8 – Environnement participant de l'expérimentation 1.

Pour la seconde partie de l'expérimentation, à savoir l'attaque de la zone de construction par des

vagues de monstres, certains changements étaient opérés dans l’environnement. Les points suivants font référence à la figure 4.9. La sous-figure 4.9a présente une *turtle*⁸, un petit robot constructeur capable d’exécuter des scripts et permettant ainsi d’automatiser la construction d’éléments. Le rôle de cette *turtle* était la construction d’une couronne au-dessus de la statue en cours de construction (sous-figure 4.9b) sertie en son centre d’un *nexus*⁹. Cet élément, dès son activation, attire les monstres alentours qui tenteront alors de le détruire (sous-figure 4.9c). Durant les cinq minutes de l’attaque, les avatars des participants risquaient fortement de mourir sous les coups de leurs ennemis, entraînant ainsi la perte des objets de leur inventaire. C’est pourquoi un distributeur d’épées était placé à côté des panneaux d’information (sous-figure 4.9d) afin de permettre aux participants de s’armer pour repartir au « cœur de la mêlée ».

(a) *Turtle*, constructeur de la couronne.

(b) Couronne, appât à monstres.



(c) Monstres attaquant la statue des participants.



(d) Distributeur d’épées en cas de mort des avatars.

FIGURE 4.9 – Éléments de la seconde partie de l’expérimentation 1.

Environnement virtuel → Partie expérimentateur : cette partie avait pour objectif de permettre à l’expérimentateur de lancer les commandes nécessaires au bon déroulement de l’expérimentation le plus rapidement possible. Nous avons alors fait en sorte de résumer ces actions à une pression sur un bouton unique (figure 4.10a). À l’appui, le bouton envoyait un signal *redstone* vers deux ensembles de *command block* (figure 4.10b) — un pour chaque participant — permettant de remplir leur inventaire des matériaux nécessaires à la construction. La commande présentée en figure 4.10c a pour effet de donner (*/give*) au participant (*Voleur2Poules*) une quantité (*64*) de blocs de laine (*35*) de couleur grise claire (*08*). Ces ensembles de *command block* permettaient aussi la téléportation des participants dans la zone de construction dans l’éventualité où la courte attente avant le lancement effectif de l’expérimentation les aurait poussés à aller explorer les environs. Le bouton général envoyait aussi un signal à deux ordinateurs⁸ utilisés comme minuteurs et permettant, l’un de piloter la *turtle*

8. Voir *mod ComputerCraft*, en annexe - section B.2.

9. Voir *mod Invasion Mod*, en annexe - Section B.2.

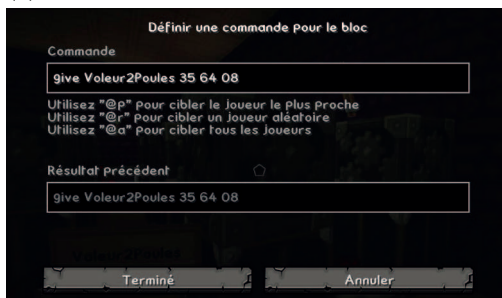
précédemment citée, l'autre de lancer le script de fin de l'expérimentation (figure 4.10d). Grâce à ces circuits de *redstone* et aux *command blocks*, l'expérimentateur était en mesure de lancer la session en quelques secondes et n'avait plus aucune action à mener avant la fin de cette dernière.



(a) Bouton de lancement de l'expérimentation.



(b) *Command blocks* dédiés à un participant.



(c) Commande de remplissage de l'inventaire.



(d) Ordinateur-minuteur exécutant le script de fin d'expérimentation.

FIGURE 4.10 – Éléments de la partie réservée à l'expérimentateur.

4.2.3 Participants

Les participants de cette première expérimentation ($N = 52$; 13 femmes, 39 hommes) étaient âgés de 19 à 27 ans ($M = 21,15$; $ÉT = 1,830$). Nous avons choisi une population homogène de jeunes adultes car l'âge des participants, conformément à Siriaraya et Siang Ang (2012), a un impact sur le sentiment de présence sociale expérimenté. Les participants étaient tous étudiants de l'*École de Design Nantes Atlantique*¹⁰, établissement où s'est tenu cette expérimentation du 16 au 20 juin 2014. Le recrutement s'est basé sur le volontariat par un courriel envoyé à tous les étudiants de l'établissement. Afin d'éviter tout biais engendré par une familiarité entre individus nous avons précisé dans le courriel de recrutement que nous désirions que les binômes ne se connaissent pas et donc que les participants ne se concertent pas afin de passer sur le même créneau qu'un ami ou un camarade de promotion. Nous avons néanmoins dû éliminer de nos résultats deux binômes présentant une forte familiarité.

Comme nous l'avons précédemment expliqué, il nous était nécessaire de connaître certaines données sur nos participants telles que leur familiarité avec le jeu vidéo en général et plus précisément avec *Minecraft*. Ces informations leur ont été demandées au moyen d'un questionnaire pré expérimentation. Tous les participants parlaient français couramment nous assurant une bonne compréhension des explications données par les expérimentateurs. Seulement deux participants ont dit ne pas du tout

10. Voir : <http://www.lecolededesign.com>

jouer au jeu vidéo, le reste des participants atteste jouer aux jeux vidéo à diverses fréquences (62% des participants jouent moins de 5 heures par semaine, 26% jouent de 5 à 15 heures par semaine et 12% jouent plus de 15 heures par semaine). Il leur a aussi été demandé leur familiarité avec les jeux vidéo multijoueurs en ligne, 36,5% des participants n'avaient aucune expérience de ce type. Les autres (63,5%) attestent jouer en ligne à diverses fréquences (48,5% des participants moins de 5 heures par semaine, 27,5% de 5 à 15 heures par semaine et 24% plus de 15 heures par semaine). Nous pouvons donc dire que nous étions en présence d'une population de joueurs plutôt expérimentés en solo mais moins familiers du jeu vidéo en ligne. Nous avons également interrogé les participants sur le degré de familiarité avec le système d'interaction, en leur demandant sur quel type de plateforme ils jouaient : la plupart d'entre eux jouent sur ordinateurs (84,6%), mais une proportion importante joue aussi sur consoles de salon (61,5%). Tous les participants ont attesté savoir comment utiliser une manette de jeu. Quant à leur connaissance préalable de *Minecraft*, 84,6% des participants connaissaient le jeu sans pour autant en avoir une très grande expérience (pratique). Dans ce pourcentage, 45,5% avaient une expérience plus ou moins grande de *Minecraft* en multijoueurs (10% moins de 10 heures, 15% de 10 à 50 heures et 75% plus de 50 heures.)

4.2.4 Mesures

Comme nous l'avons présenté dans le chapitre consacré à la méthodologie appliquée à notre étude, nous avons construit notre expérimentation autour de mesures objectives relatives à la performance de nos participants, sur un questionnaire post-expérimentation permettant de recueillir les appréciations des participants relativement aux facteurs que nous souhaitons étudier et enfin sur des données physiologiques. Malheureusement, en raison de problèmes techniques (que nous détaillerons dans la section *analyse et synthèse* de cette expérimentation, section 4.2.8 en page 111), les seuls résultats valides que nous avons obtenus ici sont ceux issus de notre questionnaire ainsi que des observations que nous avons pu mener sur l'accomplissement de la tâche demandée aux participants. Nous présenterons néanmoins la mise en place de ces mesures et les équipements y étant corrélés.

Nous allons à présent, en nous appuyant ici encore, sur la taxonomie de facteurs du modèle de Pallot *et al.* (2013), détailler l'origine dans la littérature des questions posées dans notre questionnaire et nous y adjoindrons les cas échéant les autres mesures mises en place. L'ensemble des questions est disponible en annexe B.5. Notre questionnaire était composé de questions à échelles de Likert et à échelles sémantiques bipolaires à cinq niveaux. Si ce type d'échelles présente fréquemment sept niveaux plutôt que cinq, nous justifions notre choix par un souci de compréhension et de réappropriation du questionnaire par les participants. En effet, les questions posées tendent à mesurer des phénomènes subjectifs qu'il n'est pas toujours aisé de bien appréhender. Un nombre de niveaux d'échelles plus réduit facilite le choix du positionnement. A titre d'illustration, référons-nous à la figure 4.11a : comment pouvons-nous sémantiquement décrire les niveaux deux et trois ? Le niveau deux, signifie « *Pas d'accord* » et le niveau trois « *Plutôt pas d'accord* », une nuance que nous avons estimée difficile à faire eu égard au contexte dans lequel les participants font remplir le questionnaire (après trente minutes

d’immersion qui, pour un primo-accédant à une expérience de réalité virtuelle immersive, peuvent être éprouvantes). Nous avons alors préféré l’utilisation d’échelle à cinq niveaux (figure 4.11b) ou le niveau deux signifie « *Plutôt pas d’accord* » et ou le niveau trois permet la neutralité. Il est aussi fréquent de voir dans ce type de questionnaire l’utilisation de deux questions à échelles inverses évaluant le même facteur ou la même variable. Par exemple, une question ayant pour énoncé « *Je me suis amusé* » et une autre « *Je me suis ennuyé* ». Toutefois, comme nous l’avons évoqué dans le cadre des groupes de discussions vis-à-vis du temps (voir section 3.1.2), un questionnaire comportant un trop grand nombre de questions peut ennuyer les participants qui risquent de vouloir en finir au plus vite compromettant ainsi la sincérité de leurs réponses.

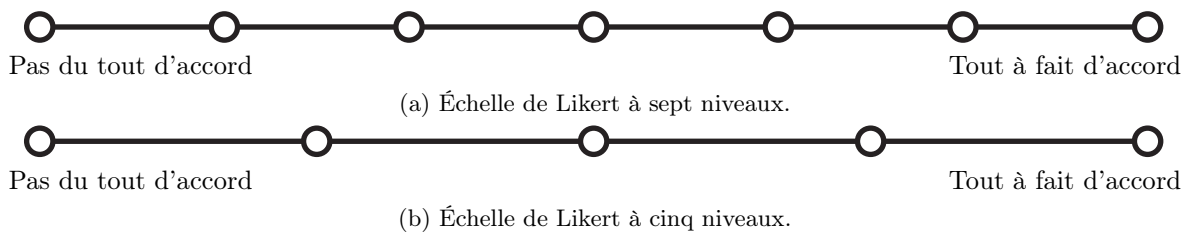


FIGURE 4.11 – Comparaison entre échelles de Likert à sept et cinq niveaux.

Qualité de Service ▶ Vivacité Sensorielle

▶ **Intrusivité** : nous avons souhaité mesurer la variable **intrusivité** que nous avons ajoutée au facteur vivacité sensorielle du modèle de qualité des DIMEs de Wu *et al.* (2009) afin de savoir si l’expérience vécue était influencée négativement par les matériels utilisés. Pour ce faire, notre questionnaire comporte cinq questions vis-à-vis de la potentielle gêne occasionnée par les matériels avec lesquels les participants étaient équipés. Voir questions 12.1 à 12.5 dans notre questionnaire en annexe B.5.

Expérience (Utilisateur) Immersive ▶ Processus Expérientiel

▶ **Présence Sociale** : afin de mesurer l’impact de nos deux contextes (avec et sans communication vocale) sur le sentiment de présence sociale, nous avons inclus dans notre questionnaire sept questions extraites du « *Networked Minds social presence questionnaire* » de Biocca *et al.* (2001). Ces questions évaluent la présence sociale selon quatre facteurs, l’attention allouée, la compréhension mutuelle, l’interdépendance comportementale et l’assistance mutuelle. Voir questions 9 et 14 à 19 dans notre questionnaire en annexe B.5.

▶ **Réponse Émotionnelle** : afin de saisir l’état émotionnel général des participants durant l’expérimentation, la première question de notre questionnaire leur demandait de se positionner sur une échelle à cinq niveaux illustrée par des schémas des émotions (proche des *emoticons*). Échelle allant de la frustration à la joie. Voir question 1 dans notre questionnaire en section B.5.

Expérience (Utilisateur) Immersive ▶ Processus Rationnel

▶ **Flow Psychologique** : afin de mesurer l’impact de nos deux contextes sur le *flow* psychologique des participants, nous avons inclus dans notre questionnaire deux questions rédigées par nos soins, l’une requérant une évaluation de l’amusement, l’autre du niveau du niveau de défi éprouvé. Voir

questions 10 et 11 dans notre questionnaire en annexe B.5.

► **Téléprésence** : afin de mesurer l'impact de nos deux contextes sur le sentiment de présence virtuelle, nous avons inclus dans notre questionnaire cinq questions issues du « *Igroup Presence Questionnaire* » de Schubert *et al.* (2001) (voir question une à cinq du IPQ¹¹). Ces questions évaluent le sentiment de présence virtuelle ainsi que deux variables de l'immersion perceptuelle de Lombard et Ditton (1997) et Biocca et Delaney (1995) qui dans notre cas font partie de la QoS : la vivacité sensorielle et le degré de contrôle. Bien qu'afférente à la QoS, leur évaluation est ici pertinente puisqu'elles sont toutes deux très largement impliquées dans le sentiment de présence virtuelle. Voir questions 2 à 6 dans notre questionnaire en annexe B.5. D'un autre côté, nos participants seront équipés d'un capteur ECG et d'un capteur GSR afin de mesurer l'occurrence du stress lors de la seconde partie de l'expérimentation (attaque du *nexus* par les monstres) afin d'attester du sentiment de présence virtuelle, en accord avec Meehan *et al.* (2002) et Pertaub *et al.* (2002). Cette mesure physiologique est complétée par une captation par casque EEG afin, là encore, d'évaluer le sentiment de présence virtuelle par comparaison avec les travaux de Clemente *et al.* (2013) qui définissent les zones du cerveau montrant une plus grande activité lorsqu'un sujet est en état de présence virtuelle.

Qualité de l'Expérience

► **Comportement Social** : nous souhaitons recueillir des informations sur les échanges vocaux entre participants. Naturellement, ces questions n'ont été soumises qu'à la moitié des participants, ceux ayant vécu l'expérience avec la possibilité de communiquer vocalement. Les réponses à ces questions ne sont donc en aucun cas utilisables comme matériel de comparaison mais nous ont permis de récolter des données afférentes aux facteurs et variables suivants :

- Fréquence des interactions sociales vocales (voir question 21 en annexe B.5) ;
- nature des interactions sociales vocales (voir questions 25, 27, et 28) ;
- variable attention allouée (voir question 26, 27 et 28) de la co-présence, facteur de la présence sociale (Biocca *et al.*, 2001) ;
- comportement empathique (voir question 23) de la QoE ou chez Biocca *et al.* (2001), de l'engagement psychologique, facteur de la présence sociale ;
- variable compréhension mutuelle (voir question 24), de l'engagement psychologique, facteur de la présence sociale (Biocca *et al.*, 2001) ;
- variable assistance mutuelle (voir question 22) de l'engagement comportemental, facteur de la présence sociale (Biocca *et al.*, 2001) ;
- réponse émotionnelle (voir question 30) de l'expérience (utilisateur) immersive (Pallot *et al.*, 2013) ;
- variable engagement (voir questions 29) de la présence virtuelle (Witmer et Singer, 1998) ;
- qualité de service (voir questions 31)(Wu *et al.*, 2009) ;
- variable degré de contrôle (voir questions 32) de la présence virtuelle (Steuer, 1992).

11. IPQ : <http://www.igroup.org/pq/ipq/download.php>

Des zones libres étaient aussi disponibles afin que les participants puissent faire part d'autres sujets de conversations qui se seraient tenus durant la session d'immersion.

► **Efficacité** : afin de mesurer l'impact de nos deux contextes sur l'efficacité des participants d'après leur perception de cette dernière, nous avons inclus dans notre questionnaire deux questions rédigées par nos soins (voir questions 7 et 8 dans notre questionnaire en annexe B.5). D'autre part, dans le but d'obtenir une évaluation quantitative objective de la performance de nos binômes de participants, nous avons effectué un comptage du nombre de blocs posés dans le temps alloué à la construction ainsi qu'une observation des erreurs éventuelles dans la construction par rapport au modèle donné.

Autres questions : une des questions de notre questionnaire portait sur l'occurrence du stress (que nous classons ici dans la catégorie *autres questions*, car ne pouvant être considéré comme un comportement ou une réponse émotionnelle ¹²). Le but était ici de confirmer les mesures physiologiques par ECG et GSR vis-à-vis du sentiment de présence virtuelle. Voir question 13 dans notre questionnaire en annexe B.5.

En toute fin de questionnaire, une page vierge était destinée à toutes les remarques dont les participants auraient voulu nous faire part. Mais, comme nous l'avons expliqué dans le chapitre méthodologie (voir section 3.1.2), sans guides ou préconisations les réponses données ne sont d'aucune aide pour la validation de notre hypothèse et ne concernent en majeure partie que le ressenti des participants vis-à-vis de l'utilisation du matériel d'affichage ou d'interaction mis en place. Certaines apportent néanmoins des éclairages intéressants qui nous ont permis des améliorations lors de la conception de notre seconde expérimentation.

Questionnaire a posteriori : un très court questionnaire de quatre questions a été envoyé aux participants quatre mois après cette expérimentation, afin de recueillir des données sur le potentiel émotionnel d'une telle expérience ainsi que sur leur niveau de satisfaction d'un point de vue hédonique (voir annexe B.6). Cet envoi tardif se justifie par le fait que, dans nos souvenirs, nous ne gardons qu'une vision générale qui représente les émotions ou les ressentis prédominants de l'expérience que nous nous remémorons.

4.2.5 Équipement des participants

***NOTA** : les explications données ici sont présentées sous la forme d'un schéma (figure 4.12) en page 99. Les chiffres entre parenthèses ponctuant la section suivante y font référence.*

Afin d'atteindre un haut niveau de QoS, nous avons souhaité utiliser l'*Oculus Rift* version *DK1* (1) comme périphérique d'affichage. Cet HMD était au moment de la conduite de cette expérimentation (juin 2014) le produit présentant le meilleur compromis prix / performances du marché, ce dernier présentait néanmoins un inconvénient majeur, une forte propension à générer de la *cyber sickness*.

12. Stress : état réactionnel de l'organisme soumis à une agression brusque. Le stress est fondamentalement différent de l'anxiété, celle-ci est une émotion alors que le stress est un mécanisme de réponse pouvant amener différentes émotions, dont l'anxiété. (Larousse 2015)

Toutefois, il occulte totalement le monde réel et malgré une résolution basse (1280 x 800 soit 640 x 800 par œil). Il permet de pouvoir bouger la tête indépendamment du corps, ce qui favorise sensiblement le sentiment de présence virtuelle (Hupont *et al.*, 2015 ; Seibert, 2014). Le passage de *Minecraft* en vision stéréoscopique spécifique à l'*Oculus Rift* s'est fait grâce à un *mod*¹³. C'est son utilisation qui nous a contraints à l'emploi de la version 1.6.2 de *Minecraft*, bien que très antérieure à la version courante au moment de la création de notre expérimentation (ver. 1.7.9) et présentant donc des contraintes supplémentaires quant à la comptabilité avec d'autres *mods* utilisés.

Pour permettre aux participants d'entendre les sons générés par l'environnement et de pouvoir communiquer vocalement — selon la condition expérimentale — ils étaient équipés de casque audio comprenant un microphone (2). *Minecraft* ne supportant pas par défaut la communication vocale, nous avons fait le choix parmi les nombreuses solutions possibles d'utiliser *TeamSpeak*¹⁴. Ce logiciel, spécifiquement développé pour ce type d'usage, ne consomme que très peu de ressources matérielles et est très stable quant à la transmission de la communication vocale. Un ordinateur différent de celui des participants a été utilisé comme serveur afin de sauvegarder les ressources des postes utilisateur (3).

L'utilisation de l'*Oculus Rift* a aussi conditionné le périphérique d'interaction à utiliser. En effet, utiliser un clavier et une souris sans pouvoir y jeter un regard en cas de perte de repères n'est pas acceptable. Nous leur avons alors préféré l'utilisation de manettes de jeu filaires (*Microsoft Xbox 360* pour *Windows*) (4) proposant une grande intuitivité quant à leur utilisation. Afin de rendre compatible *Minecraft* avec ce périphérique d'interaction, nous avons dû utiliser un logiciel externe¹⁵ permettant d'assigner les touches du clavier ou les mouvements de la souris aux commandes de la manette.

Ainsi équipés, les participants étaient en mesure d'interagir avec le même environnement virtuel (5) hébergé sur le troisième ordinateur dont nous avons précédemment parlé (serveur *TeamSpeak*)(6), via un réseau local supprimant ainsi tout problème de latence et facilitant la mise en marche et la réinitialisation du serveur à chaque session d'expérimentation (sauvegarde locale de l'environnement précédent et remplacement par un environnement vierge).

Si nous résonnons en termes de QoS, et plus particulièrement ici en termes de vivacité sensorielle, l'étendue de cette dernière n'implique que deux sens, la vue et l'ouïe. Ces deux sens sont ici totalement investis par les stimuli visuels grâce au port du HMD et du casque audio. Quant à sa profondeur, vis-à-vis de l'ouïe, l'utilisation de *TeamSpeak* et d'un réseau local assurent la bonne transmission de la communication vocale. En revanche, de la résolution assez faible de l'*Oculus Rift* — 640 x 800 par œil — résulte une visibilité des pixels. Nous savons aussi que dans sa version *DK1*, l'*Oculus Rift* présente une légère latence entre les mouvements de la tête de l'utilisateur et sa restitution dans le virtuel, deux éléments que nous savons impliqués dans le phénomène de *cyber sickness*. À l'éclairage de ces constatations, nous estimons que notre configuration offrira une QoS tout à fait satisfaisante, nous pensons néanmoins, en conséquence des défauts décrits précédemment de l'*Oculus Rift* que nous

13. Voir *mod Minecraft*, en annexe - Section B.2.

14. Voir : <https://www.teamSpeak.com/>.

15. Voir le logiciel *JoyToKey* : <http://joytokey.net/en/>. Le fichier de configuration créé pour nos expérimentations est disponible au téléchargement, voir en annexe - Section B.2.

risquons d’être confrontés à des problèmes liés à la « *cyber sickness* ». Malheureusement, ce phénomène étant très variable d’une personne à l’autre nous ne pouvons pas prédire dans quelle mesure ni avec quelle intensité le problème se présentera.

La partie inférieure (7) de la figure 4.12 présente les matériels de mesures physiologiques mis en place : le casque EEG (a) est un *EEG Headset* (aujourd’hui *Epoc+*) de la marque *Emotiv*¹⁶. Les capteurs ECG (b) et les capteurs GSR (c), proviennent de l’entreprise *Technology Ergonomics Applications*¹⁷ (TEA).

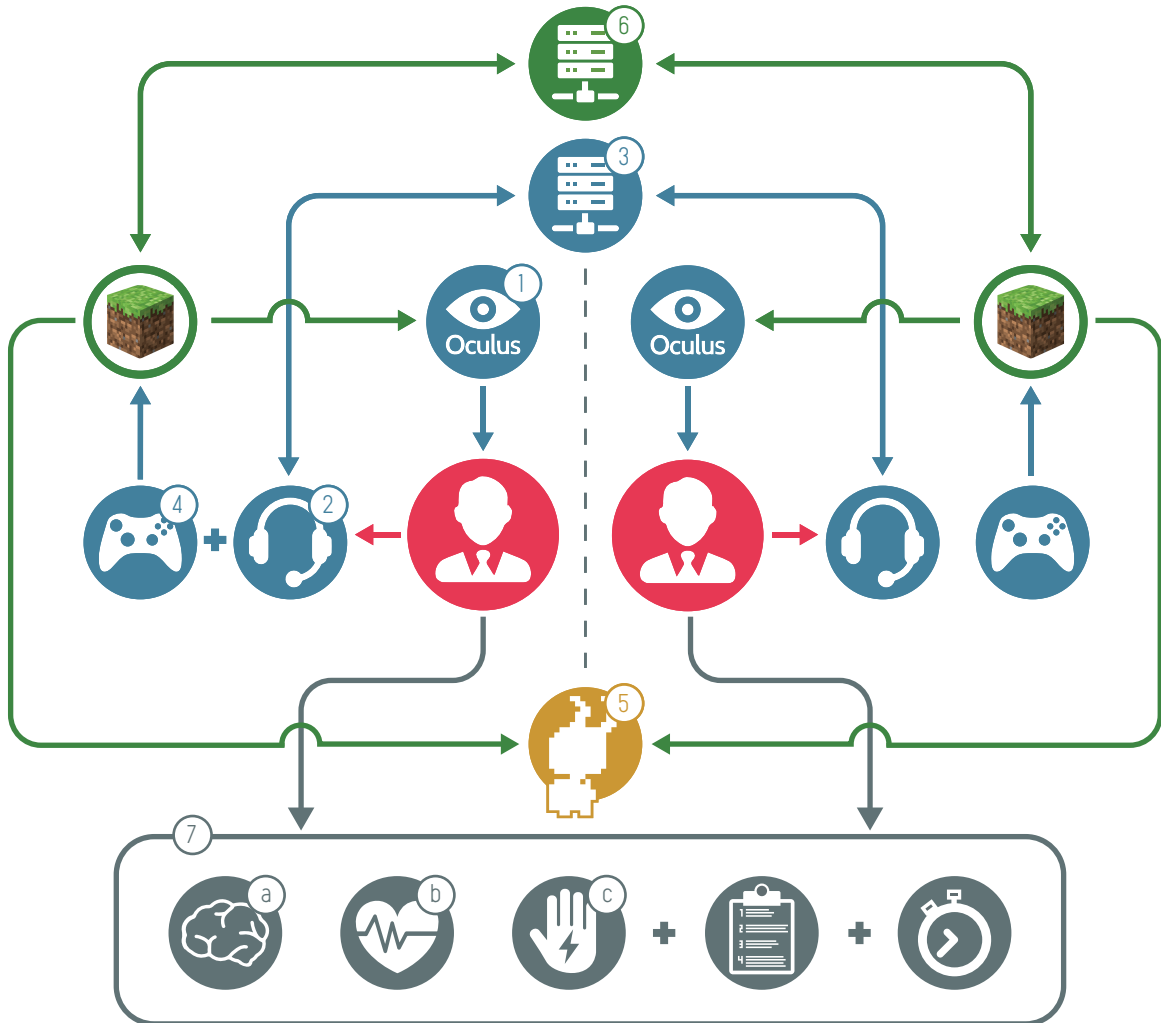


FIGURE 4.12 – Schéma de la configuration globale de l’expérimentation 1.

4.2.6 Procédure

Lors de la passation des participants, nous nous sommes en premier lieu informés sur la familiarité entre individus. Assurés de son absence, nous avons immédiatement séparé les binômes dans deux salles différentes — où nous avons préalablement installé nos postes d’expérimentation — afin qu’ils ne puissent pas communiquer a priori de la session d’immersion. Aucune information n’a été donnée sur les réels objectifs de l’expérimentation afin d’éviter d’influencer leurs réponses aux questionnaires

16. Voir site *Emotiv* : <https://emotiv.com/store/epoc-detail/>

17. Voir site *TEA* : <http://teaergo.com/site/en/products/manufacturers/tea>

ou leur état d'esprit en général. Les participants ont rempli un questionnaire pré expérimentation (voir annexe B.4) recueillant les informations citées dans la section 4.2.3 (expérience préalable de *Minecraft*, pratique du jeu vidéo, etc...). Nous avons ensuite donné aux participants des informations sur les contrôles de leur avatar grâce à la manette de jeu, les matériels dont ils allaient être équipés, le questionnaire post expérimentation (échelles de Likert et sémantiques bipolaires), l'interface et les modalités d'interaction de *Minecraft* (utilisation de l'inventaire) et la tâche qu'ils allaient devoir réaliser (le livret utilisé pour illustrer ces explications est disponible en annexe B.3). Les participants étaient ensuite équipés par les expérimentateurs puis, pour le premier équipé, mis en attente dans l'environnement. Dans la condition expérimentale où la communication vocale était autorisée, il était demandé aux participants s'ils étaient bien en mesure d'entendre leur partenaire.

L'expérimentateur pouvait alors lancer la session grâce au bouton dédié (voir section 4.2.2) sur le poste serveur. Les participants commençaient alors la réalisation des tâches demandées (voir section 4.2.1), soit, en premier lieu, la construction de la statue durant vingt-cinq minutes puis la défense du *nexus* durant cinq minutes. À la fin du temps imparti, les participants étaient téléportés dans une zone souterraine noire (sans lumière) et un message textuel leur signifiait la fin de la session d'expérimentation. Une fois déséquipés, ils étaient invités à remplir le questionnaire post expérimentation puis à quitter la salle. Là encore, malgré l'insistance des participants, aucune information sur les objectifs réels de l'expérimentation n'a été donnée. Cela, afin qu'en cas de contact avec un autre participant n'ayant pas encore pris part à sa session, aucun biais ne vienne influencer nos résultats.

4.2.7 Résultats

Nous avons, en vue de l'analyse quantitative de nos questionnaires, conduit des tests de normalité (Kolmogorov-Smirnov) sur chacune des variables obtenues (réponses aux questionnaires et observations) afin de savoir si nous pouvions utiliser ou non des tests paramétriques. Aucune des variables testées ne suivait une courbe de distribution normale. Ces résultats sont cohérents avec le nombre de participants ($N = 52$) mais surtout avec l'utilisation d'échelle de Likert à cinq niveaux seulement. Par conséquent, nous nous sommes reposés sur la famille des tests non paramétriques (Pallant, 2010). Au regard de nos conditions expérimentales : deux groupes de participants pour deux conditions d'expérimentation (voir 4.2.6), nous nous sommes appuyés sur le U-test de Mann-Whitney pour échantillons indépendants. Pour établir la signification statistique de nos résultats, nous avons utilisé le seuil de probabilité conventionnel, soit $p < 0,05$.

Nous considérons ici qu'un jeu de réponses au questionnaire est valide si les deux personnes composant un binôme sont restées immergées durant au moins dix minutes. En effet, cette durée permettait aux participants de se familiariser et d'interagir avec l'environnement virtuel, mais aussi de communiquer avec leur partenaire (si le contexte l'autorisait). Le temps moyen avant la pose d'un premier bloc (45,42 secondes) et le nombre moyen de blocs posés en dix minutes comparé au nombre moyen de blocs posés pendant la durée totale de la session (93,65 blocs contre 200,52 blocs) confirme la pertinence de ce choix.

Les résultats obtenus seront exposés suivant la même organisation utilisée dans la présentation des mesures mises en place (voir section 4.2.4), c'est-à-dire suivant l'organisation des facteurs proposés dans le modèle de qualité des DIMEs de Wu *et al.* (2009) et dans le modèle d'expérience utilisateur de (Pallot *et al.*, 2013) en découlant.

Qualité de Service ► Vivacité Sensorielle

► **Intrusivité** : cinq questions de notre questionnaire portaient sur l'intrusivité ou la potentielle gêne occasionnée par les matériels équipant les participants. Bien que nous n'ayons pas pu exploiter les données en provenance des capteurs physiologiques, il est tout de même intéressant d'exposer l'intrusivité qu'ils ont pu générer. Des données utiles non seulement pour l'analyse des résultats à suivre mais aussi pour notre seconde expérimentation. Les résultats présentés ici sont des moyennes sur l'ensemble des participants. Un U-test de Mann-Whitney pour échantillons indépendants n'a pas révélé de différence statistique significative entre les deux contextes expérimentaux. Le tableau 4.1 présente, pour chaque matériel, une liste du nombre de participants pour chaque niveau des échelles de Likert. Afin de faciliter la lecture, l'effectif le plus important est affiché en **rouge**. Dans les colonnes correspondantes aux capteurs ECG et GSR, on constate que le nombre de réponses est de 28 et non pas de 52 (nombre de participants total). En effet, nous nous sommes aperçus du mauvais fonctionnement de nos capteurs physiologiques à la moitié du temps allouée à notre expérimentation et avons décidé de ne plus en équiper les participants en raison du manque de fiabilité du matériel. Nous pouvons constater dans le tableau 4.1 que l'intrusivité des capteurs physiologiques a été évaluée comme très faible par un large pourcentage de l'échantillon en ayant été équipé. Le changement induit dans les conditions expérimentales par le choix de ne plus en équiper les participants peut donc être considéré comme marginal et sans impact sur les résultats obtenus par la suite.

Expérience (Utilisateur) Immersive ► Processus Expérientiel

► **Présence Sociale** : quel est l'impact de la communication vocale sur le sentiment de présence sociale ? Comme décrit dans la section *mesures* (section 4.2.4), l'évaluation de la présence sociale s'appuie sur quatre variables : l'attention allouée, la compréhension mutuelle, l'interdépendance comportementale et l'assistance mutuelle.

Vis-à-vis de la variable **attention allouée**, une question était posée : « *Vous avez fait preuve d'une attention particulièrement vive vis-à-vis de l'autre personne :* » (1 = Je n'ai pas été très attentif à l'autre & 5 = J'ai été très attentif à l'autre) — question 14 dans le questionnaire. Un U-Test de Mann-Whitney n'a révélé aucune différence statistique significative entre nos deux contextes expérimentaux, sans (SCV) et avec communication vocale (ACV). La figure 4.13 présente les moyennes des réponses à cette question pour les deux contextes, la faible différence entre contextes ainsi que les larges écarts-types sont cohérents avec le résultat du test statistique mené.

Échelle	Poids de l' <i>Oculus Rift</i> ¹		Motion sickness (<i>Oculus Rift</i>) ²	
	Effectifs	Pourcentage	Effectifs	Pourcentage
Pas du tout (1)	20	38,5 %	10	19,2 %
2	17	32,7 %	9	17,3 %
3	9	17,3 %	12	23,1 %
4	5	9,6 %	6	11,5 %
Beaucoup (5)	1	1,9 %	15	28,8 %
Total	52	100 %	52	100 %
Moyenne (écart-type)	2,04 (1,066)	—	3,13 (1,495)	—
—	Casque EEG ³		Capteur ECG ⁴	
Pas du tout (1)	28	53,8 %	23	82,1 %
2	6	11,5 %	4	14,3 %
3	5	9,6 %	0	0 %
4	8	15,4 %	0	0 %
Beaucoup (5)	2	3,8 %	1	3,6 %
Total	52	100 %	28	100 %
Moyenne (écart-type)	1,98 (1,315)	—	1,29 (,810)	—
—	Capteur GSR ⁵		—	
Pas du tout (1)	22	78,6 %	—	—
2	6	21,4 %	—	—
3	0	.000 %	—	—
4	0	.000 %	—	—
Beaucoup (5)	0	.000 %	—	—
Total	28	100 %	—	—
Moyenne (écart-type)	1,21 (,418)	—	—	—

Votre performance a-t-elle été affectée par le matériel dont vous étiez équipé :

¹ le poids de l'*Oculus Rift* ?

² un sentiment de malaise dû à l'*Oculus Rift* ?

³ le casque à électroencéphalogramme ?

⁴ le capteur de rythme cardiaque ?

⁵ le capteur de conductance de la peau ?

Questions 12.1 à 12.5 dans le questionnaire post-expérimentation, voir annexe B.5

TABLEAU 4.1 – Données effectives et moyennes de l'évaluation de l'intrusivité des matériels utilisés.

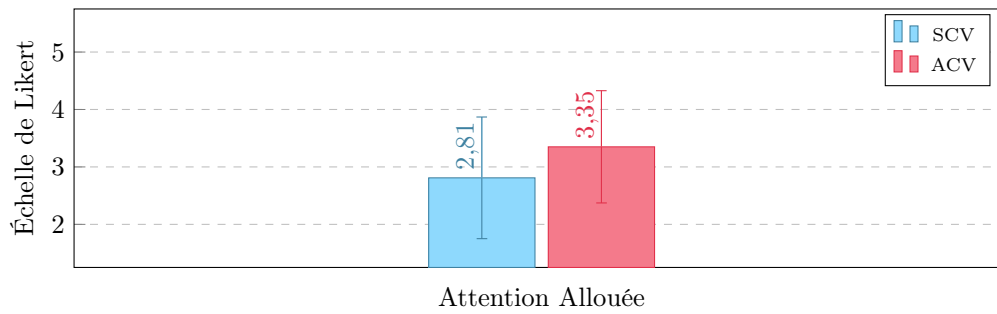


FIGURE 4.13 – Moyennes des évaluations de l'attention allouée.

Pour la variable **compréhension mutuelle**, deux questions étaient posées :

- « *L'autre personne vous a parue comprendre ce que vous désiriez faire (actions) :* » (1 = Pas du tout compris & 5 = Très bien compris) — question 15 dans le questionnaire. Un U-test de Mann-Whitney a révélé une différence statistique significative entre les contextes SCV (N = 26) et ACV (N = 26) ($U = 116$; $z = -4,172$; $p < ,0001$).
- « *J'ai compris ce que désirait l'autre personne (actions) :* » (1 = Pas du tout compris & 5 = Très bien compris) — question 16 dans le questionnaire. Un U-test de Mann-Whitney a révélé une différence statistique significative entre les contextes SCV (N = 26) et ACV (N = 26) ($U = 156$; $z = -3,453$; $p = ,001$).

La figure 4.14 présente les moyennes des réponses à ces deux questions pour les deux contextes. On constate alors que la variable compréhension mutuelle est significativement mieux évaluée dans le contexte ACV.

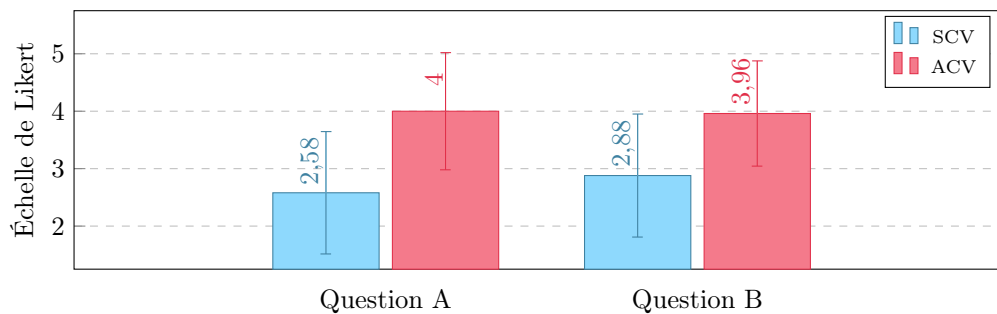


FIGURE 4.14 – Moyennes des évaluations de la compréhension mutuelle.

Vis-à-vis de la variable **interdépendance comportementale**, deux questions étaient posées :

- « *Vos actions étaient dépendantes de celles de l'autre personne :* » (1 = J'ai agi uniquement en fonction de moi & 5 = J'ai totalement agi en fonction des actions de l'autre) — question 17 dans le questionnaire. Un U-test de Mann-Whitney a révélé une différence statistique significative entre les contextes SCV (N = 26) et ACV (N = 26) ($U = 160,5$; $z = -3,393$; $p = ,001$).
- « *Votre style de jeu (par exemple, utilisation des échelles, manière de construire la statue) était une réponse directe au style de jeu de l'autre personne :* » (1 = J'ai agi de manière totalement autonome & 5 = Je me suis adapté au style de jeu de l'autre) — question 18 dans le question-

naire. Un U-test de Mann-Whitney n'a révélé aucune différence statistique significative entre les deux contextes vis-à-vis de cette question.

La figure 4.15 présente les moyennes des réponses à ces deux questions pour les deux contextes : on constate, en accord avec les tests statistiques menés que, l'interdépendance des actions est mieux évaluée dans le contexte ACV mais que la différence de contexte n'a pas d'impact significatif sur l'évaluation de l'interdépendance du style de jeu des participants.

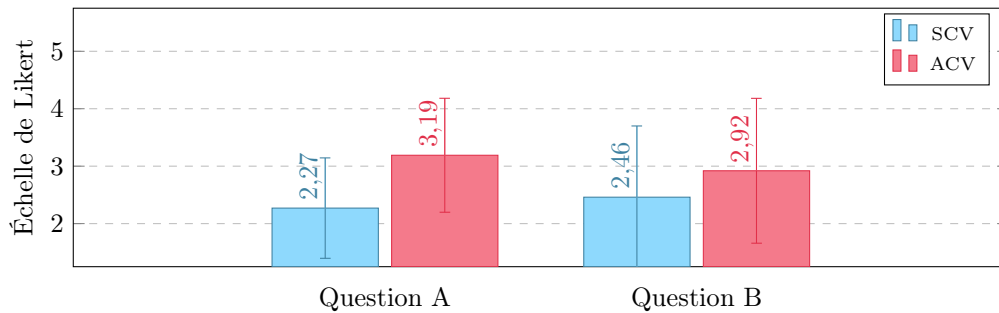


FIGURE 4.15 – Moyennes des évaluations de l'interdépendance comportementale.

Pour la variable **assistance mutuelle**, une question était posée : « *Votre partenaire vous a aidé dans l'accomplissement de la tâche :* » (1 = Pas du tout aidé & 5 = Beaucoup aidé) — question 19 dans le questionnaire. Un U-Test de Mann-Whitney n'a révélé aucune différence statistique significative entre nos deux contextes expérimentaux, SCV et ACV. La figure 4.16 présente les moyennes des réponses à cette question pour les deux contextes, la faible différence ainsi que les larges écarts-types sont cohérents avec le résultat du test statistique mené.

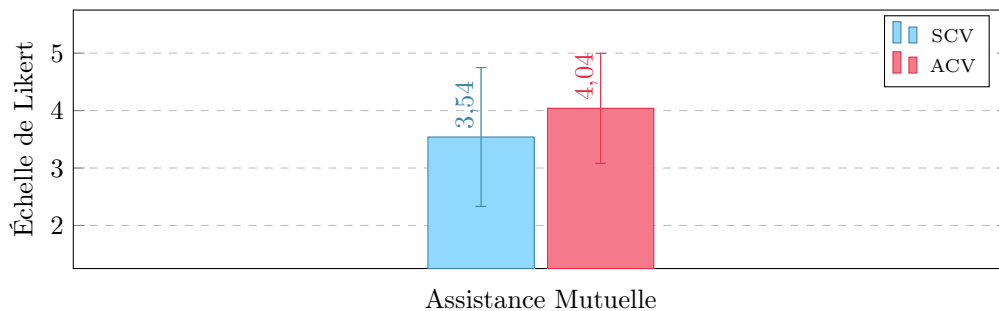


FIGURE 4.16 – Moyennes des évaluations de l'assistance mutuelle.

► **État Émotionnel :** quel est l'impact de la communication vocale sur l'état émotionnel ? Afin d'avoir un aperçu de l'**état émotionnel** général des participants, une question était posée, requérant le positionnement sur une échelle à cinq niveaux illustrée par des schémas des émotions (1 = Colère / frustration & 5 = Joie / allégresse) — question 1 dans le questionnaire. Un U-Test de Mann-Whitney n'a montré aucune différence statistique significative entre nos deux contextes expérimentaux, SCV et ACV. La figure 4.17 présente les moyennes des réponses à cette question pour les deux contextes, la faible différence ainsi que les larges écarts-types sont cohérents avec le résultat du test statistique mené. Néanmoins, nous constatons que les moyennes des évaluations tendent à montrer que l'état

émotionnel général des participants vis-à-vis de cette expérience d’immersion dans *Minecraft* était positif. Dans les deux contextes, ces moyennes sont très proches de quatre sur une échelle à cinq niveaux. Évaluation confirmée par un calcul des effectifs pour chacun des cinq niveaux : 38,5% des participants ont évalué leur état émotionnel à 4 sur une échelle à cinq niveaux et 30,8% à 5. 69,3% des participants ont donc évalué leur expérience comme positive d’un point de vue émotionnel.

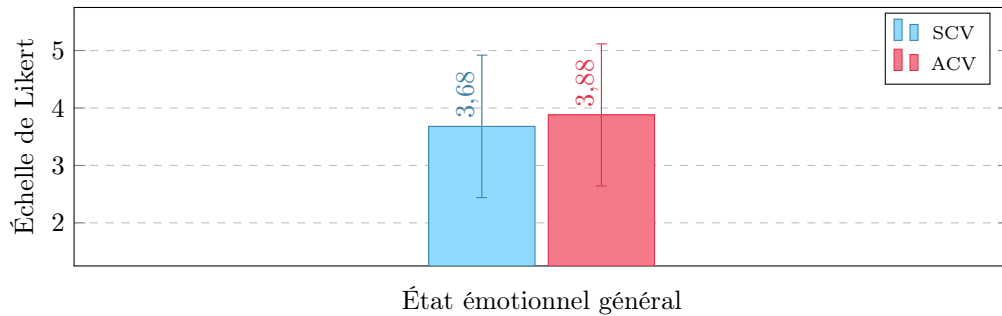


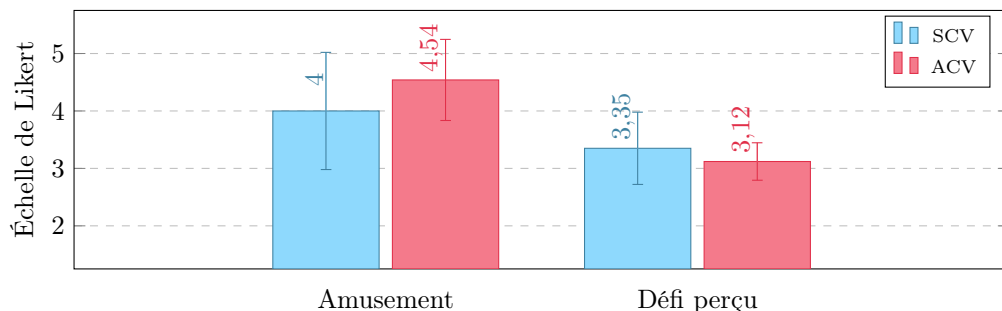
FIGURE 4.17 – Moyennes des évaluations de l’état émotionnel général

Expérience (Utilisateur) Immersive ▶ Processus Rationnel

▶ **Flow Psychologique** : quel est l’impact de la communication vocale sur le *flow* psychologique ?

Comme décrit dans la section *mesures* (section 4.2.4) l’évaluation du *flow* psychologique s’appuie sur deux variables, l’amusement (qui est la résultante de l’état de *flow*), et le niveau de défi perçu.

Vis-à-vis de la variable **amusement** du *flow* psychologique, une question était posée : « *Je me suis amusé durant cette expérience* : » (1 = Je me suis ennuyé & 5 = Je me suis amusé) — question 11 dans le questionnaire. Un U-test de Mann-Whitney a révélé une différence statistique significative entre les contextes SCV (N = 26) et ACV (N = 26) (U = 235 ; z = -2,064 ; p = ,039). Pour la variable **niveau de défi perçu**, une question était posée : « *Vous avez ressenti une forme de challenge durant cette expérience* : » (1 = Aucun challenge (trop facile) & 5 = Trop de challenge (Impossible d’accomplir les tâches demandées)) — question 12 dans le questionnaire. Un U-Test de Mann-Whitney n’a montré aucune différence statistique significative entre nos deux contextes expérimentaux, SCV et ACV. La figure 4.18 présente les moyennes des réponses à ces deux questions pour les deux contextes. On constate que les participants estiment s’être plus amusés dans le contexte ACV, mais que la différence de contexte n’a pas d’impact significatif sur l’évaluation du niveau de défi perçu, même si l’on constate une légère supériorité de son évaluation dans le contexte SCV.

FIGURE 4.18 – Moyennes des évaluations du *Flow* Psychologique.

► **Téléprésence** : quel est l'impact de la communication vocale sur le sentiment de présence virtuelle ? Comme décrit dans la section mesures (section 4.2.4), l'évaluation du sentiment de présence virtuelle s'appuie sur l'évaluation directe de ce sentiment ainsi que sur deux variables de la QoS, la vivacité sensorielle et le degré de contrôle. Deux questions étaient posées afin d'évaluer le sentiment de **présence virtuelle** :

- a) « *Dans l'environnement virtuel, vous aviez le sentiment « d'être sur place » : » (1 = Pas du tout « sur place » & 5 = Tout à fait « sur place ») — question 2 dans le questionnaire.*
- b) « *D'une certaine façon, vous aviez l'impression que le monde virtuel vous entourait : » (1 = Pas du tout entouré & 5 = Totalement entouré) — question 3 dans le questionnaire.*

Un U-test de Mann-Whitney n'a révélé aucune différence statistique significative entre les deux contextes vis-à-vis de ces deux questions. La figure 4.19 présente les moyennes des réponses à ces deux questions pour les deux contextes, et illustre la différence négligeable entre nos deux contextes vis-à-vis du sentiment de présence virtuelle.

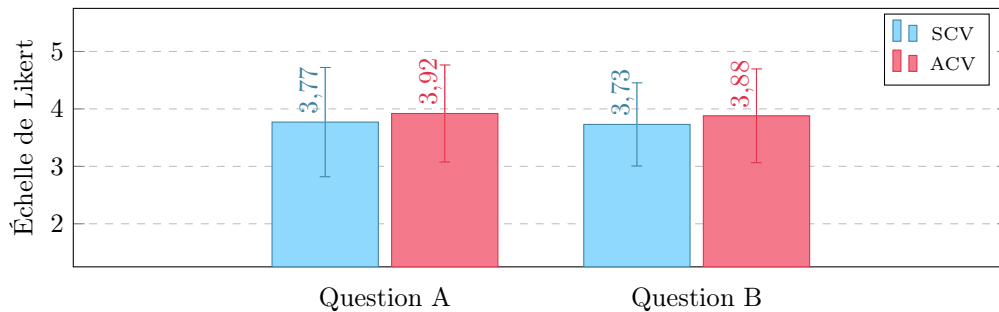


FIGURE 4.19 – Moyennes des évaluations du sentiment de présence virtuelle.

Pour l'évaluation de la **vivacité sensorielle**, deux questions étaient posées :

- a) « *Vous aviez l'impression d'être simplement en train de percevoir des images : » (1 = Pas du tout d'accord & 5 = Tout à fait d'accord) — question 4 dans le questionnaire.*
- b) « *Vous étiez toujours conscient de l'environnement réel durant l'expérience : » (1 = Pas du tout conscient & 5 = Très conscient) — question 6 dans le questionnaire.*

Un U-test de Mann-Whitney n'a révélé aucune différence statistique significative entre les deux contextes vis-à-vis de ces deux questions. La figure 4.20 présente les moyennes des réponses à ces deux questions pour les deux contextes.

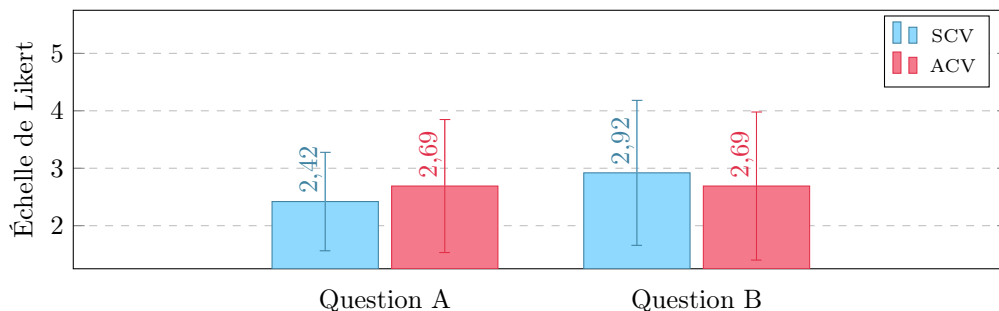


FIGURE 4.20 – Moyennes des évaluations de la vivacité sensorielle.

Vis-à-vis de la variable **degré de contrôle**, une question était posée : « *Vous avez eu la sensation d'agir dans l'environnement virtuel plutôt que d'agir sur un quelconque mécanisme à l'extérieur de celui-ci :* » (1 = J'agissais dans le monde réel & 5 = J'agissais dans le monde virtuel) — question 5 dans le questionnaire. Un U-test de Mann-Whitney a révélé une différence statistique significative entre les contextes SCV (N = 26) et ACV (N = 26) ($U = 235$; $z = -2,009$; $p = ,045$). La figure 4.21 présente les moyennes des réponses à cette question pour les deux contextes et indique que c'est dans le contexte SCV que les participants ont eu le plus le sentiment d'interagir directement avec l'environnement virtuel plutôt qu'avec un élément extérieur à ce dernier.

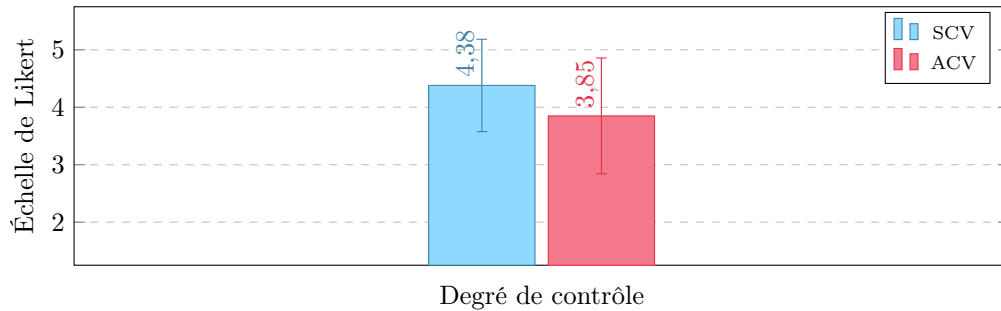


FIGURE 4.21 – Moyennes des évaluations du degré de contrôle.

Qualité de l'Expérience

► **Comportement Social :** Comme nous l'avons expliqué dans la section 4.2.4, les interactions sociales vocales n'ayant eu lieu que pour les binômes immergés dans le contexte ACV, les réponses données dans le questionnaire ne sont pas utilisables en tant que matériel de comparaison. Nous présenterons alors les moyennes des évaluations obtenues à ces questions (sur une échelle à cinq niveaux) afin d'apporter un éclairage sur les comportements sociaux des participants durant l'expérimentation.

- a) Fréquence : « *Vous avez interagi avec votre partenaire (conversation) à la fréquence suivante :* » (1 = Rarement & 5 = De nombreuses fois) — question 21 dans le questionnaire.
M = 4; ÉT = 1,216.
- b) Assistance Mutuelle (Comportement Empathique) : « *Votre partenaire vous a appris des choses durant l'expérience :* » (1 = Rien appris & 5 = Beaucoup appris) — question 22 dans le questionnaire.
M = 2,54; ÉT = 1,318.
- c) Comportement Empathique : « *Vous avez, au travers de vos conversations, ressenti l'état émotionnel de votre partenaire :* » (1 = Pas du tout & 5 = Complètement) — question 23 dans le questionnaire.
M = 2,88; ÉT = 1,296.
- d) Compréhension Mutuelle (Présence Sociale) : « *Vous avez ressenti un manque de compréhension entre vous et votre partenaire :* » (1 = Il était difficile de se comprendre & 5 = Nous nous sommes parfaitement compris) — question 24 dans le questionnaire.
M = 3,75; ÉT = 1,189.

- e) Nature (A) : « *Vos discussions étaient plutôt :* » (1 = Négatives (tension) & 5 = Positives (partage, amusement)) — question 25 dans le questionnaire.
M = 4,17 ; ÉT = ,761.
- f) Attention Allouée (Présence Sociale) : « *Le degré d'intensité de vos conversations était :* » (1 = Très calme & 5 = Totalelement enflammée) — question 26 dans le questionnaire.
M = 2,29 ; ÉT = 1,122.
- g) Nature (B) et Attention Allouée (Présence Sociale) : « *Vous êtes vous mutuellement encouragés durant l'expérience :* » (1 = Pas du tout & 5 = Beaucoup) — question 27 dans le questionnaire.
M = 2,87 ; ÉT = 1,227.
- h) Nature (C) et Attention Allouée (Présence Sociale) : « *Vous êtes vous mutuellement conseillés vis-à-vis de la tâche à accomplir :* » (1 = Pas du tout & 5 = Beaucoup) — question 28 dans le questionnaire.
M = 4,21 ; ÉT = ,884.
- i) Engagement (Présence Virtuelle) : « *À propos de quel sujet avez-vous échangé : la réalisation de la tâche demandée :* » (1 = Pas du tout & 5 = Beaucoup) — question 29 dans le questionnaire.
M = 4,67 ; ÉT = ,564.
- j) Réponse Émotionnelle : « *À propos de quel sujet avez-vous échangé : votre ressenti vis-à-vis de l'expérience en cours :* » (1 = Pas du tout & 5 = Beaucoup) — question 30 dans le questionnaire.
M = 2,67 ; ÉT = 1,373.
- k) Qualité de Service : « *À propos de quel sujet avez-vous échangé : l'environnement virtuel :* » (1 = Pas du tout & 5 = Beaucoup) — question 31 dans le questionnaire.
M = 2,13 ; ÉT = 1,191.
- l) Degré de contrôle : « *À propos de quel sujet avez-vous échangé : les modalités d'interactions (utilisation de la manette et du jeu Minecraft) :* » (1 = Pas du tout & 5 = Beaucoup) — question 32 dans le questionnaire.
M = 2,25 ; ÉT = 1,189.

► **Efficacité** : quel est l'impact de la communication vocale sur l'efficacité perçue et réelle des participants ? Vis-à-vis de l'efficacité, nous possédons deux types de données, celles issues des questionnaires, quantitatives subjectives correspondant à l'efficacité perçue par les participants, et celles, données quantitatives objectives, issues du comptage du nombre de blocs posés et du fait que les participants aient commis des erreurs ou non. Nous exposerons en premier les données provenant des questionnaires.

Pour l'évaluation de l'**efficacité perçue**, deux questions étaient posées :

- a) « *Vous avez été efficace vis-à-vis de l'accomplissement de la tâche :* » (1 = Pas du tout efficace & 5 = Très efficace) — questions 7 dans le questionnaire. Un U-test de Mann-Whitney a révélé

une différence statistique significative entre les contextes SCV ($N = 26$) et ACV ($N = 26$) ($U = 207$; $z = -2,487$; $p = ,013$).

- b) « Vous étiez à votre niveau de performance maximum : » (1 = Au début & 5 = À la fin) — question 8 dans le questionnaire. Un U-test de Mann-Whitney n’a révélé aucune différence statistique significative entre les deux contextes vis-à-vis de cette question.

La figure 4.22 présente les moyennes des réponses à ces deux questions pour les deux contextes, et illustre le fait que les participants ont estimé être plus efficaces dans le contexte ACV. Les moyennes de la question B, bien que ne présentant pas de différence statistique significative, montrent une tendance des participants à se percevoir plus efficaces après un certain temps passé en immersion (notion d’apprentissage, d’expérience).

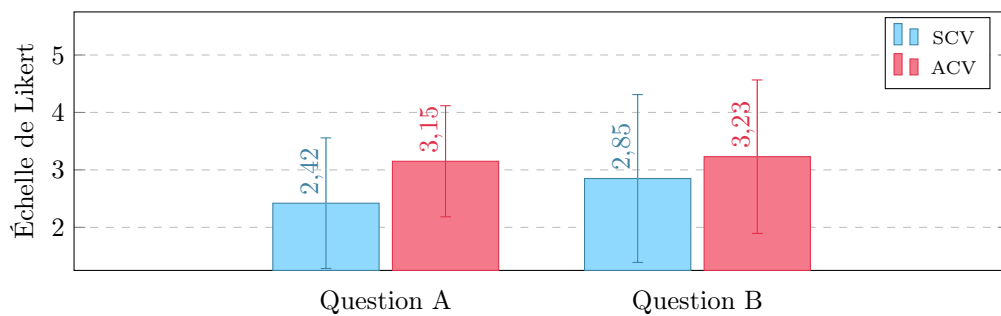


FIGURE 4.22 – Moyennes des évaluations de l’efficacité perçue.

Pour la mesure quantitative objective de l’**efficacité**, les résultats sont basés sur le nombre de blocs posés par les participants ainsi que sur le fait qu’ils aient commis des erreurs ou non. Afin de nous assurer de la robustesse de ces mesures, nous avons demandé aux participants (dans notre questionnaire pré-expérimentation) quelle était leur expérience sur *Minecraft*. Cette information nous a permis de tester l’impact de cette expérience préalable sur l’efficacité perçue et observée. Nous avons aussi examiné la distribution des participants dans nos deux contextes vis-à-vis de leur expérience préalable de *Minecraft*. Malgré le fait que nous n’ayons pas contrôlé la distribution de nos participants dans chaque contexte (les participants ayant eux-mêmes choisi le jour et l’heure de leur passage), nous avons obtenu deux groupes presque totalement homogènes (voir tableau 4.2.)

	Pas d’expérience	Moins de 10h	De 1 à 50h	Plus de 50h	Total
SCV	15	01	03	07	26
ACV	14	02	02	08	26
Total	29	03	05	15	52

TABLEAU 4.2 – Répartition des participants vis-à-vis de leur expérience préalable de *Minecraft*.

Afin d’évaluer les possibles interactions entre nos deux contextes (SCV & ACV), le nombre de blocs placés et l’expérience préalable de *Minecraft*, l’utilisation d’une analyse de la variance (ANOVA) était requise. Nous avons choisi l’analyse de la variance pour sa robustesse aux erreurs de type 1 (Winer *et al.*, 1991). L’analyse de la variance à deux facteurs n’a révélé aucune interaction entre les

deux facteurs cités ci-dessus (expérience et contexte) sur le nombre de blocs posé, nous pouvions alors focaliser notre attention sur les effets principaux entre facteurs, c'est à dire, l'impact de l'expérience préliminaire de *Minecraft* sur le nombre de blocs placé et l'impact de nos deux contextes expérimentaux sur cette même variable. Un test de Kruskal-Wallis n'a révélé aucune différence statistique significative vis-à-vis du nombre de blocs placés en fonction de l'expérience préalable de *Minecraft* (le test de Kruskal-Wallis est un test non paramétrique pour plus de deux échantillons indépendants (Pallant, 2010)). En revanche, un U-test de Mann-Whitney mené sur l'impact de nos deux contextes sur le nombre de blocs placés a révélé une différence statistique significative entre les contextes SCV ($N = 26$) et ACV ($N = 26$) ($U = 188$; $z = -2,410$; $p = ,016$). La figure 4.23 présente la comparaison du nombre de blocs placés entre les contextes SCV & ACV et montre que c'est dans le contexte ACV que les participants ont posé le plus grand nombre de blocs, avec une moyenne de 51,88 blocs posés de plus que dans le contexte SCV (nombre de blocs total de la statue : **360**). Pour finir, vis-à-vis des erreurs, nos observations montrent que dans le contexte SCV, **sept** binômes ont commis des erreurs, alors que, dans le contexte ACV seulement **deux** binômes en ont commis. Différence significative appuyée par un U-test de Mann-Whitney entre les contextes SCV ($N = 26$) et ACV ($N = 26$) ($U = 202$; $z = -2,513$; $p = ,012$).

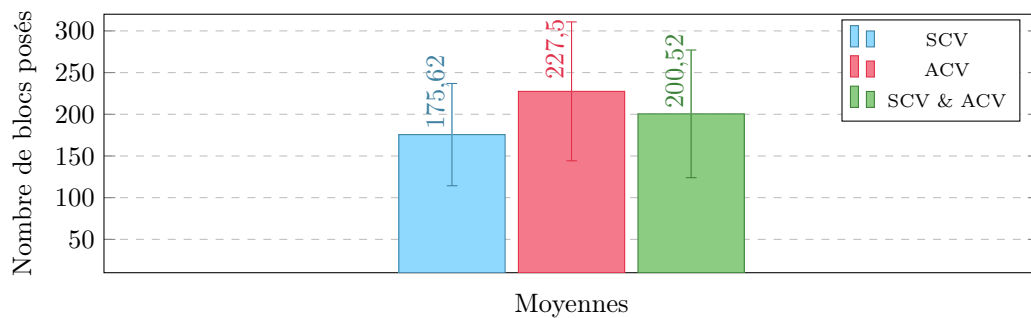


FIGURE 4.23 – Comparaison du nombre de blocs placés entre les contextes SCV & ACV.

Autres questions : vis-à-vis de l'occurrence du stress, une question était posée : « *Vous avez ressenti du stress durant votre expérience :* » (1 = J'étais serein durant toute la durée de l'expérience & 5 = J'ai ressenti beaucoup de stress) — question 13 dans le questionnaire. Un U-test de Mann-Whitney a révélé une différence statistique significative entre les contextes SCV ($N = 26$) et ACV ($N = 26$) ($U = 216$; $z = -2,340$; $p = ,019$). La figure 4.24 présente les moyennes des réponses à cette question pour les deux contextes, et illustre que les participants ont ressenti plus de stress dans le contexte SCV. Nous ne sommes malheureusement pas en mesure de mettre, en parallèle de ces réponses, les observations physiologiques issues des capteurs ECG et GSR que nous avons utilisés à cause des problèmes de fonctionnement rencontrés au cours de l'expérimentation.

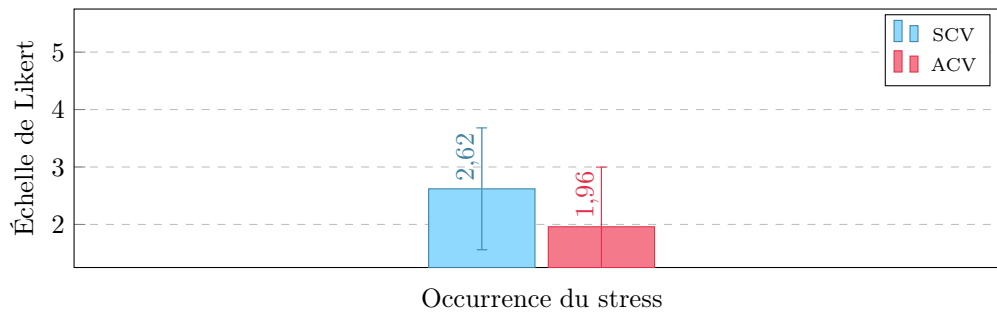


FIGURE 4.24 – Moyennes des évaluations de l'occurrence du stress.

Questionnaire a posteriori : dans le court questionnaire a posteriori (voir annexe B.6) envoyé aux participants (quatre mois après l'expérimentation), une question était posée sur l'**engagement** (présence virtuelle) : « *Quel fut votre ressenti vis-à-vis de cette expérience :* » (1 = Ennuyeuse & 5 = Captivante) — question 2 dans le questionnaire. Un U-test de Mann-Whitney a révélé une différence statistique significative entre les contextes SCV ($N = 18$) et ACV ($N = 12$) ($U = 60,500$; $z = -2,123$; $p = ,043$). Vis-à-vis des trois autres questions portant sur l'**état émotionnel**, et sur l'**intention d'utilisation** (Adoption Technologique), aucune différence statistique significative n'a été relevée. La figure 4.25 présente les moyennes des réponses à ce questionnaire pour les deux contextes. Malgré la différence statistique significative relevée vis-à-vis de la question portant sur l'engagement, ces résultats sont à considérer comme une tendance plutôt que comme de réels résultats. En effet, le nombre de participants ayant répondu à ce questionnaire est faible (SCV : 18 réponses ; ACV : 12 réponses) réduisant drastiquement la robustesse des résultats obtenus.

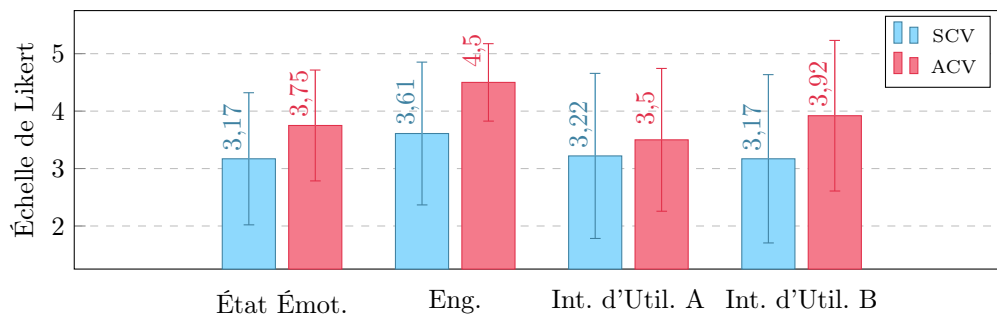


FIGURE 4.25 – Moyennes des réponses obtenues au questionnaire a posteriori.

4.2.8 Analyse et synthèse

Avant de proposer l'analyse de nos résultats, nous allons décrire les limites de ce premier travail d'expérimentation. En premier lieu, nous avons souhaité revenir sur les difficultés que nous avons rencontrées vis-à-vis des matériels utilisés. L'*Oculus Rift*, que nous avons décidé d'utiliser pour ses qualités immersives et l'intérêt qu'il suscite dans la communauté, s'est révélé très dommageable pour nos travaux. En effet, nous avons, lors de la conception de notre expérimentation, sous-estimé sa propension à générer chez nos participants un fort sentiment de malaise, ce qui a entraîné la perte de huit jeux de données. En effet, quatre de nos participants n'ont pas supporté le port du casque

immersif plus de cinq à sept minutes, entraînant ainsi leur abandon et par là même la perte des données collectées ainsi que de celles de leur binôme. Outre ces abandons, nous avons constaté au travers de l'évaluation de l'intrusivité de nos matériels qu'un bon nombre de participants ne s'était pas senti à l'aise avec le port de l'*Oculus Rift*, quinze de nos participants, soit 28,8% de notre échantillon, ont évalué le casque comme générant un fort malaise qui a perturbé leur expérience. Nous nous devons néanmoins de souligner que les conditions d'expérimentation n'étaient pas optimales, car les salles utilisées ne possédaient par de système de climatisation et il y régnait une certaine chaleur (la session d'expérimentation s'est tenue au mois de juin), que nous avons essayé de limiter au maximum (courants d'air, aération nocturne) puisqu'il était impossible, pour des raisons logistiques, de décaler la session. Cela a sans aucun doute favorisé l'apparition de la *cyber sickness*. D'un autre côté, *Minecraft* ou tout du moins son aspect cubique ainsi que la prise de hauteur imposée par notre construction verticale, ne sont sans doute pas étrangers à l'occurrence de ce désagréable sentiment. Pour ces raisons, nous avons fait le choix de ne plus utiliser l'*Oculus Rift* pour la suite de nos travaux mais de nous orienter vers un autre système de visualisation stéréoscopique, les lunettes 3D actives. Certes moins immersives, ces dernières présentent l'avantage de ne pas générer autant de *cyber sickness*. Nous reviendrons sur ce choix et ses implications dans la section concernant les matériels dédiés à notre seconde expérimentation (section 4.3.4).

Une autre limite de cette première expérimentation réside dans nos difficultés à produire les données quantitatives basées sur les mesures physiologiques. Deux problèmes principaux ont engendré cette perte de données. Le premier, une malfunction de l'un des appareils d'enregistrement nous a fait perdre la totalité des données physiologiques (ECG et GSR) d'une journée d'expérimentation pour l'un des postes utilisateurs. Le second problème auquel nous fûmes confrontés s'explique par le fait que nous ne possédions qu'une seule ceinture pectorale pour les deux capteurs ECG (utilisée de plus sur le poste où nous avons rencontré le problème d'enregistrement évoqué) : nous avons donc plusieurs fois constaté que les patchs adhésifs n'avaient pas tenu sur la peau des participants entraînant ainsi des captations incomplètes. Pour ce qui est du casque EEG, c'est une erreur dans le logiciel de configuration qui est à l'origine des problèmes rencontrés : l'indicateur de batterie présent sur le logiciel de captation (les casques n'en possédant pas) indiquait une valeur très optimiste et iréaliste. Nos tests préliminaires n'avaient pas mis en lumière ce problème, et ce n'est qu'après l'enchaînement de plusieurs passations qu'il s'est révélé. D'autre part, nous avons rencontré des difficultés vis-à-vis du positionnement du casque EEG à cause des élastiques de maintien de l'*Oculus Rift*. Nos tests préliminaires avaient permis de constater que, bien que délicate, sa mise en place était faisable. Néanmoins, lors de l'expérimentation, les élastiques de l'*Oculus Rift* combinés aux natures de cheveux très variées de nos participants ont engendré un déplacement fréquent du casque EEG, entraînant une perte de signal ou un enregistrement sur des zones non conformes au schéma à suivre. En conséquent, nous avons ici encore perdu une grande partie de nos données et n'étions donc plus en mesure de fournir un nombre de jeux de données suffisant pour permettre une évaluation correcte de notre échantillon. Nous n'avons pas été en mesure d'évaluer le sentiment de présence virtuelle par la physiologie et donc de les confronter à nos résultats quantitatifs subjectifs issus des questionnaires post expérimentation. Ces

points tendent malheureusement à limiter les conclusions que nous pourrions tirer de cette première expérimentation.

Nous estimons aussi que le nombre de participants est une des limites de notre travail. Ces derniers ayant été recrutés sur la base du volontariat et n'ayant reçu pour gratification que le plaisir de vivre l'expérience proposée, il n'a pas été aisé d'atteindre ce nombre de 52 participants. Nous avons souvent pu voir dans la littérature que pour des études empiriques menées dans le domaine de la réalité virtuelle et présentant un protocole expérimental relativement proche du nôtre (immersion courte de 10 à 45 minutes), le nombre de participants oscillait entre cinquante et soixante-dix participants. Néanmoins, la plupart de ces études ne se focalisent que sur un ou deux facteurs de l'expérience de la réalité virtuelle. Or, dans notre cas, nous avons proposé une évaluation aussi holistique que possible et nous pensons qu'un nombre de participants plus important (cent au minimum) serait très souhaitable pour ce type d'étude. Malheureusement, il reste difficile de réunir des échantillons larges sur une période courte, plus encore quand la seule rétribution des participants est le plaisir de la participation. Comme nous le verrons pour nos expérimentations suivantes, nous n'avons jamais été en mesure de réunir une telle population sur la base du volontariat.

Intéressons-nous à présent à nos résultats et à leurs implications, avant de revenir à nos hypothèses. Vis-à-vis de la présence sociale et en premier lieu de l'attention allouée, nous n'avons pas constaté de différence significative entre nos deux contextes SCV et ACV. Ce résultat est possiblement explicable par le fait que lorsque nous devons travailler de concert avec un partenaire, nous sommes dans l'obligation de lui allouer une certaine attention. Lorsqu'il est possible de communiquer verbalement, cette allocation de l'attention est transparente, automatique, lorsque notre partenaire prend la parole, nous l'écoutons, conscient que l'interaction sociale qu'il produit à notre intention est sans doute en rapport avec ce que nous sommes en train de faire ensemble. Dans le cas où la communication vocale n'est pas possible, nous ressentons tout de même le besoin de savoir ce que fait notre partenaire, s'il ne rencontre pas de difficultés ou s'il n'est pas en train de se tromper et de mettre en péril notre réussite commune, nous incitant ainsi à lui prêter attention régulièrement. La différence entre nos deux contextes vis-à-vis de l'attention allouée résiderait alors plus dans la manière d'être attentif, en écoutant ou en observant, plutôt que dans la « quantité » d'attention. La seconde variable évaluée de la présence sociale était la compréhension mutuelle : de manière assez prévisible, nous avons relevé une différence statistique significative mettant en exergue une bien meilleure compréhension mutuelle dans le contexte ACV. Cela s'explique par la nature même de l'interaction vocale qui est de transmettre un message avec une modalité commune afin que celui-ci soit compris. Ce constat s'applique aussi à la différence révélée en faveur du contexte ACV vis-à-vis de la variable interdépendance comportementale. Il est plus aisé d'influencer le comportement de l'autre lorsque l'on peut lui demander de faire telle ou telle action. Cette modalité est par ailleurs de nature à faire émerger un *leader* prenant dans une certaine mesure les rênes de la tâche à mener. Nous avons néanmoins conçu notre expérimentation afin que les deux participants doivent agir afin de pouvoir terminer la tâche dans le temps imparti, jugulant ainsi une prise de position précoce en tant que *leader*. Malgré l'impossibilité à communiquer vocalement, les participants du contexte SCV ont estimé qu'une certaine interdépendance comportementale les avait

orientés dans leur construction (moyenne de 2,27 sur 5) nous permettant de dire qu’une forme de translucence (voir section 2.1.2.3 ainsi que Erickson et Kellogg (2000)) s’est produite durant l’immersion. Même si ce phénomène ne permet pas un impact aussi direct que la communication vocale, elle a permis aux participants de faire coïncider leurs efforts par l’étude du comportement de l’autre. Il est intéressant ici de noter que, vis-à-vis de ces phénomènes d’interdépendance comportementale et de translucence, aucune différence significative n’est apparue entre nos deux contextes pour l’assistance mutuelle. Évaluée en moyenne à 4,04 sur 5 dans le contexte ACV où il était aisé de demander de l’aide, elle a tout de même obtenu une évaluation moyenne de 3,54 sur 5 dans le contexte SCV. Attestant de la capacité des participants à détecter les difficultés de leur partenaire par une lisibilité et une compréhension aisées de ses actions, puis d’y remédier, c’est un indice de plus de l’occurrence de la translucence. Au regard de ces résultats, il semblerait que, malgré l’impossibilité de communiquer vocalement, les participants au contexte SCV se soient sentis socialement présents. Une question de type « *Vous êtes-vous senti en présence d’une autre personne ?* », aurait pu nous aider à confirmer ces résultats. Nous ne l’avons néanmoins pas posée, estimant que nos participants étaient conscients avant leur immersion d’être mis en présence d’une autre personne réelle : les réponses auraient alors très certainement été positives et n’auraient pas éclairé notre évaluation. Nous pouvons néanmoins penser que dans le contexte ACV, les participants se sont sentis davantage présents socialement, du fait d’une meilleure compréhension mutuelle.

Le *flow* psychologique ou état de *flow* était évalué grâce à deux variables. L’une étant la résultante de l’état de *flow*, l’amusement et l’autre, l’une des conditions permettant par son équilibre avec les compétences de l’utilisateur d’accéder à l’état de *flow*, le niveau de défi perçu. Les participants à cette expérimentation ont, au travers de leur évaluation, montré qu’ils s’étaient plus amusés lorsqu’ils étaient en mesure de communiquer vocalement, ce qui s’explique par le fait que la possibilité d’échanger permet de pallier plus facilement aux potentielles difficultés rencontrées, réduisant ainsi les risques de frustration. Mais aussi par le fait qu’il devient alors possible de faire de l’humour, de partager son ressenti ou d’encourager son partenaire, c’est-à-dire d’avoir une activité « sociale » qui dépasse le cadre strict de la tâche à réaliser. Néanmoins, et même si la différence entre les deux contextes est significative, les participants au contexte SCV ont évalué leur niveau d’amusement à 4 sur 5 en moyenne (contre 4,54 dans le contexte ACV). L’impact de la communication vocale ne semble alors pas si important que nous le pensions de prime abord. Une tâche ou un jeu ennuyeux le restera avec ou sans communication vocale. Il faut toutefois replacer ces résultats dans leur contexte, une immersion courte proposant une tâche précise et explicite. Dans le cas d’une immersion plus longue avec une activité de type exploratoire, il est possible que les participants soient rapidement saisis par l’ennui. La communication vocale permettant alors d’échanger, palliant ainsi dans une certaine mesure à l’ennui, favoriserait sans doute d’une manière plus éloquente le contexte ACV vis-à-vis de l’évaluation de l’amusement. Nous pouvons néanmoins dire que l’interaction sociale vocale peut apporter plus d’amusement grâce aux échanges qu’elle rend possibles. Pour ce qui est du niveau de défi perçu, il est très intéressant de constater que nous n’avons pas trouvé de différence significative entre nos deux contextes. Les évaluations moyennes — 3,35 sur 5 pour SCV et 3,12 pour ACV — montrent que, même

sans la capacité de communiquer, les participants n'ont pas trouvé la tâche beaucoup plus difficile. Étant données les performances des participants (sur lesquelles nous reviendrons), moins bonnes dans le contexte SCV, et de la répartition homogène de l'expérience préalable de *Minecraft* dans nos deux échantillons, il est impossible d'affirmer que ces évaluations du niveau de défi perçu sont dues à une meilleure maîtrise de *Minecraft* par les utilisateurs en SCV et que la communication vocale a facilité la tâche du groupe ayant produit la tâche ACV. Nous pensons plutôt que ce résultat est une indication en faveur du fait que le phénomène de translucence est un très bon palliatif à l'absence de communication verbale.

Vis-à-vis du sentiment de présence virtuelle, nous n'avons identifié aucune différence significative entre nos deux contextes, que ce soit pour l'évaluation directe du sentiment de présence virtuelle ou pour l'évaluation de la vivacité sensorielle (facteur de la QoS fortement impliqué dans le sentiment de présence virtuelle). Nous savons, pourtant grâce à notre étude de la littérature, que la présence sociale impact de manière positive la présence virtuelle, et au regard de nos résultats vis-à-vis de cette dernière, nous pensons qu'en raison d'une meilleure compréhension mutuelle nos participants ACV s'étaient sentis plus socialement présents malgré le fait que nous n'ayons pas relevé de différence statistique significative. Par corrélation, une meilleure présence sociale aurait pu engendrer une augmentation de leur niveau de présence virtuelle. Néanmoins, cela ne semble pas être le cas. Cela nous indique-t-il que la communication vocale n'est pas un facteur déterminant de la présence sociale ? Mais qu'un être virtuel dont le comportement semble humain suffit à générer un sentiment de présence sociale ? Ce constat semble cohérent avec les études de Guadagno *et al.* (2007) et de Garau *et al.* (2005) démontrant l'importance du réalisme comportemental des êtres virtuels. Nous avons en revanche relevé une différence significative entre nos deux contextes vis-à-vis de la dernière variable utilisée pour évaluer le sentiment de présence virtuelle, le degré de contrôle. Les participants immergés dans le contexte SCV ont estimé qu'ils interagissaient plus avec le monde virtuel qu'avec une quelconque interface entre eux et l'environnement (4,38 sur 5 contre 3,85 en ACV). Cela pourrait s'expliquer par une meilleure maîtrise de la manette de jeu par l'échantillon immergé SCV, mais tous nos participants ayant attesté être familier de ce type de périphérique, il nous est difficile de l'affirmer. Nous sommes toutefois en droit de nous poser la question suivante. Est-il possible que le fait de communiquer vocalement crée un lien avec le monde réel, créant une forme de *Break In Presence* augmentant la conscience de l'interface par l'utilisateur ? Dans cette logique nous pouvons aussi penser que le contenu des conversations, et plus particulièrement concernant les contrôles à la manette (la fréquence de l'occurrence de ce sujet a été évaluée à 2,25 sur 5 par les participants ACV), a pu exacerber ce sentiment d'interagir avec un élément extérieur à l'environnement virtuel.

Intéressons-nous à présent aux performances perçues et objectives de nos participants. D'un point de vue subjectif, ce sont les participants ACV qui se sont perçus comme le plus efficaces (moyenne de 3,15 sur 5 contre 2,42 SCV). Cette évaluation est confirmée par nos observations des performances : nous avons mis en évidence une différence moyenne de 51,88 blocs supplémentaires posés ainsi qu'une différence quant au nombre de binômes ayant commis des erreurs dans le contexte ACV (2 contre 7 SCV). Il est donc pertinent de dire que, par leurs interactions sociales vocales, les participants

ont été en mesure d'accroître leurs performances. Ceci par le conseil, les indications, le soutien, ou par la prise de position en tant que *leader*. La question de l'impact de la nature de la tâche se pose alors. Nous étions ici dans le cadre d'une tâche très interdépendante, en collaboration, où les binômes travaillaient de concert sur le même artefact avec le même niveau d'information. Mais dans une situation différente, trouverions-nous les mêmes résultats vis-à-vis de leurs performances ? Et si nous passions d'une tâche collaborative à une tâche coopérative ? C'est-à-dire à une tâche où les deux participants travaillent chacun sur une partie distincte de la construction générale, en ayant chacun les informations propres à leur partie, mais pas celle de leur partenaire, et où la réussite du binôme passe par la réussite personnelle de chaque utilisateur (Howard *et al.*, 2014) ? Dans une telle situation, la nécessité de communiquer vocalement serait moindre, réduisant ainsi la proportion des interactions sociales vocales allouée à la tâche (qui est ici le sujet le plus récurrent, dont l'occurrence a été évaluée à 4,67 sur 5, aussi corrélé aux conseils dont l'occurrence fut évaluée à 4,21 sur 5) et laissant plus de place pour des discussions informelles ou à des encouragements (occurrence évaluée à 2,87 sur 5). Ces discussions moins focalisées sur la tâche en cours auraient-elles le même effet sur les performances de nos participants ou la différence entre les deux contextes (SCV et ACV) serait-elle drastiquement réduite ?

Pour finir, nous souhaiterions évoquer *Minecraft* et l'expérience utilisateur qu'il suscite. Outre les problèmes de *cyber sickness* que nous attribuons plus à l'*Oculus Rift* qu'à l'aspect visuel de *Minecraft*, nous avons pu constater, au travers de l'évaluation de l'état émotionnel des participants, de leur niveau de satisfaction, ainsi que de leur désir de réutilisation, que la configuration de notre expérimentation proposait une expérience amusante, intéressante et satisfaisante. Ce constat nous conforte dans l'utilisation de *Minecraft* comme terrain d'expérimentation en réalité virtuelle. Malgré son faible niveau de réalisme, les participants ont attesté d'un bon niveau de présence virtuelle et n'ont globalement pas émis de remarques indiquant que leur expérience avait été désagréable à cause de l'aspect cubique de l'environnement, ou des modalités d'interaction. Nous continuerons donc dans cette voie pour la suite de nos expérimentations, étant donné la relative facilité avec laquelle il est possible de configurer les environnements virtuels créés avec *Minecraft* pour l'atteinte de nos objectifs.

Revenons à présent sur notre hypothèse :

(H.1) *les interactions sociales vocales ont un impact bénéfique sur l'expérience utilisateur en 3D-IVE.*

Que pouvons-nous dire après ce premier travail d'expérimentation ? Nous avons révélé certains impacts que nous pouvons imputer directement au fait de pouvoir communiquer vocalement, en particulier sur l'amusement et les performances des participants, car nous avons constaté une évaluation supérieure tant subjective qu'objective de ces deux facteurs dans le contexte ACV. Deux constatations qui vont dans le sens de la validation de notre hypothèse. D'un autre côté, nous pensions découvrir le même type d'impact sur le sentiment de présence virtuelle, ce qui ne fut pas le cas. Ces constats nous mettent face au fait que l'expérience utilisateur, en raison de sa nature variable et très dépendante du contexte, reste un phénomène difficile et complexe à évaluer et plus encore à prédire. Néanmoins, après

cette première expérimentation, c'est vers la dépendance à la nature de la tâche que nous avons souhaité orienter notre travail suivant. En effet, il nous est apparu comme pertinent, dans le but de mieux cerner l'impact de l'interaction sociale vocale sur l'expérience utilisateur, de changer de situation afin de confronter nos résultats et ainsi de voir s'il est possible ou non de dégager des éléments récurrents. Grâce à l'expérience acquise, nous serons lors de la conception de cette seconde expérimentation, en mesure de lever certaines limitations rencontrées précédemment.

4.3 Seconde expérimentation

4.3.1 Objectif et activité proposée

La question soulevée dans la section analyse et synthèse de la première expérimentation, à savoir si les résultats obtenus précédemment seraient similaires pour une nature de tâche différente, a motivé la réalisation de cette deuxième expérimentation. C'est pourquoi, tout en conservant pour environnement virtuel *Minecraft*, nous avons décidé de passer d'une tâche collaborative où les deux participants travaillaient conjointement sur une même tâche, à une tâche coopérative où les deux participants travailleraient chacun sur une partie différente d'une même construction mais impliquant la réussite de chacun pour accomplir la tâche globale (Howard *et al.*, 2014). Les participants avaient alors chacun une moitié du plan de la construction (voir figure 4.26), cette division a été faite de manière à ce qu'ils puissent en fin de construction extrapoler certains détails de la partie de l'autre et ainsi lui venir en aide dans une certaine mesure. Pour les mêmes raisons que précédemment, nous avons utilisé un personnage de style *chibi pixel art* (voir figure 4.26), ici encore issu du jeu vidéo « *Tiny Tower* ». Autre différence majeure avec notre première expérimentation, la construction ne s'effectuait plus verticalement, mais horizontalement, ceci pour deux raisons. Les participants à notre première expérimentation nous ont rapporté au travers des questionnaires (partie commentaires libres) leurs difficultés vis-à-vis de l'utilisation des échelles qu'il était impératif de placer afin d'atteindre les hauteurs de la construction. D'autre part, nous avons constaté par l'observation et l'écoute des participants durant la première expérimentation que nombre d'entre eux étaient victimes d'une sensation de vertige une fois en hauteur, et soit ils s'en trouvaient perturbés et mal à l'aise, soit ils jouaient à se faire peur avec cette sensation de vide. Ce passage de la verticale à l'horizontale explique aussi l'ajout du motif autour du personnage : lors de notre première expérimentation, la statue avait deux blocs d'épaisseur amenant le nombre de blocs à un total de 360. Ici, la construction du personnage sur une unique épaisseur n'aurait représenté que 195 blocs, c'est pourquoi nous avons ajouté le motif de fond afin d'atteindre 496 blocs. Nombre qui compte tenu de la réduction de la difficulté entraînée par le passage à l'horizontale, du fait qu'ici les participants disposaient de cinq minutes de plus que dans notre première expérimentation (30 minutes contre 25 minutes) et en accord avec les tests préliminaires menés, permet d'atteindre la complémentation de la tâche dans le temps imparti. Afin d'accentuer la coopération, de faire en sorte que la tâche globale ne soit terminée que par la réussite des deux parties nous avons aussi introduit une différence dans les matériaux disponibles pour chacun des participants.

Outre les matériaux communs, les participants possédaient également deux matériaux spécifiques : on peut voir sur la figure 4.26 que le participant n°1 possédait les matériaux rouges et orange, et que le participant n°2 possédait les matériaux noirs et jaunes. Enfin, dans cette deuxième expérimentation, nous avons aussi exclu la seconde partie consistant à défendre le *nexus* de l'invasion des monstres, car nous souhaitions utiliser, comme nous l'expliquerons par la suite, les capteurs ECG et GSR à d'autres fins que la mesure de la présence par la physiologie.

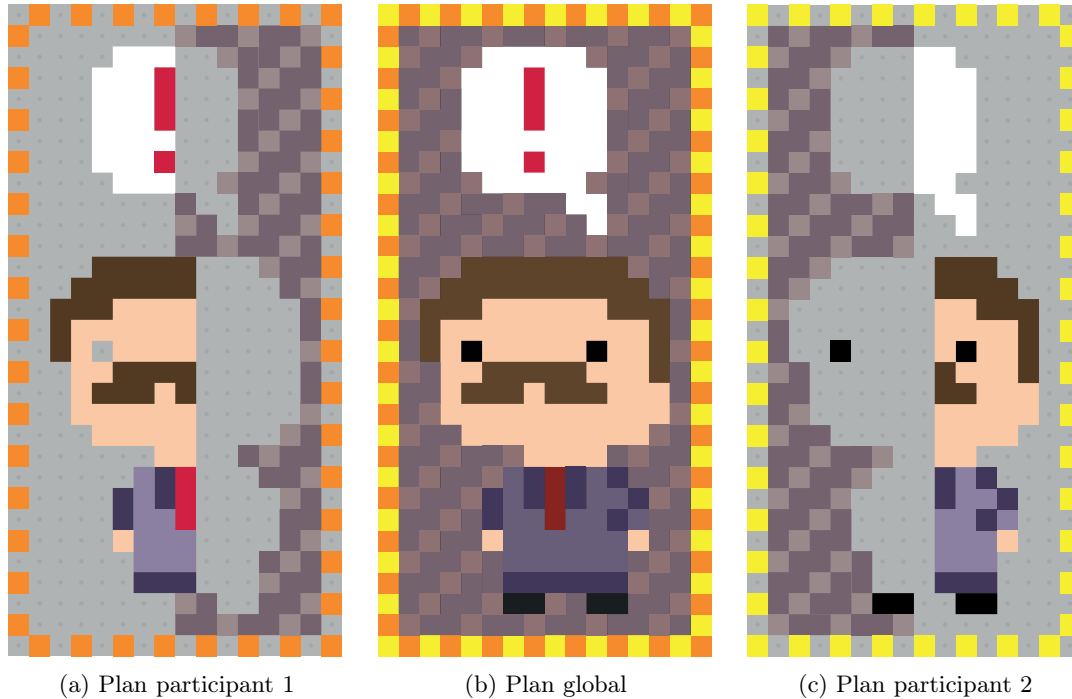
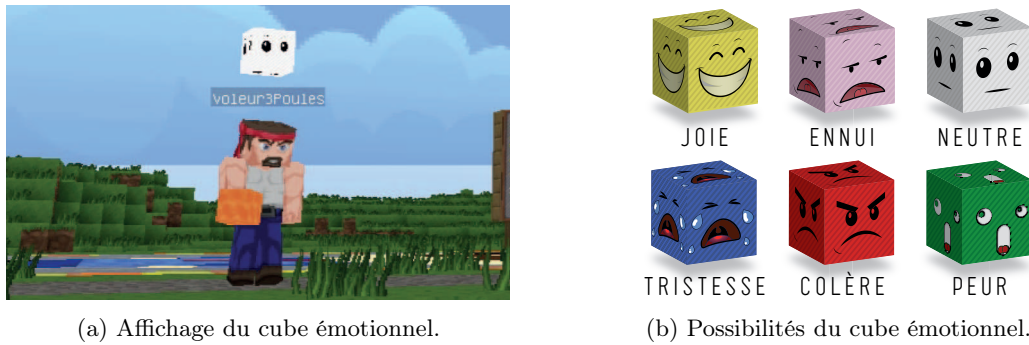


FIGURE 4.26 – Modèle du personnage à construire.

Lors de notre étude de la littérature, nous avons pu constater l'importance de la communication non verbale en tant que vecteur d'émotions, mais aussi le pouvoir de ces dernières quant à l'expérience utilisateur (augmentation de l'engagement, du sentiment de présence virtuelle). C'est pourquoi, dans cette expérimentation, nous avons voulu proposer l'ajout d'une fonctionnalité permettant l'affichage des émotions des participants. Cette fonctionnalité s'illustre par un cube (au format habituel dans *Minecraft*, $1m^3$) disposé au-dessus de l'avatar du participant (figure 4.27a) et permettant une variation sur six émotions ou état d'esprit sous la forme d'*emoticons* colorées (figure 4.27b). Le choix des couleurs pour ces différents états provient de la roue des émotions de Plutchik (1980), très largement reconnue par la communauté scientifique. Pour ce qui est des *emoticons*, elles ont été créées à partir de l'observation des visages humains corrélés à ces émotions ou états d'esprit. On constate que ces derniers sont, hormis pour la joie et la neutralité, négatifs : ce choix découle du fonctionnement du cube. En effet, afin d'afficher les émotions des participants, nous nous sommes appuyés sur la physiologie au travers du rythme cardiaque et de la conductance de la peau. En nous basant sur les travaux de Rainville *et al.* (2006) et de Nakasone *et al.* (2005) vis-à-vis des schémas physiologiques des émotions, nous avons défini des intervalles croisant rythmes cardiaques et conductance de la peau permettant la variation de la texture du cube et ainsi l'affichage de l'émotion ou de l'état d'esprit des participants

en temps réel. Nous avons éprouvé de grandes difficultés à trouver dans la littérature des données claires et factuelles vis-à-vis des variations des schémas physiologiques au regard des émotions. Les schémas que nous avons définis sont donc une réappropriation des travaux de Rainville *et al.* (2006) et de Nakasone *et al.* (2005), que nous avons fait coïncider avec l'état neutre des participants. Cet état neutre était recueilli par les capteurs ECG et GSR durant une minute en début d'expérimentation. Sur la base de ces moyennes physiologiques, nous mesurons pour nos deux capteurs les variations du rythme cardiaque et de la conductance de la peau, selon les intervalles définis, et nous affichons l'une des six émotions déterminées (voir tableau 4.3). Afin d'intégrer cette nouvelle fonctionnalité dans *Minecraft*, nous avons développé un *mod*, le *auraMod*¹⁸ (voir annexe C.1).

FIGURE 4.27 – *Mod* cube émotionnel.

Émotions	RC ¹ min (BPM)	RC max (BPM)	NCP ² min (μs)	NCP max (μs)
Ennui	$M < M + (-2)$	$M + (-2)$	$M + 0,1$	$M + 1,4$
Neutre	$M + (-1,99)$	$M + 2$	M	M
Joie	$M + 2,1$	$M + 5,63$	$M + 3,1$	$M + 3,3$
Tristesse	$M + 5,64$	$M + 7,13$	$M + 1,5$	$M + 1,9$
Colère	$M + 7,14$	$M + 10,22$	$M + 3,4$	$M + 5$
Peur	$M + 10,23$	$M + 20,5$	$M + 2$	$M + 3$

¹RC : rythme cardiaque en battement par minute.

²NCP : niveau de conductance de la peau en micro siemens.

TABLEAU 4.3 – Intervalles de mesure pour les capteurs ECG et GSR pour la définition des émotions.

Afin de déceler le potentiel impact de notre cube émotionnel (EmoCube) sur l'expérience utilisateur, ce n'est plus une expérimentation avec deux contextes, mais avec quatre que nous devons conduire.

1. Sans Communication Vocale et Sans Cube émotionnel : abrégé **SCV-SC** ;
2. Sans Communication Vocale et Avec Cube émotionnel : abrégé **SCV-AC** ;
3. Avec Communication Vocale et Sans Cube émotionnel : abrégé **ACV-SC** ;
4. Avec Communication Vocale et Avec Cube émotionnel : abrégé **ACV-AC** ;

18. *auraMod* : conception par Rémy Eynard, développement réalisé par Benjamin Poussard, ingénieur de recherche pour l'Institut Carnot - ARTS

4.3.2 Environnement Virtuel

Nous avons, lors de cette seconde expérimentation, pu utiliser une version de *Minecraft* plus récente que précédemment puisque, comme nous le verrons, nous avons préféré à l'*Oculus Rift* une paire de lunettes 3D actives *Nvidia*. La version ici utilisée est la 1.7.2. Comme précédemment, les informations concernant la configuration de l'environnement virtuel sont disponibles en annexes C.1.

De la même manière que pour notre première expérimentation, nous retrouvons la séparation de l'environnement en deux parties, l'une réservée aux participants, l'autre à l'expérimentateur. D'un point de vue général, nous avons conservé le même pack de texture (*Sphax PureBDcraft 64x64*), compte tenu des retours très positifs des participants connaissant *Minecraft* sur ce choix. Les fonctionnalités de blocage du temps et des intempéries ont aussi été conservées pour les raisons évoquées en section 4.2.2 mais l'activation est passée d'une fois par seconde à une fois toutes les dix secondes. Ce changement n'a aucun effet visible sur l'application mais permet une économie substantielle des ressources matérielles du serveur.

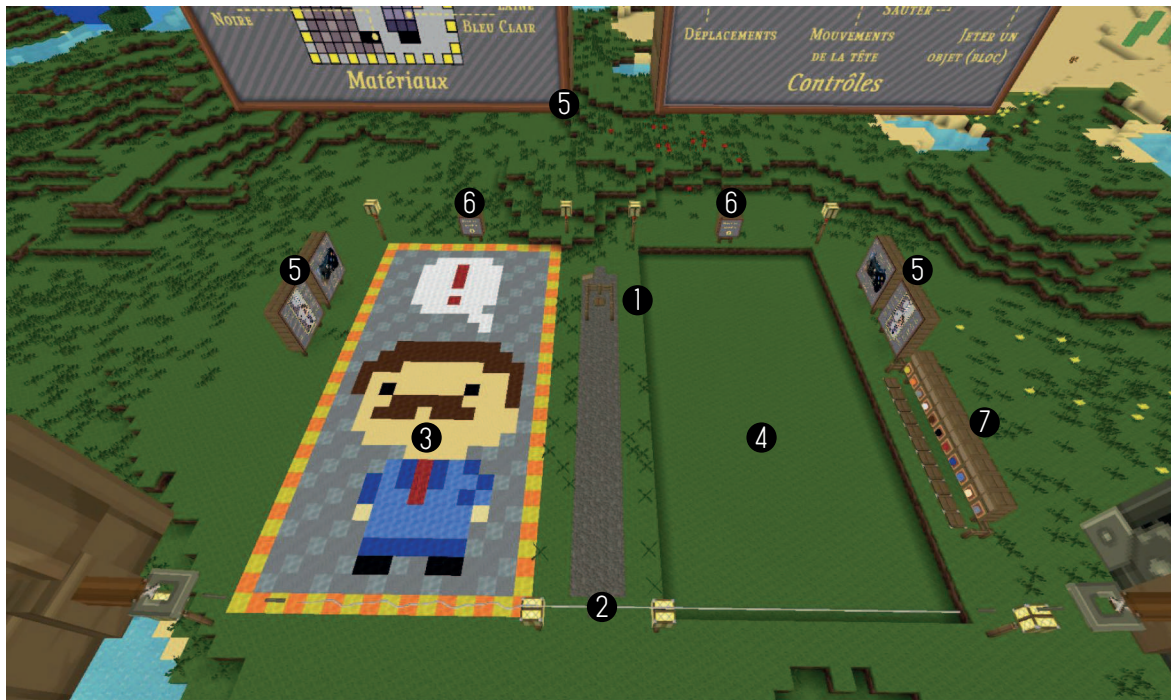
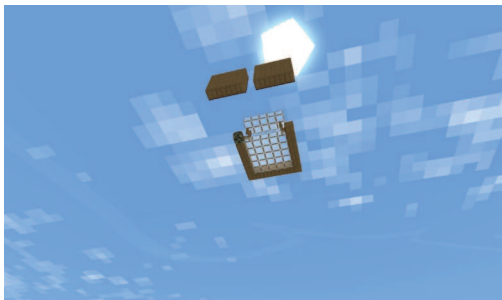


FIGURE 4.28 – Vision de l'environnement *Minecraft* depuis la plateforme d'observation.

Environnement virtuel → Partie participants : la partie participants de l'environnement virtuel s'est en revanche vue largement remodelée d'après les observations faites durant notre première expérimentation, afin d'améliorer la satisfaction quant à la participation à cette expérience, mais aussi de pallier les problèmes d'échelles et de vertige précédemment évoqués. Les chiffres suivants font référence à la figure 4.28 qui présente les éléments de l'environnement virtuel. La prise de vue de la figure 4.28 s'est faite depuis la plateforme d'observation, un nouvel élément créé en réponse au passage à une construction horizontale. En effet, étant donné la taille de la construction, 16 x 31 blocs, soit seize mètres par trente et un mètres, il était impératif de permettre aux participants de prendre du recul afin d'avoir une vue d'ensemble de leur travail. Afin d'accéder à cette plateforme

très élevée (voir figure 4.29a), une plaque de pression activant un *command block* (1) permettait de téléporter directement les participants sur cette dernière. Pour éviter une chute qui serait fatale, la seule échappatoire est une ouverture sur l'avant de la plateforme qui, au passage de l'avatar, active un fil tendu (2) lui-même raccordé à un *command block* téléportant le participant au niveau du sol. De cette plateforme, les participants avaient une vue d'ensemble du modèle taille réelle (3) présent pour les aider, ainsi que de leur zone de construction (4). On retrouve la présence des panneaux d'informations (5) (voir détails en figure 4.29b) dont la particularité vis-à-vis du plan est d'être différents selon le participant (une moitié chacun) ; on les retrouve aussi au niveau de la plateforme afin que les participants puissent faire le lien entre le plan et la vue d'ensemble s'étalant sous leurs pieds. Un panneau indiquant le haut de la construction (6) est aussi positionné, sa présence n'a néanmoins pas empêché deux binômes de commencer leur construction dans le mauvais sens. Les deux derniers éléments notables de cet environnement (7) sont un distributeur (voir détails en figure 4.29c) permettant aux participants de remplir leur inventaire lorsqu'un type de matériau venait à manquer, et un écran (voir 4.29d) permettant l'affichage d'une vidéo informative en début d'expérimentation.



(a) Plateforme d'observation, vue du sol.



(b) Panneaux d'informations.



(c) Distributeur de matériaux.



(d) Écran de présentation de l'expérimentation.

FIGURE 4.29 – Éléments de l'environnement *Minecraft*.

Environnement virtuel → Partie expérimentateur : nous ne reviendrons pas sur la partie expérimentateur de cet environnement, ce dernier fonctionnant sur les mêmes principes que celui de la première expérimentation, le lancement de tous les mécanismes nécessaires grâce à un unique bouton. Certaines modifications ont néanmoins été apportées permettant le fonctionnement de l'écran vidéo et une modification des compteurs de temps afin de correspondre à ces nouveaux environnement et tâche de construction.

4.3.3 Participants

Les participants ($N = 52$; 8 femmes, 44 hommes) étaient âgés de 17 à 44 ans ($M = 22,42$; $ÉT = 5,641$). Ici encore, la population était homogène (la valeur maximale de 44 ans étant marginale), excluant les biais liés à l'âge (Siriara et Siang Ang, 2012). Les participants étaient issus des diverses universités lavalloise (étudiants et personnels), nous les avons contactés par courriel pour leur proposer de participer à nos travaux. Afin de connaître certaines données sur nos participants telles que leur familiarité avec le jeu vidéo en général et plus précisément avec *Minecraft*, les participants ont reçu, suite à leur inscription, un questionnaire pré expérimentation par courriel. Tous les participants parlaient français couramment, nous assurant une bonne compréhension des explications données par les expérimentateurs. Seulement trois participants ont dit ne pas du tout jouer aux jeux vidéo, le reste des participants atteste jouer aux jeux vidéo à diverses fréquences (41% des participants jouent moins de 5 heures par semaine, 24,5% jouent de 5 à 15 heures par semaine et 34,5% jouent plus de 15 heures par semaine). Il leur a aussi été demandé leur familiarité avec les jeux vidéo multijoueurs en ligne : 21,2% des participants n'avaient aucune expérience de ce type. Les autres (78,8%) attestent jouer en ligne à diverses fréquences (41% des participants moins de 5 heures par semaine, 37% de 5 à 15 heures par semaine et 22% plus de 15 heures par semaine). Nous pouvons donc dire que nous étions en présence d'une population de joueurs plutôt expérimentés en solo mais aussi en ligne. Afin de connaître leur aisance avec les modalités d'interaction avec l'environnement virtuel grâce à la manette, nous avons demandé aux participants sur quel type de plateforme ils jouaient : la plupart d'entre eux jouent sur ordinateurs (88,5%), mais une proportion relativement importante joue aussi sur consoles de salon (53,8%). Tous les participants ont attesté savoir comment utiliser une manette de jeu. Vis-à-vis de leur connaissance préalable de *Minecraft*, 88,5% des participants connaissaient le jeu sans pour autant en posséder une grande expérience (pratique). Dans ce pourcentage, 72% avaient une expérience plus ou moins grande de *Minecraft* en multijoueurs (12% moins de 10 heures, 27% de 10 à 50 heures et 61% plus de 50 heures.)

4.3.4 Mesures

Pour cette seconde expérimentation, la gamme de mesures mise en place était constituée de mesures quantitatives subjectives, recueillies grâce à un questionnaire post expérimentation, évolution de celui utilisé lors de notre première expérimentation. De mesures quantitatives objectives, par l'observation des performances à l'égard de la tâche à réaliser, ainsi que par l'enregistrement de l'activité cérébrale des participants par casque EEG. Enfin, postérieurement aux sessions d'immersion, d'une évaluation qualitative grâce à la conduite de groupe de discussion. Nous verrons dans la section *analyse et synthèse* (section 4.3.8) que nous avons là encore fait face à certaines déconvenues vis-à-vis des moyens de mesures mis en place.

Nous allons à présent, toujours en nous appuyant sur la taxonomie de facteurs du modèle (Wu *et al.*, 2009) étendu par Pallot *et al.* (2013), exposer les questions contenues dans notre questionnaire ainsi

que les autres mesures mises en place (EEG, mesures quantitatives objectives de la performance). Nous serons plus succincts, étant donné que la plupart des questions sont identiques entre notre première et notre seconde expérimentation et donc proviennent des mêmes sources déjà citées. Notre questionnaire reste basé sur des questions à échelles de Likert et à échelles sémantiques bipolaires à cinq niveaux, afin de faciliter de compréhension et de réappropriation du questionnaire par les participants (détails en section 4.2.4).

Qualité de Service ▶ Vivacité Sensorielle

▶ **Intrusivité** : afin d'évaluer l'intrusivité des matériels équipant les participants et ainsi de savoir si l'expérience vécue était influencée négativement par les matériels utilisés, quatre questions étaient posées. Voir questions 12.1 à 12.4 dans notre questionnaire en annexe C.3.

Expérience (Utilisateur) Immersive ▶ Processus Expérientiel

▶ **Présence Sociale** : pour mesurer l'impact de nos deux contextes et du cube émotionnel sur le sentiment de présence sociale, nous avons inclus dans notre questionnaire six questions extraites du « *Networked Minds social presence questionnaire* » de Biocca *et al.* (2001). Ces questions évaluent la présence sociale selon quatre facteurs : l'attention allouée, la compréhension mutuelle, l'interdépendance comportementale et l'assistance mutuelle. Voir questions 9 et 13 à 17 dans notre questionnaire en annexe C.3.

▶ **Émotions Sociales ▶ Empathie** : dans le but d'évaluer l'apport de notre fonctionnalité de cube émotionnel, une question était posée au participant sur sa perception de l'état émotionnel de son partenaire. Voir question 18 dans notre questionnaire en annexe C.3.

▶ **Réponse Émotionnelle** : afin de saisir l'état émotionnel général des participants durant l'expérimentation, la première question de notre questionnaire leur demandait de se positionner sur une échelle à cinq niveaux illustrée par des schémas des émotions (proche des *emoticons*), allant de la frustration à la joie, voir question 1 dans notre questionnaire en annexe C.3. Afin de confirmer le bon fonctionnement du cube émotionnel, une série de quatre questions tendaient à évaluer les états émotionnels et leur prédominance durant l'expérimentation. Ainsi, les participants étaient invités à se positionner sur une échelle à cinq niveaux présentant des émotions ou états émotionnels contraires négatifs / positifs, par exemple 1 = Ennui et 5 = Absorption totale. Puis une échelle à cinq niveaux a évalué la prédominance de l'état décrit, selon les pôles sémantiques suivants : 1 = Momentané (moins de 5 min) et 5 = Permanent.

Expérience (Utilisateur) Immersive ▶ Processus Rationnel

▶ **Flow Psychologique** : dans le but de mesurer l'impact de nos deux contextes et du cube émotionnel sur le *flow* psychologique des participants, nous avons inclus dans notre questionnaire deux questions rédigées par nos soins. L'une requérant une évaluation de l'amusement, l'autre du niveau de défi éprouvé (voir questions 2 et 6 dans notre questionnaire en annexe C.3). Une question supplémentaire concernant la concentration était aussi présente mais nous la traitons à part car, comme décrite lors de notre revue de littérature, la concentration de l'utilisateur est nécessaire à l'apparition de *flow* mais est néanmoins considérée comme une variable de la partie immersion psychologique

du sentiment de présence virtuelle (voir Lombard et Ditton (1997) ; Biocca et Delaney (1995) dans la section 2.2.2.1 et Csikszentmihalyi (1997) dans la section 2.3.2.2). Dans le contexte présent, la question étant orientée sur la tâche plus que sur l’environnement ou l’expérience générale, il est plus cohérent de traiter les réponses engendrées en les liant au *flow* psychologique (voir question 3 dans notre questionnaire en annexe C.3).

► **Téléprésence** : pour mesurer l’impact de nos deux contextes ainsi que du cube émotionnel sur le sentiment de présence virtuelle, nous avons inclus dans notre questionnaire cinq questions, dont deux issues du « *Igroup Presence Questionnaire* », de Schubert *et al.* (2001) (voir question une et deux du IPQ¹¹). Ces questions évaluent le sentiment de présence virtuelle lui-même, une variable de l’immersion perceptuelle qui dans notre cas fait partie de la QoS, le degré de contrôle ainsi qu’une variable de l’immersion psychologique, la concentration (Lombard et Ditton (1997) et Biocca et Delaney (1995)). Voir questions 4, 5, 7, 8 et 21 dans notre questionnaire en annexe B.5. Comme précédemment nous avons mis en place une captation par casque EEG pour comparaison avec les travaux de Clemente *et al.* (2013). Cela en veillant bien sûr à ne pas reproduire les problèmes d’utilisation rencontrés précédemment.

► **Acceptabilité Technologique** ► **Utilité** : afin d’évaluer l’utilité perçue du cube émotionnel, une question était posée. Cette dernière requerrait du participant une évaluation de l’utilité perçue du cube vis-à-vis de la compréhension de l’état émotionnel de son partenaire (voir question 19 en annexe C.3). Évidemment, cette question n’était présente que dans les questionnaires des participants ayant vécu l’expérience en présence du cube émotionnel. Tout comme les questions portant sur la communication vocale, cette question ne sera pas utilisée comme matériel de comparaison.

Qualité de l’Expérience

► **Comportement Social** : ici encore, des questions réservées aux participants ayant vécu l’expérience dans le contexte ACV et permettant de recueillir des informations sur la nature des échanges vocaux étaient posées. Dans le contexte de cette expérimentation et comme ultérieurement, nous ne pouvons évidemment pas établir de comparaison entre nos deux contextes SCV et ACV, en revanche nous étions en mesure de tester les différences potentielles entre contextes ACV-SC et ACV-AC. Ces dernières sont identiques à celles de notre précédent questionnaire, nous rappelons néanmoins les facteurs évalués :

- Fréquence des interactions sociales vocales (voir question 23 en annexe B.5) ;
- nature des interactions sociales vocales (voir questions 27, 29, et 30) ;
- variable attention allouée (voir question 26, 27 et 28) de la co-présence, facteur de la présence sociale (Biocca *et al.*, 2001) ;
- comportement empathique (voir question 25) de la QoE ou chez Biocca *et al.* (2001), de l’engagement psychologique, facteur de la présence sociale ;
- compréhension mutuelle (voir question 27), de l’engagement psychologique, facteur de la présence sociale (Biocca *et al.*, 2001) ;
- variable assistance mutuelle (voir question 24) de l’engagement comportemental, facteur de la

- présence sociale (Biocca *et al.*, 2001) ;
- réponse émotionnelle (voir question 32) de l’expérience utilisateur immersive (Pallot *et al.*, 2013) ;
- variable engagement (voir question 31) de la présence virtuelle (Witmer et Singer, 1998) ;
- qualité de service (voir question 33)(Wu *et al.*, 2009) ;
- variable degré de contrôle (voir question 34) de la présence virtuelle (Steuer, 1992).

Des espaces vierges étaient aussi disponibles afin que les participants puissent faire part d’autres sujets de conversation qui se seraient tenus durant la session d’immersion.

► **Efficacité** : afin de mesurer l’impact de nos deux contextes et du cube émotionnel sur l’efficacité perçue par les participants, nous avons inclus dans notre questionnaire deux questions rédigées par nos soins (voir questions 11 et 12 dans notre questionnaire en annexe B.5). Comme précédemment, nous avons aussi évalué l’efficacité de nos binômes de manière quantitative objective. Mais lors de cette seconde expérimentation, notre analyse fut plus fine. En effet, plutôt que de nous baser sur l’environnement virtuel et donc sur la version terminée de la tâche de construction, nous avons ici analysé les vidéos capturées de l’expérience vécue par chaque participant. Cette méthode, bien que beaucoup plus fastidieuse que la précédente, nous a permis d’obtenir, outre le nombre total de blocs placés, le nombre de destructions de blocs, autrement dit de blocs mal placés puis remplacés. De ces valeurs est extrait le nombre de blocs correctement placés (nombre de blocs total - nombre de blocs mal placés) qui, mis en relation avec le nombre de blocs mal placés, nous permet d’obtenir un rapport entre le nombre d’erreurs de placement et le nombre de blocs correctement placés, autrement dit, l’efficacité du participant ou du binôme (nombre de blocs mal placés ÷ nombre de blocs bien placés). Nous avons aussi observé le temps écoulé avant la pose du premier bloc, le temps total avant complétion de la tâche (certains participants ayant terminé avant les trente minutes allouées) ainsi que le nombre d’erreurs à la fin de l’expérimentation, par comparaison avec le modèle.

► **Adoption Technologique** ► **Intention d’Utilisation** : vis-à-vis de l’intention d’utilisation, deux questions héritées du questionnaire a posteriori de la première expérimentation (voir annexe B.6) étaient présentes, voir questions 36 et 37 en annexe C.3)

Autres questions : comme précédemment, une page vierge en fin de questionnaire était destinée à toutes les remarques dont les participants auraient voulu nous faire part.

4.3.5 Équipement des participants

Les équipements utilisés dans le cadre de cette seconde expérimentation sont très similaires à ceux de la première, nous passerons donc rapidement sur cette section dans laquelle nous nous focaliserons sur les différences entre cette expérimentation et la précédente.

Comme nous avons pu le constater lors de notre première expérimentation, l’*Oculus Rift DK1*, malgré ses fonctionnalités immersives (orientation de la tête, isolation du monde réel), a causé le trouble chez un grand nombre de nos participants. Et malgré la sortie de la version *DK2* améliorée

de ce casque, nous n'avons pas souhaité reproduire l'expérience et avons préféré à l'*Oculus Rift* des lunettes 3D actives de la marque *Nvidia*¹⁹. Après plusieurs tests utilisateurs, nous avons eu de bons retours vis-à-vis du niveau de QoS offert par ce matériel et surtout du fait que ce dernier ne causait aucun ou très peu d'inconfort même après trente minutes d'immersion. Afin de faire fonctionner les lunettes 3D avec *Minecraft*, nous avons eu recours à un « *wrapper* », fichier **.dll* permettant de forcer *Java* à rendre deux images. Cet outil, combiné à un *shader* spécifiquement créé, permet d'adapter parfaitement le logiciel à l'utilisation des lunettes *Nvidia* (voir ressources en annexe C.1). Dégagés de la contrainte de l'*Oculus Rift*, nous avons pu utiliser la version 1.7.2 de *Minecraft* nous permettant ainsi l'utilisation du *mod WebDisplay* que nous détaillerons dans la section 4.3.6.

De même que précédemment, vis-à-vis du son et de la communication vocale, nous avons utilisé les mêmes matériels et logiciels. C'est-à-dire des casques audio comprenant un microphone ainsi que le logiciel *TeamSpeak*.

Nous l'avons expliqué préalablement, c'est l'utilisation de l'*Oculus Rift* qui avait conditionné notre choix de la manette comme périphérique de contrôle. Nous aurions donc pu, pour cette seconde expérimentation, faire le choix de repasser sur un mode de contrôle clavier / souris. Deux raisons nous ont fait conserver la manette : la première, nous n'avons eu aucun retour négatif vis-à-vis de ce périphérique, excepté de la part de joueurs de *Minecraft* expérimentés, ces derniers étant de fervents défenseurs du clavier et de la souris. La seconde raison découle de ces remarques : le fait que ces participants ne retrouvent pas leurs marques habituelles tend à gommer quelque peu la différence entre joueurs expérimentés et débutants. Nous avons donc conservé nos manettes de jeu filaires (*Microsoft Xbox 360* pour *Windows*) ainsi que le logiciel *JoyToKey* pour les rendre compatibles avec *Minecraft*.

Comme précédemment, les participants étaient équipés de casques EEG (*Emotiv EEG Headset*). Encore une fois, le passage à l'utilisation des lunettes 3D nous a largement facilité la tâche et nous n'avons pas ici rencontré les problèmes de positionnement engendrés par les élastiques de l'*Oculus Rift*. D'autre part, les participants étaient équipés d'un capteur ECG et d'un capteur GSR. Comme nous le verrons par la suite, afin d'évaluer son impact sur l'expérience vécue, le cube émotionnel n'était activé que pour la moitié des participants. Nous avons néanmoins équipé la totalité de ces derniers avec les capteurs ECG et GSR afin de conserver une QoS constante, la présence de ces capteurs pouvant générer des perturbations, une gêne (intrusivité) chez les participants, même si les observations effectuées lors de notre première expérimentation ont attesté que l'intrusivité des capteurs était très mineure.

4.3.6 Procédure

Lors de l'accueil des participants, nous nous sommes en premier lieu informés sur la familiarité entre individus. Assurés de son absence, nous avons immédiatement séparé les binômes dans deux salles différentes — ou nous avons préalablement installé nos postes d'expérimentation — afin qu'ils ne puissent pas communiquer a priori de la session d'immersion. Aucune information n'a été donnée

19. Voir : <http://www.nvidia.fr/object/3d-vision-technology-fr.html>

sur les réels objectifs de l'expérimentation afin d'éviter toute influence potentielle sur leurs réponses aux questionnaires ou sur leur état d'esprit. Le questionnaire pré expérimentation ayant été complété avant leur venue (sauf cas marginal), nous pouvions directement passer aux informations à donner aux participants. Lors de cette seconde expérimentation, les informations données par les expérimentateurs étaient très succinctes, car nous avons fait le choix de donner les instructions au moyen d'une vidéo informative en début d'expérimentation. Ces explications consistaient alors simplement en une explication des contrôles permettant de bouger le corps et le regard de l'avatar, ainsi que d'une description des matériels dont les participants allaient être équipés (le livret utilisé pour illustrer ces explications est disponible en annexe — section C.2). Les participants étaient ensuite équipés par les expérimentateurs. Si un participant était équipé avant son binôme, il était mis en attente dans l'environnement. Dans le contexte où la communication vocale était autorisée, il était demandé aux participants s'ils étaient bien en mesure d'entendre leur partenaire.

L'expérimentateur pouvait alors lancer la session grâce au bouton dédié (pour rappel, voir section 4.2.2) sur le poste serveur. Les participants étaient alors téléportés devant un grand écran disposé dans l'environnement virtuel (voir figure 4.29d) (écran créé grâce au *mod*, *Webdisplay*, voir annexe C.1). En effet, afin de réduire au maximum les interactions entre participants et expérimentateurs, ainsi que de favoriser leur immersion dans l'univers de *Minecraft*, 90% des informations nécessaires au bon déroulement de l'expérimentation étaient ici données directement dans *Minecraft*. Cela, grâce à une vidéo créée spécifiquement et où la personne donnant les informations était l'avatar par défaut de *Minecraft*, *Steve* (ce dernier ayant néanmoins la voix de l'auteur de ce manuscrit, ce qui a été jugé préférable à l'utilisation d'une synthèse vocale trop inhumaine). Cette vidéo (version dédiée au contexte ACV-SC) est disponible à l'adresse suivante : <http://www.remy-eynard.net/minecraft-exp/introduction.html>. Au travers de cette séquence de six minutes, les participants recevaient toutes les informations nécessaires à la compréhension de la tâche à réaliser, de l'interface de *Minecraft*, des fonctions du cube émotionnel (le cas échéant), de l'utilisation du distributeur de matériaux ainsi que de la plateforme d'observation, des modalités du questionnaire post-expérimentation et pour finir des contrôles manette de leur avatar. À la fin de la vidéo, les participants voyaient leur inventaire se remplir des matériaux nécessaires et pouvaient alors commencer leur tâche de construction (une vidéo présentant un extrait d'une des sessions est disponible à l'adresse suivante : <http://www.remy-eynard.net/minecraft-exp/extrait.html>). Trois minutes avant la fin du temps imparti, un son de cloche retentissait et un compte à rebours apparaissait sur l'écran où était précédemment affichée la vidéo. Ces trois minutes passées, les participants étaient téléportés dans une zone souterraine noire (sans lumière) et un message textuel leur signifiait la fin de la session d'expérimentation. Une fois les matériels les équipant retirés, ils étaient invités à remplir le questionnaire post-expérimentation puis à quitter la salle.

4.3.7 Résultats

Nous avons, en vue de l'analyse quantitative de nos questionnaires, conduit des tests de normalité (Kolmogorov-Smirnov) sur chacune des variables (réponses au questionnaire et observations) obte-

nues afin de savoir si nous pouvions utiliser ou non des tests paramétriques. Aucune des variables testées ne suivait une courbe de distribution normale. Ces résultats sont cohérents avec le nombre de participants ($N = 52$) mais surtout avec l'utilisation d'échelles de Likert à cinq niveaux seulement. Par conséquent, nous nous sommes reposés sur la famille des tests non paramétriques (Pallant, 2010). Au regard de nos conditions expérimentales, à savoir quatre groupes de participants pour quatre conditions d'expérimentation (voir 4.3.6), nous nous sommes appuyés sur le U-test de Mann-Whitney pour échantillons indépendants afin de comparer nos contextes expérimentaux deux à deux comme illustré en figure 4.30. Pour établir la signification statistique de nos résultats, nous avons utilisé le seuil de probabilité conventionnel, soit $p < 0,05$. La figure 4.30 présente les cinq tests statistiques menés afin d'obtenir une vision globale des différences possibles entre nos différents contextes et donc de l'impact de la communication vocale ainsi que du cube émotionnel.

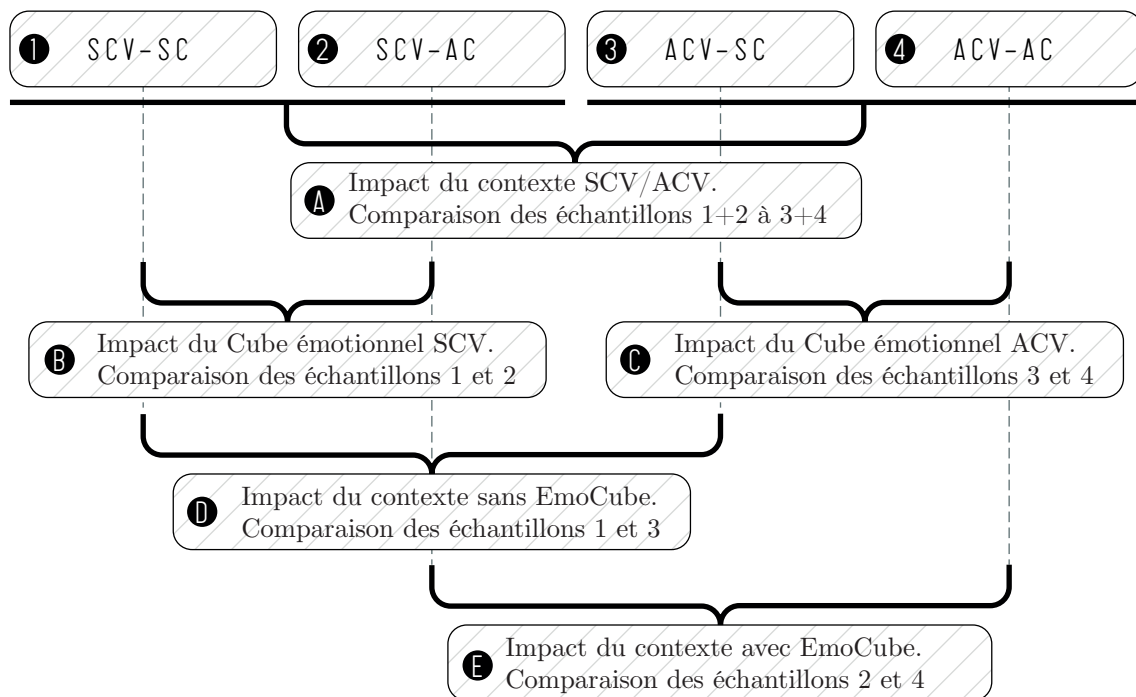


FIGURE 4.30 – Comparaisons statistiques menées pour l'expérimentation 2.

En comparaison avec notre précédente expérimentation, l'utilisation des lunettes 3D actives à permis d'éviter l'arrêt prématuré de certains participants causé par la *cyber sickness* (*Oculus Rift*). Toutes les données recueillies sans distinction remplissent les conditions pour être analysées.

Les résultats obtenus seront exposés suivant la même organisation précédemment utilisée afin de présenter les mesures mises en place (voir section 4.3.4), c'est-à-dire suivant l'organisation des facteurs proposés dans le modèle de qualité des DIMEs de Wu *et al.* (2009) et dans le modèle d'expérience utilisateur de (Pallot *et al.*, 2013) en découlant.

Qualité de Service ▶ Vivacité Sensorielle

► **Intrusivité :** quatre questions de notre questionnaire portaient sur l'intrusivité ou la potentielle gêne occasionnée par les matériels équipant les participants. Les résultats présentés ici sont des moyennes sur l'ensemble des participants, des U-test de Mann-Whitney pour échantillons indépen-

dants n'ayant pas révélé de différence statistique significative vis-à-vis des comparaisons A, B, C, D et E décrites en figure 4.30. Le tableau 4.4 présente, pour chaque matériel, une liste du nombre de participants pour chaque niveau des échelles de Likert. Afin de faciliter la lecture, l'effectif le plus grand est affiché en **rouge**.

	Lunettes 3D ¹		Casque EEG ²	
Échelle	Effectifs	Pourcentage	Effectifs	Pourcentage
Pas du tout (1)	29	55,8 %	36	73,5 %
2	15	28,8 %	8	16,3 %
3	3	5,8 %	5	10,2 %
4	5	9,6 %	0	0 %
Beaucoup (5)	0	0 %	0	0 %
Total	52	100 %	49	100 %
Moyenne (écart-type)	1,69 (,961)	—	1,37 (,668)	—
	Capteur ECG ³		Capteur GSR ⁴	
Pas du tout (1)	44	84,6 %	35	67,3 %
2	3	5,8 %	10	19,2 %
3	2	3,8 %	4	7,7 %
4	2	3,8 %	3	5,8 %
Beaucoup (5)	1	1,9 %	0	0 %
Total	52	100 %	52	100 %
Moyenne (écart-type)	1,33 (,879)	—	1,52 (,874)	—

Votre performance a-t-elle été affectée par le matériel qui vous équipait :

¹ les lunettes 3D ?

² le casque à électroencéphalogramme ?

³ le capteur de rythme cardiaque ?

⁴ le capteur de conductance de la peau ?

Questions 12.1 à 12.4 dans le questionnaire post-expérimentation, voir annexe C.3

TABEAU 4.4 – Données effectives et moyennes de l'évaluation de l'intrusivité des matériels utilisés.

Expérience (Utilisateur) Immersive ► Processus Expérientiel

► **Présence Sociale** : quel est l'impact de l'interaction sociale vocale et de l'affichage des émotions de l'utilisateur sur le sentiment de présence sociale ? Comme décrit dans la section Mesures (section 4.3.4), l'évaluation de la présence sociale s'appuie sur quatre variables : l'attention allouée, la compréhension mutuelle, l'interdépendance comportementale et l'assistance mutuelle.

Vis-à-vis de la variable **attention allouée**, une question était posée : « Vous avez fait preuve d'une attention particulièrement vive vis-à-vis de l'autre personne : » (1 = Je n'ai pas été très attentif à l'autre & 5 = J'ai été très attentif à l'autre) — question 13 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 et ont donné les résultats suivants :

- A) Entre les contextes SCV ($N = 26$) et ACV ($N = 26$), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 172,5$; $z = -3,215$; $p = ,001$);
- B) entre les contextes SCV-SC ($N = 12$) et SCV-AC ($N = 12$), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 41,5$; $z = -2,399$; $p = ,016$);
- C) entre les contextes ACV-SC ($N = 14$) et ACV-AC ($N = 14$), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative;
- D) entre les contextes SCV-SC ($N = 12$) et ACV-SC ($N = 14$), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 29$; $z = -2,969$; $p = ,003$);
- E) entre les contextes SCV-AC ($N = 12$) et ACV-AC ($N = 14$), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative.

La figure 4.31 présente les moyennes des réponses à cette question pour les comparaisons statistiques menées : nous pouvons constater que vis-à-vis des comparaisons A et D, c'est dans le contexte ACV que l'attention allouée est la mieux évaluée. Pour la comparaison B nous pouvons constater que c'est lors de la présence du cube émotionnel que l'attention allouée est la mieux évaluée.

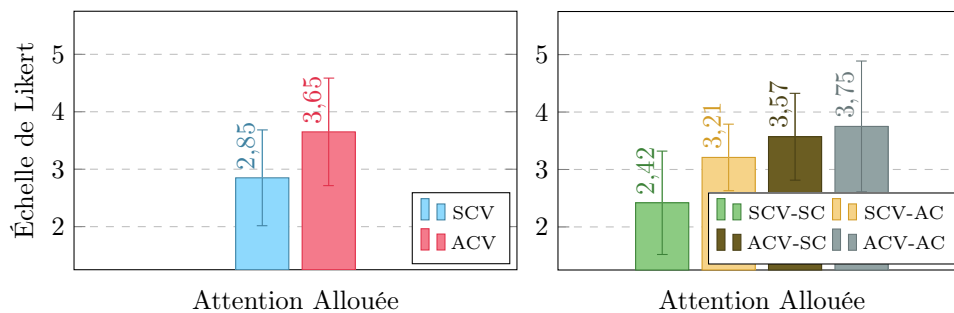


FIGURE 4.31 – Moyennes des évaluations de l'attention allouée.

Pour la variable **compréhension mutuelle**, deux questions étaient posées : « *L'autre personne vous a paru comprendre ce que vous désiriez faire (actions) :* » (1 = Pas du tout compris & 5 = Très bien compris) — question 14 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 et ont donné les résultats suivants :

- A) Entre les contextes SCV ($N = 26$) et ACV ($N = 26$), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 97$; $z = -4,734$; $p < ,001$);
- B) entre les contextes SCV-SC ($N = 12$) et SCV-AC ($N = 12$), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative;
- C) entre les contextes ACV-SC ($N = 14$) et ACV-AC ($N = 14$), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative;
- D) entre les contextes SCV-SC ($N = 12$) et ACV-SC ($N = 14$), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 28$; $z = -3,094$; $p = ,002$);
- E) entre les contextes SCV-AC ($N = 12$) et ACV-AC ($N = 14$), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 28$; $z = -3,094$; $p = ,002$).

La figure 4.32 présente les moyennes des réponses à cette question pour les comparaisons statistiques menées.

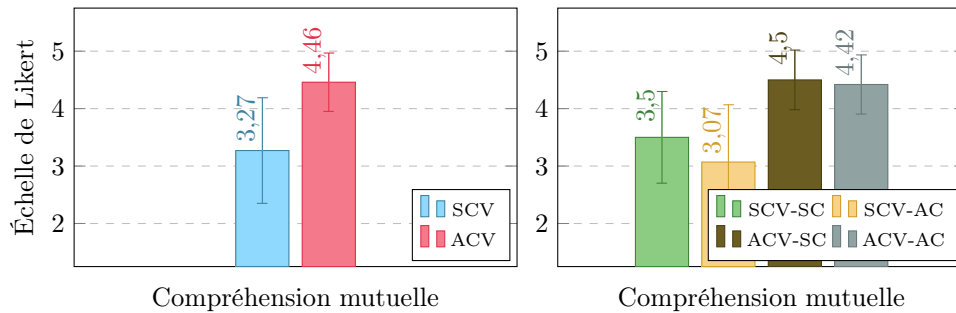


FIGURE 4.32 – Moyennes des évaluations de la compréhension mutuelle - Question 14.

Pour la question: « *J'ai compris ce que désirait l'autre personne (actions):* » (1 = Pas du tout compris & 5 = Très bien compris) — question 15 dans le questionnaire — des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 et ont donné les résultats suivants :

- A) Entre les contextes SCV (N = 26) et ACV (N = 26), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 137,5$; $z = -3,486$; $p < ,001$) ;
- B) entre les contextes SCV-SC (N = 12) et SCV-AC (N = 12), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative ;
- C) entre les contextes ACV-SC (N = 14) et ACV-AC (N = 14), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative ;
- D) entre les contextes SCV-SC (N = 12) et ACV-SC (N = 14), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 36$; $z = -2,595$; $p = ,009$) ;
- E) entre les contextes SCV-AC (N = 12) et ACV-AC (N = 14), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 31$; $z = -2,874$; $p = ,004$).

La figure 4.33 présente les moyennes des réponses à cette question pour les comparaisons statistiques menées.

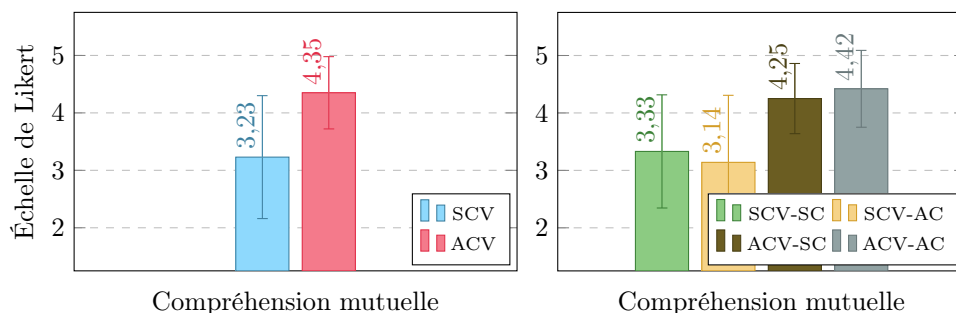


FIGURE 4.33 – Moyennes des évaluations de la compréhension mutuelle - Question 15.

Au regard des moyennes présentées dans les figures 4.32 et 4.33, nous pouvons constater que c'est

dans le contexte ACV que la compréhension mutuelle est la mieux évaluée, sans impact notable du cube émotionnel.

Vis-à-vis de l'**interdépendance comportementale**, deux questions étaient posées : « *Vos actions étaient dépendantes de celles de l'autre personne :* » (1 = J'ai agi uniquement en fonction de moi & 5 = J'ai totalement agi en fonction des actions de l'autre) — question 16 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 et ont donné les résultats suivants :

- A) Entre les contextes SCV (N = 26) et ACV (N = 26), le U-test de Mann-Whitney a révélé une différence statistique significative (U = 222 ; z = -2,230 ; p = ,026) ;
- B) entre les contextes SCV-SC (N = 12) et SCV-AC (N = 12), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative ;
- C) Entre les contextes ACV-SC (N = 14) et ACV-AC (N = 14), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative ;
- D) entre les contextes SCV-SC (N = 12) et ACV-SC (N = 14), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative ;
- E) entre les contextes SCV-AC (N = 12) et ACV-AC (N = 14), le U-test de Mann-Whitney a révélé une différence statistique significative (U = 37,5 ; z = -2,496 ; p = ,013).

La figure 4.34 présente les moyennes des réponses à cette question pour les comparaisons statistiques menées.

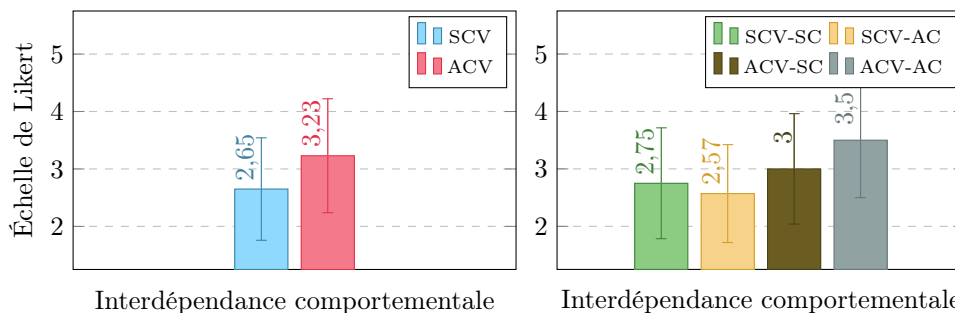


FIGURE 4.34 – Moyennes des évaluations de l'interdépendance comportementale - Question 16.

Pour la question : « *Votre style de jeu (par exemple : manière de construire la statue, récupération de matériaux) était une réponse directe au style de jeu de l'autre personne :* » (1 = J'ai agi de manière totalement autonome & 5 = Je me suis adapté sur le style de jeu de l'autre) — question 17 dans le questionnaire, des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 et ont donné les résultats suivants :

- A) Entre les contextes SCV (N = 26) et ACV (N = 26), le U-test de Mann-Whitney a révélé une différence statistique significative (U = 235,5 ; z = -1,959 ; p = ,050) ;
- B) entre les contextes SCV-SC (N = 12) et SCV-AC (N = 12), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative ;

- C) entre les contextes ACV-SC (N = 14) et ACV-AC (N = 14), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative ;
- D) entre les contextes SCV-SC (N = 12) et ACV-SC (N = 14), le U-test de Mann-Whitney n'a révélé aucune différence statistique significative ;
- E) entre les contextes SCV-AC (N = 12) et ACV-AC (N = 14), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 43$; $z = -2,225$; $p = ,036$).

La figure 4.35 présente les moyennes des réponses à cette question pour les comparaisons statistiques menées.

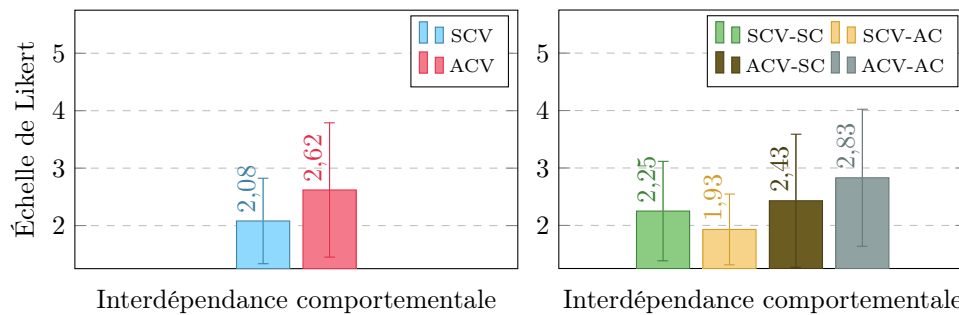


FIGURE 4.35 – Moyennes des évaluations de l'interdépendance comportementale - Question 17.

Au regard des figures 4.34 et 4.35, nous pouvons constater que, pour les deux questions ci-dessus, lorsque qu'une différence statistique significative est révélée, c'est dans le contexte ACV que l'interdépendance comportementale est la mieux évaluée. Là encore, la présence ou l'absence du cube émotionnel ne montre pas d'impact significatif.

Pour la variable **assistance mutuelle**, une question était posée : « *Votre partenaire vous a aidé dans l'accomplissement de la tâche :* » (1 = Pas du tout aidé & 5 = Beaucoup aidé) — question 9 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 mais n'ont révélé aucune différence statistique significative. La figure 4.36 présente néanmoins les moyennes des réponses à cette question pour chaque contexte expérimental.

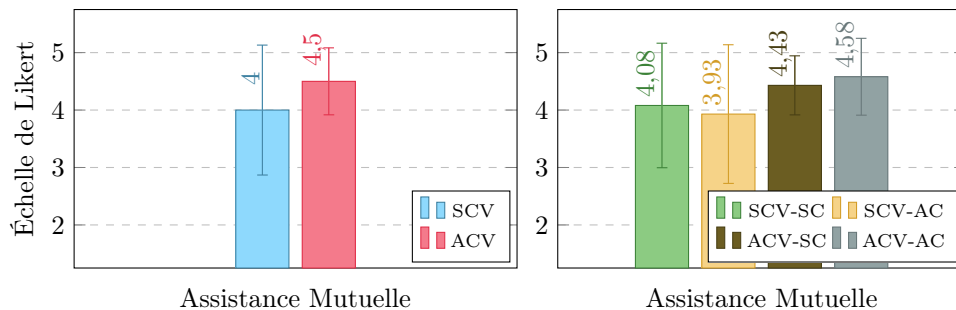


FIGURE 4.36 – Moyennes des évaluations de l'assistance mutuelle.

► **Émotions Sociales ► Empathie** : quel est l'impact de l'interaction sociale vocale et de l'affichage des émotions de l'utilisateur sur l'empathie ? Afin d'évaluer l'**empathie** entre les participants, une question était posée : « *Vous avez ressenti l'état émotionnel de votre partenaire :* » (1 = Pas du

tout & 5 = Oui complètement) — question 18 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 et ont donné les résultats suivants :

- A) Entre les contextes SCV (N = 26) et ACV (N = 26), le U-test de Mann-Whitney n’a révélé aucune différence statistique significative ;
- B) entre les contextes SCV-SC (N = 12) et SCV-AC (N = 12), le U-test de Mann-Whitney n’a révélé aucune différence statistique significative ;
- C) entre les contextes ACV-SC (N = 14) et ACV-AC (N = 14), le U-test de Mann-Whitney n’a révélé aucune différence statistique significative ;
- D) entre les contextes SCV-SC (N = 12) et ACV-SC (N = 14), le U-test de Mann-Whitney a révélé une différence statistique significative ($U = 41,5$; $z = -2,286$; $p = ,027$) ;
- E) entre les contextes SCV-AC (N = 12) et ACV-AC (N = 14), le U-test de Mann-Whitney n’a révélé aucune différence statistique significative.

La figure 4.37 présente les moyennes des réponses à cette question pour les comparaisons statistiques menées.

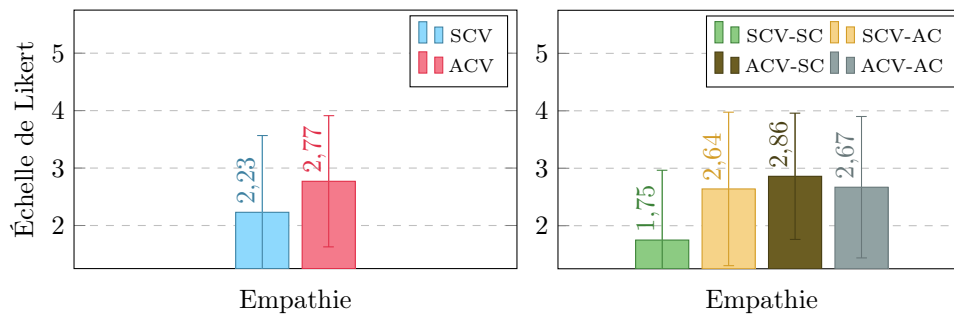


FIGURE 4.37 – Moyennes des évaluations de l’empathie.

En accord avec la figure 4.37, nous pouvons constater que la différence statistique révélée implique les deux contextes SCV et ACV mais sans la présence du cube émotionnel. Ce résultat ainsi que le fait que nous n’ayons pas différence significative sur le test entre contexte SCV et ACV globaux atteste de l’absence d’impact global du contexte ainsi que du cube émotionnel sur l’empathie dans le cadre de cette expérimentation.

► **Réponse Émotionnelle :** quel est l’impact de l’interaction sociale vocale et de l’affichage des émotions de l’utilisateur sur l’état émotionnel ? Vis-à-vis de la **réponse émotionnelle** et en premier lieu de son évaluation par les schémas (emoticons), une question était posée : « *Votre ressenti global vis-à-vis de cette expérience :* » (1 = Colère / frustration & 5 = Joie / allégresse) — question 1 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 mais n’ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

La figure 4.38 présente les moyennes des réponses à cette question pour chaque contexte expérimental.

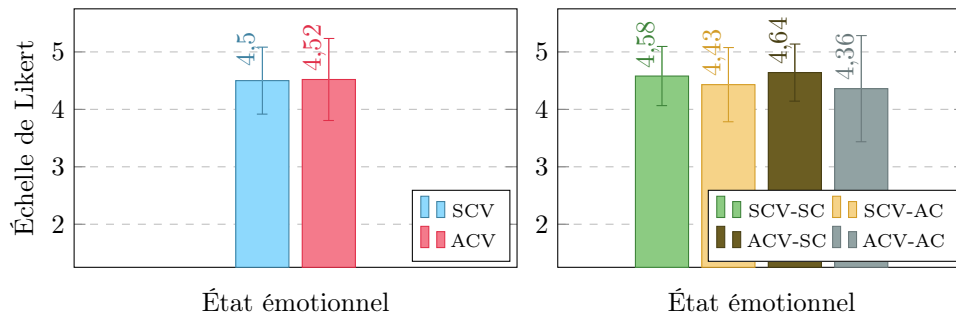


FIGURE 4.38 – Moyennes des évaluations de l'état émotionnel - Question 1.

En accord avec la figure 4.38 et malgré l'absence de différence statistique significative, nous pouvons constater que d'une manière générale, le ressenti des participants vis-à-vis de cette expérimentation est très positif.

Vis-à-vis de la confirmation du fonctionnement du cube émotionnel au travers des questions requérant l'évaluation des états émotionnels au cours de l'expérience et de la dominance de ces derniers (question 20), la présentation des données ne présente pas d'intérêt ici. En effet, elles n'apportent aucune indication utile à notre investigation de l'expérience utilisateur. Nous reviendrons néanmoins sur ce point dans la section *Analyse et synthèse* (4.3.8) de cette expérimentation.

Expérience (Utilisateur) Immersive ► Processus Rationnel

► **Flow Psychologique** : quel est l'impact de l'interaction sociale vocale et de l'affichage des émotions de l'utilisateurs sur le *Flow* Psychologique ? Pour le **Flow Psychologique**, deux questions étaient posées, la première évaluant le **niveau de challenge perçu** : « Vous avez ressenti une forme de challenge durant cette expérience : » (1 = Aucun challenge (trop facile) & 5 = Trop de challenge (Impossible d'accomplir les tâches)) — question 2 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 mais n'ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

La figure 4.39 présente les moyennes des réponses à cette question pour les comparaisons statistiques menées. Ces moyennes nous permettront d'estimer si les participants étaient en état de *flow* (bon équilibre entre niveau de défi perçu et compétence perçue, compétence perçue que nous pouvons estimer grâce au questionnaire pré-expérimentation).

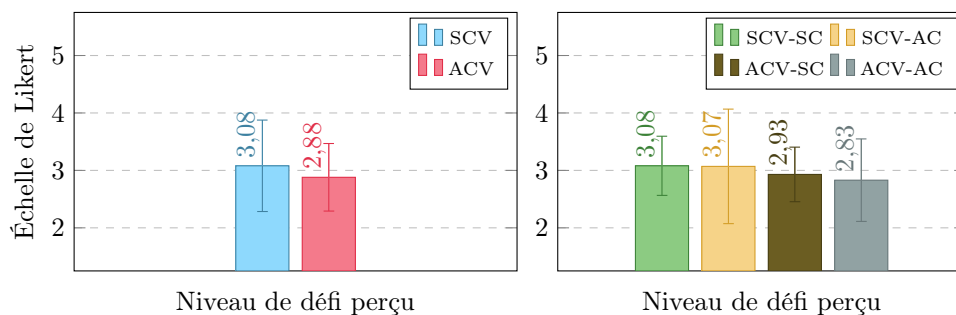


FIGURE 4.39 – Moyennes des évaluations du niveau de défi perçu - Question 2.

La seconde question évaluait l'**amusement**, résultant de l'état de *flow* : « *Je me suis amusé durant cette expérience :* » (1 = Je me suis ennuyé & 5 = Je me suis amusé) — question 6 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 mais n'ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

Néanmoins et pour la même raison que précédemment, la figure 4.40 présente les moyennes des réponses à cette question.

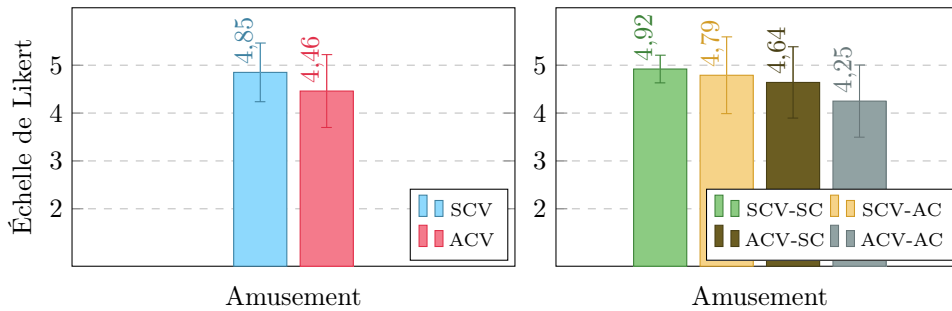


FIGURE 4.40 – Moyennes des évaluations de l'amusement - Question 6.

Nous pouvons constater sur la figure 4.40 que le niveau d'amusement de nos participants vis-à-vis de l'expérience vécue est très bien évalué, la moyenne la plus basse étant de 4,25 sur 5 dans le contexte ACV-SC.

Pour finir avec l'évaluation du *flow*, une question évaluait une seconde résultante de cet état, la **concentration** : « *Vous étiez focalisé sur la tâche à accomplir :* » (1 = Pas de tout & 5 = Oui complètement) — question 3 dans le questionnaire. Des U-Test de Mann-Whitney ont été conduits pour chacune des comparaisons décrites en figure 4.30 mais n'ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

La figure 4.41 présente les moyennes des réponses à cette question.

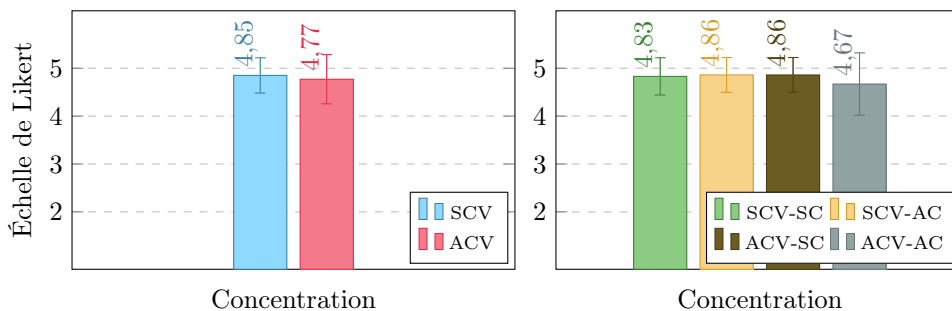


FIGURE 4.41 – Moyennes des évaluations de la concentration - Question 3.

Au regard des figures 4.39, 4.40 et 4.41, nous pouvons constater que les trois variables évaluées, niveau de challenge perçu, amusement et concentration, ont été très bien évaluées par nos participants, attestant qu'ils étaient généralement en état de *flow*. Néanmoins, en l'absence de différences statistiques significatives pour l'ensemble des comparaisons menées, nous ne pouvons que constater que ni nos

contextes SCV et ACV, ni le cube émotionnel, ne semblent avoir d'impact sur les variables testées.

► **Téléprésence** : quel est l'impact de l'interaction sociale vocale et de l'affichage des émotions de l'utilisateur sur le sentiment de présence virtuelle ? Vis-à-vis du sentiment **présence virtuelle**, cinq questions étaient posées, deux d'entre elles évaluant directement le sentiment de **présence virtuelle**, les deux suivantes, le **degré de contrôle** des participants sur l'environnement virtuel et la dernière la **concentration** des participants. Les deux questions évaluant directement le sentiment de présence virtuelle étaient les suivantes : « Dans l'environnement virtuel, vous aviez le sentiment « d'être sur place, présent » : » (1 = Pas du tout « sur place » & 5 = Tout à fait « sur place ») — question 7 dans le questionnaire — et « D'une certaine façon, vous aviez l'impression que le monde virtuel vous entourait » (1 = Pas du tout entouré & 5 = Totalemt entouré) — question 8 dans le questionnaire. Les U-Test de Mann-Whitney conduits pour chacune des comparaisons décrites en figure 4.30 n'ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

Les figures 4.42 et 4.43 présentent néanmoins les moyennes des réponses à ces questions pour les comparaisons statistiques menées. Ces moyennes nous permettront d'estimer le niveau de présence virtuelle éprouvé par nos participants.

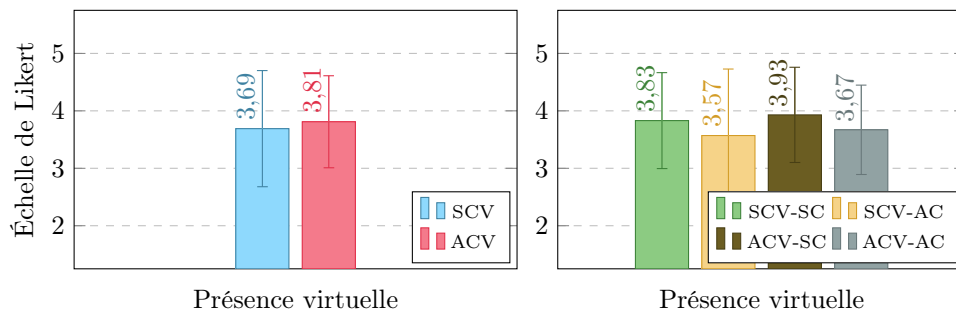


FIGURE 4.42 – Moyennes des évaluations du sentiment de présence virtuelle - Question 7.

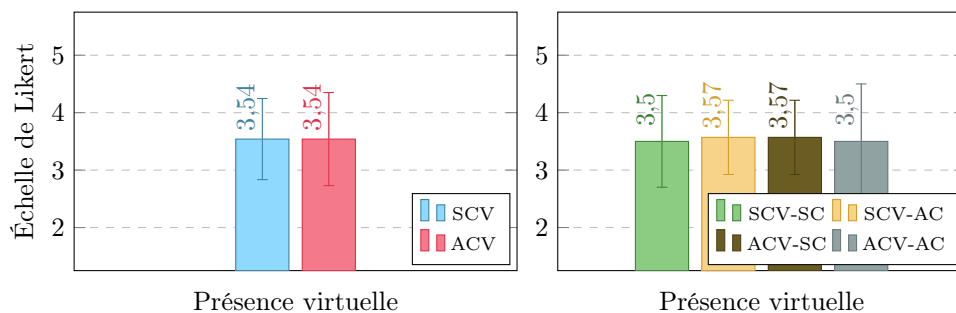


FIGURE 4.43 – Moyennes des évaluations du sentiment de présence virtuelle - Question 8.

En accord avec les figures 4.42 et 4.43, nous pouvons constater que la plus petite évaluation moyenne du sentiment de présence virtuelle est de 3,5 sur 5, attestant du bon niveau de présence virtuelle.

Pour évaluer le **degré de contrôle**, deux questions étaient posées : « Vous vous êtes sentis maître de vos mouvements, de vos actions dans l'environnement virtuel » : » (1 = Pas du tout & 5 = Oui

complètement) — question 4 dans le questionnaire — et : « *Votre avatar répondait bien aux actions que vous désiriez faire* » (1 = Pas du tout & 5 = Oui complètement) — question 5 dans le questionnaire. Les U-Test de Mann-Whitney conduits pour chacune des comparaisons décrites en figure 4.30 n'ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

Les figures 4.44 et 4.45 présentent néanmoins les moyennes des réponses à ces questions pour les comparaisons statistiques menées, permettant là encore une estimation du niveau de présence virtuelle éprouvé par les participants.

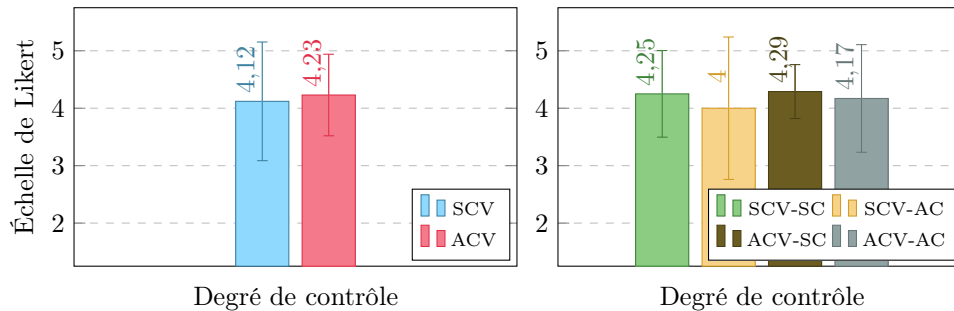


FIGURE 4.44 – Moyennes des évaluations du degré de contrôle - Question 4.

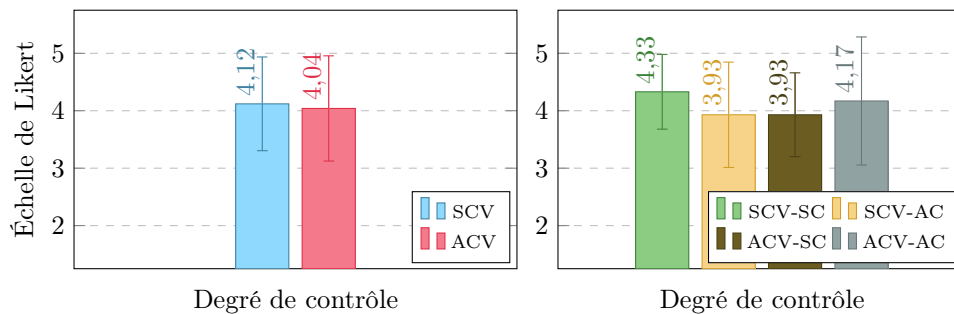


FIGURE 4.45 – Moyennes des évaluations du du degré de contrôle - Question 5.

En accord avec les figures 4.44 et 4.45, nous pouvons constater que la plus petite évaluation moyenne du niveau de contrôle est de 3,93 sur 5, attestant du très bon de degré de contrôle des participants sur leurs avatars et sur l'environnement virtuel.

Pour évaluer la **concentration**, une question était posée : « *Quel était votre sentiment au sujet de cette expérience* » (1 = Ennuyeux & 5 = Captivant) — question 21 dans le questionnaire. Les U-Test de Mann-Whitney conduits pour chacune des comparaisons décrites en figure 4.30 n'ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

La figure 4.46 présente néanmoins les moyennes des réponses à cette question pour les comparaisons statistiques menées. Ces moyennes nous permettront d'estimer le niveau de concentration éprouvé par nos participants.

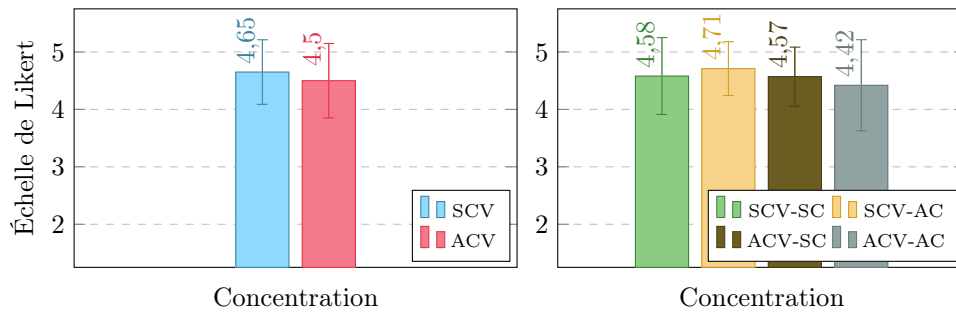


FIGURE 4.46 – Moyennes des évaluations de la concentration.

Au regard de la figure 4.46, nous pouvons constater que la plus petite évaluation moyenne de la concentration est de 4,42 sur 5, attestant du très bon niveau de concentration des participants lors de l'expérimentation.

Vis-à-vis des mesures EEG, n'ayant rencontré aucune difficulté face à de l'utilisation des casques, et les mesures affichées lors de l'enregistrement nous semblant cohérentes avec ce que nous avons pu voir dans d'autres travaux ou tutoriels, nous pensions que notre prise de mesures était correcte. En dépit de cela, lors de l'analyse des données recueillies, nos mesures présentaient un très grand nombre d'artefacts qu'il est possible de supprimer par l'utilisation d'algorithmes dédiés à cet usage. Malgré nos nombreuses tentatives et en suivant les sources que nous avons pu trouver vis-à-vis de ces démarches, nous ne sommes pas arrivés à supprimer de manière convaincante ces artefacts, ne nous permettant pas une comparaison avec les travaux de Clemente *et al.* (2013) comme nous désirions le faire. Les causes possibles de cet échec seront exposées dans la section *Analyse et synthèse* de cette expérimentation (section 4.3.8).

► **Acceptabilité Technologique** ► **Utilité** : afin d'évaluer l'utilité perçue du cube émotionnel, une question était posée : « *Le « cube émotionnel » vous a aidé à comprendre l'état émotionnel de votre partenaire :* » (1 = Pas du tout & 5 = Oui complètement) — question 19 dans le questionnaire. Le U-Test de Mann-Whitney conduit afin de tester les différences potentielles entre les contexte SCV-AC et ACV-AC n'a révélé aucune différence statistique significative.

La figure 4.47 présente les moyennes des réponses à cette question pour la comparaison statistique menée. Ces moyennes nous permettront d'évaluer l'utilité perçue du cube émotionnel par les participants.

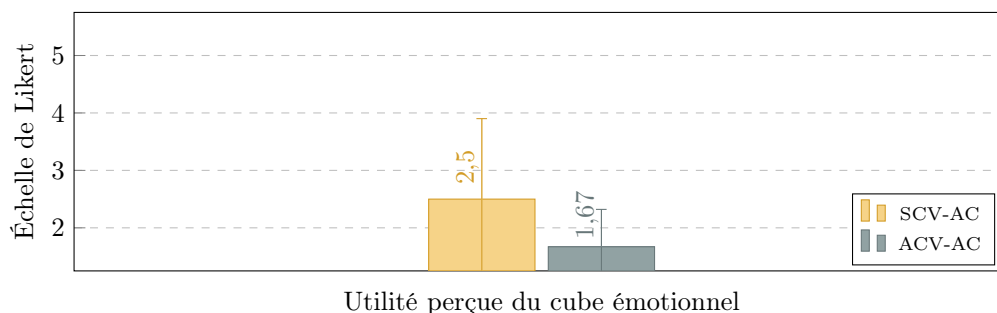


FIGURE 4.47 – Moyennes des évaluations de l'utilité perçue du cube émotionnel.

En accord avec la figure 4.47, nous pouvons constater que le cube émotionnel n'a pas été perçu comme réellement utile vis-à-vis de la perception de l'état émotionnel du partenaire d'expérimentation.

Qualité de l'Expérience

► **Comportement Social :** quel est l'impact de l'interaction sociale vocale et de l'affichage des émotions de l'utilisateur sur le comportement social ? Lors de notre première expérimentation, les questions portant sur les échanges verbaux n'avaient pu être utilisées comme matériel de comparaison. Dans le cadre de cette seconde expérimentation, nous étions en mesure de comparer les contextes ACV-SC et ACV-AC, afin de déterminer si la présence du cube émotionnel a pu influencer les sujets des conversations qui se sont tenues au cours de l'expérimentation.

Les U-Test de Mann-Whitney menés dans ce but ont révélé une différence significative sur deux questions : « *Sur quels thèmes avez-vous échangé ? L'environnement virtuel :* » (1 = Pas du tout & 5 = Beaucoup) — question 33 dans le questionnaire ($U = 36$; $z = -2,591$; $p = ,013$). Et : « *Sur quels thèmes avez-vous échangé ? Les modalités d'interaction (utilisation de la manette et du jeu Minecraft) :* » (1 = Pas du tout & 5 = Beaucoup) — question 34 dans le questionnaire ($U = 31,5$; $z = -2,763$; $p = ,005$). Afin de savoir dans lequel des deux contextes ACV-SC ($N = 14$) ou ACV-AC ($N = 14$) ces deux sujets ont été le plus évoqués, les moyennes des réponses à ces questions ont été calculées :

- Question 33 : ACV-SC : $M = 1,64$ ÉT = ,745 et ACV-AC : $M = 2,75$ ÉT = 1,138 ;
- question 34 : ACV-SC : $M = 1,93$ ÉT = ,829 et ACV-AC : $M = 3,58$ ÉT = 1,505.

On constate alors que c'est dans le contexte ACV-AC que les deux sujets, environnement virtuel et modalités d'interaction, ont été le plus évoqués selon les participants.

Intéressons-nous à présent aux moyennes globales de ces questions portant sur les interactions sociales vocales.

- a) Fréquence : « *Vous avez interagi avec votre partenaire (conversation) à la fréquence suivante :* » (1 = Rarement & 5 = De nombreuses fois) — question 23 dans le questionnaire.
 $M = 4$; ÉT = ,938.
- b) Assistance Mutuelle (Comportement Empathique) : « *Votre partenaire vous a appris des choses durant l'expérience :* » (1 = Rien n'appris & 5 = Beaucoup appris) — question 24 dans le questionnaire.
 $M = 2,31$; ÉT = 1,158.
- c) Comportement Empathique : « *Vous avez au travers de vos conversations ressenti l'état émotionnel de votre partenaire :* » (1 = Pas du tout & 5 = Complètement) — question 25 dans le questionnaire.
 $M = 3,42$; ÉT = 1,027.
- d) Compréhension Mutuelle (Présence Sociale) : « *Vous avez ressenti un manque de compréhension entre vous et votre partenaire :* » (1 = Il était difficile de se comprendre & 5 = Nous nous sommes parfaitement compris) — question 26 dans le questionnaire.
 $M = 4,35$; ÉT = ,745.

- e) Nature (A): « *Vos discussions étaient plutôt :* » (1 = Négatives (tension) & 5 = Positives (partage, amusement)) — question 27 dans le questionnaire.
M = 4,54; ÉT = ,647.
- f) Attention Allouée (Présence Sociale): « *Le degré d'intensité de vos conversations était :* » (1 = Très calme & 5 = Totalement enflammée) — question 28 dans le questionnaire.
M = 2,15; ÉT = ,952.
- g) Nature (B) et Attention Allouée (Présence Sociale): « *Vous êtes-vous mutuellement encouragés durant l'expérience :* » (1 = Pas du tout & 5 = Beaucoup) — question 29 dans le questionnaire.
M = 2,81; ÉT = 1,226.
- h) Nature (C) et Attention Allouée (Présence Sociale): « *Vous êtes-vous mutuellement conseillés vis-à-vis de la tâche à accomplir :* » (1 = Pas du tout & 5 = Beaucoup) — question 30 dans le questionnaire.
M = 3,88; ÉT = ,952.
- i) Engagement (Présence Virtuelle): « *À propos de quel sujet avez-vous échangé : la réalisation de la tâche demandée :* » (1 = Pas du tout & 5 = Beaucoup) — question 31 dans le questionnaire.
M = 4,58; ÉT = ,578.
- j) Réponse Émotionnelle: « *À propos de quel sujet avez-vous échangé : votre ressenti vis-à-vis de l'expérience en cours :* » (1 = Pas du tout & 5 = Beaucoup) — question 32 dans le questionnaire.
M = 2,27; ÉT = 1,185.
- k) Qualité de Service: « *À propos de quel sujet avez-vous échangé : l'environnement virtuel :* » (1 = Pas du tout & 5 = Beaucoup) — question 33 dans le questionnaire.
M = 2,15; ÉT = 1,084.
- l) Degré de contrôle: « *À propos de quel sujet avez-vous échangé : les modalités d'interaction (utilisation de la manette et du jeu Minecraft) :* » (1 = Pas du tout & 5 = Beaucoup) — question 34 dans le questionnaire.
M = 2,69; ÉT = 1,436.

Nous reviendrons sur ces résultats dans la section *Analyse et synthèse* (section 4.3.8.)

► **Efficacité :** quel est l'impact de l'interaction sociale vocale et de l'affichage des émotions de l'utilisateur sur l'efficacité perçue et réelle ? Afin d'évaluer l'**efficacité perçue** des participants, deux questions étaient posées: « *Vous avez été efficace vis-à-vis de l'accomplissement de la tâche :* » (1 = Pas du tout efficace & 5 = Très efficace) — question 10 dans le questionnaire — et: « *Vous avez eu le sentiment de travailler ensemble sur la tâche* » (1 = Pas du tout ensemble & 5 = En synergie totale) — question 11 dans le questionnaire. Les U-Test de Mann-Whitney conduits pour chacune des comparaisons décrites en figure 4.30 n'ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

Les figures 4.48 et 4.49 présentent les moyennes des réponses à ces questions pour les comparaisons statistiques menées. Ces moyennes nous permettront une comparaison avec les mesures quantitatives

objectives de l'efficacité afin d'obtenir des indications sur les différences potentielles entre l'estimation de leur efficacité par les participants et leur efficacité réelle.

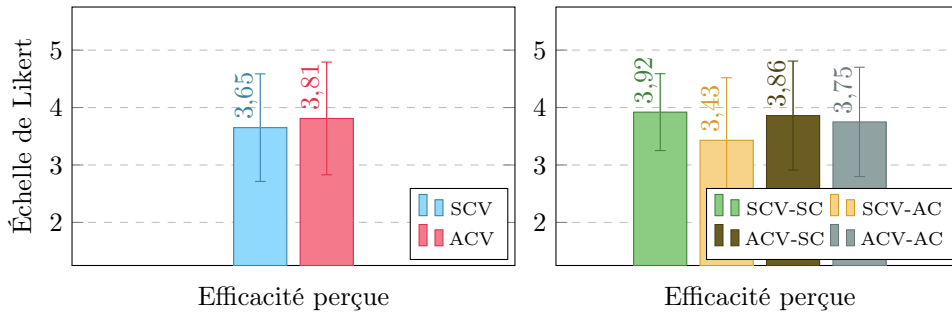


FIGURE 4.48 – Moyennes des évaluations de l'efficacité perçue - Question 10.

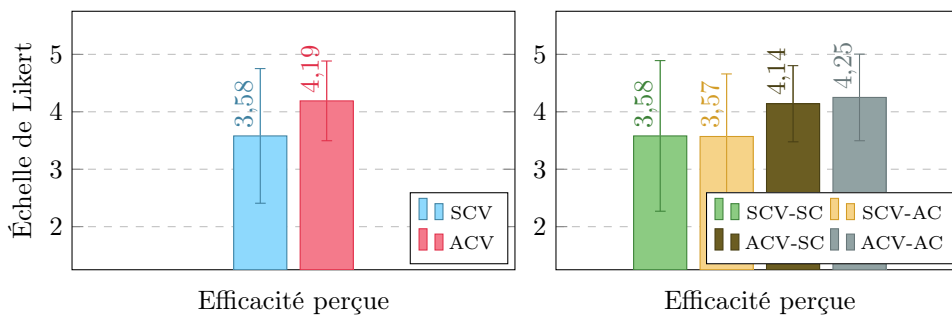


FIGURE 4.49 – Moyennes des évaluations de l'efficacité perçue - Question 11.

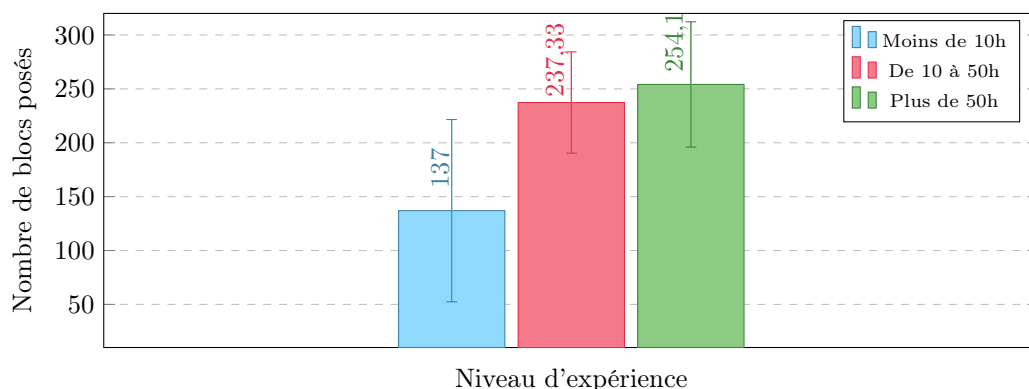
Vis-à-vis de l'évaluation quantitative objective de l'efficacité des participants, nous avons recueilli les données suivantes : le temps passé avant la pose du premier bloc, le nombre de blocs placés, le nombre de blocs détruits, les rapports indicatifs d'efficacité individuelle et par binôme, le temps de complétion de la tâche ainsi que le nombre d'erreurs restantes sur la construction à la fin de l'expérimentation. Afin de nous assurer de la robustesse de ces mesures, nous avons demandé aux participants (dans notre questionnaire pré-expérimentation) quelle était leur expérience sur *Minecraft*. Cette information nous permet de tester l'impact de cette expérience préalable sur l'efficacité perçue et observée. Nous avons aussi examiné la distribution des participants dans nos quatre contextes vis-à-vis de leur expérience préalable de *Minecraft*. Malgré le fait que nous n'ayons pas contrôlé la distribution de nos participants dans chaque contexte (les participants ayant eux-mêmes choisi le jour et l'heure de leur passage), nous avons obtenu des groupes relativement homogènes (voir tableau 4.5.)

Afin d'évaluer les possibles interactions entre nos quatre contextes (SCV-SC, SCV-AC, ACV-SC & ACV-AC) et le niveau d'expérience préalable de *Minecraft* sur les variables observées de l'efficacité, l'utilisation d'analyses de la variance (ANOVA) était requise. Ces analyses de la variance n'ont révélé aucune interaction entre nos contextes et le niveau d'expérience préalable de *Minecraft* sur les variables observées de l'efficacité. Nous pouvions alors focaliser notre attention sur les effets principaux entre facteurs, c'est-à-dire l'impact du niveau d'expérience préalable de *Minecraft* sur les variables observées de l'efficacité et l'impact de nos quatre contextes expérimentaux sur ces mêmes variables. L'analyse de la variance ayant mis en exergue un effet principal du niveau d'expérience préalable de *Minecraft*

	Pas d'expérience	Moins de 10h	De 1 à 50h	Plus de 50h	Total
SCV-SC	05	01	03	03	12
SCV-AC	03	00	03	06	12
ACV-SC	05	01	01	07	14
ACV-AC	06	02	02	04	14
Cumul SCV	08	01	06	09	24
Cumul ACV	11	03	03	11	28
Total cumul	19	04	09	20	52

TABLEAU 4.5 – Répartition des participants vis-à-vis de leur expérience préalable de *Minecraft*.

sur le nombre de blocs placés, un test de Kruskal-Wallis a été mené vis-à-vis de ces deux variables, et a confirmé cet impact général ($\chi^2 = 6,499$; $p = ,039$). Nous avons alors mené un test U de Mann-Whitney pour analyser plus finement les différences entre groupes relatifs à l'expérience préalable, sur le nombre de blocs posés. Une différence significative est observée entre le groupe jouant moins de 10 heures et celui jouant plus de 50 heures ($U = 7$; $z = -2,556$; $p = ,011$). Afin de comprendre cet impact, les moyennes du nombre de blocs placés pour chaque niveau d'expérience préalable de *Minecraft* ont été calculées et sont présentées en figure 4.50. Nous pouvons constater sur cette figure une très large différence ($>$ à 100 blocs) entre les participants, dont les niveaux d'expérience préalable de *Minecraft* se situent de moins de 10 heures à plus de 50 heures (malgré la différence affichée sur la figure 4.50 entre les groupes dont l'expérience préalable de *Minecraft* se situe de moins de 10 heures à de 10 à 50 heures, aucune différence statistique significative n'a été relevée, indubitablement, cela est dû au large écart type que présente le groupe dont l'expérience préalable est de moins de 10 heures vis-à-vis du nombre de blocs posé). Prenant en compte qu'une fois un certain niveau atteint les différences possibles se voient réduites, la différence entre les deux échantillons les plus expérimentés est bien moindre (≈ 17 blocs).

FIGURE 4.50 – Moyennes du nombre de blocs placés vis-à-vis du niveau d'expérience préalable de *Minecraft*.

Nous avons par la suite continué notre analyse de l'impact de nos contextes SCV et ACV, ainsi que du cube émotionnel sur les variables observées de l'efficacité. Les U-Test de Mann-Whitney conduits

pour chacune des comparaisons décrites en figure 4.30 ont révélé les différences significatives suivantes :

- Entre les contextes SCV et ACV : Une différence statistique significative a été révélée sur le temps passé avant la pose du premier bloc ($U = 216$; $z = -2,055$; $p = ,040$). Afin de savoir dans lequel des deux contextes testés le temps passé avant la pose du premier bloc fut le plus long, les moyennes ont été calculées : SCV : $M = 217,81$ (secondes) ; $\text{ÉT} = 65,961$ et ACV : $M = 222,77$ (secondes) ; $\text{ÉT} = 79,914$.
- Entre les contextes ACV-SC et ACV-AC : Une différence statistique significative a été révélée sur le temps de complétion de la tâche ($U = 8$; $z = -2,849$; $p = ,003$). Afin de savoir dans lequel des deux contextes testés le temps de complétion de la tâche fut le plus long, les moyennes ont été calculées : ACV-SC : $M = 1669,90$ (secondes) ; $\text{ÉT} = 127,198$ et ACV-AC : $M = 1456,25$ (secondes) ; $\text{ÉT} = 171,844$.

Aucune autre différence statistique significative n'a été relevée vis-à-vis des observations faites sur le temps passé avant la pose du premier bloc, sur le temps de complétion de la tâche et sur le rapport d'efficacité pour les comparaisons décrites en figure 4.30.

► **Adoption Technologique** ► **Intention d'Utilisation** : quel est l'impact de l'interaction sociale vocale et de l'affichage des émotions de l'utilisateur sur l'adoption technologique ?

Vis-à-vis de l'**intention d'utilisation** de notre configuration expérimentale par les participants, deux questions étaient posées : « *Souhaiteriez-vous utiliser à nouveau cette application (tâche comprise) dans un contexte personnel* » : » (1 = Pas du tout & 5 = Oui évidemment) — question 36 dans le questionnaire — et : « *Recommanderiez-vous l'utilisation de cette application à d'autres personnes* » (1 = Pas du tout & 5 = Oui évidemment) — question 38 dans le questionnaire. Les U-Test de Mann-Whitney conduits pour chacune des comparaisons décrites en figure 4.30 n'ont révélé aucune différence statistique significative dans aucun des cas pour la variable testée.

Les figures 4.51 et 4.52 présentent les moyennes des réponses à ces questions pour les comparaisons statistiques menées. Ces moyennes permettent d'évaluer la qualité de notre configuration expérimentale, l'intention d'utilisation (future) étant un bon indicateur de plaisir éprouvé étant donné qu'ici l'activité ne présentait aucune utilité pour les participants si ce n'est celle de passer un bon moment.

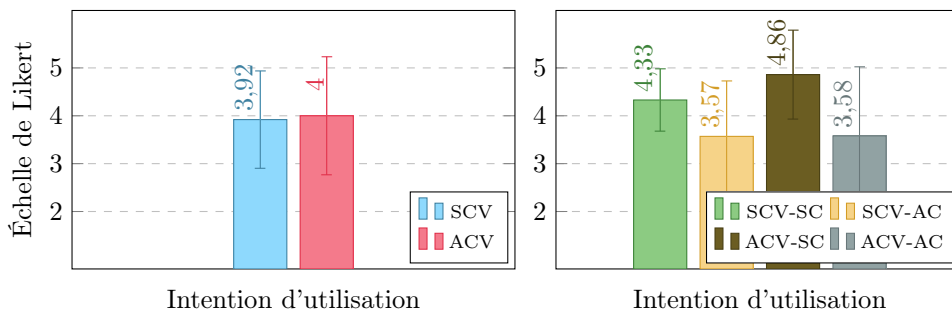


FIGURE 4.51 – Moyennes des évaluations de l'intention d'utilisation - Question 38.

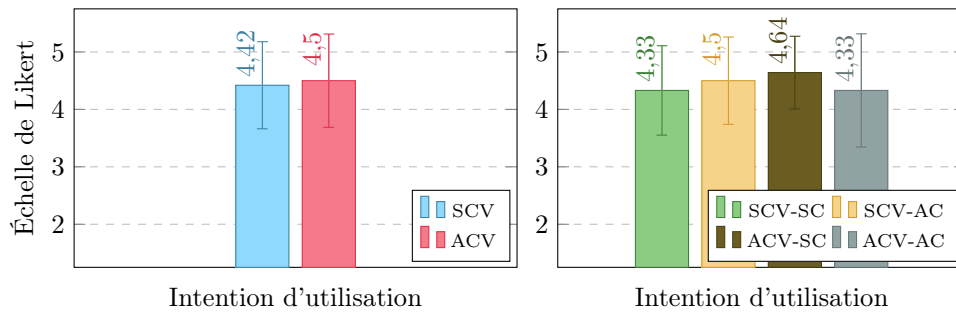


FIGURE 4.52 – Moyennes des évaluations de l'intention d'utilisation - Question 38.

Les figures 4.51 et 4.52 indiquent que les participants ont apprécié l'expérience que leur a procurée cette expérimentation : ils seraient en majorité prompt à refaire cette expérience et même à la conseiller à d'autres personnes.

4.3.8 Analyse et synthèse

Pour ce second travail d'expérimentation, et grâce à l'expérience de la phase précédente, nous avons fait en sorte de ne pas rencontrer les mêmes écueils. Les changements techniques (QoS) les plus importants furent l'abandon de l'*Oculus Rift* au profit des lunettes 3D actives. Cette modification a permis à tous les participants de compléter leur session d'expérimentation sans pour autant avoir une influence néfaste sur le sentiment de présence virtuelle. Nous pouvons constater dans le rapport de l'évaluation de l'intrusivité des matériels utilisés (voir tableau 4.4) que 73,5% de nos participants ont estimé les lunettes 3D comme non intrusives, et que la note maximale est de trois sur cinq, un rapport presque inverse à l'évaluation de l'*Oculus Rift*. Cela a également facilité l'utilisation des casques EEG, bien qu'en définitive, nous le verrons par la suite, les résultats n'en furent pas plus probants. Deuxièmement, le passage d'une construction verticale à une construction horizontale nous a permis d'éviter les sensations de vertige — bien que celles-ci eusses sans doute été moindres avec les lunettes 3D — ainsi que les frustrations ressenties par les participants à l'encontre de l'utilisation des échelles et des chutes trop fréquentes qu'elles engendraient. Nous avons amélioré l'expérience globale, la rendant plus fluide, plus agréable. En premier lieu, par l'information des participants directement dans l'environnement virtuel grâce à la vidéo prévue à cet effet, leur permettant de se sentir immergés dans une application qualitative conçue pour eux et pas dans le but unique de recueillir des données. La tâche, outre le changement d'axes de la construction, s'est aussi vue améliorée, par l'inclusion du distributeur de matériaux ajoutant un élément d'interaction à l'environnement et rendant l'activité moins « robotique ». La plateforme d'observation, elle aussi, en un sens, interactive par la manière d'y accéder, avait pour objectif d'ajouter un élément ludique mais aussi de réduire la frustration des participants quant à la difficulté de faire coïncider leurs plans respectifs avec leur construction. Nous souhaitions rendre cet environnement plus accessible aux personnes n'ayant aucune connaissance de *Minecraft* (plus nombreuses lors de cette seconde expérimentation), fait dont nous avons eu la confirmation lors d'échanges informels avec nos participants après leur passation.

Malgré ces modifications et améliorations, notre travail présente certaines limites. La première est liée, comme précédemment, au nombre de participants : malgré nos soixante-dix-sept inscrits, nous n'avons pu réunir que cinquante-deux participants. En effet, nous avons fait face à une vague de désistements nous ayant été notifiés ou non, représentant 19% de notre effectif initial. Il est donc fort à propos lors de la mise en place de tels travaux de prévoir un recrutement 20% supérieur au nombre de participants souhaité. Nous ne reviendrons pas sur les implications d'un nombre de participants réduit, ces dernières ayant été exposées en section 4.2.8. Néanmoins, dans le cadre de cette seconde expérimentation, cela se montre encore plus limitant sachant que nous n'avons plus deux mais quatre contextes soit quatre groupes auxquels nous aurions désiré attribuer vingt participants au minimum.

Malgré nos efforts et notre réussite apparente lors de la collecte de données grâce aux casques EEG, nous n'avons pas réussi à analyser les données, ces dernières présentant des artefacts de mouvement — ce qui est tout à fait normal — que nous n'avons su supprimer. Nous avons pourtant suivi les indications à notre disposition mais sans succès malgré nos tentatives répétées. Nous devons imputer cet échec à notre manque d'expérience vis-à-vis de ce type de matériel. La persévérance n'étant pas toujours un bon palliatif à la connaissance, nous avons pris conscience de la nécessité d'être formé à ce type d'outil pour en tirer des résultats probants. Nous sommes en l'occurrence extrêmement déçus de cet échec, car nous croyons fermement aux potentiels de ce type de technologies dans l'étude de phénomènes mentaux tels que la présence virtuelle.

Vis-à-vis du cube émotionnel, nous sommes face à un doute : nous avons constaté, lors du visionnage des enregistrements vidéo produits au cours des sessions d'expérimentation, que ce dernier ne variait que peu, et la plupart du temps uniquement entre la neutralité et l'ennui. La question se pose alors de savoir si cela provenait plutôt de l'expérience proposée, n'étant pas suffisamment génératrice d'émotions pour susciter des changements physiologiques, ou si la conception de notre cube émotionnel, et plus particulièrement de la définition des schémas de variation, n'était pas correcte. Nous pensons que plusieurs facteurs sont impliqués : notre activité était sans doute trop « fade » d'un point de vue émotionnel et d'un autre côté, la minute définie pour l'établissement de l'état neutre nous semble avec le recul insuffisante. Les schémas physiologiques utilisés demanderaient une validation empirique grâce à un support proposant l'alternance d'émotions fortes par exemple des films d'horreur ou romantiques sur une large population (afin de s'assurer d'avoir suffisamment de participants sensibles à ce type de stimuli émotionnels). Nous avons, dans notre questionnaire, mis en place quatre questions visant à nous apporter une évaluation du fonctionnement du cube par comparaison avec nos observations. Ces questions requéraient le positionnement sur une échelle de Likert entre deux émotions ou état d'esprits opposés (par exemple, ennui / absorption totale) ainsi qu'une évaluation de la durée de cette émotion ou état d'esprit durant l'expérimentation. Lors de l'analyse des données recueillies, nous nous sommes aperçus que cette conception n'était pas la plus adaptée. En effet, les réponses obtenues ici oscillent sur nos échelles à cinq niveaux en moyenne entre 3,5 et 4,5 pour toutes les émotions ou états d'esprit, et vis-à-vis de la durée, entre 3,5 et 3,8 en moyenne. Cela pose un problème de précision, il est en effet possible que les participants se soient à un moment ennuyés puis se soient totalement impliqués dans la tâche. Il aurait été intéressant de privilégier l'utilisation du *Self-Assessment Manikin questionnaire* de

Bradley et Lang (1994) (voir en annexe, fig A.4) plutôt qu’une opposition entre deux émotions ou état d’esprit qui ne permet pas de traduire une variation d’état émotionnel au cours de l’expérience. De plus, les participants ont évalué le cube comme peu utile quant à la perception de l’état émotionnel de leur partenaire. Nous pensons que cela provient en partie de leur constat du peu de variations et donc d’un désintérêt pour le cube après quelques minutes d’immersion. Cette fonctionnalité demanderait, pour une utilisation future, une validation des schémas physiologiques, la conception d’une activité spécifique (génératrice d’émotions) ainsi que la définition de questions plus pertinentes quant à son évaluation.

Nous n’avons, ici encore, pas pu tenir de groupes de discussion, nos participants n’ayant en grande majorité pas répondu à l’appel à participation. Nous les avons pourtant informés de notre besoin de les revoir dans ce cadre, mais seulement cinq personnes nous ont répondu. Refusant de nous passer de groupes de discussion une fois de plus, l’aspect qualitatif nous apparaissant comme primordial pour notre étude empirique, nous avons mis sur pied une nouvelle expérimentation, complémentaire à celle-ci, qui sera décrite dans la suite de ce manuscrit.

Focalisons-nous à présent sur les résultats obtenus lors de cette seconde expérimentation, tout d’abord pour la présence sociale. Lors de notre analyse statistique de l’évaluation de l’attention allouée, nous avons mis en avant des différences statistiques significatives entre les contextes SCV-SC et SCV-AC. Nous pensons que cette différence en faveur d’un contexte proposant le cube émotionnel peut s’expliquer par un effet de curiosité vis-à-vis de cette fonctionnalité. En effet, les participants étant informés de sa présence et de son utilité, il est très probable que cela ait suscité chez eux une certaine curiosité, pour percevoir les variations décrites lors de la vidéo introductive à l’expérimentation. Pour les deux autres cas, différence entre les contextes SCV et ACV ainsi qu’entre SCV-SC et ACV-SC, nous trouvons, contrairement à notre première expérimentation, une meilleure évaluation de l’attention allouée dans le contexte ACV. Nous avons précédemment expliqué l’absence de différence par la nécessité des binômes SCV de prêter attention à leur partenaire afin de s’assurer de l’avancement de la tâche ou pour vérifier que ce dernier ne rencontrait pas de problèmes. Ici, la tâche coopérative et non plus collaborative tend à réduire ce besoin, chacun étant en mesure de contribuer à la tâche suivant son plan, sans dépendre de son partenaire. Néanmoins, dans le contexte ACV, comme nous l’avons précédemment expliqué, l’allocation de l’attention est naturelle, à la prise de parole de l’un des protagonistes, l’autre lui prête attention. C’est donc par la nature de la tâche que nous pensons pouvoir expliquer cette différence d’attention allouée en faveur du contexte ACV, un premier indice de la prépondérance de la situation sur l’expérience utilisateur. La variable compréhension mutuelle montre ici des résultats similaires à ceux de notre première expérimentation. Les trois tests statistiques proposant les comparaisons des contextes SCV et ACV (globaux, avec et sans la présence du cube) montrent une évaluation supérieure de la compréhension mutuelle dans les contextes ACV, fait que nous pouvons expliquer par la nature de l’interaction sociale vocale dont le but premier est de permettre la transmission d’un message et la compréhension de ce dernier par notre interlocuteur, plus facilement que par l’entremise du comportement ou du paralangage. Il en va de même pour l’interdépendance comportementale pour laquelle nous avons révélé le même type de différence entre

contextes SCV et ACV, là encore en faveur du contexte ACV. Le fait de communiquer vocalement facilite le passage de l'information et autorise donc la transmission d'ordres ou de conseils influençant le comportement. A contrario de ces différences entre contextes vis-à-vis de la compréhension mutuelle et de l'interdépendance comportementale, dans le cas de l'assistance mutuelle, la différence (non significative) est très ténue. Ainsi les participants dans le contexte ACV l'ont évalué en moyenne à 4,5 sur 5 et à 4 pour le contexte SCV. Bien que la coopération ne demande théoriquement pas d'aide pour accomplir la tâche dans son intégralité, nous avons fait en sorte de créer une intrication forte entre les deux parties de la construction (l'un étant assigné à la construction du motif de fond gauche et à la moitié de statue droite, et inversement). Cette conception permet un travail divisé au début de l'expérimentation mais, plus la construction avance, plus les deux participants doivent faire coïncider leurs efforts, et doivent faire appel à la communication vocale ou à la translucence. Ce dont attestent les résultats de l'évaluation de la variable assistance mutuelle. Nous discuterons plus loin dans ces lignes l'impact de translucence sur l'efficacité dans le contexte SCV, en la comparant avec le contexte ACV.

Dans notre questionnaire, une question évaluait l'empathie en demandant aux participants s'ils avaient perçu les émotions de leur partenaire. Une différence statistique significative a été relevée entre les contextes SCV-SC et ACV-SC. Sur cette question, étant donnés les contextes comparés, il est évident que le cube émotionnel n'est en rien impliqué. De plus, la comparaison entre les contextes SCV et ACV globaux, comprenant donc un plus grand nombre d'évaluations, n'a pas montré de différence statistique significative. Néanmoins, si nous devons donner une hypothèse quant à la différence entre les contextes SCV-SC et ACV-SC, il semble pertinent de dire que la communication vocale par les variations de l'intonation, du débit, ou des termes utilisés, offre une compréhension des émotions comme nous avons pu le décrire au cours de notre état de l'art (voir section 2.1.1.1).

Vis-à-vis du *flow* psychologique, et selon les trois variables évaluées, niveau de challenge perçu, concentration et amusement, aucune différence significative n'a été relevée. Lors de notre première expérimentation, la différence au niveau de l'amusement entre les contextes ACV et SCV était déjà faible, bien que significative. Ici la différence est encore plus ténue et étonnamment, si nous devons comparer les moyennes obtenues ici, c'est dans le contexte SCV que l'amusement présente une meilleure évaluation moyenne (SCV 4,85 sur 5 et ACV 4,46). Nous pensions pourtant que, la tâche requérant moins d'interactions entre participants, ceux immergés SCV auraient tendance à éprouver une forme de solitude, voir d'ennui, et ce n'est semble-t-il pas le cas. Cela atteste encore une fois de l'importance de l'activité proposée, communication vocale ou non, qui semble en définitive plus importante qu'un fort sentiment de présence sociale dans le contexte que nous avons mis en place. Peut-être avons-nous fait trop efforts pour rendre la tâche réalisable et satisfaisante dans les deux contextes. Mais proposer une activité que nous aurions sue comme ayant une tendance ennuyeuse sans communication verbale, n'aurait-elle pas trop orienté les résultats, entamant ainsi l'égalité de condition que requiert l'étude empirique ?

Comme pour faire écho aux constats fait vis-à-vis du *flow*, l'évaluation du sentiment de présence

virtuelle n'a pas montré non plus de différence statistique significative. Les variations entre contextes sont très faibles (inférieure à 0,15 sur 5 en moyenne), et ce pour toutes les variables évaluées (présence virtuelle, vivacité sensorielle, degré de contrôle et concentration). Néanmoins, ces évaluations indiquent un très bon niveau de présence virtuelle chez les utilisateurs, nous pouvons donc en tirer un constat positif mais qui néanmoins n'apporte pas d'éléments de réponse à notre hypothèse. La configuration mise en place, le trio *Minecraft*, lunettes 3D et manette de jeu, permet d'atteindre un haut niveau de présence virtuelle malgré le manque de réalisme visuel du logiciel utilisé, et le fait que les lunettes 3D n'occulent pas le monde réel. Il faut aussi prendre en compte l'évaluation du *flow* que nous savons comme incident sur le sentiment de présence virtuelle : l'activité proposée a beaucoup plu aux participants favorisant ainsi leur engagement et par corrélation leur sentiment de présence virtuelle.

Nous avons, comme lors de notre première expérimentation, inclus dans notre questionnaire des questions permettant le recueil d'informations vis-à-vis des interactions sociales vocales et plus précisément des thèmes ayant été abordés par les participants au cours de leur immersion. Précédemment, nous ne pouvions en aucun cas mener de comparaison sur la base de ces données, mais dans le contexte de cette seconde expérimentation, nous étions en mesure de comparer les contextes ACV-SC et ACV-AC. Grâce au cube émotionnel, nous pensions voir une différence quant à l'occurrence du thème état émotionnel, néanmoins ce ne fut pas le cas. Nous pouvons y voir une confirmation du fonctionnement contestable de notre cube émotionnel, ou le fait que, même si le cube a parfaitement fonctionné, il n'est pas forcément simple d'échanger sur les émotions et plus encore quand nous sommes immergés dans un environnement virtuel avec un inconnu. Deux thèmes ont davantage occupé les discussions entre participants dans le contexte comportant le cube émotionnel : l'environnement virtuel et les modalités d'interaction. Si nous nous référons à ces sujets, il est très difficile d'extraire un quelconque constat vis-à-vis de l'implication du cube émotionnel dans ces différences. En effet les deux sujets étant très factuels, il semble probable qu'une différence d'expérience ou de facilité d'utilisation chez les participants ait engendré ces différences. Comme précédemment, nous constatons que le thème le plus abordé fut la tâche à accomplir, ce qui nous apparaît comme normal étant donné que les participants avaient pour objectif premier la complétion de cette dernière.

Abordons maintenant l'efficacité perçue et objective de nos participants. Cette seconde expérimentation, par le passage d'une tâche collaborative à coopérative, était en partie conçue afin de confronter les résultats avec ceux de notre première expérimentation, qui montraient une différence importante entre l'efficacité objective atteinte dans les deux contextes, cela avec une nette supériorité du contexte ACV. La question était alors posée de savoir si, dans une tâche requérant moins d'entente entre les participants, les conseils et autres encouragements émis vocalement suffiraient à atteindre une efficacité supérieure. Du point de vue des participants, et donc de l'évaluation subjective de leur efficacité, aucune différence statistique significative n'a été relevée. Et pour ce qui est des observations que nous avons menées, nous avons tout d'abord mis en exergue une influence de l'expérience préalable de *Minecraft* sur l'efficacité. Néanmoins, étant donnée la répartition relativement homogène de nos échantillons quant à ce facteur et l'absence d'interaction de nos contextes et de l'expérience préalable de *Minecraft* sur l'efficacité, nous pouvons ignorer ce constat pour nous intéresser uniquement à l'impact de nos

contextes sur l'efficacité. Une différence significative a été relevée vis-à-vis du temps passé avant la pause du premier bloc, ce temps étant plus long dans le contexte ACV. Cela aurait pu s'expliquer par le fait que les participants pouvant communiquer vocalement ont pris le temps de se concerter sur la stratégie à adopter avant de démarrer. Toutefois à l'étude des moyennes, on constate que la différence réelle n'est que de cinq secondes. Une autre différence statistique significative est apparue vis-à-vis du temps de complétion de la tâche que nous estimons être un très bon indicateur de la performance. Toutefois, cette différence fut relevée entre les contextes ACV-SC et ACV-AC, il nous semble donc hasardeux de dire que la présence du cube émotionnel ait pu réduire le temps de complétion de la tâche. Nous considérons alors que les performances de nos participants sont très proches entre les contextes SCV et ACV, ce qui tend à confirmer notre hypothèse selon laquelle l'impact des interactions sociales vocales est soumis à la nature de la tâche proposée, la simple possibilité de promulguer des conseils ou des encouragements, ne semble pas suffisant pour améliorer les performances des participants. Il faut toutefois être vigilant quant à ce constat puisqu'il s'applique à un contexte bien précis, celui de la coopération. De plus il est fort probable que nos participants ne se soient que peu encouragés en raison de leur manque de familiarité, certains étant sans doute timides par nature, il est aussi possible que le fait de communiquer au travers de l'environnement virtuel les ait rendus plus réservés.

Vis-à-vis de la satisfaction hédonique générale, mais aussi des intentions de réutilisation (ces deux facteurs étant intimement liés), nous pensions, comme nous l'avons déjà évoqué vis-à-vis de l'amusement, que les participants SCV auraient moins apprécié l'expérience en raison de la nature de la tâche n'invitant pas à l'échange, même non verbal (par le comportement). Encore une fois, nos résultats démentent totalement ce postulat. Que ce soit pour l'évaluation du sentiment hédonique global vis-à-vis de la tâche, de la possibilité de reproduire l'expérience ou de la conseiller à une personne tierce, les moyennes des réponses dépassent 4 sur 5 (exception faite d'un 3,92 vis-à-vis de l'intention de réutilisation pour le contexte SCV). Il semble donc que, communication verbale ou non, l'expérience utilisateur proposée par notre expérimentation était très satisfaisante.

Revenons maintenant à notre hypothèse :

H.1 : *Les interactions sociales vocales ont un impact bénéfique sur l'expérience utilisateur en 3D-IVE.*

Si nous nous référons aux interprétations que nous avons pu faire dans cette section, les facteurs importants de l'expérience utilisateur, le *flow* psychologique, la présence virtuelle et l'efficacité, ne semblent que peu ou pas influencés par la communication vocale, dans la situation d'une tâche coopérative. Bien sûr, le manque de résultats qualitatifs nous interdit d'affiner ces résultats et nous invite à la réserve. Néanmoins, il semblerait que les interactions sociales vocales n'aient un impact sur l'expérience utilisateur que dans une situation où elles sont requises afin d'atteindre le but fixé. Évidemment, ce but peut être uniquement hédonique ou social, ce qui expliquerait pourquoi nous trouvons de tels résultats lorsque, dans le contexte du jeu vidéo, certaines études montrent que les interactions sociales mènent à plus d'amusement et plus de satisfaction hédonique par la mise en valeur de nos qualités au travers de l'autre (Cole et Griffiths, 2007 ; Ducheneaut *et al.*, 2006). Cet état de fait souligne bien la

dépendance à la situation de l'expérience utilisateur et la complexité de son étude, confirmant ainsi notre seconde hypothèse (**H.2**). Nous sommes néanmoins satisfaits des résultats obtenus, ces derniers étant en contradiction avec les conceptions que nous pouvions nous en faire, montrant ainsi l'intérêt de nos hypothèses pouvant sous certains aspects sembler évidentes. L'objectif de notre expérimentation complémentaire, grâce aux groupe de discussions, est apporter de nouveaux éléments nous permettant de mieux comprendre les différences de résultats entre nos deux expérimentations.

4.4 Expérimentation complémentaire

4.4.1 Objectifs de l'expérimentation

L'objectif principal de notre troisième expérimentation est de compléter les deux expérimentations quantitatives précédentes par l'obtention de résultats qualitatifs au travers de groupes de discussions. Ainsi, nous avons modifié notre protocole expérimental de manière à ce que les participants soient en mesure de comparer eux-mêmes les deux contextes expérimentaux (avec et sans communication vocale) ainsi que les deux environnements virtuels (et tâches) précédemment utilisés (au regard des résultats précédents, nous n'avons pas ici réutilisé le cube émotionnel). Nous avons alors défini trois échantillons de participants : le premier vivrait l'expérience dans l'environnement de la première expérimentation (construction verticale) et accomplirait deux sessions, l'une dans le contexte ACV puis SCV. Le second échantillon serait immergé dans l'environnement de l'expérimentation 2 (construction horizontale) et de la même manière, ferait deux sessions, une pour chaque contexte. Pour finir, le troisième échantillon ferait lui quatre sessions, une pour chaque environnement virtuel et dans ces deux cas, pour chaque contexte. Il a été décidé de toujours débiter par le contexte ACV car, bien que ce choix puisse sembler faciliter la tâche des participants, des tests préliminaires ont montré que si les participants vivaient l'expérience une fois sans pouvoir communiquer verbalement, ils étaient significativement moins enclins à communiquer lors de la seconde session. En raison de son protocole différent, cette expérimentation est considérée comme complémentaire aux deux précédentes, avec pour objectif une mise en perspective des résultats antérieurs par une approche qualitative.

4.4.2 Environnement virtuel

Dans le cadre de cette expérimentation complémentaire, les deux environnements virtuels précédemment décrits pour les expérimentations une et deux ont été utilisés. Pour plus de précisions, se référer aux sections concernées (sections 4.2.2 & 4.3.2).

4.4.3 Participants

Les participants ($N = 24$; 3 femmes, 21 hommes) étaient âgés de 21 à 24 ans ($M = 22,25$; $ET = ,897$). Ici encore, la population était homogène, excluant les biais liés à l'âge (Siriaraya et Siang Ang,

2012). Les participants étaient des étudiants de Master au sein de l'*École Nationale Supérieure d'Arts et Métiers* ; ils ont été invités à participer à cette expérimentation dans le cadre de deux semaines d'initiation à la recherche au cours de leur seconde année de cycle Master. Seulement un participant a dit ne pas du tout jouer aux jeux vidéo, le reste des participants atteste jouer aux jeux vidéo à diverses fréquences (27,3% des participants jouent moins de 5 heures par semaine, 22,7% jouent de 5 à 15 heures par semaine et 50% jouent plus de 15 heures par semaine). Il leur a aussi été demandé leur familiarité avec le jeu en mode multijoueur, là encore un seul participant a dit ne pas du tout jouer au jeux vidéo en ligne (un autre n'a pas répondu à cette question), le reste des participants atteste jouer en ligne à diverses fréquences (47,8% des participants moins de 5 heures par semaine, 26,1% de 5 à 15 heures par semaine et 26,1% plus de 15 heures par semaine) nous pouvons donc dire, que nous étions en présence d'une population de joueurs plutôt expérimentés tant en solo qu'en ligne. Afin d'être sûr que la plupart des participants seraient à l'aise avec les modalités d'interactions avec l'environnement, il leur a été demandé sur quel type de plateforme ils jouaient : la plupart d'entre eux jouent sur ordinateurs (91,7%), une proportion bien moindre sur consoles (25%), néanmoins tous les participants savaient comment utiliser une manette de jeu.

4.4.4 Mesures

Afin, ici encore, de croiser les données (quantitatives et qualitatives), nous avons produit un questionnaire court en adéquation avec le guide utilisé pour la conduite des groupes de discussion. Ce questionnaire s'est vu décliné en deux versions, en effet nous avons pour cette expérimentation trois groupes (ou échantillons) de participants, chacun composé de 8 participants soit 4 binômes par échantillon. Les deux premiers groupes étaient immergés soit dans l'environnement issu de l'expérimentation 1 (tâche collaborative & construction verticale) soit dans celui issu de l'expérimentation 2 (tâche coopérative & construction horizontale). Ces deux groupes ont été soumis aux deux contextes (SCV & ACV) et ont rempli la même version du questionnaire présentant pour chaque question deux échelles de Likert, une pour chaque contexte (SCV et ACV). Le troisième groupe a lui été immergé successivement dans les deux environnements virtuels et pour chacun d'eux dans les deux contextes, en conséquence, la seconde version du questionnaire comportait quatre échelles de Likert pour chaque question. Basées sur les modèles d'expérience utilisateur de Wu *et al.* (2009) et Pallot *et al.* (2013), les mesures mises en place suivent la taxonomie de facteurs proposés par ces modèles. Questionnaire et guide des groupes de discussion sont disponibles en annexes D.2 et D.3).

Expérience (Utilisateur) Immersive ► Processus Expérientiel

► **Présence Sociale :** Afin d'évaluer l'impact de nos deux contextes (SCV et ACV), ainsi que la différence entre nos environnements virtuels (pour l'échantillon 3) sur le sentiment de présence sociale, quatre questions étaient posées : deux questions portant sur la variable attention allouée au comportement et aux interactions vocales, les deux autres sur l'impact de ces dernières. Ces questions sont dérivées des facteurs attention allouée et interdépendance comportementale décrits par Biocca *et al.* (2001) dans leurs travaux sur le « *Networked Minds Questionnaire* » et sur la présence sociale.






Expérience (Utilisateur) Immersive ▶ Processus Rationnel

▶ **Flow Psychologique :** Afin d'évaluer l'impact de nos deux contextes (SCV et ACV), ainsi que la différence entre nos environnements virtuels (pour l'échantillon 3) sur la variable amusement du *flow* psychologique, nous avons ajouté à notre questionnaire deux questions : l'une, basée sur les constats de Danckert et Allman (2005) et de Ijsselsteijn *et al.* (2001) vis-à-vis de la distorsion du temps ressentie en fonction du niveau d'amusement ou d'ennui, et l'autre question requérant l'évaluation de manière explicite du niveau d'amusement ressenti vis-à-vis de cette expérimentation.

▶ **Téléprésence :** Afin d'évaluer l'impact de nos deux contextes (SCV et ACV), ainsi que la différence entre nos environnements virtuels (pour l'échantillon 3) sur le sentiment de présence virtuelle, nous avons établi deux questions : la première évalue le sentiment de présence virtuelle et est extraite comme précédemment du questionnaire IPQ¹¹ de (Steuer, 1992), la seconde rédigée, par nos soins, requiert l'évaluation du sentiment d'incarnation du participant par son avatar (voir incarnation en section 2.2.2.1).

▶ **Efficacité :** Afin d'évaluer l'impact de nos deux contextes (SCV et ACV), ainsi que la différence entre nos environnements virtuels (pour l'échantillon 3) sur l'efficacité perçue des participants, nous avons ajouté à notre questionnaire deux questions héritées de nos précédents questionnaires.

Lors des groupes de discussion, nous sommes revenus sur les facteurs et variables précédemment cités et avons ajouté un temps de discussion vis-à-vis du sentiment hédonique global généré par l'expérience vécue (au travers de l'adoption technologique). La liste suivante présente succinctement les thèmes abordés et leur ordre ainsi que le temps alloué aux discussions sur ce thème. Le détail de ce guide est présenté en annexe D.3) :

1. Le phénomène de translucence :  environ 15 minutes ;
2. l'efficacité :  environ 15 minutes ;
3. le sentiment de présence virtuelle :  environ 15 minutes ;
4. la qualité de service :  environ 5 minutes ;
5. l'acceptation de la technologie :  environ 5 minutes.

4.4.5 Équipement des participants

Comme lors de la seconde expérimentation, nous avons ici utilisé des lunettes 3D actives de la marque *Nvidia* ainsi que des casques audio équipés de microphones pour la communication vocale. Là encore, pour les raisons précédemment évoquées d'accessibilité, nous avons utilisé des manettes de jeu (manettes de jeu filaire de *Microsoft Xbox 360* pour *Windows*).

4.4.6 Procédure

Il nous était ici impossible de contrôler les échanges préalables entre les participants comme nous avons pu le faire précédemment, étant donné que ces derniers faisaient tous partie de la même pro-

motion d'étudiants. Aucune information n'a été donnée aux participants sur les objectifs réels de l'expérimentation afin d'éviter toute influence potentielle sur leurs réponses aux questionnaires ou sur leur état d'esprit durant l'immersion. Un questionnaire pré-expérimentation recueillant les informations présentées dans la section 4.4.3 ainsi que le niveau d'expérience sur *Minecraft* a été rempli par tous les participants quelques jours avant l'expérimentation.

Avant le début de l'immersion, les participants ont été informés du type de questionnaire qu'ils allaient devoir remplir (basé sur des échelles de Likert) ainsi que du fait qu'un groupe de discussion se tiendrait le soir même ou le lendemain de leur passation. Ils ont ensuite reçu les explications d'usage en accord avec la procédure des expérimentations 1 et 2.

Comme décrit dans la section 4.4.1, l'objectif était, au travers des groupes de discussion, d'obtenir une analyse qualitative de nos travaux afin de pouvoir les mettre en perspective avec les résultats précédents. De ce fait et comme rapidement décrit dans la section ??, au lieu de ne faire l'expérimentation que dans un seul des deux contextes (avec ou sans communication vocale), les participants ont été divisés en trois échantillons chacun composé de 8 participants soit 4 binômes par échantillon :

- **Échantillon 1** : environnement virtuel issu de l'expérimentation 1, immersion ACV puis SCV ;
- **échantillon 2** : environnement virtuel issu de l'expérimentation 2, immersion ACV puis SCV ;
- **échantillon 3** : environnement virtuel issu de l'expérimentation 1 et environnement virtuel issu de l'expérimentation 2, respectivement, immersion ACV puis SCV.

Les participants se connaissaient tous préalablement à l'expérimentation, nous les avons donc laissés choisir eux-mêmes leurs créneaux de passage, car nous ne pouvions éviter les possibles biais induits par cette familiarité interne au groupe. Ils n'ont néanmoins pas été informés sur les modalités de l'expérimentation (environnements virtuels et contextes).

Il était demandé aux participants de produire la tâche de construction correspondante à l'environnement virtuel dans lequel ils étaient immergés (voir section 4.2.6 & 4.3.6). Néanmoins, une modification a été apportée à la tâche de l'expérimentation 1 : la seconde partie de l'expérimentation a été supprimée, pour deux raisons. Comme nous n'avons pas mis en place de mesure physiologique, son intérêt devenait moindre. De plus, elle aurait risqué de parasiter les réponses lors du groupe de discussion, particulièrement lors de celui concernant l'immersion consécutive dans les deux environnements. Quel que soit l'environnement virtuel dans lequel les participants étaient immergés, la durée était la même, trente minutes à partir de l'apparition des matériaux dans l'inventaire. À la fin de l'immersion, chaque participant a rempli un questionnaire post-expérimentation à échelles de Likert sur cinq points, composé de dix questions évaluant les facteurs et variables présentés en section 4.4.4. Les questionnaires post-expérimentation sont disponibles en annexe D.2.

4.4.7 Résultats

Nous avons, en vue de l'analyse quantitative de nos questionnaires, conduit des tests de normalité (Kolmogorov-Smirnov) sur chacune des variables obtenues afin de savoir si nous pouvions utiliser ou

non des tests paramétriques. Aucune des variables testées ne suivait une distribution normale. Ces résultats sont cohérents avec le nombre réduit de participants ($N = 24$) ainsi qu'avec l'utilisation d'échelles de Likert à cinq niveaux seulement. Comme pour nos deux expérimentations précédentes, ce choix était motivé par les besoins d'une bonne compréhension et appropriation du questionnaire par les participants. Par conséquent, nous nous sommes reposés sur la famille des tests non paramétriques (Pallant, 2010). Au regard de nos conditions expérimentales et selon l'étude à réaliser, nous nous sommes appuyés sur le test de rang signé de Wilcoxon pour échantillons liés, sur le U-test de Mann-Whitney pour échantillons indépendants ainsi que sur le test de Friedman. Pour établir la signification statistique de nos résultats, nous avons utilisé le seuil de probabilité conventionnel soit $p < 0,05$.

Étant donné le faible nombre de participants à cette expérimentation, nous assumons le fait qu'une analyse statistique des résultats n'apporte pas de garanties de robustesse, c'est pourquoi les résultats quantitatifs présentés ici sont à considérer comme des indications que nous pourrions mettre en perspective avec les résultats qualitatifs de nos groupes de discussion. Dans cette logique, nous présenterons en plus des différences statistiques significatives le cas échéant, les moyennes obtenues pour chaque question et dans chacune des conditions afin de mettre en exergue les orientations générales de nos participants. Enfin, nous présenterons les résultats obtenus en groupe de discussion vis-à-vis du facteur courant. Les résultats seront présentés de la manière suivante c'est à dire, pour chaque facteur évalué, les résultats seront découpés selon les échantillons de participants ainsi que selon les différents contextes de notre expérimentation. Au regard des effets d'une expérience préliminaire sur le sentiment de présence décrit par Freeman *et al.* (1999) et potentiellement sur l'efficacité, nous considérons qu'il est préférable de ne pas comparer les échantillons 1 & 2 à l'échantillon 3 ; cela implique la présentation des résultats de la manière suivante :

A) Échantillon de participants 1 & 2 :

- Comparaison entre les contextes ACV et SCV lors de l'immersion dans l'environnement virtuel issu de l'expérimentation 1 (EV.1) ou de l'expérimentation 2 (EV.2) ;
- comparaison entre EV.1 et EV.2.

B) Échantillon de participants 3 :

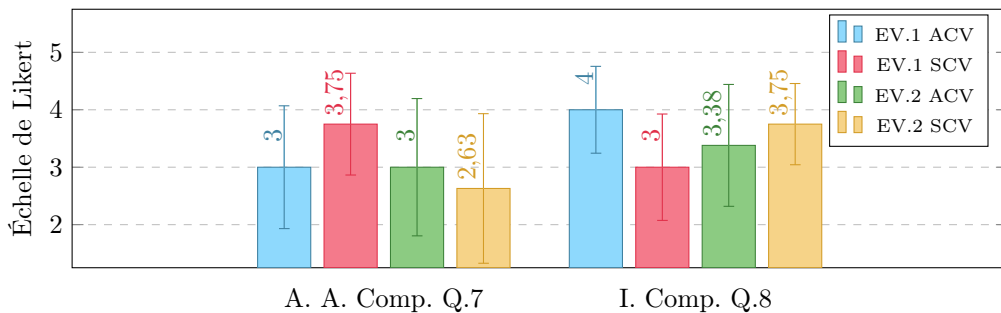
- Comparaison entre les contextes ACV et SCV lors de l'immersion dans EV.1 et dans EV.2 ;
- comparaison entre EV.1 et EV.2.

Expérience (Utilisateur) Immersive ► Processus Expérientiel

► **Présence Sociale :** Quatre questions de notre questionnaire mesuraient le sentiment de présence sociale des participants lors de l'expérimentation : deux portant sur le comportement, autrement dit une forme de communication non verbale et les deux autres sur la communication vocale. Ces deux dernières questions n'ont été posées que dans le cas où la communication vocale était autorisée lors de l'expérimentation et ne peuvent donc être utilisées que pour une comparaison entre environnements virtuels EV.1 et EV.2. Pour chacun des échantillons de participants, nous avons en premier lieu fait une comparaison entre les contextes ACV et SCV vis-à-vis de l'attention allouée au comportement et de l'impact du comportement sur les actions. Puis nous nous sommes intéressés à ces mêmes

variables (attention et impact) pour la communication vocale (l'attention allouée à la communication verbale et impact de la communication vocale sur les actions). Enfin, pour compléter notre analyse, nous avons, pour chaque échantillon et dans le contexte autorisant la communication vocale, produit une comparaison entre l'attention allouée au comportement et à la communication vocale, puis entre l'impact du comportement sur les actions et celui de la communication vocale. Cela, afin d'obtenir des indications sur lequel, du comportement ou de la communication vocale suscite une plus grande attention et entraîne un plus grand impact sur les actions, autrement dit, qui du vocal ou du non verbal suscite une plus grande attention et un plus grand impact.

Vis à vis de l'**attention allouée au comportement** et de l'**impact du comportement sur les actions**, deux questions étaient posées : « *Avez-vous fait preuve d'une attention particulière vis-à-vis du comportement de l'autre personne (déplacement, méthode utilisée pour poser les blocs, etc...) :* » (Pour 1 = Je n'ai pas été très attentif à l'autre & 5 = J'ai été très attentif à l'autre) — question 7 dans le questionnaire — et : « *Comment estimeriez-vous l'impact de ce comportement sur vos actions :* » (Pour 1 = Cela m'a perturbé ; 3 = Aucun impact & 5 = Cela m'a aidé, inspiré) — question 8 dans le questionnaire. Un test des rangs signés de Wilcoxon pour échantillons liés n'a révélé aucune différence statistique significative ni entre nos deux contextes dans aucun des échantillons de participants ni entre les deux environnements virtuels dans le cadre de l'échantillon 3.



A. A. Comp. : attention allouée au comportement ;

I. Comp. : impact du comportement ;

FIGURE 4.53 – Moyennes des évaluations de l'attention allouée au comportement et de l'impact du comportement pour les échantillons 1 & 2.

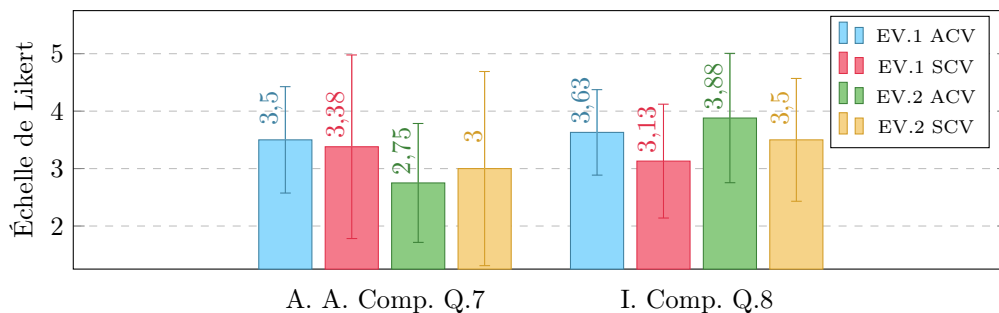
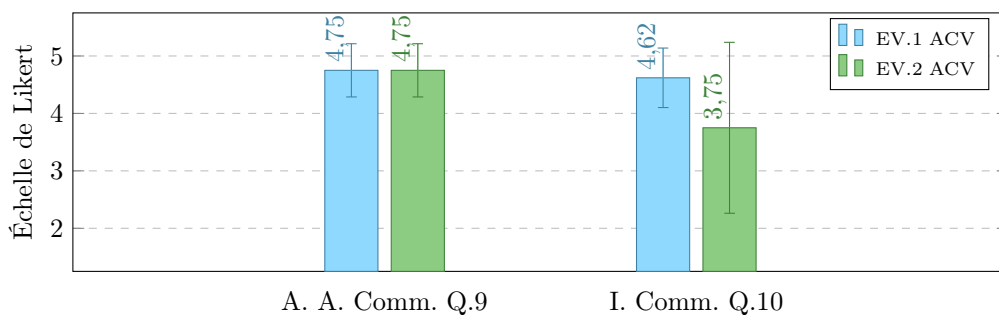


FIGURE 4.54 – Moyennes des évaluations de l'attention allouée au comportement et de l'impact du comportement pour l'échantillon 3.

Etant donnée la variabilité des résultats présentés en figures 4.53 et 4.54, aucune tendance ne se

dessine réellement, il est donc difficile ici de déduire un quelconque constat sur la base de ces moyennes.

Pour l'**attention allouée à la communication vocale** et l'**impact de la communication vocale sur les actions**, deux questions étaient posées : « *Avez-vous été attentif à ce qu'a pu vous dire l'autre personne :* » (Pour 1 = Je n'ai pas été très attentif à nos conversations & 5 = J'ai été très attentif à ce qui se disait) — question 9 dans le questionnaire — et : « *Comment estimeriez-vous l'impact de ces échanges sur vos actions :* » (Pour 1 = Cela m'a perturbé ; 3 = Aucun impact & 5 = Cela m'a aidé, inspiré) — question 10 dans le questionnaire. Ces deux questions n'étant posées que dans le contexte supportant la communication vocale, un test de rang signé de Wilcoxon pour échantillons liés a été mené dans le cadre de l'échantillon 3, mais n'a révélé aucune différence statistique significative entre les deux environnements virtuels.



A. A. Comm. : attention allouée à la communication vocale ;

I. Comm. : impact de la communication vocale.

FIGURE 4.55 – Moyennes des évaluations de l'attention allouée et de l'impact de la communication vocale pour les échantillons 1 & 2.

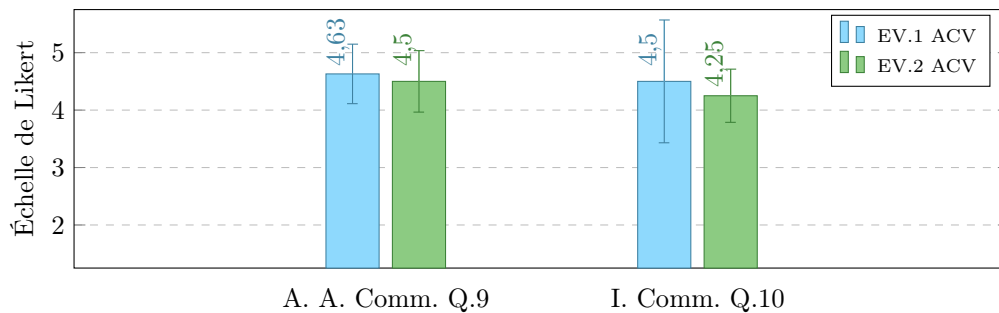


FIGURE 4.56 – Moyennes des évaluations de l'attention allouée et de l'impact de la communication vocale pour l'échantillon 3.

En accord avec les figures 4.55 et 4.56, et malgré l'absence de différences statistiques significatives, nous pouvons constater que l'impact de la communication vocale sur les actions des participants est légèrement mieux évalué dans l'EV.1 pour tous les échantillons de participants.

Nous comparons à présent l'attention allouée au comportement et son impact à l'attention allouée à la communication vocale et son impact, ici la comparaison se base sur les moyennes des évaluations faites par les participants.

Afin de savoir lequel de la communication vocale ou du comportement a le plus suscité d'attention

chez les participants, les moyennes des réponses ont été calculées pour chaque échantillon et environnement virtuel vis à vis des questions : « *Avez-vous fait preuve d'une attention particulière vis-à-vis du comportement de l'autre personne (déplacement, méthode utilisée pour poser les blocs, etc...) ?* » — question 7 et « *Avez-vous été attentif à ce qu'a pu vous dire l'autre personne :* » — question 9.

- **Échantillon 1 :** Attention allouée à la communication vocale $M = 4$ et attention allouée au comportement $M = 3$.
- **Échantillon 2 :** Attention allouée à la communication vocale $M = 4,75$ et attention allouée au comportement $M = 3$.
- **Échantillon 3 - EV.1 :** Attention allouée à la communication vocale $M = 4,63$ et attention allouée au comportement $M = 3,50$.
- **Échantillon 3 - EV.2 :** Attention allouée à la communication vocale $M = 4,50$ et attention allouée au comportement $M = 2,75$.

Afin de savoir lequel de la communication vocale ou du comportement a le plus d'impact sur les actions des participants, les moyennes des réponses ont été calculées pour chaque échantillon et environnement virtuel vis à vis des questions : « *Comment estimeriez-vous l'impact de ce comportement sur vos actions :* » (Pour 1 = Cela m'a perturbé ; 3 = Aucun impact & 5 = Cela m'a aidé, inspiré) — question 8 et « *Comment estimeriez-vous l'impact de ces échanges sur vos actions :* » (Pour 1 = Cela m'a perturbé ; 3 = Aucun impact & 5 = Cela m'a aidé, inspiré) — question 10.

- **Échantillon 1 :** Impact de la communication vocale $M = 4,62$ et impact comportement $M = 4,75$.
- **Échantillon 2 :** Impact de la communication vocale $M = 3,75$ et impact comportement $M = 3,38$.
- **Échantillon 3 - EV.1 :** Impact de la communication vocale $M = 4,5$ et impact comportement $M = 3,63$.
- **Échantillon 3 - EV.2 :** Impact de la communication vocale $M = 4,25$ et impact comportement $M = 3,88$.

Nota : Dans les histogrammes suivants, l'ordre numérique des questions a été modifié afin de faciliter la comparaison des questions traitant de l'attention allouée et de l'impact de la communication vocale. Ainsi vous trouverez la question 7 au côté de la question 9 et la question 8 au côté de la question 10.

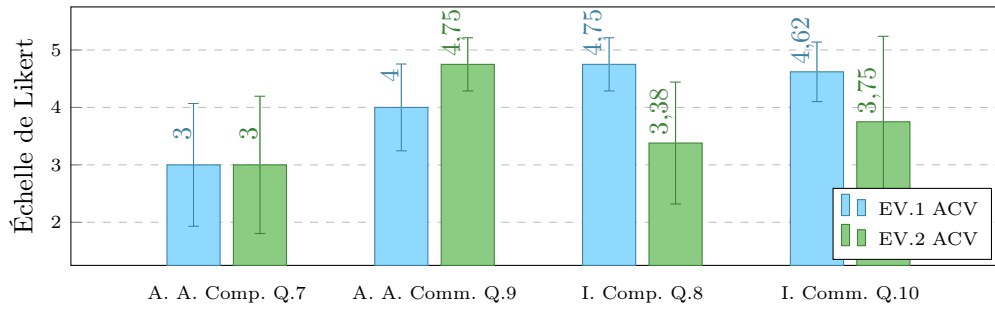


FIGURE 4.57 – Moyennes des évaluations de l’attention allouée et de l’impact de la communication vocale et du comportement pour les échantillons 1 & 2 pour comparaison.

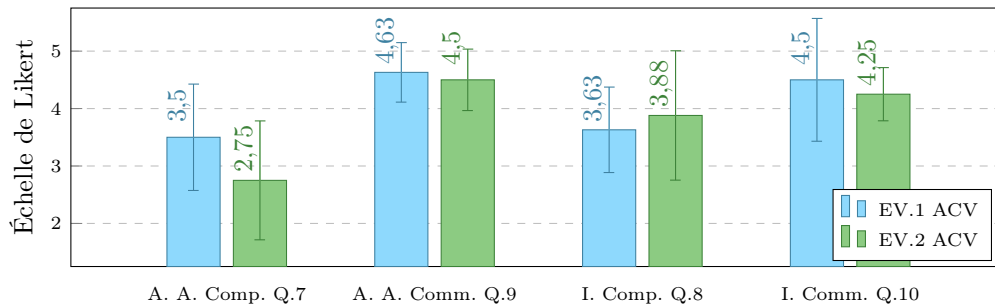


FIGURE 4.58 – Moyennes des évaluations de l’attention allouée et de l’impact de la communication vocale et du comportement pour l’échantillon 3 pour comparaison.

On constate globalement que la communication vocale suscite une plus grande attention chez les participants que le comportement. Dans le cadre de nos trois échantillons, les impacts de la communication vocale et du comportement sont très proches et relatifs aux deux environnements virtuels, difficile donc d’attribuer un impact plus fort à l’un ou à l’autre. Cela permet néanmoins de constater que c’est dans l’EV.1 (tâche collaborative) que l’impact de la communication vocale et du comportement est généralement le mieux évalué.

Lors des groupes de discussion, nous nous sommes particulièrement intéressés, en plus de la présence sociale, à un phénomène non verbal lié à la présence de l’autre et à son impact sur le comportement. Le phénomène de « translucence » (voir section 2.1.2.3), ou l’adaptation vis-à-vis du comportement d’une tierce personne (ici le partenaire d’expérience).

- Certains participants ont trouvé que l’expérience préalable dans le contexte supportant la communication vocale les avait aidé dans leur adaptation au comportement de leur partenaire lors de la seconde session. En effet, la première session leur a permis de saisir dans une certaine mesure la manière de penser et d’agir de leur partenaire vis-à-vis de la tâche, facilitant ainsi leur compréhension du comportement de ce dernier lors de la session sans communication vocale (*deux occurrences dans l’échantillon 1*).
- Les participants s’accordent sur le fait que s’adapter au comportement de l’autre sans communication vocale n’est pas chose aisée mais que, même si cela demande des efforts supplémentaires, ils sont dans la plupart des cas arrivé à leurs fins (*consensus dans les 3 échantillons*).

- Certains participants n’ont pas du tout prêté attention au comportement de leur partenaire, ils ont attesté du fait que cela était lié à la confiance qu’ils leur portaient quant à leurs compétences dans *Minecraft* (deux occurrences dans les échantillons 2 & 3).
- Dans le contexte SCV et vis-à-vis du type de tâche proposé, les participants ont constaté que la translucence se mettait en place de manière naturelle non pas vis-à-vis du comportement direct de l’autre avatar mais plutôt vis-à-vis de ses actions passées en d’autres termes vis-à-vis des blocs déjà posés et surtout de la méthode employée pour le faire. À propos de la tâche de l’EV.2 (collaboration) : *« J’ai vu qu’il avait commencé à faire le tour de ma partie du personnage avec les blocs dédiés au fond. J’ai trouvé ça pertinent et ai donc fait de même, ce qui nous a facilité la tâche pour le personnage et je pense a réduit le nombre de nos erreurs potentielles »* (trois occurrences dans l’échantillon 2).
- Outre cette adaptation, les participants ont constaté que leur attention (dans le contexte SCV) se tournait vers leur partenaire dès que celui-ci ne produisait plus le moindre son, plus de son de pose de bloc en l’occurrence. Ils se détournaient à ce moment-là de la tâche, allaient voir de plus près afin d’essayer de comprendre pourquoi ce dernier n’émettait plus de son (trois occurrences dans l’échantillon 1). Conformément à cela, l’émission de sons par le partenaire a eu tendance à rassurer les participants en leur faisant comprendre que si leur partenaire avançait sur la tâche c’est que tout devait bien aller pour lui (voir citation dans le paragraphe suivant) (deux occurrences dans l’échantillon 1 & 3).
- Ces émissions sonores non verbales sont aussi fortement impliquées dans le sentiment de présence sociale lorsqu’il est impossible de communiquer vocalement. Dans notre contexte d’immersion dans *Minecraft*, ces indications de présence non verbales sonores sont par exemple le bruit que fait un participant en posant, en détruisant un bloc ou en chutant. *« Lorsque j’entendais des bruits m’indiquant qu’il était en train de poser des blocs, cela voulait pour moi dire qu’il allait bien. En revanche, c’est lors des silences que je me disais qu’il rencontrait sans doute un problème ! J’allais donc voir et essayer de comprendre ce qu’il se passait »*. (deux occurrences dans les échantillons 1 & 2.)
- Dans le contexte ACV, il a été constaté un effacement de la personne au profit de l’avatar, il est intéressant de noter que la première occurrence de ce constat a été faite par un participant ne pratiquant que peu le jeu en ligne : *« Au bout de cinq minutes de jeu, j’avais envie de l’appeler Voleur2Poules²⁰ et non plus Michel. C’est uniquement parce que je le connais bien que j’ai continué à l’appeler par son prénom »*. Les joueurs très familiers du jeu en ligne ont confirmé ce constat en disant ne pas l’avoir évoqué, car cela leur semblait absolument évident dans le cadre d’un jeu multijoueur (trois occurrences dans les échantillons 2 & 3).
- Les participants ont été interrogés sur les différences qu’ils avaient ressenties vis-à-vis de l’attention allouée au comportement ainsi qu’à la communication vocale et à l’impact de ces derniers

20. Lors de nos expérimentations, nous avons utilisé les licences *Minecraft* personnelles du doctorant afin de ne pas en faire acheter par le laboratoire, *Voleur2Poules* et *Voleur3Poules* sont les pseudonymes liés à ces licences, ils étaient affichés au-dessus de la tête des avatars. L’utilisation d’un mod pour cacher les pseudonymes a été tentée mais entraînait en conflit avec un autre mod important pour le fonctionnement de l’expérimentation et a donc dû être abandonnée.

sur leur propre comportement. Il leur est apparu comme évident que c'est quand ils ne pouvaient pas se parler qu'ils faisaient le plus attention au comportement de l'autre de manière générale (*deux occurrences dans les échantillons 1 & 3*). Néanmoins, certains participants sont allés à l'encontre de ce constat arguant que dans leur cas ils avaient alloué plus d'attention au comportement de l'autre lorsqu'ils pouvaient se parler, car il était possible de joindre l'action à la parole par exemple en se positionnant à un emplacement précis et en disant au partenaire de venir poser tel ou tel bloc à cet endroit (*deux occurrences dans les échantillons 1 & 3*). Pour expliquer le premier constat ci-dessus, les participants ont indiqué que lorsqu'il est possible de se parler, on définit des rôles et on est alors en droit d'attendre que le partenaire fasse le travail qui lui est attribué. Lorsque la communication vocale n'est pas autorisée, cette mise en place stratégique étant impossible, il est nécessaire de regarder de temps en temps ce que fait le partenaire afin d'adapter son comportement vis-à-vis de ce dernier pour être efficace (*trois occurrences dans les échantillons 1 & 2*). D'un autre côté, aller voir ce que fait son partenaire permet de se rassurer soit même sur l'avancement de la tâche et sur le fait que son partenaire va bien, qu'il ne rencontre pas de problèmes majeurs demandant une aide quelconque (*deux occurrences dans les échantillons 1*). Les participants se sont aussi accordés sur le fait que, lorsqu'il est possible de se parler, ils avaient tendance à travailler sur la même partie ou, dans le cadre de l'EV.2, sur des parties très proches l'une de l'autre. Comme s'ils essayaient de rester à portée de voix (or aucun système de spatialisation du son n'a été mis en place). À l'inverse, lorsque la communication vocale n'était pas possible, s'ils se retrouvaient à travailler proches l'un de l'autre, cela était fortuit ou du moins lié à la tâche et non pas à une volonté de leur part (*quatre occurrences dans les échantillons 1,2 & 3*).

- Les participants ont été invités à partager la nature des interactions vocales émises durant l'expérimentation. Il en ressort qu'ils ont communiqué quasi exclusivement à propos des tâches à accomplir (*cinq occurrences dans les échantillons 1 & 3*). C'est donc la planification et les directives (afin de corriger les erreurs par exemple) qui ont le plus animé les conversations (*trois occurrences dans les échantillons 2 & 3*) ainsi que de manière ponctuelle des sollicitations visant à savoir si tout se passait bien, si aucune difficulté n'est rencontrée. Ces sollicitations furent plus présentes dans le cadre de l'immersion dans l'EV.2 donc en coopération. Les participants expliquent cela par le fait qu'en coopération, la nécessité de parler pour avancer sur la tâche est moindre, on a donc plus tendance à « meubler » les silences en demandant si tout va bien (*deux occurrences dans l'échantillon 3*).

Deux questions de notre questionnaire mesuraient la variable amusement du *flow* psychologique ressentie par les participants lors de l'expérimentation. La première au travers de la perception du temps de l'utilisateur : « *Subjectivement, comment estimeriez-vous la durée de l'expérience :* » (Pour 1 = Plutôt courte & 5 = Plutôt longue) — question 5 dans le questionnaire. La seconde requérait l'évaluation directe du niveau d'amusement éprouvé : « *Vous êtes-vous amusés durant cette expérience :* » (Pour 1 = Non, pas du tout & 5 = Oui, beaucoup) — question 6 dans le questionnaire. Un test des rangs signés de Wilcoxon pour échantillons liés a révélé une différence statistique significative entre

les contextes ACV et SCV sur l'échantillon 1 lors de l'immersion dans l'EV.1 et sur l'échantillon 3 lors de l'immersion dans les deux environnements virtuels vis-à-vis de la question 6.

- **Échantillon 1 - EV.1 :** Entre les contextes ACV ($N = 8$) et SCV ($N = 8$) : $Z = -2,121$, $p = ,034$.
- **Échantillon 3 - EV.1 :** Entre les contextes ACV ($N = 8$) et SCV ($N = 8$) : $Z = -2,428$, $p = ,015$.
- **Échantillon 3 - EV.2 :** Entre les contextes ACV ($N = 8$) et SCV ($N = 8$) : $Z = -2,428$, $p = ,015$.

Afin de savoir dans lequel des deux contextes l'amusement fut le mieux évalué par les participants, les moyennes des réponses ont été calculées pour chaque contexte. Les moyennes suivantes représentent le positionnement moyen des participants sur l'échelle de Likert à cinq niveaux :

- **Échantillon 1 - EV.1 :** ACV, $M = 4,5$ et SCV, $M = 3,75$.
- **Échantillon 3 - EV.1 :** ACV, $M = 4,25$ et SCV, $M = 2,88$.
- **Échantillon 3 - EV.2 :** ACV, $M = 4,25$ et SCV, $M = 3$.

Aucune différence statistique significative n'a été relevée vis-à-vis de la question 5 que ce soit entre les deux contextes ACV et SCV, qu'entre l'EV.1 et l'EV.2 dans le cadre de l'échantillon 3. Les figures 4.59 et 4.60 présentent l'ensemble des moyennes dans les différents contextes expérimentaux pour ces deux questions.

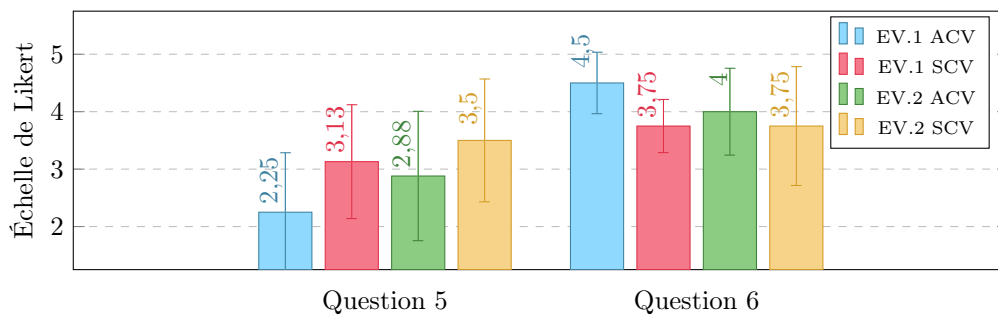


FIGURE 4.59 – Moyennes des évaluations de l'amusement pour les échantillons 1 & 2.

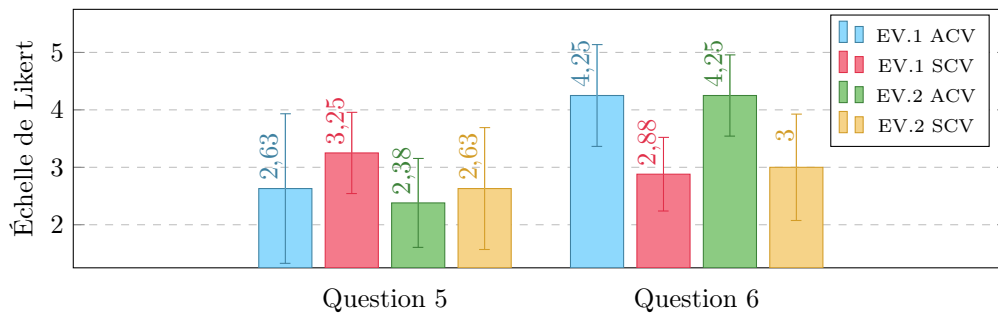


FIGURE 4.60 – Moyennes des évaluations de l'amusement pour l'échantillon 3.

On constate ici que, quel que soit l'échantillon ou l'environnement concerné, l'expérience est évaluée

comme semblant plus courte et plus amusante dans le contexte supportant la communication vocale. Dans les ensembles de données ne présentant pas de différences statistiques significatives, les différences sont plus ténues, mais présentes.

Lors des groupes de discussion, la question de savoir pourquoi les participants ont éprouvé plus d’amusement dans l’un ou l’autre des contextes ou environnements a été posée. Il en ressort que :

- En premier lieu, tous les participants attestent s’être amusés, voire s’être beaucoup amusés (*consensus global dans les échantillons 1,2 & 3*).
- Les participants s’accordent sur le fait que l’on s’amuse plus quand on peut communiquer vocalement. Ils ont néanmoins des difficultés à exprimer pourquoi de manière objective, simplement, le fait de pouvoir parler, de partager l’expérience en temps réel et de travailler ensemble les amuse (*cinq occurrences dans les échantillons 1,2 & 3*).
- Il est intéressant de noter que les participants n’ont à aucun moment dit qu’ils s’étaient moins amusés lorsqu’ils ne pouvaient pas se parler. En revanche, ils ont clairement exposé la frustration ressentie quand le besoin ou l’envie de parler émerge. Le lien négatif entre frustration et amusement a alors été évoqué (*quatre occurrences dans les échantillons 1,2 & 3*). D’un autre côté, le fait de ne pas pouvoir se parler amène de nouveaux éléments d’amusement inhérents au contexte. Par exemple, le bruit produit par la chute de leur partenaire et l’impossibilité de lui demander la raison de sa chute ou de savoir si son avatar est mort ou non les ont beaucoup fait rire. Le fait de voir leur camarade faire une erreur, de venir la corriger pour ensuite la voir reproduite de nouveau a quelque chose de ridicule et d’absurde qui génère de l’amusement autant que de la frustration. « À un moment, je me suis dit, mais il est bête ou il le fait exprès ! On va jamais s’en sortir... Mais dans ce contexte, c’était plus drôle que navrant » (*quatre occurrences dans les échantillons 1,2 & 3*).
- Un effet transversal de la tâche coopérative (EV.2) a été évoqué. Ce type de tâche répartissant le travail de manière précise donne moins l’impression de travailler à deux, et donc, en opposition avec ce que nous avons évoqué plus haut, cela ne favorise pas l’amusement (*trois occurrences dans l’échantillon 3*).
- Tout comme dans le cadre de la présence virtuelle, les participants estiment que le challenge proposé par les tâches intégrées aux environnements virtuels a eu une influence positive sur leur amusement. Cela expliquerait pour eux le fait que, malgré la frustration de ne pouvoir se parler, le fait de vouloir faire mieux dans la seconde session de chaque environnement leur a permis de tout de même s’amuser (*quatre occurrences dans les échantillons 1 & 2*).
- Les qualités inhérentes de l’environnement virtuel sont aussi perçues comme ayant un impact positif sur l’amusement. Ici, les participants ont jugé l’EV.2 comme étant plus beau, plus fini, comportant plus de fonctionnalités pour leur faciliter le travail, cela a rendu leur expérience moins complexe et donc plus amusante. « La plateforme en hauteur, c’est énorme en comparaison de l’environnement un ! Ça permet d’avoir une vision d’ensemble super facilement et en plus son usage est fun et bien pensé » (*trois occurrences dans l’échantillon 3*).

► **Téléprésence** : Deux questions de notre questionnaire mesuraient la présence virtuelle ressentie par les participants durant leur expérience. La première requérant l'évaluation directe du sentiment de présence virtuelle : « *Dans l'environnement virtuel, vous aviez le sentiment « d'être sur place » :* » (1 = Pas du tout « *sur place* » & 5 = Tout à fait « *sur place* ») — question 1 dans le questionnaire. Et la seconde portait sur l'incarnation par l'avatar : « *Avez-vous eu le sentiment d'être incarné dans votre avatar :* » (1 = Non, pas du tout & 5 = Oui, je me suis totalement transposé dans l'avatar) — question 2 dans le questionnaire. Un test des rangs signés de Wilcoxon pour échantillons liés n'a révélé aucune différence statistique significative entre nos deux contextes ACV et SCV dans aucun des échantillons de participants ni entre les deux environnements virtuels dans le cadre de l'échantillon 3.

Les figures 4.61 et 4.62 présentent les moyennes obtenues pour ces questions afin de nous permettre de juger du niveau de présence virtuelle ressenti durant l'expérimentation.

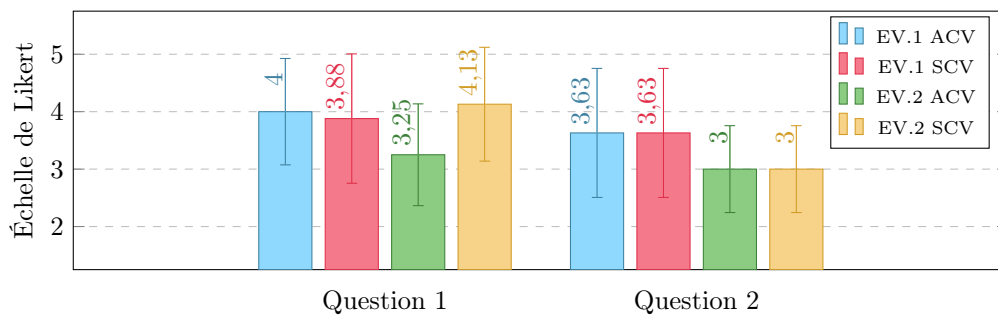


FIGURE 4.61 – Moyennes des évaluations de la présence virtuelle pour les échantillons 1 & 2.

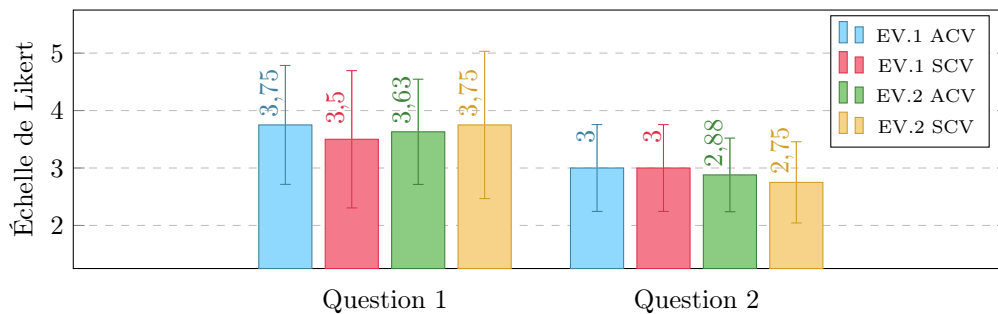


FIGURE 4.62 – Moyennes des évaluations de la présence virtuelle pour l'échantillon 3.

On constate ici que les variations dans l'évaluation moyenne du sentiment de présence virtuelle sont très faibles, voire inexistantes. Il est intéressant de noter que dans le cadre de l'EV.2 et vis-à-vis de la question 1, c'est dans le contexte SCV que les participants ont le mieux évalué le sentiment de présence virtuelle contrairement à l'EV.1 qui présente le résultat inverse avec une différence bien moindre.

Lors des groupes de discussion, les participants ont émis les constats suivants vis-à-vis du sentiment de présence virtuelle ressenti (on notera que certains de ces constats sont contradictoires, ce qui résulte des différentes personnalités ou expériences préliminaires du jeu vidéo ou de la réalité virtuelle) :

— Le sentiment de présence virtuelle est plus fort lorsque l'on ne communique pas, le fait de parler

avec l'autre personne entraînant un décrochage vis-à-vis de la réalité virtuelle en apportant un élément du réel (*six occurrences dans les échantillons 1 & 2*).

- Le fait de parler fait que l'on se soucie plus de l'autre, de son avancement sur la tâche demandée, des potentielles difficultés qu'il peut rencontrer. Cela tend à réduire le sentiment de présence puisque l'on pense à autre chose, on est moins concentré. Le même constat a aussi été fait, mais à l'inverse en arguant que l'on est plus présent lorsque l'on ne peut pas se parler, car cela accroît la concentration sur la tâche et ainsi le sentiment de présence virtuelle (*trois occurrences dans l'échantillon 1*).
- A contrario, certains participants trouvent que le sentiment de présence est plus fort lorsque l'on peut communiquer, le fait de se parler et de regarder l'avatar de l'autre lors de cette communication donne de la vie à l'environnement en le rendant socialement actif (*deux occurrences dans l'échantillon 2*).
- Dans la même logique, certains estiment que le fait de se parler donne vie à l'avatar, on est alors certains de ne pas être en face d'un agent virtuel, mais face à un humain virtualisé, ce qui permet à l'utilisateur de se dire, « S'il est virtualisé, alors moi aussi » ce qui transpose un peu plus le participant dans le monde virtuel (*deux occurrences dans l'échantillon 2*).
- L'expérience passée du jeu vidéo a une influence sur le sentiment de présence virtuelle en fonction du profil utilisateur. Les joueurs axés sur le jeu solo ont ressenti un meilleur sentiment de présence virtuelle lorsqu'il n'était pas possible de communiquer, à l'inverse des joueurs plutôt axés sur le multijoueurs (*trois occurrences dans l'échantillon 1*).
- Les participants ont mis en exergue un effet de l'expérience préliminaire lors de leur seconde session (SCV). Ils ont perçu un niveau de défi plus élevé, car désireux d'améliorer leur performance précédente, et ils estiment que le challenge joue un rôle dans le sentiment de présence virtuelle ressenti (*trois occurrences dans les échantillons 2 & 3*).
- D'un point de vue purement technique, les participants ont attesté que le matériel de visualisation choisi (lunettes 3D *Nvidia*) était très enclin à générer un bon sentiment de présence virtuelle, car proposant une visualisation stéréoscopique de bonne qualité et ne présentant pas les problèmes de « *cyber sickness* » fréquemment rencontrés avec les casques immersifs de type *Oculus Rift* (*deux occurrences dans l'échantillon 3*). De plus, les lunettes permettent à l'utilisateur de garder conscience de son corps, ce qui a été jugé comme moins perturbant que de ne pas avoir de corps du tout lors d'une immersion par casque immersif (ce qui reste aujourd'hui fréquent mais n'est néanmoins pas le cas dans *Minecraft*) (*une occurrence dans l'échantillon 2*).

► **Efficacité :** Deux questions de notre questionnaire mesuraient l'efficacité perçue par les participants vis-à-vis de la tâche qu'il leur était demandé de produire : « *Vous avez été efficace vis-à-vis de l'accomplissement de la tâche* : » (1 = Pas du tout efficace & 5 = Très efficace) — question 3 dans le questionnaire — et : « *Vous avez eu le sentiment de travailler ensemble sur la tâche* » (1 = Pas du tout ensemble & 5 = En synergie totale) — question 4 dans le questionnaire. Un test des rangs signés de Wilcoxon pour échantillons liés a révélé une différence statistique significative entre les contextes ACV et SCV sur les échantillons 1 & 3 lors de l'immersion dans l'EV.1 vis-à-vis de la question 4.

- **Échantillon 1** : Entre les contextes ACV (N = 8) et SCV (N = 8) : $Z = -2,598$, $p = ,009$.
- **Échantillon 3 — EV.1** : Entre les contextes ACV (N = 8) et SCV (N = 8) : $Z = -2,379$, $p = ,017$.

Afin de savoir dans lequel des deux contextes l'efficacité perçue fut la mieux évaluée par les participants, les moyennes des réponses ont été calculées pour chaque contexte. Les moyennes suivantes représentent le positionnement moyen des participants sur l'échelle de Likert à cinq niveaux.

- **Échantillon 1** : ACV, $M = 4,63$ et SCV, $M = 3,25$.
- **Échantillon 3 — EV.1** : ACV, $M = 4,63$ et SCV, $M = 2,63$.

Aucune différence statistique significative n'a été révélée vis-à-vis de la question 3 que ce soit entre les deux contextes ACV et SCV, ou entre les deux environnements dans le cadre de l'échantillon 3.

Les figures 4.63 et 4.64 présentent les moyennes obtenues pour ces questions.

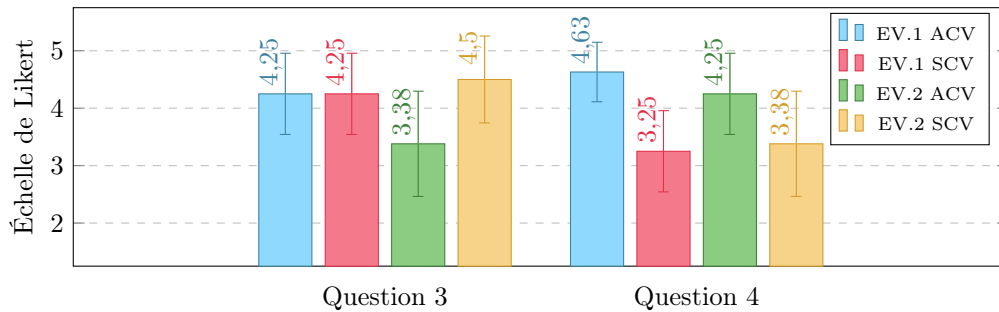


FIGURE 4.63 – Moyennes des évaluations de l'efficacité pour les échantillons 1 & 2.

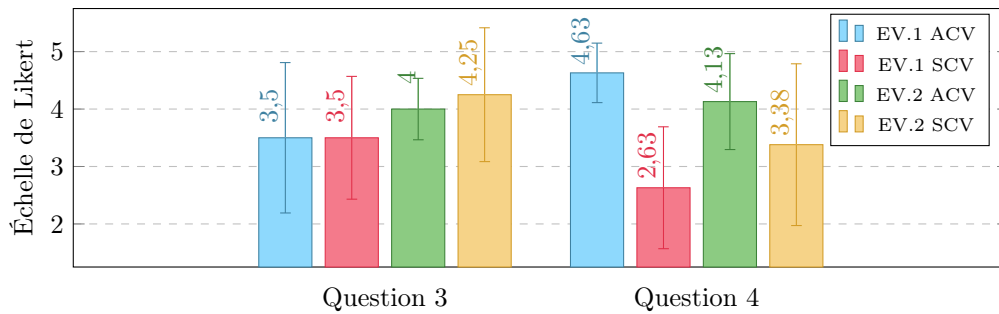


FIGURE 4.64 – Moyennes des évaluations de l'efficacité pour l'échantillon 3.

On constate ici que l'efficacité perçue au travers de l'interaction avec l'autre (question 4) est bien mieux évaluée lorsqu'il est possible de communiquer, et ce, quel que soit l'environnement d'immersion. Toutefois, c'est lors de l'immersion dans l'EV.1 (tâche collaborative) que les participants se sont sentis le plus efficaces.

Il ressort des groupes de discussion que les participants se sont perçus comme étant plus efficaces lorsqu'ils ne pouvaient pas communiquer vocalement (*sept occurrences dans les échantillons 1,2 & 3*). Des hypothèses expliquant le phénomène ont été émises :

- Une grande partie des participants s'accorde sur le fait que l'expérience préliminaire, avoir à faire la même tâche deux fois, rend la seconde occurrence plus simple à réaliser sans pour autant

garantir une meilleure performance ou une réduction des erreurs (*quatre occurrences dans les échantillons 1,2 & 3*). Plus encore quand la première fois il est possible de se parler puisque cela permet la mise en place d'une stratégie qui sera réutilisée lors de la seconde session (*trois occurrences dans les échantillons 1,2 & 3*).

- La nature de la tâche, collaborative ou coopérative, a été jugée comme très influente sur l'efficacité. Les participants se sont trouvés plus efficaces en coopération (EV.2), car ils ont remarqué que cela permettait une meilleure répartition du travail, chacun travaillant sur sa partie, alors qu'en collaboration ils avaient plutôt tendance à travailler ensemble sur une même partie de la tâche. « *C'est plus facile en coopération, car on ne se marche pas dessus, pas besoin de dire à l'autre de bouger pour poser un bloc puisque l'on ne bosse pas sur le même morceau* » (*deux occurrences dans l'échantillon 3*). Ils ont aussi constaté qu'en coopération (EV.2) et dans le cadre expérimenté d'une tâche interdépendante, ils étaient plus attentifs à leur propre travail sachant que leurs erreurs pouvaient être très influentes sur le travail de leur partenaire (*deux occurrences dans l'échantillon 3*). Dans le cadre de la tâche collaborative (EV.1), un facteur jugé comme facilitant l'efficacité est le fait de pouvoir corriger les erreurs de l'autre participant sans avoir à le lui dire (ce qui était aussi tout à fait possible dans la tâche coopérative, mais aucune indication n'a été donnée aux participants sur leur droit ou non de modifier le travail de leur partenaire, la grande majorité a estimé que cela leur était interdit) (*deux occurrences dans l'échantillon 3*).
- Des différences notables ont été constatées quant au rôle sur l'efficacité de la possibilité de communiquer vocalement. Ainsi, lorsqu'il est possible de communiquer, il est envisageable pour l'un des participants de prendre une position de maître d'œuvre. D'un point haut lui permettant de voir l'ensemble de l'ouvrage, il est en mesure de donner des directives (*quatre occurrences dans les échantillons 1 & 3*). Ce positionnement en tant que *leader* est lié à l'expérience préliminaire du partenaire sur *Minecraft* : à l'énoncé de cette prise de position, deux des quatre binômes présents ont répondu qu'aucun d'eux ne s'était positionné en *leader*, car ils avaient toute confiance en leur partenaire sachant que ce dernier était au moins aussi expérimenté qu'eux sur *Minecraft*. Dans le cas des binômes ayant vu l'émergence d'un *leader*, au moins l'un des participants était peu familier du jeu utilisé. Outre ce positionnement de maître d'œuvre, il est plus aisé, lorsque l'on peut communiquer vocalement, de signaler une erreur à son partenaire, et en retour de ce signalement, d'obtenir une réponse immédiate qui permet de savoir instantanément s'il a compris, s'il va faire la correction ou pourquoi il a commis cette erreur (*trois occurrences dans les échantillons 1 & 2*). Néanmoins, l'écueil dans lequel sont tombés bon nombre de binômes est celui de perdre du temps en communications stériles. Non pas que ces dernières soient décorréliées de la tâche en cours, mais plutôt par une forme de conflit émergeant du fait que chacun veut imposer sa vision, sa manière de faire ce qui donne lieu à des débats parfois de plusieurs minutes nuisant grandement à l'avancement de la tâche : « *Régis et moi avons tous les deux une forte personnalité, du coup, on s'est pas mal «frité» sur les erreurs commises voulant attribuer la faute à l'autre ! Ce qui nous a fait perdre pas mal de temps au final...* » (*cinq occurrences dans les échantillons 2 & 3*).

- En définitive, un consensus est établi sur le fait que, malgré les difficultés éprouvées pour se synchroniser lorsqu'il est impossible de communiquer vocalement (difficultés pouvant générer des erreurs), les participants se sont sentis plus efficaces dans ce contexte. Avec toutefois une nuance très intéressante, « *On est plus efficace sans se parler pour faire le gros du travail, mais dès qu'il faut corriger les erreurs, ça devient un véritable calvaire* » (quatre occurrences, puis consensus lors du bilan dans les échantillons 2 & 3).

► **Ressenti hédonique global :** Lors des groupes de discussion, nous avons voulu connaître le ressenti de nos participants vis-à-vis de l'expérimentation menée. Des questions qui n'ont pas soulevé la moindre discussion mais ont donné des réponses unanimes. À savoir s'ils étaient prêts à refaire cette expérience, tous s'accordent sur un « oui » à la condition de changer la tâche à accomplir. Quant à leur engagement vis-à-vis de ces tâches, force est de constater une très grande implication de leur part, une vraie envie de réussir, d'aller jusqu'au bout. Nous leur avons aussi demandé ce qu'ils diraient à une personne leur demandant si elle devait venir participer à cette expérience. Là encore, la réponse est unanime, « *Vas-y sans hésitation* ». Ils ont trouvé très intéressant de se voir privés de communication vocale et de voir qu'ils étaient tout de même capables d'accomplir les tâches demandées et même d'être aussi, voire plus, efficace que lorsqu'il était possible de communiquer.

4.4.8 Analyse et synthèse

Afin d'éviter d'être redondants dans notre exposé, nous avons jugé préférable de traiter les explications et implications des résultats associées à cette expérimentation complémentaire directement par la comparaison de nos trois phases d'expérimentation dans le chapitre suivant (chapitre 5). En effet, les groupes de discussion corroborent ou dédisent certains constats que nous avons pu faire précédemment, il est donc pertinent de traiter ces derniers résultats d'un point de vue global. Nous ne présenterons ici que des limites de notre expérimentation complémentaire.

La première limite de cette expérimentation complémentaire est bien sûr le nombre de participants s'élevant seulement à vingt-quatre. Nous n'avions pas prévu de faire cette expérimentation, pensant être en mesure de conduire nos groupes de discussion avec les participants de la seconde expérimentation, ce qui ne fut pas le cas comme nous l'avons précédemment expliqué. Mais, après analyse de nos résultats, nous avons besoin d'éclaircissements, qui devaient être fournis par des résultats qualitatifs. Étant très conscients du temps nécessaire pour réunir des participants, additionné au temps de passage de l'expérimentation, nous ne pouvions nous permettre de relancer un tel processus dans le temps qu'il nous restait pour mener à bien notre travail de thèse. Nous avons alors dû composer avec la population disponible : les étudiants du Master hébergé dans notre structure d'accueil. Ce choix nous amène à une seconde limite ou tout du moins un biais que nous ne pouvions éviter : la familiarité non seulement entre les participants, mais aussi avec les expérimentateurs. Facteurs que nous savons comme influant sur les réponses à un questionnaire, la population sondée tendant souvent de manière inconsciente à orienter ses réponses en fonction de son affinité avec le ou les expérimentateurs. Fort heureusement, les étudiants ne s'étaient jamais intéressés à nos travaux et ne connaissaient donc pas la finalité de

notre étude ni ce que nous désirions observer de leurs comportements lors de l'expérimentation.

La troisième limite réside dans l'exposition consécutive aux deux contextes ACV puis SCV. Cette double exposition est limitante vis-à-vis de l'évaluation de l'efficacité, car il est évident que la seconde fois, une même tâche sera plus aisée. A contrario, l'impact peut être néfaste sur l'amusement, le fait de reproduire la tâche jugulant l'effet de découverte. Qui plus est, l'ordre d'immersion, ACV puis SCV, risque de réduire l'amusement, les participants se sentant frustrés de ne plus pouvoir communiquer. Toutefois, cette double exposition est aussi la force de cette expérimentation complémentaire puisqu'elle a permis aux participants de s'exprimer sur la comparaison des deux contextes ainsi que des deux environnements.

Le dernier point que nous estimons comme limitant est que nous avons conduit les groupes de discussion le même jour que les sessions d'expérimentation. Bien qu'il soit difficile de trouver dans la littérature une information précise quant au temps à laisser entre session d'immersion et groupes de discussion, nous pensons qu'il est pertinent de laisser aux participants le temps de « digérer » leur expérience afin d'avoir une vision présentant un minimum de recul quant à cette dernière lors du groupe de discussion. D'un autre côté, le discours inverse pourrait être tenu arguant que les réponses « à chaud » sont sans doute plus sincères.

Malgré ses limites, cette expérimentation est bien complémentaire aux deux autres, et les informations recueillies lors des groupes de discussion ont apporté des éclaircissements sur nos résultats précédents. Nous présenterons nos réflexions vis-à-vis de ces derniers dans le chapitre 5 en corrélation avec les résultats obtenus lors de nos deux premières expérimentations.

Chapitre 5

Discussion

« Néanmoins, l'avidité qui consume l'esprit non critique à l'égard de ce qu'il s'imagine certitude ou finalité le pousse à se délecter d'ombre. »

Eric Temple Bell
Mathématicien et écrivain - 1883-1960

Sommaire

5.1	Présence sociale	172
5.2	Flow Psychologique (Engagement)	174
5.3	Présence Virtuelle	176
5.4	Performances & Efficacité	177
5.5	Adoption technologique (ressentit hédonique global)	179
5.6	Schémas récapitulatifs des résultats	179
5.7	Tableau récapitulatif des résultats	181

Nous allons, au sein de ce chapitre, mettre en parallèle les résultats de nos trois expérimentations. L'objectif est de mettre en exergue les éventuelles constantes dans nos résultats quant à l'impact de la communication vocale sur les facteurs évalués de l'expérience utilisateur. Mais aussi, d'en confronter les différences afin de comprendre ce qui a pu les engendrer. Ce travail d'analyse nous permettra par la suite dans le chapitre 6 d'exposer nos conclusions quant aux réponses apportées à nos questions de recherche ainsi qu'à la validation ou à l'infirmerie de nos hypothèses.

Nous suivrons ici la trame précédemment utilisée, consistant à suivre la taxonomie de facteurs du modèle de qualité de DIMEs de Wu *et al.* (2009) et du modèle d'expérience utilisateur de Pallot *et al.* (2013) en étant issu.

5.1 Présence sociale

Nous avons, au cours de nos trois expérimentations, évalué la présence sociale au travers de quatre variables : l'attention allouée, la compréhension mutuelle, l'interdépendance comportementale et l'assistance mutuelle. Lors de notre première expérimentation, nous avons constaté que l'évaluation de l'attention allouée ne présentait pas de différence statistique significative entre nos deux contextes. Nous avons attribué cette absence au fait que dans le cadre d'une tâche collaborative, il était nécessaire d'allouer de l'attention à l'autre, qu'il soit possible de communiquer vocalement ou non, afin de s'assurer de l'avancement de la tâche commune. Par conséquent, la différence ne se situait pas dans la quantité d'attention allouée mais plutôt sur ce sur quoi elle se portait, les échanges vocaux ou le comportement. Lors de la seconde expérimentation, ce postulat s'est en un sens vérifié, puisque nous avons relevé une plus grande allocation de l'attention dans le contexte ACV. Cela est tout à fait cohérent avec le passage à une tâche coopérative, car cette dernière requiert moins d'attention quant aux actions du partenaire du fait de la répartition explicite du travail. Les participants SCV ont ainsi moins prêté attention au comportement de leur binôme. Dans le cas du contexte ACV, l'attention s'allouant automatiquement dès la prise de parole, la seule chose qui aurait pu la faire décroître aurait été le mutisme de l'un des participants. Ces constats sont soutenus par l'expérimentation complémentaire pour laquelle nous avons comparé l'attention allouée à la communication vocale et celle allouée au comportement, ainsi que l'impact de ces derniers sur les participants (interdépendance comportementale). Nous avons alors constaté que dans tous les cas, c'est la communication verbale qui suscite le plus d'attention, ce qui est en accord avec la structure du cerveau comme nous l'avons vu dans notre revue de littérature (Delamarre, 2011) : nous faisons en premier lieu attention au langage verbal.

Lors des groupes de discussion, les participants ont confirmé cette différence vis-à-vis de l'attention allouée au comportement en expliquant que lorsqu'il est possible de se parler, une répartition des tâches se mettait en place : on attend alors de son partenaire qu'il fasse ce qui lui a été assigné, et lui faisant confiance, on ne regarde pas en permanence son avancement. De plus, si le partenaire rencontre un problème il est en mesure de l'exprimer. En revanche, lorsque la communication vocale est interdite, la seule manière de suivre l'avancement de la tâche ou l'apparition de difficultés consiste à regarder

son partenaire en action. En complément du suivi des actions, les participants ont expliqué lors des groupes de discussions leur utilisation d'éléments sonores non vocaux pour réguler leur allocation de l'attention. En effet, les actions de poser un bloc ou de le détruire produisant des sons spécifiques, leur absence signalait que le partenaire n'était plus en activité, indiquant ainsi qu'il était peut-être face à un problème. Un impact de la familiarité sur l'attention allouée a aussi été évoquée. Les participants à cette dernière expérimentation se connaissaient, certains d'entre eux étaient amis et étaient conscients du niveau d'expérience de *Minecraft* de leur partenaire. Dans le cas de participants plutôt expérimentés, la connaissance de ce niveau de compétence tendait à réduire l'attention allouée, car les partenaires se faisaient confiance sachant que leur binôme était tout à fait à même d'accomplir la tâche demandée. Ce croisement des résultats nous permet de dire que l'allocation de l'attention est naturelle, automatique lorsqu'il est possible de communiquer vocalement, mais que s'il est impossible de se parler, l'attention est allouée au comportement en fonction du besoin inhérent à la tâche à accomplir en s'aidant de la présence d'autres facteurs non verbaux comme les sons émis par les actions entreprises. Dans les deux cas, les participants sont en mesure de capter les messages vocaux ou comportementaux de leur partenaire, bien que — comme décrit lors des groupes de discussions — cela demande un effort plus grand lorsque la communication vocale n'est pas possible.

Vis-à-vis de la compréhension mutuelle, les résultats des deux premières expérimentations montrent que c'est dans le contexte ACV que la compréhension mutuelle est la meilleure. Ce qui était attendu en raison de la nature même du langage vocal, dont l'objectif et la transmission d'un message selon des modalités communes à l'émetteur et au récepteur, permettant ainsi sa compréhension plus facilement que par le comportement ou le paralangage. Cet écart du niveau de compréhension mutuelle a impliqué des différences significatives entre les contextes SCV et ACV pour l'interdépendance comportementale. Ceci s'explique aisément par le fait que le langage vocal, le fait de pouvoir donner des indications, des explications permet une plus grande influence sur les actions de l'autre, imposant ainsi une plus grande interdépendance entre les participants. Néanmoins, les participants des deux expérimentations dans le contexte SCV ont aussi attesté d'une interdépendance comportementale, montrant que l'étude du comportement de l'autre, l'écoute des indications non verbale permet d'adapter son propre comportement en fonction de l'autre (translucence). Cette constatation du fait que la translucence permet de pallier l'absence de communication verbale est confortée par les résultats de l'expérimentation complémentaire, les moyennes des réponses aux questionnaires n'ont pas montré de différence importante entre l'impact de la communication vocale et l'impact du comportement. Il est toutefois nécessaire de rappeler le nombre réduit de participants ainsi que l'exposition consécutive ACV puis SCV ayant pu avoir un impact sur ces résultats. Lors des groupes de discussion, les participants ont apporté une explication quant au fonctionnement de l'interdépendance comportementale dans le contexte SCV : outre l'observation des actions de leur partenaire, leur comportement translucent était surtout généré par les actions passées de leur binôme, autrement dit par les blocs déjà posés, que ce soit pour la tâche collaborative ou coopérative. C'est en étudiant le travail déjà accompli qu'ils étaient le plus en mesure d'adapter leur comportement afin de faire coïncider leurs efforts.

Un autre indice de l'efficacité de la translucence réside dans les évaluations de la variable assistance

mutuelle. Nous n'avons relevé aucune différence statistique significative entre les contextes SCV et ACV vis-à-vis de cette variable, et ce pour les deux expérimentations. Les participants ont donc estimé que leur partenaire avait été d'une grande aide, et cela même dans le cas où il leur était impossible de communiquer.

La présence sociale est donc influencée par la communication vocale sur plusieurs aspects, en premier lieu elle réduit les efforts à fournir pour être attentif à son partenaire. Puis elle permet une bien meilleure compréhension mutuelle facilitant ainsi l'interdépendance comportementale par la promulgation d'indications et de conseils. Néanmoins, en l'absence de communication vocale, les utilisateurs sont tout de même capables de s'influencer mutuellement et de s'aider par la translucence, afin d'atteindre l'objectif fixé. Il est important de noter que la situation (nature de la tâche), selon qu'elle demande un fort niveau de coopération ou non, a un impact notable sur l'attention allouée. En revanche, nous n'avons pas relevé les mêmes effets sur les variables interdépendance comportementale et assistance mutuelle, mais nous pensons que cela est en grande partie dû au fait que la tâche coopérative, bien que présentant une division du travail explicite, restait très interdépendante en raison de sa conception. D'après les résultats obtenus, nous pouvons dire que la communication vocale améliore la présence sociale mais qu'il n'y a pas une absence totale d'attention et d'interaction entre utilisateurs lorsqu'elle n'est pas possible, la translucence prenant alors le relais. En conclusion vis-à-vis de la présence sociale, nous citerons un des participants des groupes de discussion (ayant produit la tâche coopérative) dont les propos attestent de l'occurrence de la présence sociale lors de notre expérimentation par la conservation d'un comportement naturel montrant un effet de la communication vocale sur les distances sociales (Yee *et al.*, 2007 ; Bailenson *et al.*, 2001b) : « *J'ai constaté que lorsque l'on pouvait se parler, je restais très souvent à proximité de lui, comme si je voulais rester à portée de voix, alors que je savais très bien que cela n'avait pas d'influence* ».

5.2 Flow Psychologique (Engagement)

Portons à présent notre analyse sur le *flow* psychologique, évalué grâce à deux variables, le niveau de défi perçu, qui par son équilibre avec le niveau de compétence perçue de l'utilisateur permet l'occurrence de l'état de *flow* (nous ne considérons pas le niveau de compétence perçue comme une variable dans nos évaluations étant donné que dans nos trois expérimentations, les échantillons sondés étaient homogènes quant à leur d'expérience préalable de *Minecraft*) et l'amusement, variable résultante de l'état de *flow* comme décrit par Sweetser et Wyeth (2005). Vis-à-vis du niveau de défi perçu — évalué seulement dans les deux expérimentations principales — nous n'avons relevé aucune différence significative entre les contextes SCV et ACV. Nos échantillons étaient, comme nous l'avons déjà précisé, homogènes quant à leur expérience préalable de *Minecraft*, nous ne pouvons donc imputer cette absence de différence au fait que les participants immergés SCV aient eu plus de facilité égard à une expérience préalable plus importante. Deux hypothèses peuvent expliquer ce constat : nos tâches répondaient bien à notre exigence de faisabilité tant ACV que SCV, et deuxièmement, les comportements translucents ont suffi à rendre la tâche SCV réalisable. Même si, comme nous le verrons par la

suite, une différence d'efficacité a été relevée dans la première expérimentation, une grande partie des participants a réussi à compléter ou a presque complété la tâche demandée, jugeant leurs prestations satisfaisantes dans les deux contextes ils ont estimé le niveau de challenge de l'activité de manière suffisamment similaire pour que nous ne détectons pas de différence statistique significative. Nous ne pouvons donc pas démontrer d'impact de la communication vocale sur le niveau de défi perçu, néanmoins, cette variable étant très corrélée à l'activité proposée, cela s'applique spécifiquement à nos expérimentations. Vis-à-vis de l'amusement, et de manière globale pour les deux expérimentations principales, les participants ont estimé s'être plus amusé dans le contexte ACV. C'est aussi le cas pour l'expérimentation complémentaire où nous avons demandé, en plus de l'évaluation directe de l'amusement, une évaluation de la durée perçue de l'expérimentation que nous savons être un très bon indicateur de l'amusement (Danckert et Allman, 2005), le temps semblant se contracter lorsque l'on passe un bon moment. Les participants ont estimé que le temps passait plus vite lors de la session ACV, ce qui va dans le sens de nos constatations précédentes. Toutefois, il faut ici prendre en compte la double exposition, supprimant l'effet de découverte et engendrant plus d'ennui du fait de la reproduction de la même tâche, ainsi que l'ordre de cette double exposition, les participants ayant d'abord vécu l'expérience ACV puis SCV.

Lors des groupes de discussion, nous avons demandé aux participants pourquoi ils pensaient s'être plus amusés dans l'un ou l'autre des deux contextes. Ils n'ont pas été en mesure de répondre de manière précise, mais tous se sont accordés sur le fait qu'ACV, l'expérience était plus amusante sans pouvoir extraire le moindre élément de justification. Nous pensons que cela est dû à la possibilité ACV de faire des remarques, de l'humour, d'entendre l'autre rire, tous ces éléments de conversations non corrélées directement à la tâche en cours, mais qui ajoutent de la vie, du tonus à l'expérience vécue. Entre aussi en ligne de compte la frustration quant à la résolution d'un problème, une erreur de placement de bloc par exemple. Sans communication vocale, impossible de dire à son partenaire qu'il est en train de se fourvoyer, il faut donc stopper ce que l'on est en train de faire, pour aller le voir et tenter de lui faire comprendre son erreur. Lors des groupes de discussion, cette problématique a été jugée comme très frustrante, et la frustration étant l'ennemi de l'amusement, en résulte une baisse de ce dernier. Néanmoins, que ce soit dans les évaluations par questionnaire ou lors des groupes de discussion, les participants ont tous attesté s'être bien, voire beaucoup amusés lors de cette expérience, en expliquant même que le fait de ne pas pouvoir se parler amène de nouveaux éléments d'amusement auxquels on ne fait pas attention ACV. Par exemple, le bruit produit par la chute du partenaire sans pouvoir la commenter, les a grandement amusés, des « conflits » de conception (modifications d'une même partie par l'un puis par l'autre plusieurs fois de suite) les ont tout d'abord frustrés puis amusés devant le ridicule de la situation. Enfin, l'importance de l'environnement et de la tâche proposée a été évoquée. Ainsi l'environnement proposant la tâche coopérative a été jugé, lors des groupes de discussion, comme étant plus amusant. Les participants ont affirmé que cet environnement était plus abouti, présentant plus de fonctionnalités et d'interactivité palliant ainsi la frustration que pouvait présenter la construction verticale de l'EV.1. La tâche coopérative a aussi été jugée comme plus complexe en raison de l'intrication entre les deux parties du plan, et par conséquent plus amusante,

car présentant un plus haut niveau de défi. D'un autre côté, certains l'ont jugé plus ennuyeuse à cause de la division du travail imposée.

Ces observations démontrent que la communication vocale n'a pas d'effet notable sur l'état de *flow*. En revanche, elle possède un large impact bénéfique sur l'amusement, mais il est toutefois impossible de définir une expérience en 3D-IVEs comme amusante uniquement parce qu'elle supporte la communication vocale. D'autres facteurs entrent en jeu, et ont plus d'impact que la communication verbale, comme en témoigne le niveau d'amusement global dans nos trois expérimentations. L'environnement virtuel, les fonctionnalités et interactions qu'il offre, son niveau de qualité général ainsi que la ou les activités qu'il propose, sont très voire plus importants que la communication vocale sur le sentiment d'amusement. En quelques mots, une activité ennuyeuse et un environnement ne plaisant pas aux utilisateurs resteront mal évalués en termes d'amusement, que les utilisateurs soient en mesure de se parler ou non.

5.3 Présence Virtuelle

Malgré nos difficultés lors de nos tentatives d'évaluation de la présence virtuelle de manière objective (physiologie), les résultats quantitatifs subjectifs de nos trois expérimentations présentent un ensemble cohérent. Ainsi, nous avons constaté que le niveau de présence virtuelle ressenti par nos participants ne variait que peu entre les contextes SCV et ACV, cela vis-à-vis de l'évaluation directe de sentiment de présence virtuelle, de la vivacité sensorielle ainsi que de la concentration. Nous avons en revanche mis en évidence, lors de la première expérimentation, une différence significative sur la variable degré de contrôle évaluant la mesure dans laquelle les participants étaient conscients d'interagir avec un périphérique hors de l'environnement virtuel plutôt que directement avec ce dernier. Nous avons alors émis l'hypothèse que la communication vocale créait un lien avec le monde réel. Nous n'avons pas relevé de résultats similaires lors de la seconde expérimentation mais, lors des groupes de discussion, des ressentis très différents en fonction des personnalités et des affinités avec le jeu vidéo solo et multijoueurs sont apparues. Ces différences nous indiquent que le sentiment de présence virtuelle semblait bien plus corrélé aux aspects personnels de l'utilisateur plutôt qu'à la possibilité d'échanger vocalement. Ainsi, certains de nos participants se sont sentis plus présents sans communication vocale, expliquant qu'ils étaient beaucoup plus habitués à jouer seuls et que par conséquent les discussions avec leur partenaire avaient tendance à les faire « décrocher » en réduisant leur concentration et par corrélation leur sentiment de présence virtuelle. D'un autre côté, les joueurs pratiquant régulièrement le jeu vidéo multijoueurs ont eux estimé s'être sentis plus présents indiquant que la présence de l'autre donnait vie à l'environnement virtuel en le rendant socialement actif. Ils ont estimé que le fait de regarder l'avatar de l'autre lors de leurs échanges vocaux renforçait ce sentiment d'être virtuellement présent par une forme de transfert « *S'il est virtualisé, alors moi aussi* ». Il a aussi été mis en évidence que lors de la seconde session de chaque échantillon (SCV), ils s'étaient sentis plus présents, expliquant cela par une augmentation du niveau de défi ayant engendré une plus grande concentration. Pour en revenir au degré de contrôle, nous avons aussi émis l'hypothèse que

le contenu des conversations pouvait avoir un effet sur la présence virtuelle. Cette hypothèse était extraite de notre constat de la récurrence d'un thème dans les conversations entre participants : les modalités d'interaction (un impact du panneau d'information est envisageable). Il est probable que lors d'une immersion ACV, si les discussions s'orientent vers les modalités d'interaction, les possibles incohérences de l'environnement virtuel ou sur ce que peut percevoir un participant du monde réel, cela réduise le sentiment de présence virtuelle. Ceci est une vision cohérente avec le continuum de la présence de Pillai *et al.* (2013).

Nos résultats indiquent un effet mineur de la possibilité de communiquer vocalement, en revanche nous sommes convaincus que le contenu des conversations joue un grand rôle dans le sentiment de présence virtuelle comme nous l'avons indiqué. Il faut aussi prendre en compte les différentes sensibilités des utilisateurs ainsi bien sûr que les variables de la QoS et du *flow* psychologique ayant un fort impact sur le sentiment de présence virtuelle. Le problème mis en évidence ici quant aux préconisations que nous pourrions proposer vis-à-vis de la conception d'applications de réalité virtuelle, est qu'il est difficile d'orienter les conversations des utilisateurs, la seule manière est de leur proposer un environnement virtuel répondant à leurs attentes tant perceptuelles (cohérence, vivacité sensorielle,...) qu'hédoniques ainsi que des interfaces de contrôle aussi intuitives que possible.

5.4 Performances & Efficacité

Revenons à présent sur les résultats tant subjectifs que quantitatifs de l'évaluation des performances de nos participants. Dans notre première expérimentation, les participants se sont évalués comme plus performants ACV. Cela a été confirmé par nos observations affichant une supériorité de l'échantillon immergé ACV s'élevant à 51,88 blocs posés en plus. De plus, nous avons constaté que seulement deux binômes sur treize avaient commis des erreurs contre sept sur treize SCV. Au regard de ces résultats, nous avons alors soulevé la question de l'impact de la nature de la tâche, car dans cette première expérimentation la tâche était collaborative, les deux participants d'un binôme possédant le même niveau d'information (totalité du plan) ainsi que les mêmes matériaux. Afin de répondre à cette question, notre seconde expérimentation proposait une tâche coopérative, où les deux participants d'un binôme possédaient chacun une moitié du plan, mais aussi présentaient quelques dissemblances vis-à-vis des matériaux en leur possession. Les résultats obtenus, conformément à nos attentes, ont montré que dans ce cadre nécessitant moins d'efforts conjoints (bien que les deux parties du plan soient étroitement intriquées), non seulement il n'y avait plus de différence statistique significative entre les évaluations de l'efficacité perçue dans les contextes SCV et ACV, mais aussi que cette absence de différence se retrouvait dans nos observations objectives. Cette différence entre les deux tâches s'est vue confirmée par l'expérimentation complémentaire où les participants se sont, là aussi, évalués comme plus efficaces ACV dans l'EV.1 mais pas dans l'EV.2. Nous avons alors demandé lors des groupes de discussion pourquoi les participants auraient pu être plus performants dans l'un ou l'autre des contextes ou sur l'une ou l'autre des tâches. Il est nécessaire de préciser ici que les réponses obtenues lors des groupes de discussion ont indubitablement été influencées par la répétition de la même tâche

deux fois de suite. Nous avons donc informé les participants de ce biais et leur avons demandé d'en tenir compte dans leurs observations. Une partie des participants a expliqué avoir atteint une meilleure efficacité grâce à la prise de position du plus expérimenté en tant que *leader*, donnant des directives et définissant une stratégie (EV.1). D'un autre côté, certains participants ont expliqué ne pas du tout avoir vu l'émergence d'un maître d'œuvre, car ils étaient conscients de la compétence avancée de leur binôme sur *Minecraft*, et avaient confiance en lui quant à l'accomplissement de la tâche. A contrario, d'autres participants ont estimé être plus efficaces SCV pour deux raisons, la première est que le fait de ne pas pouvoir se parler favorise la concentration sur la tâche (même constat que pour la présence virtuelle). La seconde est que, dans le cas de l'émergence d'un *leader*, il arrive que des désaccords apparaissent, chacun tentant de faire prévaloir son point de vue sur la stratégie à adopter. Les participants ont estimé que cet écueil conflictuel leur avait fait perdre beaucoup de temps. Vis-à-vis de la différence entre les deux tâches, la coopération est apparue comme permettant d'atteindre une meilleure efficacité en raison tout d'abord de la division claire des tâches, alors qu'en collaboration, plus encore SCV, les participants avaient tendance, selon leurs termes, à « *se marcher dessus* », à se gêner. L'intrication profonde des deux parties de la tâche coopérative a eu pour effet de rendre les participants plus vigilants quant aux erreurs qu'ils pouvaient commettre, conscients que celles-ci impacteraient fortement leur partenaire sans que ce dernier ne puisse se référer au plan pour s'assurer de qui était dans l'erreur. Enfin, un consensus fut établi dans l'échantillon ayant expérimenté les deux environnements, sur le fait que l'efficacité était fonction d'une interaction entre la communication vocale et la granularité de la tâche. Il ont expliqué, en comparant leurs sessions successives, être plus efficaces SCV pour faire le « *gros œuvre* » au début de la session, mais qu'avec l'avancement de la tâche, régler les points de détails, résoudre les erreurs, devenait très difficile sans la possibilité de se parler.

Nous devons souligner que les constatations faites ici ne concernent que les tâches que nous avons proposées, et qui demandaient une forme plus ou moins importante de travail en commun. Les résultats seraient sans doute très différents en autorisant la communication vocale entre deux personnes ne travaillant pas du tout sur la même tâche. En témoignent par exemple des concepts tels que le *Taylorisme* pour lequel le bavardage est une source de perte de productivité. Toutefois, nos résultats indiquent que l'efficacité est impactée par la communication vocale mais que cet impact dépend aussi fortement de la nature et de la granularité de la tâche, de la personnalité des personnes en présence, de leur propension à prendre une position de leader ou d'exécutant, ainsi que de la confiance mise en son partenaire vis-à-vis de son niveau de compétences. Dans la conception d'une application de réalité virtuelle où l'on souhaiterait atteindre la plus grande efficacité possible, une étude poussée des tâches à accomplir est donc primordiale. Dire qu'autoriser la communication vocale entre participants immergés dans un environnement virtuel va leur permettre d'être plus efficaces représente une simplification trompeuse des phénomènes que nous avons observés au travers de ces trois expérimentations.

5.5 Adoption technologique (ressentit hédonique global)

Reste à présent le potentiel effet de la communication vocale sur la satisfaction des participants : outre l'évaluation de l'amusement, nous avons, lors de nos trois expérimentations, demandé une évaluation de la variable intention d'utilisation (facteur adoption technologique) au travers de deux questions. L'une demandant si les participants seraient d'accord de refaire l'expérience à laquelle ils avaient participé, la seconde s'ils la conseilleraient à une tierce personne. Nous n'avons à aucun moment décelé, vis-à-vis de ces questions, de différence statistique significative entre les contextes ACV et SCV, les utilisateurs faisant montre d'un grand enthousiasme, estimant qu'ils avaient vraiment passé un bon moment et que bien sûr ils seraient prêts à conseiller l'expérience à d'autres personnes. Il a été expliqué lors des groupes de discussion que le fait d'être privé de communication vocale sur le type de tâche proposée avait été jugé vraiment intéressant dans le sens où cette expérience leur avait montré que, même si cela leur était apparu comme difficile de prime abord, ils avaient été capables de produire la tâche en ne se basant que sur les actions et le comportement de leur partenaire.



Pour conclure ce chapitre *discussion*, les figures 5.1 et 5.2 propose une vision schématique récapitulative de nos résultats. La figure 5.1 présente les impacts (positifs ou négatifs) de l'interaction sociale vocale sur les facteurs de l'expérience utilisateur en 3D-IVE. La figure 5.2 quant à elle expose les interactions entre les autres facteurs de l'expérience utilisateur que nous avons pu observer au cours de notre étude au travers de nos résultats. Ces deux figures exposent les liens complexes qui unissent les facteurs de l'expérience utilisateur créant une toile d'interaction qui reste malgré nos efforts encore ardue à déchiffrer. Pour cette raison et afin d'apporter une vision synthétique de nos travaux, le tableau 5.1 (découpé en trois parties) propose une liste des facteurs de l'expérience utilisateur évalués au cours de nos travaux. Pour chaque facteur, est décrit l'impact de l'interaction sociale vocale sur ce dernier et sur les variables le composant. Les facteurs ou variables ayant un impact sur le facteur ou la variable couramment décrit, ainsi que les impacts qu'il ou elle engendre sur d'autres facteurs ou variable du modèle du Wu *et al.* (2009) étendu par Pallot *et al.* (2013). Nous reviendrons sur ces représentations au cours du chapitre *conclusion* (chapitre 6) car nous les considérons comme un des apports de notre travail de thèse, mais aussi comme un outil nous permettant d'avoir une vision globale quant à la validation ou à l'infirmité de nos hypothèses **H1** & **H2**.

5.6 Schémas récapitulatifs des résultats

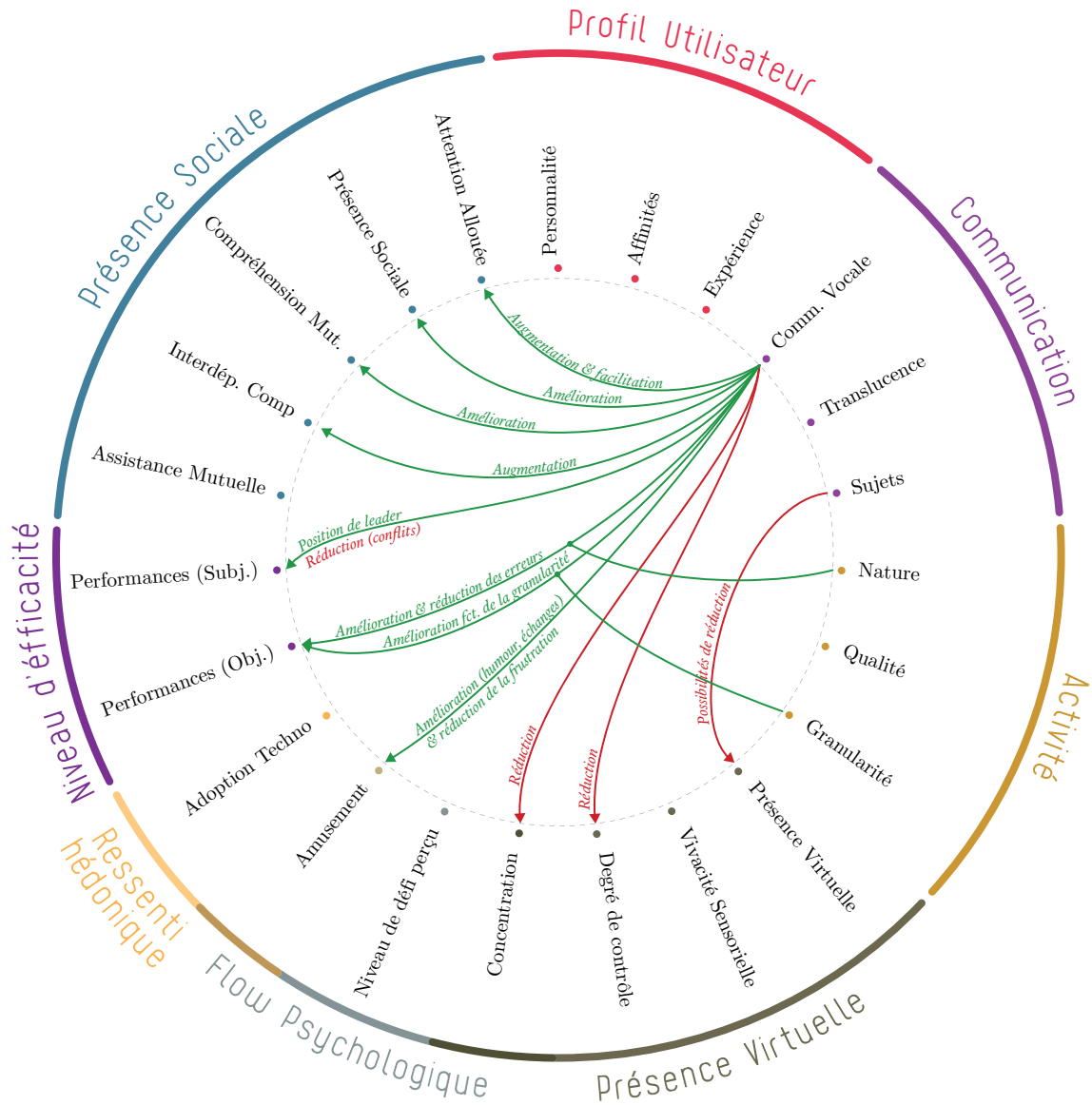


FIGURE 5.1 – Résumé schématique de l'impact de la communication vocale sur l'expérience utilisateur en 3D-IVE.

Flèches vertes : → = Impacts positifs de l'interaction sociale vocale.

Flèches rouges : → = Impacts négatifs de l'interaction sociale vocale.

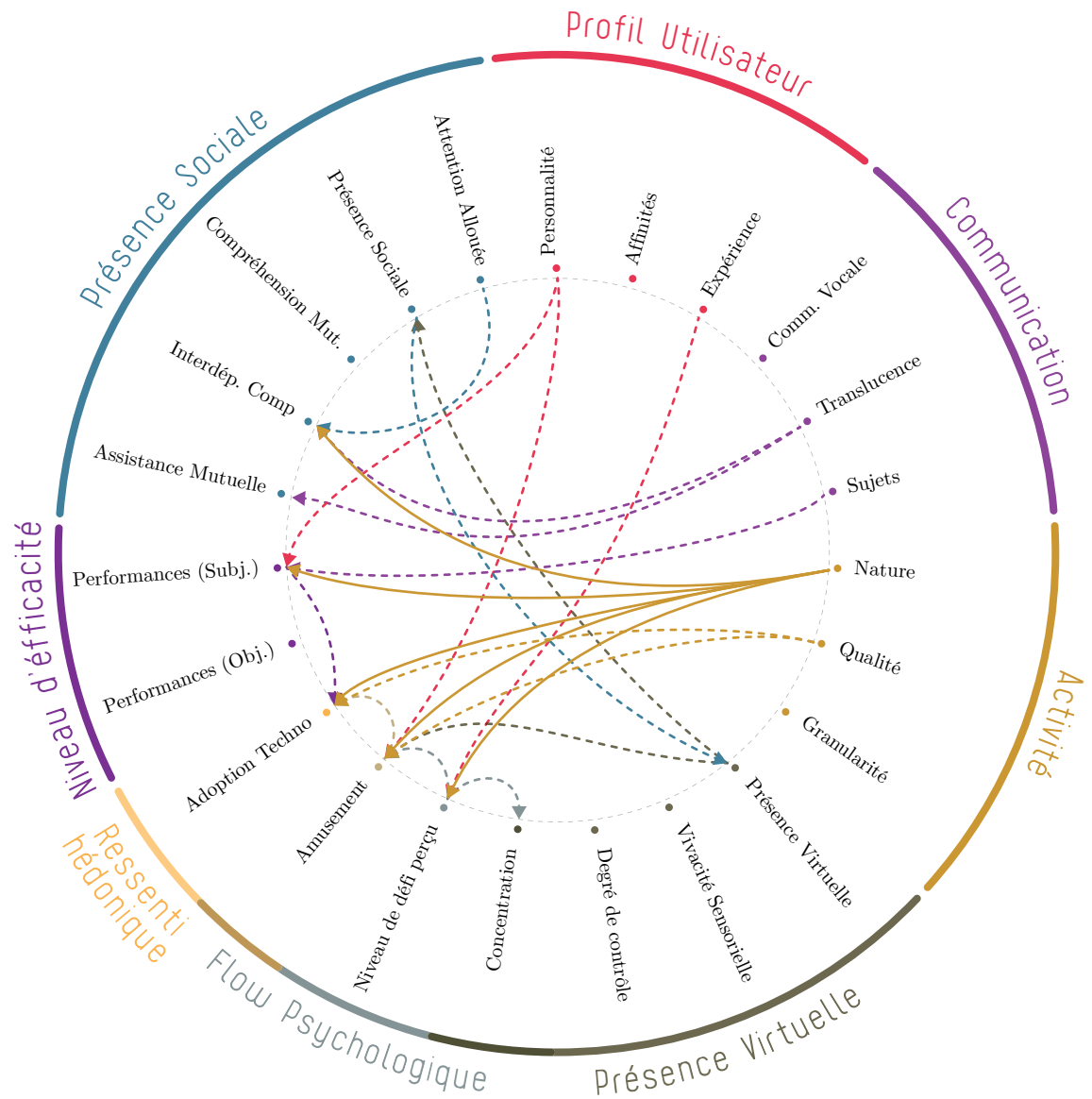


FIGURE 5.2 – Résumé schématique des impacts inter facteur de l'expérience utilisateur en 3D-IVE.

5.7 Tableau récapitulatif des résultats

TABLEAU 5.1 – Récapitulatif des résultats de l'étude - Partie 1

		Facteur	Variable	Impact de l'ISV	Est impacté par...	Impacte...
Expérience (Utilisateur) Immersive	Processus expérientiel	Présence Sociale		► Amélioration.	► Attention allouée ; ► compréhension mutuelle ; ► interdépendance comportementale ; ► assistance mutuelle.	► Présence virtuelle.
			Attention allouée	► Augmentation (par automatisme ISV) ; ► réduction de l'effort d'attention.	► Nature de la tâche ; ► émissions sonores du partenaires (régulation) ; ► familiarité (augmente ou réduit l'attention allouée en fonction de la confiance attribuée au partenaire).	► Interdépendance comportementale.
			Compréhension mutuelle	► Amélioration.	► Translucence (palliatif à l'absence d'ISV)	► Interdépendance comportementale.
			Interdépendance comportementale	► Augmentation.	► Translucence (basée sur l'observation du comportement et sur l'observation des actions passées) ; ► nature de l'activité.	► N/A
			Assistance mutuelle	► Pas d'impact décelé (en raison de la translucence).	► Translucence (basée sur l'observation du comportement et sur l'observation des actions passées).	► N/A
	Processus rationnel	Flow psychologique		► Pas d'impact décelé sur la variable évaluée et conditionnant l'état de flow (niveau de défi perçu).	► Niveau de défi perçu ; ► niveau de compétence perçu ; ► concentration (variable de la présence virtuelle).	► Amusement (plaisir).
			Niveau de défi perçu	► Pas d'impact décelé (en raison de la translucence).	► Expérience préalable ; ► nature de la tâche.	► Amusement (plus grand est le défi et, dans la mesure où il reste réalisable plus grand et le plaisir de le relever).

TABLEAU 5.2 – Récapitulatif des résultats de l'étude - Partie 2

		Facteur	Variable	Impact de l'ISV	Est impacté par...	Impacte...
Expérience (Utilisateur) Immersive	Processus rationnel	Flow psychologique	Amusement (résultante du flow).	<ul style="list-style-type: none"> ► Amélioration par l'humour, l'échange ; ► amélioration par réduction de la frustration éventuelle liée à l'activité. 	<ul style="list-style-type: none"> ► Nature de la tâche ; ► facteurs de la QoS (interactivité, qualité visuelle) ; ► personnalité de l'utilisateur. 	<ul style="list-style-type: none"> ► Adoption technologique ; ► acceptation technologique ; ► ressentit hédonique globale.
				<ul style="list-style-type: none"> ► Réduction très mineure (degré de contrôle) ; ► possibilité de réduction par les sujets de conversation pointant les défauts de l'environnement (modalités d'interactions, bugs, etc...) 	<ul style="list-style-type: none"> ► Présence sociale (en fonction de l'affinité de l'utilisateur avec les plateformes de jeu-vidéo solo / multijoueur) ; ► sujets de conversation. 	<ul style="list-style-type: none"> ► Amusement (plaisir) ; ► ressentit hédonique globale.
		Présence virtuelle	Vivacité Sensorielle	► Pas d'impact décelé	► N/A	► N/A
			Concentration	► Pas d'impact décelé	► Niveau de défi perçu (flow psychologique)	► N/A
			Degré de contrôle	► Réduction (raccrochage à la réalité primaire par l'ISV).	► N/A	► N/A
	QoE	Perform.	Objectives	<ul style="list-style-type: none"> ► Augmentation des performances (fonction de la nature de l'activité) ; ► réduction du nombre d'erreur. 	► Nature de l'activité.	► N/A

TABLEAU 5.3 – Récapitulatif des résultats de l'étude - Partie 3

Qualité de l'expérience	Facteur	Variable	Impact de l'ISV	Est impacté par...	Impacte...
	Performances	Subjectives	<ul style="list-style-type: none">► Réduction engendrée par une baisse de la concentration ;► amélioration par la prise de position en tant que <i>leader</i> de l'un des utilisateurs ;► réduction par la confrontation des stratégies quant à l'activité (perte de temps).	<ul style="list-style-type: none">► Nature de l'activité ;► concentration ;► prise de position en tant que leader de l'un des utilisateurs (elle même impacté par la familiarité entre utilisateurs) ;► sujet de conversation (planification, stratégie) ;► personnalité de l'utilisateur ;► interaction entre ISV et granularité de la tâche ;	<ul style="list-style-type: none">► Ressentit hédonique globale (satisfaction, adoption technologique).
	Adoption Technologique	Intention d'utilisation	<ul style="list-style-type: none">► Pas d'impact décelé (les contextes ACV & SCV montrant un intérêt différent mais néanmoins très satisfaisant).	<ul style="list-style-type: none">► Nature de l'activité ;► qualité de l'environnement virtuel.	<ul style="list-style-type: none">► N/A

Chapitre 6

Conclusion

*« Les hommes passent leur vie à la
recherche d’eux-mêmes. On n’arrive jamais
à une conclusion définitive en ce
domaine. »*

Audur Ava Olafsdottir
Femme écrivain islandaise (1958 -)

Sommaire

6.1	Introduction	186
6.2	Réponses aux questions de recherche et validation des hypothèses . . .	187
6.3	Contributions à la connaissance	191
6.4	Limites	192
6.5	Travaux futurs	193

Ce chapitre récapitule les objectifs de notre étude, propose un retour sur la validation ou l’infirmité de nos hypothèses, apporte des réponses à nos questions de recherche ainsi qu’une vision globale de l’apport du travail entrepris. Nous y détaillerons nos contributions au corpus de connaissances du domaine et plus particulièrement aux modèles de qualité des DIMEs de Wu *et al.* (2009) et de l’expérience utilisateur de Pallot *et al.* (2013). Pour finir, il présente les enseignements à tirer de ce travail et propose des axes de recherche pour de futurs travaux.

6.1 Introduction

L’objectif premier de ce travail de thèse était d’approfondir notre connaissance quant aux potentiels impacts de l’interaction sociale vocale sur l’expérience utilisateur lors d’une immersion partagée dans un 3D-IVE. Cet objectif fut défini sur la base du constat qu’aujourd’hui, les interactions sociales sont omniprésentes dans une large proportion des facettes de nos vies. Des actions que nous menions jusqu’alors en solitaire, par exemple faire son jogging, se rendre à un entretien d’embauche ou encore lire, sont devenues sociales par l’adoption massive des réseaux sociaux. De nombreuses études menées sur l’impact du réseau internet sur la socialisation ont depuis vu le jour, en particulier en lien avec le jeu vidéo en ligne, démontrant les effets de telles expériences sur les compétences sociales ou simplement sur le plaisir et la satisfaction éprouvée (Cole et Griffiths, 2007 ; Ducheneaut et Moore, 2005 ; Sweetser et Wyeth, 2005). Les technologies de la réalité virtuelle, quant à elles, ne sont pas dédiées uniquement au loisir, de plus en plus de champs d’application émergent au fil des évolutions technologiques et de la vulgarisation de ces dernières. Ainsi, la médecine du corps ou de l’esprit, la formation, l’éducation, la conception de produits industriels sont autant de domaines où les interactions sociales sont soit requises, soit souhaitables.

La conception de ces applications de réalité virtuelle passe aujourd’hui de plus en plus par l’expertise de l’expérience utilisateur. Basée sur nombre de modèles (Morville, 2004 ; Hassenzahl *et al.*, 2000), elle propose l’étude de *« l’ensemble des réponses et des perceptions d’une personne résultant de l’usage ou de l’anticipation de l’usage d’un produit, d’un service ou d’un système »* (ISO 9241-11, 2010). Malgré de très nombreuses études sur le sujet, nous avons constaté une lacune quant à l’étude des interactions et impacts entre interactions sociales et expérience utilisateur de la réalité virtuelle. De ce constat a émergé le but principal de ce travail de thèse : mettre en exergue les leviers, liés aux interactions sociales vocales, à mettre en place pour offrir une expérience utilisateur satisfaisante dans les 3D-IVEs (**O.1**). Pour ce faire, les objectifs secondaires suivants ont été définis : apporter une meilleure compréhension des facteurs liés aux interactions sociales vocales impactant l’expérience utilisateur (**O.2**) ; mesurer certains facteurs de l’expérience utilisateur de manière objective (**O.3**).

Notre démarche était non seulement de contribuer à une meilleure connaissance générale de l’expérience utilisateur, mais surtout, d’obtenir une meilleure compréhension des liens entre interactions sociales vocales et expérience utilisateur, puis de proposer des préconisations en vue d’orienter la conception d’applications de réalité virtuelle incluant la communication vocale en adéquation avec les

besoins des utilisateurs et en concordance avec les objectifs des concepteurs.

6.2 Réponses aux questions de recherche et validation des hypothèses

En accord avec nos objectifs, nous avons exprimé deux questions de recherche :

Q.1 : *les interactions sociales vocales ont-elles un impact sur l'expérience utilisateur lors de l'immersion de plusieurs utilisateurs dans un 3D-IVE partagé ?* et ;

Q.2 : *cet impact est-il fortement soumis à la nature de la tâche, de l'activité proposée ?*

La question **Q.1** proposait l'investigation de l'impact de l'interaction sociale vocale sur l'expérience utilisateur lors d'une immersion partagée en 3D-IVE. Conscient de la nature variable de l'expérience utilisateur, cette première question se devait d'être complétée par une interrogation (**Q.2**) vis-à-vis de l'importance de la situation sur ce potentiel impact. En effet, sans tester les réponses à notre première question dans une situation différente, nous n'aurions pas été en mesure d'extraire d'éventuelles constantes quant aux impacts décelés. Toutefois, malgré les variations intégrées lors de notre étude empirique, nos résultats, et les constats formulés, restent liés aux configurations expérimentales proposées. Nous reviendrons sur ce point dans la partie *limites* de ce chapitre (section 6.4).

Afin de répondre à la question **Q.1**, nous l'avons divisée en trois questions opérationnelles. Les deux premières (**Q.1.1** et **Q.1.2**) consistaient à déterminer les possibilités de mesure de l'expérience utilisateur et, au regard de sa nature subjective, d'étudier les possibilités de mesures objectives de certains facteurs. Au travers de notre revue de littérature de l'expérience utilisateur basée sur le modèle de Wu *et al.* (2009) et de son extension par Pallot *et al.* (2013) ainsi que des phénomènes fondamentaux de la réalité virtuelle que sont la présence virtuelle, la présence sociale et le *flow* psychologique, nous avons constaté les nombreux points communs entre les facteurs de ces modèles et la littérature étudiant les phénomènes engendrés par la réalité virtuelle. Ce rapprochement nous a permis d'établir une subdivision de ces facteurs selon des variables plus aisément mesurables. Nous avons alors, afin d'établir notre méthodologie expérimentale, fait la revue des mesures envisageables pour ces variables et avons constaté que certaines d'entre elles, comme la présence virtuelle, pouvaient être mesurées de manière objective par la physiologie en étudiant l'occurrence de réponses physiologiques dues au stress dans l'environnement virtuel (Meehan *et al.*, 2002). Ces réactions, par leur similarité avec le réel, attestent du sentiment de présence. D'autres mesures objectives peuvent être établies par l'observation, plus particulièrement en ce qui concerne les performances et l'efficacité des utilisateurs au regard d'une tâche définie. En marge de ces mesures quantitatives objectives, nous avons étudié les questionnaires existants (données quantitatives subjectives) mesurant les variables liées au facteur de l'expérience utilisateur, comme par exemple le *IPQ* de Schubert *et al.* (2001) pour la présence virtuelle ou le *Networked Mind Questionnaire* de Biocca *et al.* (2001). Enfin, afin d'approcher une vision aussi holistique que possible de l'expérience utilisateur par nos résultats, nous avons proposé la mise en place de mesures qualitatives au moyen de groupes de discussion permettant l'évaluation de tous les

facteurs et variables souhaités. En définitive, à la question **Q.1.1** : *quels sont les facteurs mesurables de l'expérience utilisateur*, en accord avec la question **Q.1.2** : *est-il possible de mesurer certains de ces facteurs de manière objective*, ainsi qu'avec les modèles sur lesquels nous basons notre travail, nous répondrons *tous*. Néanmoins, la manière de les mesurer et les données collectés sont différentes. Mais, par une triangulation de ces différentes sources de données, il est possible d'obtenir une vision à la fois précise et globale de l'expérience utilisateur.

Nous nous étions aussi questionnés sur la possibilité d'une relation d'ordre entre les facteurs de l'expérience utilisateur (**Q.1.1.1**). Nos recherches bibliographiques ainsi que notre étude empirique nous ont montré que cette relation existe. Elle s'illustre parfaitement dans le modèle de qualité des DIMEs de Wu *et al.* (2009) (figure 2.15) et dans son extension par le modèle d'expérience utilisateur de Pallot *et al.* (2013) (figure 2.17). Les facteurs de la QoS sont très importants et conditionnent fortement l'expérience utilisateur : si une application n'atteint pas un niveau de QoS satisfaisant, il est très peu probable que l'expérience utilisateur en découlant satisfasse l'utilisateur. De même, les facteurs de l'expérience utilisateur immersive conditionnent la QoE. Plus qu'une relation d'ordre il faut ici parler de relation de causalité, la défaillance d'un facteur en amont entraînant une défaillance généralisée en aval. Nous avons aussi posé la question des corrélations inter facteur (**Q.1.1.2**), de l'impact d'un facteur sur l'autre. Comme l'illustre le schéma récapitulatif 5.2, nous pouvons répondre par la positive. Plus encore, nous pouvons dire que l'expérience utilisateur est constituée de ces interactions entre facteurs. Prenons l'exemple de l'amusement (degré de plaisir, résultante du *flow*) : il possède un fort impact sur le sentiment de présence virtuelle et est lui-même très fortement impacté par la nature de la tâche et la qualité générale de l'environnement virtuel. On constate aussi des impacts entre variables au sein d'un même facteur : dans le cas de la présence sociale, la compréhension mutuelle va avoir un impact sur l'interdépendance comportementale. En effet, il est plus difficile pour un utilisateur d'en influencer un autre s'ils ne sont pas en mesure de partager sur des modalités communes.

Nous nous sommes enfin interrogés sur les composantes de l'interaction sociale vocale (**Q.1.3**). Outre le sujet de conversation et les termes utilisés, qui sont néanmoins les deux composantes prépondérantes de l'interaction sociale vocale, notre étude de la littérature nous a permis de comprendre l'importance des composantes dites non verbales du langage vocal : l'intensité, le volume, le débit, l'intonation et le ton sont autant de facteurs importants vis-à-vis de la compréhension et les réactions induites par l'émission d'un message. Cette investigation nous a été utile dans un but de compréhension générale de notre sujet, mais nous avons rapidement compris que l'étude de l'interaction vocale, facteur par facteur, et de son impact sur l'expérience utilisateur en 3D-IVE, ne pourrait être traitée dans le cadre de nos travaux. En effet, une telle problématique demanderait plus qu'un travail de thèse pour être considérée de manière exhaustive. Nous avons alors traité l'interaction sociale vocale comme un ensemble que nous avons opposé, lors de notre étude empirique, à son absence. Quant à savoir s'il existe une relation d'ordre au sein de l'interaction sociale (**Q.1.3.1**), nous avons mis en évidence le fait que l'interaction sociale vocale est toujours traitée en premier par le récepteur du message, ce qui est attesté par les travaux de Delamarre (2011) ainsi que, dans notre étude empirique, par les comparaisons menées vis-à-vis de l'attention allouée et de l'impact sur le comportement entre

communication vocale et comportement (non verbal).

Enfin, nous pouvons répondre à nos questions de recherche.

Q.1 : *les interactions sociales vocales ont-elles un impact sur l'expérience utilisateur lors de l'immersion de plusieurs utilisateurs dans un 3D-IVE partagé ?*

Nous avons, par notre étude empirique, pu confronter deux contextes, l'un présentant la possibilité d'échanger vocalement l'autre non. Les résultats obtenus ont mis en exergue les impacts de l'interaction sociale vocale sur l'expérience utilisateur. Ainsi nous avons pu observer que le fait de pouvoir communiquer vocalement au sein d'un 3D-IVE avait pour effet d'améliorer le sentiment de présence sociale en réduisant les efforts à fournir afin de prêter attention à son partenaire, en améliorant la compréhension mutuelle et ainsi en augmentant l'interdépendance comportementale. Nous avons également décelé une amélioration de l'efficacité perçue et objective de nos participants lors du travail sur une tâche collaborative, ainsi qu'un accroissement de l'amusement que nous pensons dû à la possibilité avec communication vocale de faire des remarques, de l'humour, rendant l'expérience vécue plus vivante, plus dynamique, ainsi qu'à une réduction de la frustration rencontrée dans le cas de la nécessité de résoudre un problème sans pouvoir interagir vocalement. Par ailleurs, nous avons aussi constaté que les sujets de conversation pouvaient avoir un impact néfaste sur ces facteurs. Nous n'avons, dans notre étude, par nos mesures quantitatives subjectives relevé cet impact que sur la présence virtuelle au travers de sa variable degré de contrôle. Cela, en comparant les réponses aux questionnaires avec l'occurrence du thème de conversation corrélé. Autrement dit, nos participants avaient plus le sentiment d'interagir avec un périphérique externe à l'environnement plutôt que directement avec ce dernier et nous pensons fermement que cela est dû au fait qu'ils aient beaucoup échangé sur les modalités d'utilisation de ce périphérique. Ainsi, nous pouvons aisément affirmer que si l'un des participants dit « *Ce n'est vraiment pas très joli...* », « *Tu as vu ici, il y a un défaut, ça bug...* » ou encore « *Bon sang qu'est-ce que je m'ennuie...* », il est très probable que cela fasse décroître les sentiments liés aux facteurs que concernent ces sujets (ici la présence virtuelle par la qualité de l'environnement, et le *flow* psychologique par l'amusement et l'engagement). Par nos évaluations qualitatives, le même type d'effet négatif a été décelé sur l'efficacité perçue par les participants, ces derniers ayant exprimé le fait qu'en fonction de leur personnalité des conflits de vision eu égard à la stratégie à adopter pouvaient nuire à leur progression dans l'activité proposée. Toutefois, ces constatations sont soumises à la situation, en l'occurrence dans notre étude empirique, à la nature de la tâche. Nous avons, en comparant nos deux expérimentations, l'une proposant une tâche collaborative, l'autre une tâche coopérative, relevé un impact moindre de l'interaction sociale vocale sur l'efficacité dans le cadre de la tâche coopérative. Cela, tant sur l'efficacité perçue par les participants que sur celle observée au travers de leurs performances objectives. Nos résultats quantitatifs montrent aussi un effet similaire, bien que réduit, sur l'amusement et l'interdépendance comportementale. Ces éléments nous permettent de valider partiellement notre hypothèse **H.1 :** *les interactions sociales vocales ont un impact bénéfique sur l'expérience utilisateur en 3D-IVE*. Nous parlons ici de validation partielle car, en accord avec le schéma récapitulatif 5.1, nous pouvons constater qu'en majorité les impacts de

l'interaction sociale vocale sur l'expérience utilisateur sont positifs. Néanmoins et comme nous l'avons exposé précédemment dans ce paragraphe, les sujets de conversation et plus précisément le pointage de difficultés d'interactions ou de lacunes dans la qualité de l'environnement virtuel peuvent entraîner des impacts négatifs (pour ce qui est de nos relevés) sur les performances et la présence virtuelle.

Portons à présent notre attention sur notre seconde hypothèse afin par la suite de répondre à notre seconde question de recherche (**Q.2**) :

H.2 : *L'impact de l'interaction sociale vocale sur l'expérience utilisateur en 3D-IVE est dépendant de la nature de l'activité proposée.*

Nous avons observé, en particulier lors des groupes de discussions, l'importance de la nature de l'activité sur l'expérience utilisateur en 3D-IVE de manière générale. Ainsi nous savons qu'elle possède un fort impact sur le ressenti hédonique global, mais aussi sur la présence sociale au travers de sa variable interdépendance comportementale. Quant à confirmer le fait que l'impact de l'interaction sociale vocale sur l'expérience utilisateur en 3D-IVE est dépendant de la nature de l'activité proposée, nous ne pouvons réellement répondre de manière positive. En effet, une interaction existe entre la nature de l'activité et le besoin de communication vocale qu'elle implique conditionnant ainsi l'impact de cette dernière sur l'expérience utilisateur comme nous avons pu le constater en comparant activité collaborative et coopérative. Observation qui a été confirmée par les remarques émises par les participants aux groupes de discussion vis-à-vis de la granularité de la tâche. Plus cette dernière est fine, plus les échanges de messages nécessaires à sa complétion deviennent précis en conséquence. Ce qui implique la nécessité d'une compréhension mutuelle forte et immédiate, variable que nous savons largement favorisée par la communication vocale. La réponse à apporter alors est : l'impact de l'interaction sociale vocale sur l'expérience utilisateur en 3D-IVE est dépendante de la nature de l'activité en fonction de la nécessité de communiquer induite par cette dernière.

Notre seconde question de recherche : **Q.2 :** *cet impact est-il fortement soumis à la nature de la tâche, de l'activité proposée ?* questionnait de manière sous-jacente la possibilité de généralisation, de l'extension à d'autres contextes de nos résultats quant à l'impact de l'interaction sociale vocale sur l'expérience utilisateur en 3D-IVE. Nous l'avons exposé, l'expérience utilisateur est très dépendante de la situation, de la nature de la tâche et de la qualité de l'environnement. De plus, les modalités d'interaction peuvent créer des différences importantes entre deux expériences semblant pourtant très proches. Finalement, la nature subjectivement variable d'un individu à l'autre de l'expérience utilisateur est aussi largement impactée par les facteurs intrinsèques de l'utilisateur ; c'est-à-dire, ses affinités, sa personnalité ainsi que ces expériences préalables. Toutes ces dépendances, ainsi que les variations qu'elles induisent, rendent difficile la mise en évidence de constantes vis-à-vis de l'impact de l'interaction sociale vocale sur tel ou tel facteur de l'expérience utilisateur. Certes, notre étude présente des résultats significatifs et récurrents, plus particulièrement vis-à-vis de l'amusement et de la présence sociale, mais nous devons être conscients que ces derniers ont été observés dans le contexte d'un environnement virtuel accueillant deux utilisateurs sur une tâche commune et donc ne sont pas forcément généralisables à toute situation et environnement virtuel. Qui plus est, eu égard au manque

de travaux focalisés sur des problématiques proches de la nôtre, mis en évidence lors de notre phase d'état de l'art nous ne sommes pas en mesure de proposer une comparaison de notre étude avec la littérature. Ne pouvant corroborer nos résultats par ce biais, nous nous garderons d'en proposer une généralisation ou un transfert sur d'autres contextes que celui décrit dans nos travaux.

6.3 Contributions à la connaissance

Notre travail de thèse n'a pas fait évoluer de manière directe la conception globale de l'expérience utilisateur en 3D-IVE, nous ne proposons pas de nouveau modèle car nous estimons que ceux sur lesquels nous avons basé notre étude sont suffisamment pertinents. Nous n'avons pas, en effet, décelé d'incohérence quant à la fiabilité de ces modèles au cours de nos travaux. En revanche, au-delà d'avoir apporté une synthèse sur l'expérience utilisateur, par notre état de l'art et bien que les liens entre les phénomènes fondamentaux engendrés par la réalité virtuelle et les modèles d'expérience utilisateur décrits étaient déjà présents, nous pensons que les taxonomies détaillées des variables régissant ces facteurs et les possibilités de mesurer ces derniers (voir tableau récapitulatif 3.2) seront d'une grande aide pour les chercheurs désireux de mener des travaux sur l'expérience utilisateur en 3D-IVE avec une vision holistique.

Notre apport principal réside dans les résultats de notre étude empirique, nous avons démontré que les interactions sociales vocales avaient bien un impact sur l'expérience utilisateur. De plus, grâce à la taxonomie de variables employée, nous avons été en mesure de situer ces impacts. Nous avons aussi pu observer des impacts entre facteurs non directement corrélés aux interactions sociales vocales. Ces résultats apportent des connaissances quant à l'expérience utilisateur en 3D-IVE d'un point de vue plus général. Au travers des schémas 5.1 et 5.2, ainsi que du tableau récapitulatif 5.1 nous proposons une synthèse de nos résultats simple et rapide à appréhender.

Quant à offrir une aide à la conception d'application de réalité virtuelle, nous estimons que nos travaux apportent, par les constats extraits de nos résultats, une forme de préconisation sur l'utilisation des interactions sociales vocales. En effet, nous avons par exemple mis en évidence que dans le but d'atteindre un très haut niveau de concentration et d'engagement dans une tâche partagée précise et d'une faible complexité (par exemple dans le cadre d'une formation métier), il est préférable de ne pas autoriser la communication vocale. En revanche, si le but à atteindre est plus hédonique, notre étude démontre que le fait de pouvoir interagir vocalement apporte indéniablement plus de plaisir par l'amusement par la sociabilisation et le partage d'émotions qu'elle propose. Enfin, nous ne devons pas oublier le phénomène de translucence qui, comme nous l'avons constaté, permet d'atteindre des niveaux de performance très honorables. De plus, son utilisation pour l'entraînement au travail en équipe, dans un environnement virtuel, est très intéressante, puisqu'il provoque un simple glissement d'attention momentané et demande aux utilisateurs d'être plus naturellement réceptifs aux comportements des autres sans avoir à mobiliser de ressource cognitive.

Enfin, et même si cela est moins corrélé à l'expérience utilisateur, nous avons pu valider la propen-

sion de *Minecraft* à proposer une expérience immersive très satisfaisante. Ce que nous n'estimions pas être une évidence en raison de son aspect visuel très spécifique et pouvant sembler assez caricatural. La mise en place de nos expérimentations dans *Minecraft* nous permet aussi d'attester de la flexibilité du logiciel et de la facilité de l'implémentation de fonctionnalités utiles (nous pensons particulièrement aux *mods* *ComputerCraft* et *WebDisplay*). Nous ne saurions donc qu'encourager son utilisation dans le cadre d'explorations / expérimentations, car elle permet un gain de temps et d'effort important, et, en raison des possibilités de création presque illimitées, elle permet aussi une grande variété de tâches très personnalisables.

6.4 Limites

Nous avons exposé, dans les sections *analyse et synthèse* respectives de nos trois expérimentations, leurs limites. Toutefois, il est important de revenir sur ces dernières, d'un point de vue plus général car elles ont aussi vocation à être vues comme des préconisations pour la conception de futures études empiriques sur l'expérience utilisateur en 3D-IVE.

En premier lieu, nous avons pu observer la nature extrêmement variable de l'expérience utilisateur et les difficultés que pose l'extraction de constantes des impacts d'un ou de plusieurs facteurs précis. L'une des approches possibles pour pallier ces difficultés serait de mener les études sur de plus larges échantillons que ceux que nous avons pu mobiliser. Nous pensons que dans le cadre d'une expérimentation de même type que les nôtres, comparant deux contextes il serait souhaitable d'avoir quarante participants par contexte au minimum, soit quatre-vingts participants en tout.

D'autre part, et nous l'avons largement évoqué, les résultats obtenus dans le cadre de ce travail de thèse quant à l'impact des interactions sociales vocales sur l'expérience utilisateur en 3D-IVE sont restreints au contexte précis défini dans nos expérimentations. Nous sommes convaincus qu'ils pourraient être très différents si menés dans un environnement ou sur une tâche différente. Par exemple, dans un jeu de type « balle aux prisonniers », bien que nous puissions tout à fait parler d'efficacité (nombre de touches, précision), de *flow* psychologique, ainsi que de présence virtuelle et sociale, la nature de la tâche est telle que les réactions à la communication vocale pourraient être très différentes.

Pour les raisons évoquées de variabilité, de dépendance au contexte, ainsi qu'aux facteurs internes à la personne, nous avons pu constater que l'expérience utilisateur est constituée d'un ensemble de phénomènes complexes à étudier ; et nous pensons que cette complexité explique pourquoi les études rapportées sont pour la plupart focalisées sur peu de facteurs ; ce qui nous a empêchés de corroborer notre étude par la comparaison avec une littérature adoptant une approche holistique telle que nous la proposons.

6.5 Travaux futurs

Durant ces trois années, de nombreuses idées ont émergé quant aux autres travaux pouvant mener à une meilleure compréhension de l'impact des interactions sociales vocales sur l'expérience utilisateur au sein d'un 3D-IVE. Nous avons précédemment évoqué la « balle aux prisonniers », car nous pensons que ce type d'activité permettrait d'évaluer, de la même manière que nous l'avons fait, la différence entre contextes SCV et ACV pour deux utilisateurs, mais cette fois dans un cadre les plaçant non pas en collaboration ou en coopération, mais en opposition. De plus, les règles de ce type d'activité permettraient aussi de tester les mêmes variables avec beaucoup plus de participants, opposant ainsi des équipes de deux, trois, quatre participants, ou plus. Cet ajout de joueurs offrirait aussi la possibilité d'étudier les phénomènes internes à une équipe et liés à l'impact de l'interaction sociale vocale. Dans cette même logique, il serait aussi très intéressant d'étudier ces mêmes phénomènes dans le contexte des jeux de stratégie temps-réel (RTS, par exemple « *Starcraft II* »). L'occurrence de la translucence en équipe, et sa différence avec l'organisation mise en place lorsqu'il est possible de communiquer vocalement, nous semble être un terrain d'investigation prometteur dans le cadre de notre problématique.

À l'avenir, il serait aussi pertinent d'étudier plus avant les possibilités de communication non verbale émotionnelle au sein d'un environnement virtuel. Nous l'avons vu au cours de notre revue de littérature, la restitution des expressions faciales, des mouvements et autres indicateurs de l'ordre du paralangage, n'est pas toujours compatible avec l'utilisation de périphériques de réalité virtuelle tels que les casques immersifs. Notre essai d'apporter, grâce au cube émotionnel, une communication non verbale émotionnelle grâce à la physiologie nous semble intéressant vis-à-vis de la puissance des émotions face à l'expérience utilisateur. Bien que l'exécution de cette tentative dans nos expérimentations ne fut pas à la hauteur de nos attentes, nous pensons qu'il serait intéressant d'étudier les manières d'apporter des éléments émotionnel dans les expériences de réalité virtuelle par des moyens détournés, comme par exemple notre cube émotionnel, et d'en explorer l'apport.

La compréhension plus holistique dans un environnement virtuel de l'expérience utilisateur est un défi scientifique sur lequel il est aujourd'hui primordial de travailler, car l'accès du plus grand nombre à la réalité virtuelle et augmentée est en marche : sous peu, une part de plus en plus importante de la population sera équipée de diverses manières (lunette de réalité augmentée, casque immersif, cube immersif personnel, tablette, smartphone). En considérant le fait que le partage et les échanges sociaux distants ne vont faire que croître en même temps que l'accessibilité de la réalité virtuelle, nous sommes aujourd'hui convaincus que le domaine étudié dans cette thèse va prendre de l'ampleur. Nous espérons, malgré les limites et problèmes auxquels nous avons dû faire face, avoir fait quelques pas dans la direction d'une meilleure compréhension de l'impact des interactions sociales et de l'expérience utilisateur en générale.

Bibliographie

Most played PC games 2015 / Statistic, 2015.

Mary P AIKEN et Mike J BERRY : Posttraumatic stress disorder: possibilities for olfaction and virtual reality exposure therapy. *Virtual Reality*, 19(2):95–109, 2015.

Icek AJZEN : The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2):179–211, décembre 1991.

Michael ARGYLE : *Bodily communication*. Routledge, 2 édition, 1988.

Michael ARGYLE et Janet DEAN : Eye-contact, distance and affiliation. *Sociometry*, 28(3):289–304, septembre 1965.

Michael ARGYLE, Luc LEFEBVRE et Mark COOK : The meaning of five patterns of gaze. *European Journal of Social Psychology*, 4(2):125–136, juin 1974.

Jeremy BAIENSON, Kayur PATEL, Alexia NIELSEN, Ruzena BAJSCY, Sang-Hack JUNG et Gregorij KURILLO : The Effect of Interactivity on Learning Physical Actions in Virtual Reality. *Media Psychology*, 11(3):354–376, septembre 2008a.

Jeremy N BAIENSON, Andrew C BEALL et Jim BLASCOVICH : Gaze and task performance in shared virtual environments. *The Journal of Visualization and Computer Animation*, 13(5):313–320, décembre 2002.

Jeremy N BAIENSON, Andrew C BEALL, Jim BLASCOVICH, Mike RAIMUNDO et Max WEISBUCH : Intelligent Agents Who Wear Your Face: Users' Reactions to the Virtual Self. In *Third International Workshop on Intelligent Virtual Agents*, pages 86–99. Springer-Verlag, septembre 2001a.

Jeremy N BAIENSON, Andrew C BEALL et Jack M LOOMIS : Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Human Communication Research*, 31(4):511–537, octobre 2005a.

Jeremy N BAIENSON, Andrew C BEALL, Jack M LOOMIS et Jim BLASCOVICH : Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(4):428–441, août 2004.

Jeremy N BAIENSON, Jim BLASCOVICH et Andrew C BEALL : Interpersonal distance in immersive virtual environments. *Personality and Social psychology Bulletin*, 29(7):819–833, juillet 2003.

Jeremy N BAIENSON, Jim BLASCOVICH, Andrew C BEALL et Jack M LOOMIS : Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 10(6):583–598, décembre 2001b.

Jeremy N BAIENSON, Kim SWINTH, Crystal HOYT, Susan PERSKY, Alex DIMOV et Jim BLASCOVICH : The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 14(4):379–393, août 2005b.

Jeremy N BAIENSON et Nick YEE : Digital Chameleons Automatic Assimilation of Nonverbal Gestures in Immersive Virtual Environments. *Psychological science*, 16(10):814–819, 2005.

- Jeremy N BAIENSON, Nick YEE, Dan MERGET et Schroeder RALPH : The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interactions. *Presence: Teleoperators and Virtual Environments*, 15(4):359–372, septembre 2006.
- Jeremy N BAIENSON, Nick YEE, Kayur PATEL et Andrew C BEALL : Detecting digital chameleons. *Computers in Human Behavior*, 24(1):66–87, janvier 2008b.
- Rosa M BAÑOS, Cristina BOTELLA et Mariano ALCANIZ : Immersion and emotion: their impact on the sense of presence. *CyberPsychology & Behavior*, 7(6):734–741, février 2005.
- Woodrow BARFIELD et Claudia HENDRIX : The effect of update rate on the sense of presence within virtual environments. *Virtual Reality*, 1(1):3–15, 1995.
- Thomas BAUMGARTNER, Dominique SPECK, Denise WETTSTEIN, Ornella MASNARI, Gian BEELI et Lutz JÄNCKE : Feeling Present in Arousing Virtual Reality Worlds: Prefrontal Brain Regions Differentially Orchestrate Presence Experience in Adults and Children. *Frontiers in Human Neuroscience*, 2, 2008.
- Thomas BAUMGARTNER, Lilian VALKO, Michaela ESSLEN et Lutz JÄNCKE : Neural Correlate of Spatial Presence in an Arousing and Noninteractive Virtual Reality: An EEG and Psychophysiology Study. *CyberPsychology & Behavior*, 9(1):30–45, février 2006.
- Zane L BERGE : Computer Conferencing and the On-Line Classroom. *International Journal of Educational Telecommunications*, 3(1):3–21, 1997.
- Frank BIOCCA et Ben DELANEY : *Immersive virtual reality technology*. L. Erlbaum Associates Inc., juin 1995.
- Frank BIOCCA et Chad HARMS : Defining and measuring social presence: Contribution to the networked minds theory and measure. In *5th International Workshop on Presence*, pages 7–36, 2002.
- Frank BIOCCA, Chad HARMS et Jenn GREGG : The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th International Workshop on Presence 2001*, 2001.
- Alexis K BLAU : The Development of Jealousy. *Developmental Psychology*, 23(2):233–340, 2010.
- Margaret M BRADLEY et Peter J LANG : Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, mars 1994.
- Willem-Paul BRINKMAN, Dwi HARTANTO, Ni KANG, Daniel de VLIEGHER, Isabel L KAMPMANN, Nexhmedin MORINA, Paul G M EMMELKAMP et Mark NEERINCX : A virtual reality dialogue system for the treatment of social phobia. In *2012 ACM annual conference extended abstracts*, page 1099, New York, New York, USA, 2012. ACM Press.
- Tyler J BURLEIGH, Jordan R SCHOENHERR et Guy L LACROIX : Does the uncanny valley exist ? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*, 29(3):1–18, mai 2013.
- Miroslav CAVRAG, Guillaume LARIVIERE, Ana-Maria CRETU et Stephane BOUCHARD : *Interaction with virtual spiders for eliciting disgust in the treatment of phobias*. IEEE, 2014.
- Sylvain CHAGUE et Caecilia CHARBONNIER : Digital cloning for an increased feeling of presence in collaborative virtual reality environments. In *6th International Conference on D Body Scanning Technologies*, octobre 2015.
- Chuan-Yu CHANG, Chuan-Wang CHANG, Jun-Ying ZHENG et Pau-Choo CHUNG : Physiological Emotion Analysis Using Support Vector Regression. *Neurocomputing*, 110, juin 2013.
- Tania L CHARTRAND et John A BARGH : The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, 1999.
- Li-Keng CHENG, Ming-Hua CHIENG et Wei-Hua CHIENG : Measuring virtual experience in a three-dimensional virtual reality interactive simulator environment: a structural equation modeling approach. *Virtual Reality*, 18(3):173–188, janvier 2014.

- Luca CHITTARO et Riccardo SIONI : Exploring Eye-Blink Startle Response as a Physiological Measure for Affective Computing. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 227–232. IEEE, 2013.
- Ana Paula CLAUDIO, Maria Beatriz CARMO, Ines LAUREANO GOMES, Francisco ESTEVES et Augusta GASPAS : Sense of presence inside a feared (virtual) tunnel. *2015 10th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–7, 2015.
- Miriam CLEMENTE, Alejandro RODRÍGUEZ, Beatriz REY et Mariano ALCÁÑIZ : Measuring Presence During the Navigation In a Virtual Environment Using EEG. *Annual Review of Cybertherapy and Telemedicine 2013*, 191:136–140, 2013.
- Helena COLE et Mark D GRIFFITHS : Social interactions in massively multiplayer online role-playing gamers. *CyberPsychology & Behavior*, 10(4):575–583, août 2007.
- Christian COLLET, Evelyne VERNET-MAURY, Georges DELHOMME et André DITTMAR : Autonomic nervous system response patterns specificity to basic emotions. *Journal of the Autonomic Nervous System*, 62(1):45–57, décembre 1997.
- Juliet CORBIN et Anselm STRAUSS : Basics of qualitative research: Grounded theory procedures and techniques. *Basics of qualitative research: Grounded Theory procedures and techniques*, 41, 1990.
- John W CRESWELL : *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications, Inc, 3 édition, 2009.
- Mihaly CSIKSZENTMIHALYI : *Finding flow: The psychology of engagement with everyday life*. Basic Books, 1997.
- James A DANCKERT et Ava-Ann A ALLMAN : Time flies when you're having fun: Temporal estimation and the experience of boredom. *Brain and Cognition*, 59(3):236–245, décembre 2005.
- Fred D DAVIS : Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3):319–340, 1989.
- Fred D Jr. DAVIS : *A technology acceptance model for empirically testing new end-user information systems : theory and results*. Thèse de doctorat, Massachusetts Institute of Technology, 1985.
- Celso M de MELO, Jonathan GRATCH et Peter J CARNEVALE : The Effect of Agency on the Impact of Emotion Expressions on People's Decision Making. juin 2013.
- Cécile DELAMARRE : *Alzheimer et communication non verbale*. Dunod, 2011.
- Norman K Ed DENZIN et Yvonna S Ed LINCOLN : *Handbook of qualitative research*. Sage Publications, Inc, 1994.
- John V DRAPER, David B KABER et John M USHER : Speculations on the Value of Telepresence. *CyberPsychology & Behavior*, 2(4):349–362, 1999.
- Nicolas DUCHENEAUT et Robert J MOORE : More than just 'XP': learning social skills in massively multiplayer online games. *Interactive Technology and Smart Education*, 2(2):89–100, 2005.
- Nicolas DUCHENEAUT, Nicholas YEE, Eric NICKELL et Robert J MOORE : Alone together?: exploring the social dynamics of massively multiplayer online games. In *SIGCHI Conference on Human Factors in Computing Systems 2006*, pages 407–416, New York, New York, 2006.
- Nathaniel DURLACH et Richard HELD : Telepresence. *Presence: Teleoperators and Virtual Environments*, 1(1):109–112, 1992.
- Nathaniel DURLACH et Mel SLATER : Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators and Virtual Environments*, 9(2):214–217, avril 2000.
- Paul EKMAN, Wallace V FRIESEN et Phoebe ELLSWORTH : *Emotion in the Human Face*. Guidelines for Research and an Integration of Findings. 1972.
- Seymour EPSTEIN : Cognitive-Experiential Self-Theory. In *Advanced Personality*, pages 211–238. Springer US, Boston, MA, 1998.

- Thomas ERICKSON et Wendy A KELLOGG : Social translucence: an approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction*, 7(1):59–83, mars 2000.
- Anna FELNHOFER, Oswald D KOTHGASSNER, Mareike SCHMIDT, Anna-Katharina HEINZLE, Leon BEUTL, Helmut HLAVACS et Ilse KRYSPIN-EXNER : Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. *International journal of human-computer studies*, 82:48–56, octobre 2015.
- Klint FINLEY : Minecraft mod teaches kids how to mod Minecraft, août 2014. URL <http://goo.gl/0vY9yq>.
- Jodi FORLIZZI et Katja BATTARBEE : Understanding experience in interactive systems. pages 261–268. ACM, New York, New York, USA, août 2004.
- Ana Carol Pontes de FRANÇA et Marcelo Márcio SOARES : Dialogical Self on Virtual Reality Systems: Presence and Embodiment in Human Situated Interaction. In *6th International Conference on Applied Human Factors and Ergonomics*, pages 6444–6450, 2015.
- Daniel FREEMAN, Katherine PUGH, Angus ANTLEY, Mel SLATER, Paul BEBBINGTON, Matthew GIT-TINS, Graham DUNN, Elizabeth KUIPERS, David FOWLER et Philippa GARETY : Virtual reality study of paranoid thinking in the general population. *The British Journal of Psychiatry*, 192(4):258–263, avril 2008.
- Jonathan FREEMAN, S E AVONS, Ray MEDDIS et Don E PEARSON : Using behavioral realism to estimate presence: A study of the utility of postural responses to motion stimuli. *Presence: Teleoperators and Virtual Environments*, 9(2):149–164, avril 2000.
- Jonathan FREEMAN, S E AVONS, Don E PEARSON et Wijnand A IJSSELSTEIJN : Effects of Sensory Information and Prior Experience on Direct Subjective Ratings of Presence. *Presence: Teleoperators and Virtual Environments*, 8(1):1–13, février 1999.
- Jonathan FREEMAN, Jane LESSITER, Edmund KEOGH et Jules DAVIDOFF : A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, 10(3):282–297, juin 2001.
- Chris FRITH et Uta FRITH : Theory of mind. *Current Biology*, 15(17):R644–R645, septembre 2005.
- Yoren GAFFARY : *Communication kinesthésique des émotions dans un contexte d'interaction homme-machine*. Thèse de doctorat, Université Paris Sud - Paris XI, juin 2015.
- Andrea GAGGIOLI, Marta BASSI et Antonella DELLE FAVE : Quality of experience in virtual environments. pages 122–132. Being There: Concepts, effects and measurement of user presence in synthetic environments, 2003.
- Gregory GAHM, Greg REGER, Mary V INGRAM, Mark REGER et Albert RIZZO : A Multisite, Randomized Clinical Trial of Virtual Reality and Prolonged Exposure Therapy for Active Duty Soldiers with PTSD. Rapport technique ADA599211, Geneva Foundation Tacoma WA, décembre 2013.
- Maia GARAU : *The impact of avatar fidelity on social interaction in virtual environments*. Thèse de doctorat, University of London, London, United Kingdom, 2003.
- Maia GARAU, Mel SLATER, David-Paul PERTAUB et Sharif RAZZAQUE : The Responses of People to Virtual Humans in an Immersive Virtual Environment. *Presence: Teleoperators and Virtual Environments*, 14(1):104–116, février 2005.
- Azucena GARCIA-PALACIOS, H G HOFFMAN, Albert CARLIN, Thomas FURNESS et Cristina BOTELLA : Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour Research and Therapy*, 40(9):983–993, septembre 2002.
- Alexandra Livia GEORGESCU, Bojana KUZMANOVIC, Daniel ROTH, Gary BENTE et Kai VOGLEY : The Use of Virtual Characters to Assess and Train Non-Verbal Communication in High-Functioning Autism. *Frontiers in Human Neuroscience*, 8, 2014.
- Erik GESLIN : *Processus d'induction d'émotions dans les environnements virtuels et le jeu vidéo*. Thèse de doctorat, Ecole nationale supérieure d'arts et métiers - ENSAM, juin 2013.

- Nadine GLAS et Catherine PELACHAUD : Definitions of Engagement in Human-Agent Interaction. *In 6th International Conference on Affective Computing and Intelligent Interaction*, pages 944–949, septembre 2015.
- Erving GOFFMAN : *Behavior in Public Places*. Simon and Schuster, janvier 1963.
- Rosanna E GUADAGNO, Jim BLASCOVICH et Jeremy N BAIENSON : Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, 10(1), 2007.
- Antal HAANS et Wijnand IJSSELSTEIJN : Mediated social touch: a review of current research and future directions. *Virtual Reality*, 9(2-3):149–159, décembre 2005.
- Antal HAANS et Wijnand A IJSSELSTEIJN : Embodiment and telepresence: Toward a comprehensive theoretical framework. *Interacting with Computers*, 24(4):211–218, juillet 2012.
- Edward Twitchell HALL : *The hidden dimension*. Doubleday & Co, 1966.
- Béatrice S HASLER, Gilad HIRSCHBERGER, Tal SHANI-SHERMAN et Doron A FRIEDMAN : Virtual Peacemakers: Mimicry Increases Empathy in Simulated Contact with Virtual Outgroup Members. *Cyberpsychology, Behavior, and Social Networking*, 17(12):766–771, décembre 2014.
- Marc HASSENZAHL, Axel PLATZ, Michael BURMESTER et Katrin LEHNER : *Hedonic and ergonomic quality aspects determine a software's appeal*. ACM, New York, New York, USA, avril 2000.
- Elaine HATFIELD, John T CACIOPPO et Richard L RAPSON : Emotional Contagion. *Current Directions in Psychological Science*, 2(3):96–100, juin 1993.
- Elaine HATFIELD, John T CACIOPPO et Richard L RAPSON : *Emotional contagion*. Cambridge university press, 1994.
- Carrie HEETER : Being there: the subjective experience of presence. *Presence: Teleoperators and Virtual Environments*, 1(2):262–271, mai 1992.
- Robert A HINDE : *Non-verbal Communication*. Cambridge University Press, novembre 1975.
- Donna L HOFFMAN et Thomas P NOVAK : “Emergent Consumers” Can Help Develop Successful Future Ideas. Rapport technique, juillet 2008.
- Mohammed E HOQUE, Matthieu COURGEON, Jean-Claude MARTIN et Bilge MUTLU : MACH: My Automated Conversation coach, juin 2013. URL <http://goo.gl/D424Kl>.
- Eben HOWARD, Clint COOPER, Mike P WITTIE, Steven SWINFORD et Qing YANG : Cascading impact of lag on Quality of Experience in cooperative multiplayer games. *2014 13th Annual Workshop on Network and Systems Support for Games (NetGames)*, pages 1–6, 2014.
- Matthew HUDSON, Paul CAIRNS et A Imran NORDIN : Familiarity in Team-Based Online Games: The Interplay Between Player Familiarity and the Concepts of Social Presence, Team Trust, and Performance. *In Advances in Visual Informatics*, pages 140–151. Springer International Publishing, Cham, novembre 2015.
- Chelsea Marie HUGHES : A measure of social behavior in team-based, multiplayer online games: the sociality in multiplayer online games scale (SMOG). 2015.
- Isabelle HUPONT, Joaquin GRACIA, Luis SANAGUSTIN et Miguel Angel GRACIA : How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift. *In 2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, pages 1–6. IEEE, mai 2015.
- Wijnand IJSSELSTEIJN, Ilse BIERHOFF et Yvonne Slangen-de KORT : Duration Estimation and Presence . *In 4th International Workshop on Presence 2001*, 2001.
- Wijnand A IJSSELSTEIJN et Giuseppe RIVA : Being There: The experience of presence in mediated environments. pages 3–16. Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments (Emerging Communication: Studies in New Technologies and Practices in Communication, Vol. 5), 2003.

- ISO 9241-11 : Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV) - Partie 11: Lignes directrices relatives à l'utilisabilité, mars 1998.
- ISO 9241-11 : Ergonomie de l'interaction homme-système - Partie 210: Conception centrée sur l'opérateur humain pour les systèmes interactifs, mars 2010.
- ISO/IEC 25010 : Génie du logiciel - Qualité des produits - Partie 1: Modèle de qualité, juin 2001.
- ISO/IEC 25010 : Ingénierie des systèmes et du logiciel - Exigences de qualité et évaluation des systèmes et du logiciel (SQuaRE) - Modèles de qualité du système et du logiciel, mars 2011.
- Gunnar JOHANSSON : Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2):201–211, juin 1973.
- Ankur KALRA et Karrie KARAHALIOS : TextTone: Expressing Emotion Through Text. In *Human-Computer Interaction - INTERACT 2005*, pages 966–969. Springer Berlin Heidelberg, Berlin, Heidelberg, septembre 2005.
- Anu KANKAINEN : *Thinking model and tools for understanding user experience related to information appliance product concepts*, volume 118 de *Acta polytechnica Scandinavica. Ma, Mathematics and computing series*. Alto University Library, septembre 2002.
- Konstantina KILTENI, Raphaëla GROTEN et Mel SLATER : The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):372–387, 2012.
- Taeyong KIM et Frank BIOCCA : Telepresence via Television: Two Dimensions of Telepresence May Have Different Connections to Memory and Persuasion. *Journal of Computer-Mediated Communication*, 3(2):0–0, juin 1997.
- KOICHI : Mining For Japanese Gold: The Professor Who Teaches Japanese Through Minecraft, mai 2014. URL <http://goo.gl/5bDBBd>.
- Jeppe KOMULAINEN, Jari TAKATALO, Miikka LEHTONEN et Göte NYMAN : Psychologically structured approach to user experience in games. In *5th Nordic conference on Human-computer interaction: building bridges*, pages 487–490, New York, New York, USA, octobre 2008. ACM.
- Joung Huem KWON, John POWELL et Alan CHALMERS : How level of realism influences anxiety in virtual reality environments for a job interview. *International journal of human-computer studies*, 71(10):978–987, octobre 2013.
- Mats LARSSON, Nancy L PEDERSEN et Håkan STATIN : Associations between iris characteristics and personality in adulthood. *Biological Psychology*, 75(2):165–175, mai 2007.
- Effie Lai-Chong LAW, Virpi ROTO, Marc HASSENZAHL, Arnold P O S VERMEEREN et Joke KORT : Understanding, scoping and defining user experience. In *SIGCHI Conference on Human Factors in Computing Systems 2009*, pages 719–728, New York, New York, USA, 2009. ACM Press.
- Chung K LEE, S K YOO, Yoon J PARK, Nam Hyun KIM, Kee Sam JEONG et ByungChae LEE : Using Neural Network to Recognize Human Emotions from Heart Rate Variability and Skin Resistance. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 5523–5525. IEEE, 2006.
- Kwan Min LEE : Presence, Explicated. *Communication Theory*, 14(1):27–50, février 2004.
- Rensis LIKERT : A technique for the measurement of attitudes. *Archives of Psychology*, 1932.
- Y LIU, James M RITCHIE, Theodore LIM, Zoe KOSMADOUDI, Aparajithan SIVANATHAN et Raymond C W SUNG : A fuzzy psycho-physiological approach to enable the understanding of an engineer's affect status during CAD activities. *Computer-Aided Design*, 54:19–38, septembre 2014.
- Yisi LIU, Olga SOURINA et Minh Khoa NGUYEN : Real-Time EEG-Based Human Emotion Recognition and Visualization. In *2010 International Conference on Cyberworlds*, pages 262–269. IEEE, octobre 2010.
- Matthew LOMBARD et Theresa DITTON : At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2):0, septembre 1997.

- Joe LUDWIG : Lessons learned porting Team Fortress 2 to Virtual Reality, mai 2013. URL <http://goo.gl/wzblKi>.
- Jean-Luc LUGRIN, Johanna LATT et Marc Erich LATOSCHIK : Avatar Anthropomorphism and Illusion of Body Ownership in VR. In *IEEE VR 2015*, 2015.
- Pengcheng LUO, Victor NG-THOW-HING et Michael NEFF : An Examination of Whether People Prefer Agents Whose Gestures Mimic Their Own, juin 2013. URL <http://goo.gl/rYBuxj>.
- Abraham H MASLOW : A theory of human motivation. *Psychological Review*, 50(4):370–396, juillet 1943.
- Alison MCMAHAN : *Immersion, engagement and presence*. The video game theory reader, routledge member of the taylor and francis group édition, 2003.
- Michael MEEHAN, Brent INSKO et Mary WHITTON : Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics*, 21(3):645–652, juillet 2002.
- Marvin MINSKY : Telepresence. *OMNI magazine*, 2:45–51, 1980.
- Hamed MONKARESI et Rafael A CALVO : Using Remote Heart Rate Measurement for Affect Detection. In *Workshop Affect, Compagnon Artificiel, Interaction*, 2014.
- Masahiro MORI, Karl F MACDORMAN et Norri KAGEKI : The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100, juin 2012.
- Peter MORVILLE : User Experience Design, juin 2004. URL <http://goo.gl/HFXQJx>.
- Klara NAHRSTEDT, Ahsan AREFIN, Raoul RIVAS, Pooja AGARWAL, Zixia HUANG, Wanmin WU et Zhenyu YANG : QoS and resource management in distributed interactive multimedia environments. *Multimedia Tools and Applications*, 51(1):99–132, janvier 2011.
- Arturo NAKASONE, Helmut PRENDINGER et Mitsuru ISHIZUKA : Emotion Recognition from Electromyography and Skin Conductance. *5th International Workshop on Biosignal Interpretation*, pages 219–222, 2005.
- András NELTZ : Beautiful Minecraft Village Helps Soldier Deployed in Afghanistan, juillet 2014. URL <http://goo.gl/E9I0Pk>.
- Paul NORRIS et Seymour EPSTEIN : The investigation of some fundamental issues concerning rational-analytical and intuitive-experiential thinking styles with a short form of the rational-experiential inventory. *Manuscript submitted for publication*, 2002.
- Paul NORRIS et Seymour EPSTEIN : An Experiential Thinking Style: Its Facets and Relations With Objective and Subjective Criterion Measures. *Journal of Personality*, 79(5):1043–1080, octobre 2011.
- Thomas P NOVAK et Donna L HOFFMAN : The fit of thinking style and situation: New measures of situation-specific experiential and rational cognition. *Journal of Consumer Research*, 36(1):56–72, 2009.
- Kristine L NOWAK et Frank BIOCCA : The effect of the agency and anthropomorphism on users’ sense of telepresence, copresence, and social presence in virtual environments. 12(5):481–494, octobre 2003.
- Julie PALLANT : *SPSS Survival Manual*. A Step by Step Guide to Data Analysis Using SPSS. SPSS Survival Manual, novembre 2010.
- Marc PALLOT : *Collaborative distance: investigating issues related to distance factors affecting collaboration performance*. Thèse de doctorat, Nottingham University Business School, Nottingham, United Kingdom, juillet 2011.
- Marc PALLOT, Rémy EYNARD, Benjamin POUSSARD, Olivier CHRISTMANN et Simon RICHIR : Augmented Sport: Exploring Collective User Experience. In *2013 Virtual Reality International Conference*, Laval, mars 2013.

- Miles L PATTERSON, Anthony WEBB et Warren SCHWARTZ : Passing Encounters: Patterns of Recognition and Avoidance in Pedestrians. *Basic and Applied Social Psychology*, 24(1):57–66, 2002.
- Markus PERSSON : Clearing up the world size math., mars 2010. URL <http://goo.gl/ZaufrN>.
- David-Paul PERTAUB, Mel SLATER et Chris BARKER : An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience. *Presence: Teleoperators and Virtual Environments*, 11(1):68–78, février 2002.
- Panagiotis C PETRANTONAKIS et Leontios J HADJILEONTIADIS : Emotion Recognition From EEG Using Higher Order Crossings. *IEEE Transactions on Information Technology in Biomedicine*, 14(2):186–197, octobre 2009.
- Dominique PICARD : Transition et ritualité dans l'interaction sociale. *Connexions*, 76(2):81–93, 2001.
- Marie-Louise PIERSON : *L'intelligence relationnelle*. Éditions d'Organisation, 2nd édition, 2003.
- Jayesh S PILLAI, Colin SCHMIDT et Simon RICHIR : Achieving presence through evoked reality. *Frontiers in Psychology*, 4(86), 26 2013.
- Heng Yu PING, Lili Nurliyana ABDULLAH, Alfian Abdul HALIN et Puteri Suhaiza SULAIMAN : A Study of Physiological Signals-based Emotion Recognition Systems. *International Journal of Computers & Technology*, 11(1):2189–2196, septembre 2013.
- Lukasz PIWEK, Lawrie S MCKAY et Frank E POLLICK : Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition*, (130):271–277, janvier 2014.
- Robert PLUTCHIK : *Emotion: A Psychoevolutionary Synthesis*. Longman Higher Education, mars 1980.
- Nilesh U POWAR : *An Approach for the Extraction of Thermal Facial Signatures for Evaluating Threat and Challenge Emotional States*. Thèse de doctorat, University of Dayton, 2013.
- Robert R PROVINE, Robert J SPENCER et Darcy L MANDELL : Emotional Expression Online Emoticons Punctuate Website Text Messages. *Journal of Language and Social Psychology*, 26(3):299–307, septembre 2007.
- PSYTECH : PHOBOS | Anxiety Management Virtual Reality Platform for CBT and research, 2015. URL <http://goo.gl/o6QkTg>.
- Fatemeh RABIEE : Focus-group interview and data analysis. *Proceedings of the Nutrition Society*, 63(04):655–660, mars 2007.
- Pierre RAINVILLE, Antoine BECHARA, Nasir NAQVI et Antonio R DAMASIO : Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, 61(1):5–18, juillet 2006.
- Magnus REVANG : User Experience Project: The User Experience Wheel, avril 2007. URL <http://goo.gl/nXxkKv>.
- Iman Mohammad REZAZADEH, Xiangyu WANG, Mohammad FIROOZABADI et Mohammad Reza HASHEMI GOLPAYEGANI : Using affective human-machine interface to increase the operation performance in virtual construction crane training system: A novel approach. *Automation in Construction*, 20(3):289–298, mai 2011.
- Phattanon RHIENMORA, Peter HADDAWY, Siriwan SUEBNUKARN et Matthew N DAILEY : Intelligent dental training simulator with objective skill assessment and feedback. *Artificial Intelligence in Medicine*, 52(2):115–121, juin 2011.
- Giuseppe RIVA, Fabrizia MANTOVANI, Claret Samantha CAPIDEVILLE, Alessandra PREZIOSA, Francesca MORGANTI, Daniela VILLANI, Andrea GAGGIOLI, Cristina BOTELLA et Mariano ALCANIZ : Affective Interactions Using Virtual Reality: The Link between Presence and Emotions. *CyberPsychology & Behavior*, 10(1):45–56, février 2007.
- Jean-Marc ROBERT et Annemarie LESAGE : From usability to user experience with interactive systems. *Handbook of Human-Machine Interaction*, 2010.

- Alejandro RODRÍGUEZ, Beatriz REY et Mariano ALCANIZ : Evaluating virtual reality mood induction procedures with portable EEG devices. *Annual Review of Cybertherapy and Telemedicine* 2013, 191:131–135, 2013.
- Alejandro RODRÍGUEZ, Beatriz REY, Miriam CLEMENTE, Maja WRZESIEN et Mariano ALCANIZ : Assessing brain activations associated with emotional regulation during virtual reality mood induction procedures. *Expert Systems with Applications*, 42(3):1699–1709, février 2015.
- Cristian RUSU, Virginica RUSU, Silvana RONCAGLIOLO et Carina GONZÁLEZ : Usability and User Experience: What Should We Care About ? *International Journal of Information Technologies and Systems Approach (IJITSA)*, 8(2):1–12, juillet 2015.
- Mhd Yamen SARAJI, Charith Lasantha FERNANDO, M FURUKAWA, Kouta MINARNIZAWA et Susumu TACHI : Real-time egocentric superimposition of operator's own body on telexistence avatar in virtual environment. In *2013 23rd International Conference on Artificial Reality and Telexistence*, pages 35–39. IEEE, 2013.
- Mhd Yamen SARAJI, Charith Lasantha FERNANDO, Kouta MINAMIZAWA et Susumu TACHI : Mutual hand representation for telexistence robots using projected virtual hands. In *6th Augmented Human International Conference*, pages 221–222, New York, New York, USA, mars 2015. ACM.
- MHD Yamen SARAJI, Yusuke MIZUSHINA et Charith L FERNANDO : Enforced telexistence. *ACM SIGGRAPH* 2014, 2014.
- Ayse Pinar SAYGIN, Thierry CHAMINADE, Hiroshi ISHIGURO, Jon DRIVER et Chris FRITH : The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive and Affective Neuroscience*, 7(4):nsr025–422, avril 2011.
- Dominique L SCAPIN, Bernard SENACH, Brigitte TROUSSE et Marc PALLOT : User Experience: Buzzword or New Paradigm ? In *Fifth International Conference on Advances in Computer-Human Interactions 2012*, Valencia, Spain, 2012.
- Wilbur SCHRAMM : *How communication works*, volume 3. The process and effects of mass communication, 1954.
- Thomas SCHUBERT et Frank FRIEDMANN : Embodied presence in virtual environments. pages 268–278. Visual representations and interpretations, 1999.
- Thomas SCHUBERT, Frank FRIEDMANN et Holger REGENBRECHT : Decomposing the sense of presence: Factor analytic insights. In *2nd International Workshop on Presence 1999*, janvier 1999.
- Thomas SCHUBERT, Frank FRIEDMANN et Holger REGENBRECHT : The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281, juin 2001.
- Thomas SCHUBERT, Holger REGENBRECHT et Frank FRIEDMANN : Real and illusory interactions enhance presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 11(4):425–434, août 2002.
- Ulrike SCHULTZE : Embodiment and presence in virtual worlds: a review. *Journal of Information Technology*, 25:434–449, 2010.
- Gary E SCHWARTZ et Daniel A WEINBERGER : Cardiovascular Differentiation of Happiness, Sadness, Anger, and Fear Following Imagery and Exercise1. *Psychosomatic Medicine*, 1981.
- Jonmichael SEIBERT : An exploratory study on virtual reality head mounted displays and their impact on player presence. Mémoire de D.E.A., mai 2014.
- Jun'ichiro SEYAMA et Ruth S NAGAYAMA : The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces. *Presence: Teleoperators and Virtual Environments*, 16(4):337–351, juillet 2007.
- Thomas B SHERIDAN : Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments*, 1(1):120–126, 1992.

- Dong-Hee SHIN, Frank BIOCCA et Hyunseung CHOO : Exploring the user experience of three-dimensional virtual learning environments. *Behaviour & Information Technology*, 32(2):203–214, février 2013.
- Daniel SHORT : Teaching Scientific Concepts Using a Virtual World–Minecraft. *Teaching Science*, 58(3):55–58, septembre 2012.
- Marko SIITONEN : *Social interaction in online multiplayer communities*, volume 74. Jyväskylä studies in humanities, 2007.
- Rajita SINHA, Willam R LOVALLO et Oscar A PARSONS : Cardiovascular differentiation of emotions. *Psychosomatic Medicine*, 54(4):422, 1992.
- Panote SIRIARAYA et Chee SIANG ANG : Age differences in the perception of social presence in the use of 3D virtual world for social interaction. *Interacting with Computers*, 24(4):280–291, 2012.
- Mel SLATER : Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 8(5):560–565, mars 1999.
- Mel SLATER : Presence and the sixth sense. *Presence: Teleoperators and Virtual Environments*, 11(4):435–439, août 2002.
- Mel SLATER : Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3549–3557, 2009.
- Mel SLATER, Beau LOTTO, Maria Marta ARNOLD et Maria V SANCHEZ-VIVES : How we experience immersive virtual environments: the concept of presence and its measurement. *Anuario de psicología/The UB Journal of Psychology*, 40(2):193–210, 2009.
- Mel SLATER, David-Paul PERTAUB, Chris BARKER et David M CLARK : An Experimental Study on Fear of Public Speaking Using a Virtual Environment. *CyberPsychology & Behavior*, 9(5):627–633, octobre 2006.
- Mel SLATER et Anthony STEED : A virtual presence counter. *Presence: Teleoperators and Virtual Environments*, 9(5):413–434, octobre 2000.
- Mel SLATER, Martin USOH et Anthony STEED : Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 3(2):130–144, 1994.
- Mel SLATER et Sylvia WILBUR : A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 6(6):603–616, 1997.
- Matthew J SMITH, Emily J GINGER, Katherine WRIGHT, Michael A WRIGHT, Julie Lounds TAYLOR, Laura Boteler HUMM, Dale E OLSEN, Morris D BELL et Michael F FLEMING : Virtual Reality Job Interview Training in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 44(10):2450–2463, mai 2014.
- Mohammad SOLEYMANI, Sadjad ASGHARI-ESFEDEN, Maja PANTIC et Yun FU : Continuous emotion detection using EEG signals and facial expressions. In *2014 IEEE International Conference on Multimedia and Expo*, pages 1–6. IEEE, janvier 2014.
- Jason STEADMAN, Chelsea BOSKA, Crystal LEE et Xuan-Shi LIM : Using Popular Commercial Video Games in Therapy with Children and Adolescents. *Journal of Technology in Human Services*, 32(3):201–219, 2014.
- Gregory STEEL et Scott L JONES : Using virtual environments for synchronous online courses. *Journal of Teaching and Learning with Technology*, 2(1):56–61, juin 2013.
- Michael STENGEL, Steve GROGORICK, Elmar EISEMANN, Martin EISEMANN et Marcus MAGNOR : An Affordable Solution for Binocular Eye Tracking and Calibration in Head-mounted Displays. In *ACM Multimedia Conference 2015*, 2015.
- Jonathan STEUER : Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4):73–93, décembre 1992.

- David W STEWART et Prem N SHAMDASANI : *Focus Groups: Theory and Practice*, volume 20. SAGE Publications, Inc, 3 édition, 2015.
- Suphat SUKAMOLSON : Fundamentals of quantitative research. Rapport technique, 2007.
- Penelope SWEETSER et Peta WYETH : GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment*, 3(3):3–3, juillet 2005.
- THE 3DLIVE CONSORTIUM : 3D Live European Project, 2015. URL <http://3dliveproject.eu/wp/>.
- Stefano VALENZI, Tanvir ISLAM, Peter JURICA et Andrzej CICHOCKI : Individual Classification of Emotions Using EEG. *Journal of Biomedical Science and Engineering*, 2014(08):604–620, juin 2014.
- Jan B F van ERP et Alexander TOET : Social Touch in Human–Computer Interaction. *Front. Digit. Humanit.*, 2(3):603, mai 2015.
- Marc VASS, John M CARROLL et Clifford A SHAFFER : *Supporting creativity in problem solving environments*. ACM, New York, New York, USA, octobre 2002.
- P WATZLAWICK, J H BEAVIN et D D JACKSON : *Une logique de la communication*. Points édition, 1972.
- Wanhui WEN, Guangyuan LIU, Nanpu CHENG, Jie WEI, Pengcheng SHANGGUAN et Wenjin HUANG : Emotion Recognition Based on Multi-Variant Correlation of Physiological Signals. *Affective Computing, IEEE Transactions on*, 5(2):126–140, 2014.
- B J WINER, D R BROWN et K M MICHELS : *Statistical Principles in Experimental Design*. McGraw Hill, New York, 3 édition, 1991.
- Bob G WITMER et Michael J SINGER : Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, juin 1998.
- Wanmin WU, Ahsan AREFIN, Raoul RIVAS, Klara NAHRSTEDT, Renata SHEPPARD et Zhenyu YANG : Quality of experience in distributed interactive multimedia environments: toward a theoretical framework. In *ACM Multimedia Conference 2009*, pages 481–490, New York, New York, USA, 2009. ACM Request Permissions.
- Ning YANG, Rajani MURALEEDHARAN, Johannah KOHL, Ilker DEMIRKOL, Wendi HEINZELMAN et Melissa STURGE-APPLE : Speech-based emotion classification using multiclass SVM with hybrid kernel and thresholding fusion. In *2012 IEEE Workshop on Spoken Language Technology Workshop*, pages 455–460, Miami, Florida, décembre 2012.
- Nick YEE : Motivations for play in online games. *CyberPsychology & Behavior*, 9(6):772–775, décembre 2006.
- Nick YEE, Jeremy N BAIENSON, Mark URBANEK, Francis CHANG et Dan MERGET : The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments. *CyberPsychology & Behavior*, 10(1):115–121, février 2007.

Annexe A

Figures supplémentaires

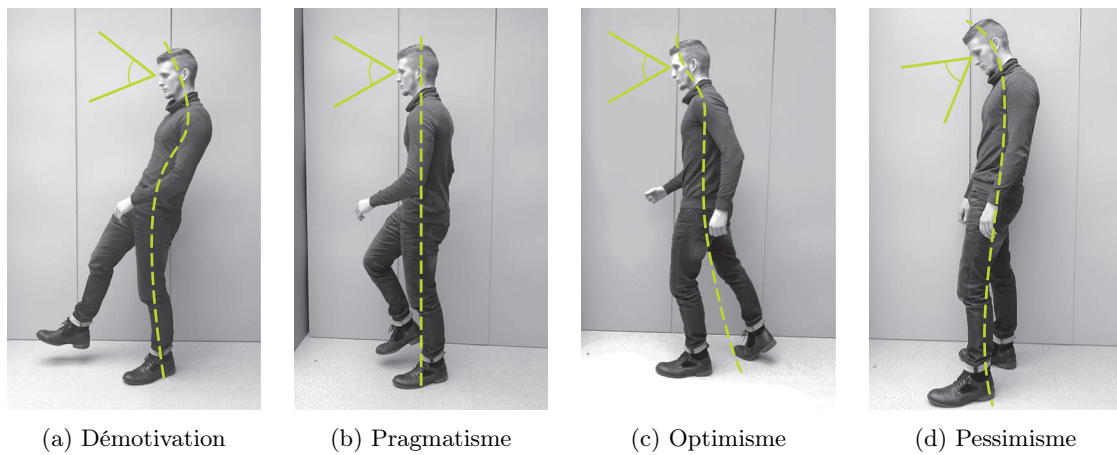


FIGURE A.1 – Différentes démarches ou attitudes de marche.

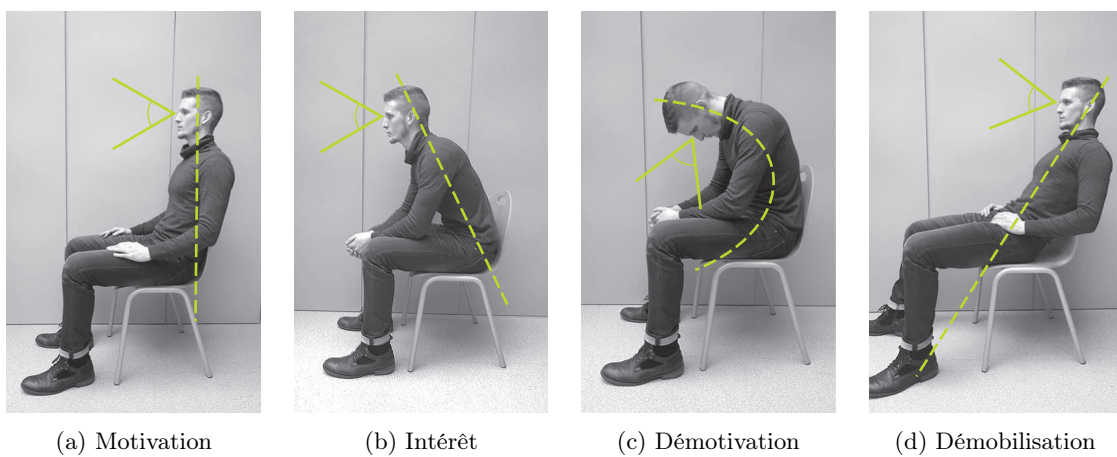


FIGURE A.2 – Différentes attitudes en position assise.

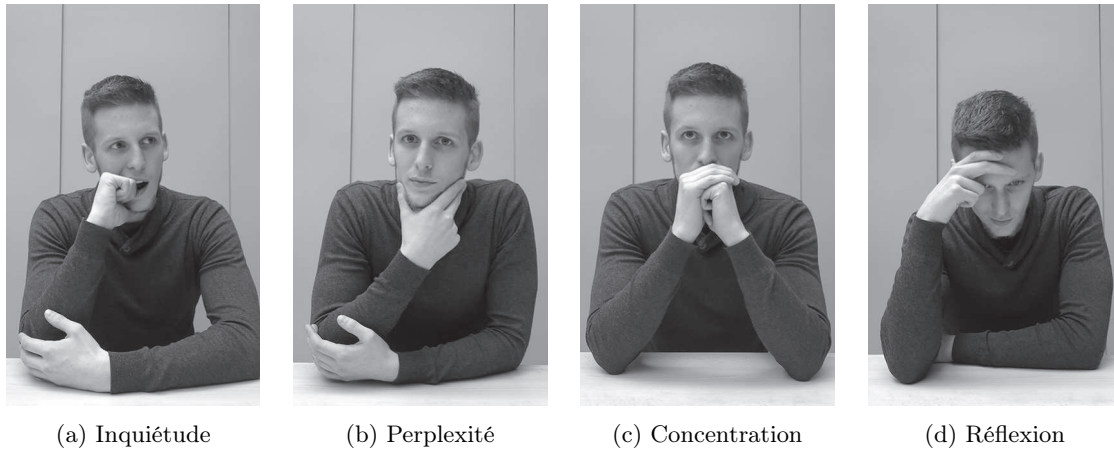


FIGURE A.3 – Différents gestes associés à leur état d'esprit.

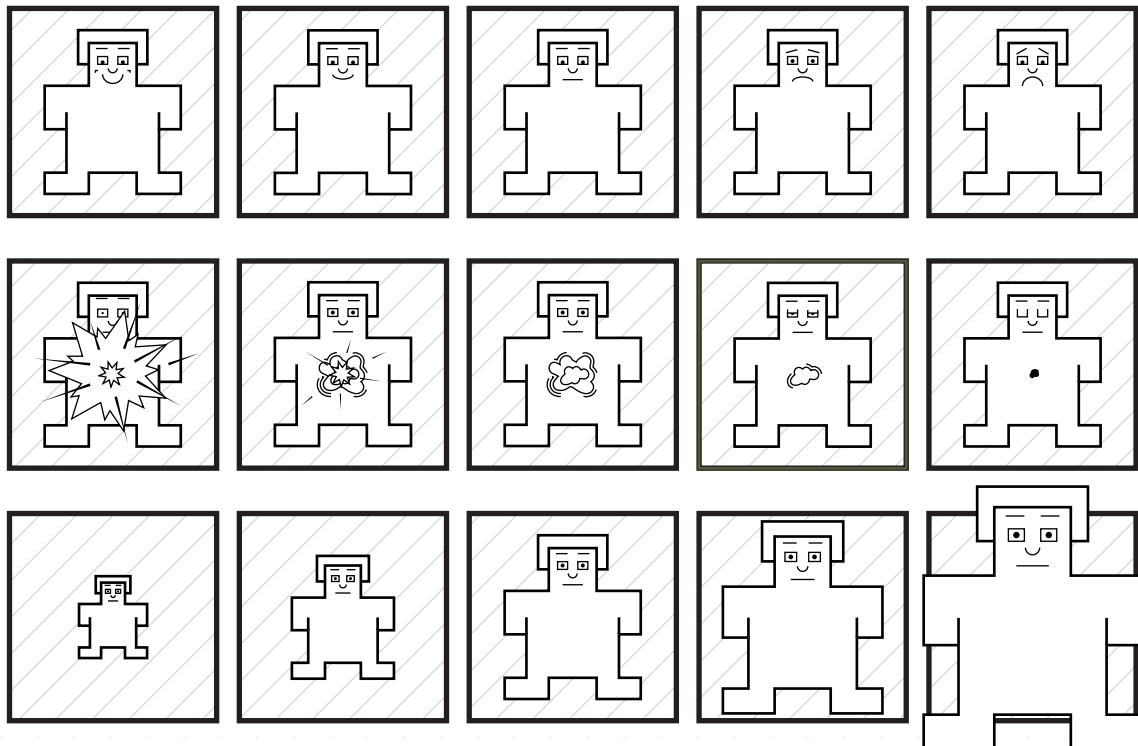


FIGURE A.4 – Le questionnaire *Self-Assessment Manikin* (SAM) utilisé pour évaluer les émotions selon leurs valence (haut), stimulation (milieu), et dominance (bas) Bradley et Lang (1994).

Annexe B

Annexes à l'expérimentation 1

B.1 Environnement virtuels



FIGURE B.1 – Panneaux d'informations dans l'environnement *Minecraft* de l'expérimentation 1.



FIGURE B.2 – Système de blocage du temps et des intempéries.

B.2 Configuration Minecraft

Logiciels externes à *Minecraft* nécessaire à la plateforme d'expérimentation :

- a) **TeamSpeak** : logiciel permettant la communication vocale au travers d'un réseau local ou du réseau internet. Deux versions du logiciel sont nécessaires, une version *serveur*, à installer sur le poste dédié à cet effet, ainsi que la version *client* pour les deux postes participants. La configuration est très simple, il suffit aux deux logiciels *clients* de se connecter à l'IP du logiciel *serveur*. Cette IP étant celle du poste *serveur* lui-même.

↓ **Logiciel** : <https://www.teamspeak.com/downloads>

- b) **JoyToKey** : logiciel permettant de transmettre à la manette de jeu les touches du clavier désirées. Le fichier de configuration *JoyToKeyMinecraft.cfg* disponible au téléchargement à la suite doit être placé dans le dossier d'installation du logiciel soit ... \JoyToKey_en\ . La configuration nommée *Minecraft* apparaîtra alors automatiquement à l'ouverture du logiciel.

↓ **Logiciel** : <http://joytokey.net/en/download>

↓ **Fichier de configuration :**

<http://remy-eynard.net/minecraft-exp/Ressources/JoyToKeyMinecraft.cfg>

Ressources pour la configuration de *Minecraft* - *Client* :

- a) Il n'est normalement plus nécessaire d'installer *Java* avant *Minecraft*, ce dernier embarquant les bibliothèques nécessaires directement dans son logiciel d'installation. Toutefois, en cas de malfonction, installer *Java* en prenant garde à la configuration système (32 ou 64 bits).

↓ **Bibliothèque *Java*** : <https://www.java.com/fr/>

- b) Télécharger *Minecraft*, il est nécessaire d'être dépositaire d'une licence (19,95€.) En l'occurrence, dans le cadre de notre configuration, trois licences sont nécessaires, une pour chaque participant, et une pour le poste *serveur* de l'expérimentateur.

↓ **Logiciel** : <https://minecraft.net/download>

- c) Installer *Minecraft*, puis le lancer. Sur le *launcher* présenté en figure B.3, cliquer sur le bouton « *Edit Profil* » en bas à gauche de la fenêtre. Puis dans le menu déroulant « *User version :* » choisir la version de *Minecraft* souhaitée. Ce choix doit se faire en fonction des *mods* à utiliser. Pour rappel, afin d'utiliser l'*Oculus Rift*, nous avons utilisé le *mod Minecraft* alors disponible uniquement pour la version 1.6.2 (aujourd'hui compatible avec la version 1.7.10). Puis lancer *Minecraft* pour le fermer une fois arrivé sur le menu. Ce lancement est un simple prérequis pour la suite des actions à mener.

- d) Télécharger et installer *Minecraft Forge*. *Minecraft Forge* est une bibliothèque de fonctionnalités nécessaire à l'utilisation de très nombreux *mod*, notamment *Minecrafter*. Choisir la version compatible avec la version de *Minecraft* choisi précédemment. *Minecraft Forge* propose un logiciel d'installation ou sera demandé de choisir entre version *client* et *serveur*. Ici, installer la version *client* et conserver le logiciel d'installation pour le poste *serveur*.

↓ **Logiciel** : <http://files.minecraftforge.net/>

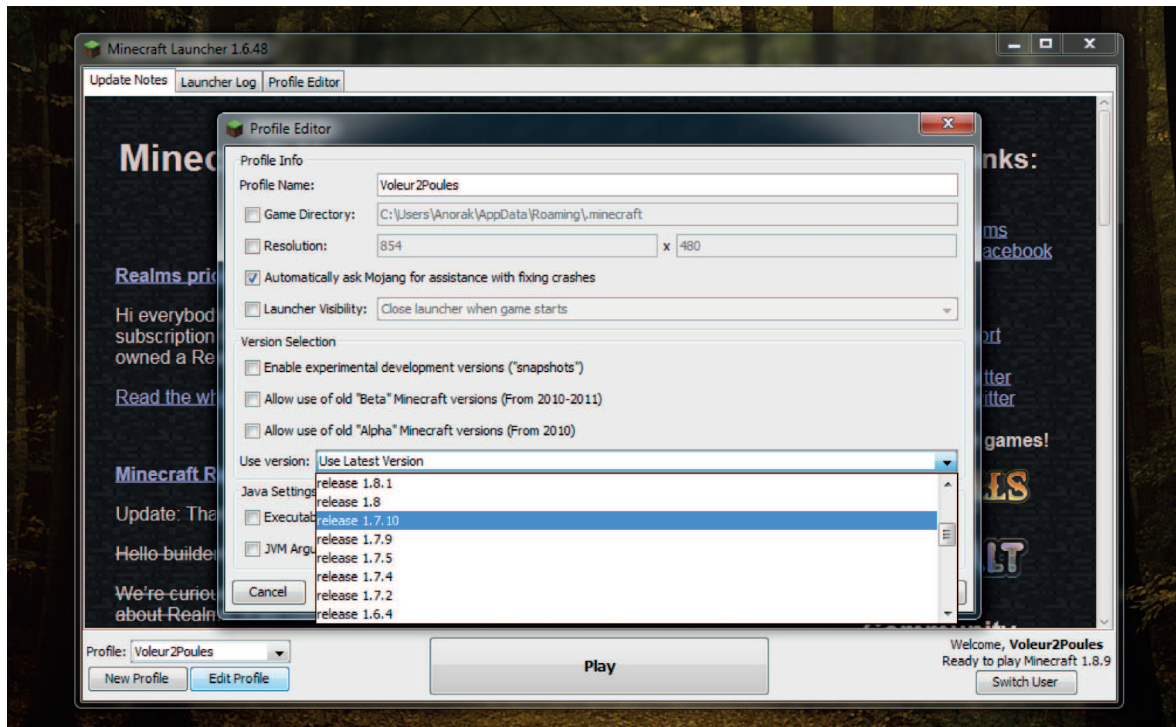


FIGURE B.3 – Minecraft launcher, choix de la version

- e) Télécharger et installer *Optifine*. *Optifine* est un *mod* utilitaire ajoutant de nombreuses fonctionnalités de configuration et d'optimisation à *Minecraft*. Son usage n'est ici pas obligatoire, mais en fonction de la machine utilisée et du fait de l'utilisation de la stéréoscopie, il peut permettre de sauvegarder des ressources matérielles et ainsi de gagner quelques précieuses images par seconde (pour afficher les IPS dans *Minecraft*, presser la touche F3). Ici encore, attention à la version de *Minecraft* utilisée, nous préconisons l'utilisation de la version standard du *mod* nommée C1.

↓ **Mod :** <http://optifine.net/downloads>

- f) Télécharger et installer *Minecrist*. *Minecrist* est le *mod* permettant l'utilisation de *Minecraft* avec l'*Oculus Rift*. Plusieurs versions sont disponibles, en fonction des versions de *Minecraft* mais aussi du support des *Razer Hydra*¹, qui en l'occurrence ne nous intéresse pas. Ici encore, un logiciel d'installation facile est proposé. Il faut être vigilant lors de l'installation à cocher la case « *Install with Forge* » (la version de *Forge* installée apparaît dans le menu déroulant jouxtant la case). Lors du prochain lancement de *Minecraft*, afin que *Minecrist* fonctionne, il faudra choisir le profil lui correspondant comme nous l'avons vu en figure B.3, soit ici : *release minecrist-x.x.x-bxxx-forge-nohydra*.

↓ **Mod :** <http://www.mtbs3d.com/phpbb/viewtopic.php?f=172&t=17489?1#p124997>

- g) Télécharger et installer *ComputerCraft*. *ComputerCraft* est le *mod* ajoutant les ordinateurs utilisés pour chronométrer le temps d'immersion et terminer la session par la téléportation des participants, ainsi que la *turtle* utilisée pour placer le *nexus* nécessaire à la seconde partie de l'expérimentation. Ici encore, choisir la version compatible avec la version de *Minecraft* utilisée. L'installation est ici manuelle, le dossier du *mod* devra être placé dans

1. *Razer Hydra* : Périphérique de contrôle spécifiquement développé pour l'utilisation des casques immersifs. Voir : <http://www.razer.com/support/gaming-controllers/razer-hydra/>

C:\Users\Utilisateur\AppData\Roaming\.minecraft\mods\.

↓ **Mod :** <http://www.computercraft.info/download/>

- h) Télécharger et installer *Invasion*. *Invasion* est le *mod* ajoutant les fonctionnalités liées au *nexus* et permettant la tenue de la seconde partie de l'expérimentation. La procédure d'installation est la même que pour *ComputerCraft*.

↓ **Mod :** <http://www.minecraft-france.fr/mod-invasion-mod-1-7-10/>

- i) Télécharger et installer le pack de textures *Sphax Pure BDCraft*. Le pack de textures est disponible en plusieurs résolutions différentes allant de 16 x 16 pixels à 512 x 512 pixels, il est primordial de choisir en fonction des ressources matérielles disponibles. Tout comme pour les *mods*, le dossier du pack de textures devra être placé dans le dossier : ...\.minecraft\resourcepacks\.

↓ **Pack de textures :** <http://bdcraft.net/purebdcraft-minecraft>

- j) Remplacer un fichier du pack de textures *Sphax Pure BDCraft*. Afin de permettre l'affichage de nos panneaux d'informations, il suffit de remplacer un fichier du pack de textures par un fichier modifié par nos soins. Fichier à remplacer :

...\.minecraft\resourcepacks\BD Craft\assets\minecraft\textures\painting\paintings_kristoffer_zetterstrand.png.

↓ **Fichier modifié :** <http://remy-eynard.net/minecraft-exp/Ressources/Plans.zip>

- k) Lancer *Minecraft* en prenant soin de bien sélectionner le bon profil. Cliquer sur « *Multijoueur* » puis « *Connexion rapide* ». Entrer l'adresse du serveur dont nous allons à présent voir la configuration. Les postes participants sont correctement configurés pour notre première expérimentation.

Ressources pour la configuration de *Minecraft* - Serveur :

- a) Télécharger le fichier **.jar* permettant l'exécution d'un serveur *Minecraft*. Placer ce fichier dans un dossier dédié (l'emplacement n'a pas d'importance) puis l'exécuter une fois et fermer le programme.

↓ **Logiciel :** <https://minecraft.net/download>

- b) La configuration d'un serveur *Minecraft* passe par l'édition d'un fichier présent dans le dossier du serveur. Le fichier *server.properties* est à éditer avec n'importe quel logiciel d'édition de texte (tel que *Notepad ++* par exemple). C'est dans ce fichier qu'il est possible de définir le mode jeu (survie ou créatif), la difficulté, l'apparition des monstres ou non, etc... Et surtout l'IP du serveur. Pour plus de facilité, le fichier *server.properties* que nous avons utilisé est disponible au téléchargement en fin de ce paragraphe, la seule variable à modifier pour qui voudrait reproduire notre configuration est l'adresse IP du serveur, voir la ligne *server-ip=—.—.—.—*.

↓ **Fichier :** <http://remy-eynard.net/minecraft-exp/Ressources/server.properties>

- c) Installer la version serveur de *Minecraft Forge*. Même procédure que pour le point d. de la section précédente, il faut néanmoins cette fois sélectionner la version serveur et donner le chemin du dossier où est installé le serveur.

- d) Installer les *mods* sur le serveur. Copier tous les *mods* précédemment installés sur les postes clients dans le dossier ... *AppData\Roaming\minecraft\mods* dans le dossier ... *\mods* du serveur.
- e) Installer un environnement virtuel sur le serveur. Copier le contenu du dossier téléchargé dans le dossier ... *\world* du serveur.
↓ **Dossier :** <http://remy-eynard.net/minecraft-exp/Ressources/EV-1.zip>
- f) Installer la version client de *Minecraft* sur le poste serveur afin de permettre à l'expérimentateur de lancer l'expérimentation. Suivre les points de la section précédente en occultant le point f. concernant le *mod Minecraft*.
- g) Lancer le serveur en cliquant sur le fichier *forge-(version)-universal.jar*.
- h) Lancer la version client de *Minecraft* et se connecter au serveur en suivant la procédure du point k. de la section précédente puis en appuyant sur « *Rejoindre* ».
- i) Une fois connecté à l'environnement virtuel, deux commandes permettent d'accéder directement aux zones participant et expérimentateur de l'environnement. Pour ce faire, ouvrir la zone de discussion en pressant la touche « *T* » et entrer les commandes suivantes : Pour rejoindre la zone de construction, */tp 197 65 28*. Pour la zone de contrôle */tp 155 48 144*.
- j) Connecter les deux postes participants au serveur en suivant les instructions du point k. de la section précédente, puis en appuyant sur « *Rejoindre* ».

B.3 Livret d'informations aux participants

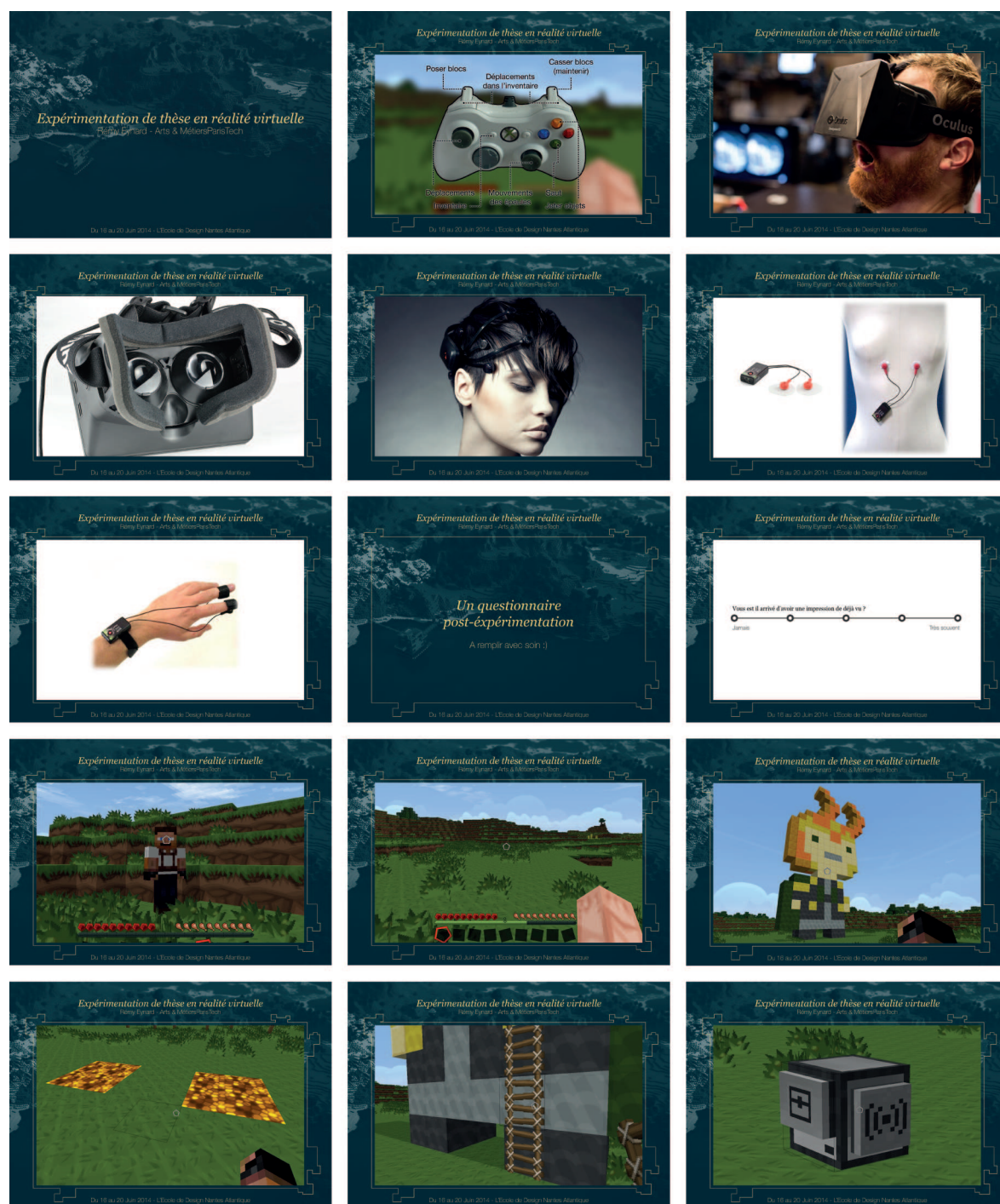


FIGURE B.4 – Livret d'informations aux participants - Expérimentation 1 - partie 1.

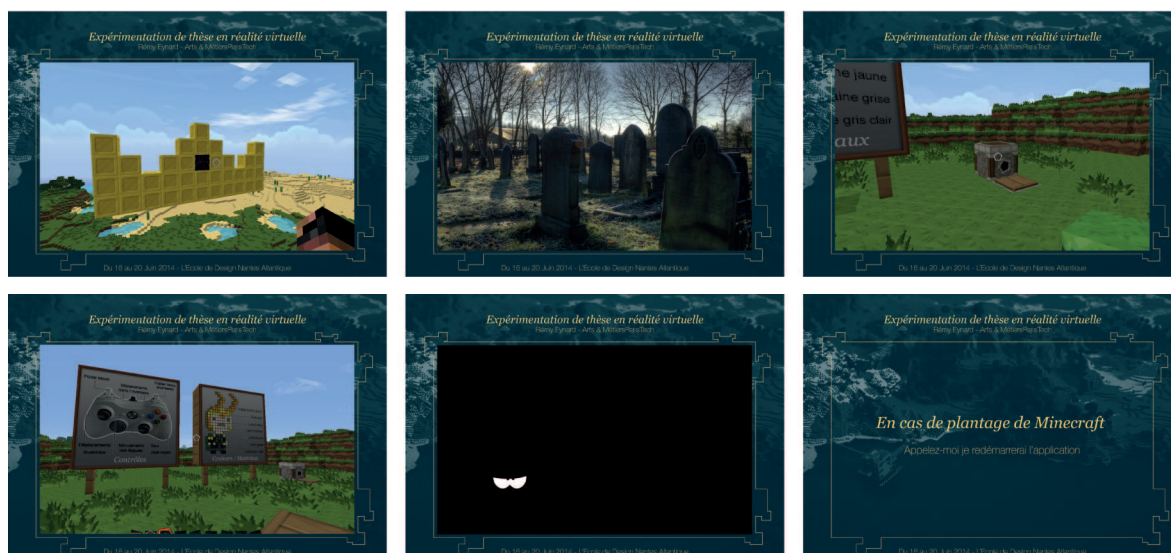


FIGURE B.5 – Livret d'informations aux participants - Expérimentation 1 - partie 2.

B.4 Questionnaire Pré Expérimentation

Voir en page suivante.

Questionnaire

Pré expérimentation

À propos de vous :

Nom :

Prénom :

Âge :

Genre : ☐ Femme ☐ Homme

Francophone confirmé : ☐ Oui ☐ Non

Êtes-vous sujet au mal des transports : ☐ Oui ☐ Non

Avez-vous déjà porté un casque immersif comme « l'Oculus Rift » : ☐ Oui ☐ Non

Si oui avez-vous ressenti un malaise lors de cette utilisation : ☐ Oui ☐ Non

Jouez-vous aux jeux vidéo (toutes plateformes) : ☐ Oui ☐ Non

Si oui à quelle fréquence (par semaine) : ☐ - de 5 heures ☐ de 5 à 15 heures ☐ plus de 15 heures

Avez vous déjà joué à des jeux en ligne de type MMORPG (toutes plateformes) : ☐ Oui ☐ Non

Si oui à quelle fréquence (par semaine) : ☐ - de 5 heures ☐ de 5 à 15 heures ☐ plus de 15 heures

Sur quel type de support jouez-vous :

☐ Ordinateur ☐ Console de salon ☐ Tablette ☐ Téléphone mobile

Connaissez-vous le jeu *Minecraft* : ☐ Oui ☐ Non

Si oui avez-vous déjà joué en multijoueurs : ☐ Oui ☐ Non

Si oui pouvez-vous grossièrement estimer votre temps de jeu total :

☐ - de 10 heures ☐ de 10 à 50 heures ☐ + de 50 heures

Réservé à l'encadrant de l'expérimentation :

Date : ☐ 16/06 ☐ 17/06 ☐ 18/06 ☐ 19/06 ☐ 20/06

Heure :

Numéro de session d'expérimentation :

Numéro de la salle d'expérimentation : ☐ 01 ☐ 02

☐ O.C

☐ L.B

☐ R.E

B.5 Questionnaire Post Expérimentation

Les questions 20 à 33 n'étaient présentes que dans la version du questionnaire remplie pour les participants ayant passé l'expérimentation dans le contexte ACV. Pour faciliter la lecture, les numéros des questions concernées sont en **orange et gras**.

Numéro de session d'expérimentation :

Numéro de la salle d'expérimentation : ☐ 01 ☐ 02

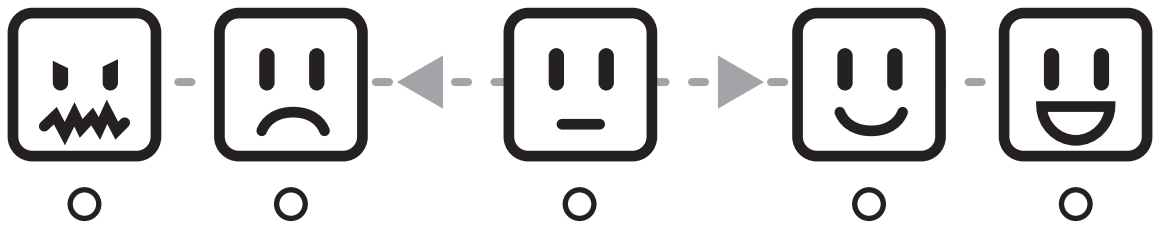
Nom :

Prénom :

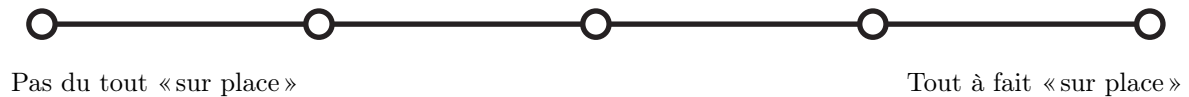
Questionnaire

Post-expérimentation

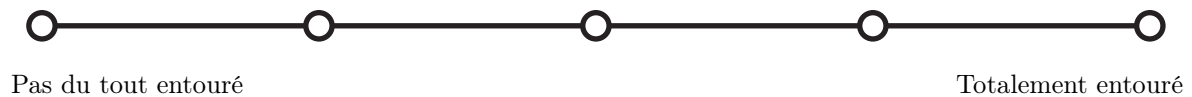
1. Votre ressenti global vis-à-vis de cette expérience :



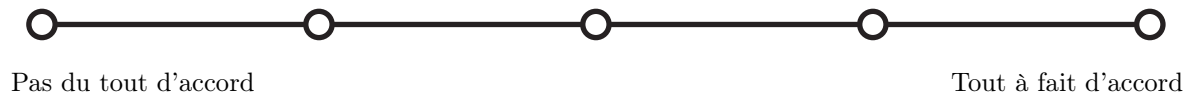
2. Dans l'environnement virtuel, vous aviez le sentiment « d'être sur place, présent » :



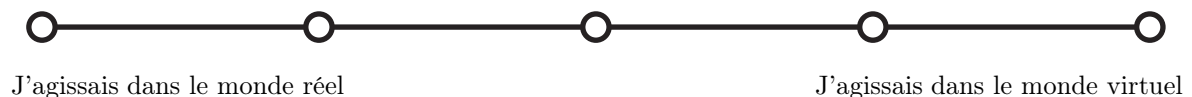
3. D'une certaine façon, vous aviez l'impression que le monde virtuel vous entourait :



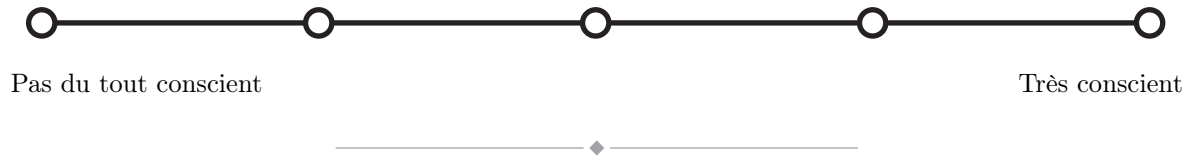
4. Vous aviez l'impression d'être simplement en train de percevoir des images :



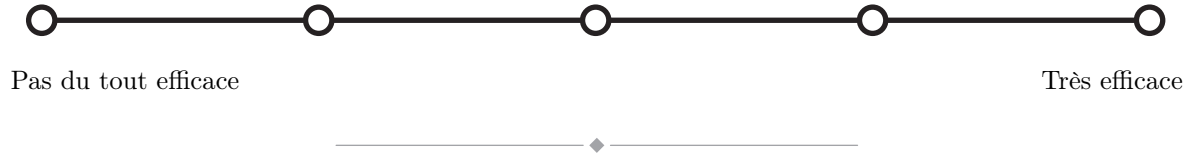
5. Vous avez eu la sensation d'agir dans l'environnement virtuel plutôt que d'agir sur un quelconque mécanisme à l'extérieur de celui-ci :



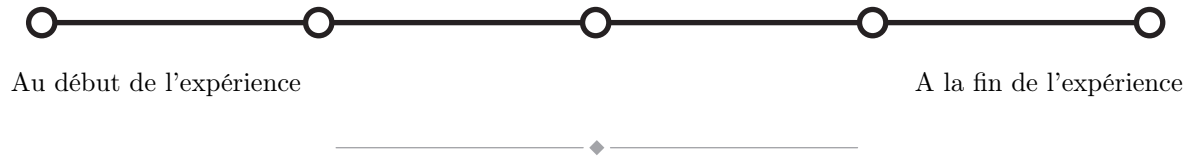
6. Vous étiez toujours conscient de l'environnement réel durant l'expérience ?



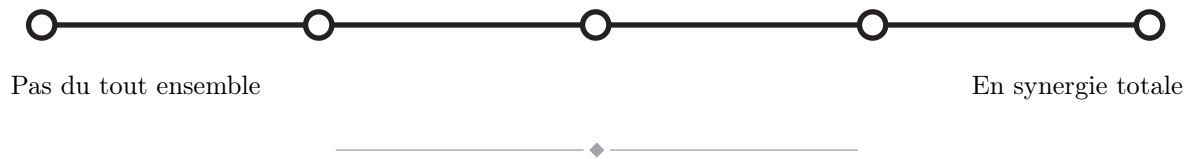
7. Vous avez été efficace vis-à-vis de l'accomplissement de la tâche :



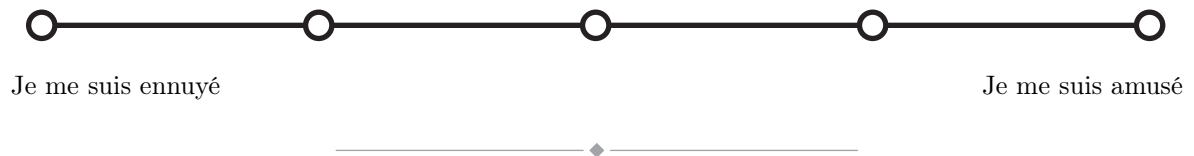
8. Vous étiez à votre niveau de performance maximum :



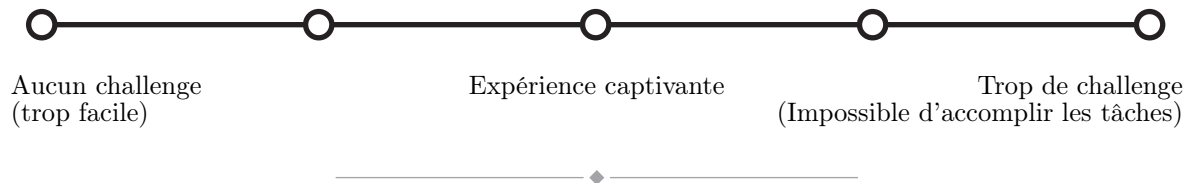
9. Vous avez eu le sentiment de travailler avec votre partenaire sur la tâche :



10. Je me suis amusé durant cette expérience :

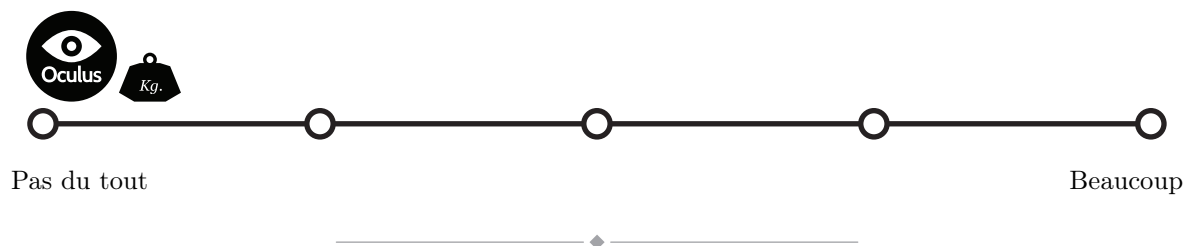


11. Vous avez ressenti une forme de challenge durant cette expérience :

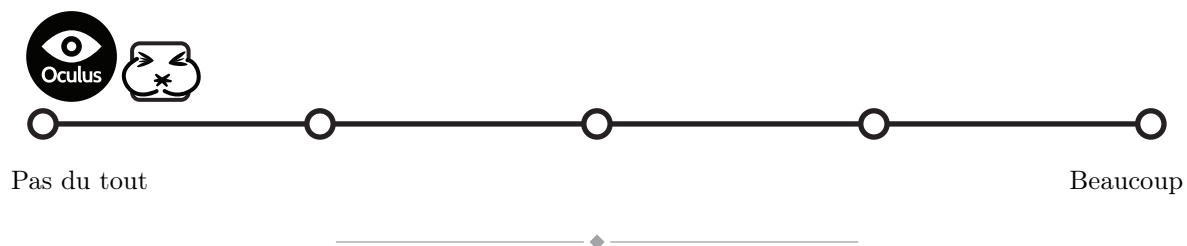


12. Votre performance a-t-elle été altérée par le matériel dont vous étiez équipé :

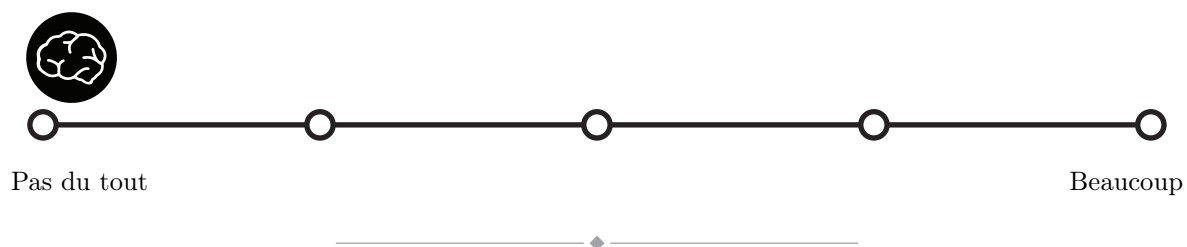
12.1 Le poids de l'Oculus Rift :



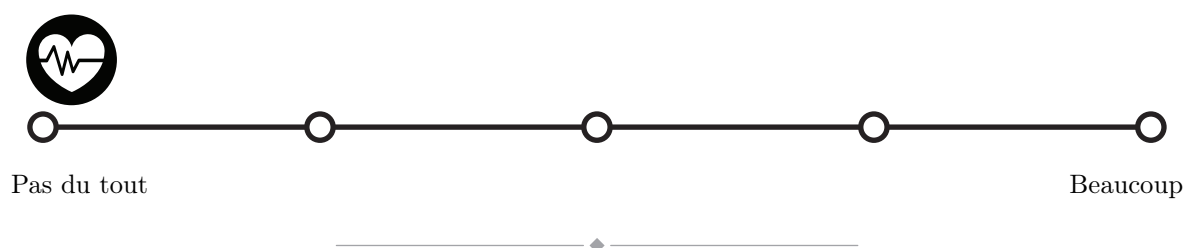
12.2 Une sensation de la mal aise due à l'Oculus Rift :



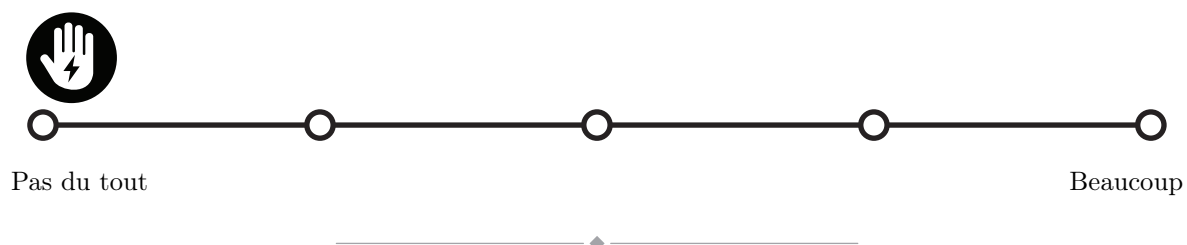
12.3 Le casque à électroencéphalogramme :



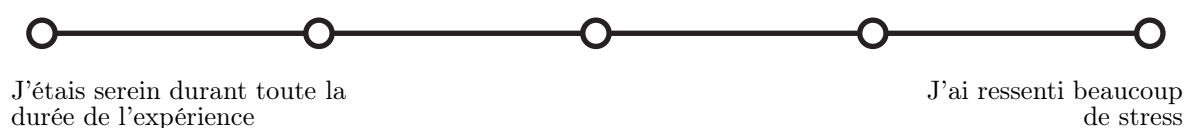
12.4 Le capteur de rythme cardiaque :




12.5 Le capteur de conductance de la peau :



13. Vous avez ressenti du stress durant votre expérience :




14. Vous avez fait preuve d'une attention particulièrement vive vis-à-vis de l'autre personne :




Je n'ai pas été très attentif à l'autre J'ai été très attentif à l'autre

15. L'autre personne vous a paru comprendre ce que vous désiriez faire (actions) :




Pas du tout compris Très bien compris

16. J'ai compris ce que désirait l'autre personne (actions) :




Pas du tout compris Très bien compris

17. Vos actions étaient dépendantes de celles de l'autre personne :




J'ai agi uniquement en fonction de moi J'ai totalement agi en fonction des actions de l'autre

18. Votre style de jeu (par exemple : utilisation des échelles, manière de construire la statue) était une réponse directe au style de jeu de l'autre personne :



J'ai agi de manière totalement autonome Je me suis adapté sur le style de jeu de l'autre

19. Votre partenaire vous a aidé dans l'accomplissement de la tâche :

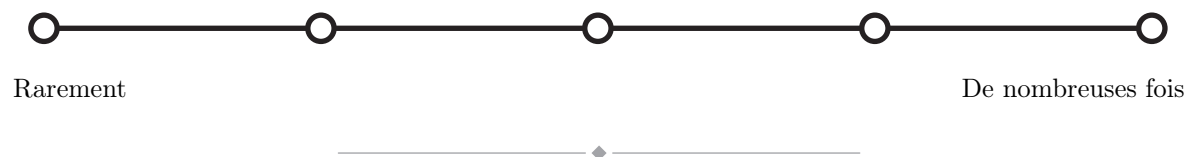


Pas du tout aidé Beaucoup aidé

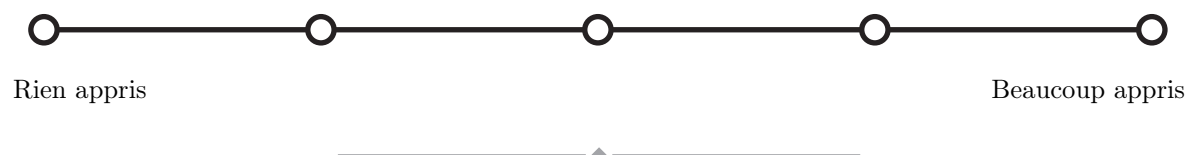
20. Vous avez interagi avec votre partenaire (conversation) : ☐ Oui ☐ Non

Si oui :

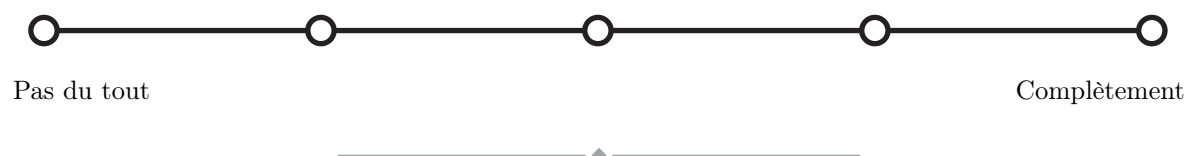
21. Vous avez interagi à la fréquence suivante :



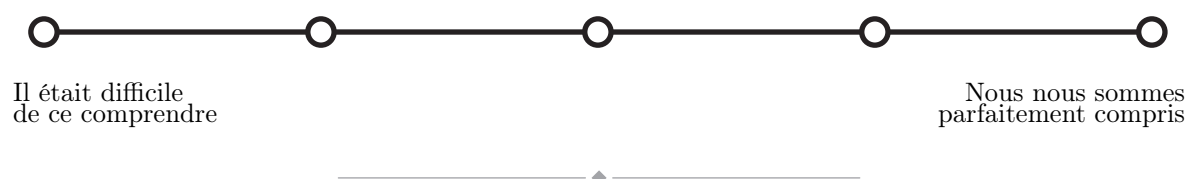
22. Votre partenaire vous a appris des choses durant l'expérience :



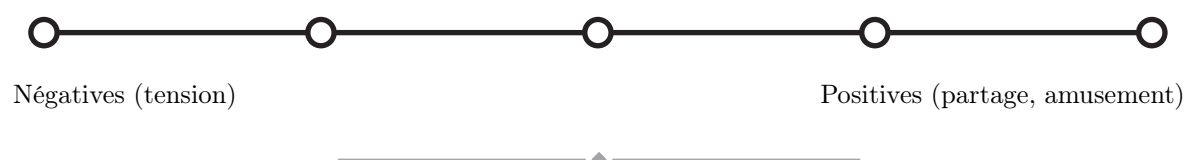
23. Vous avez, au travers de vos conversations, ressenti l'état émotionnel de votre partenaire :



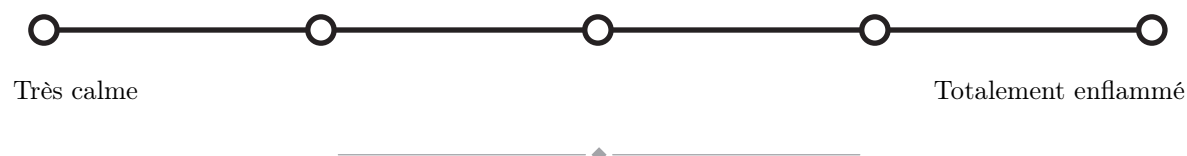
24. Vous avez ressenti un manque de compréhension entre vous et votre partenaire :



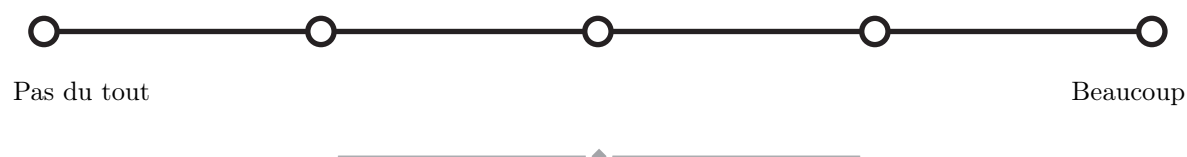
25. Vos discussions étaient plutôt :




26. Le degré d'intensité de vos conversations était :




27. Vous vous êtes mutuellement encouragés durant l'expérience :




28. Vous vous êtes mutuellement conseillés vis-à-vis de la tâche à accomplir :



Pas du tout Beaucoup




Sur quels thèmes avez-vous échangés ?


29. La réalisation de la tâche demandée :


Pas du tout Beaucoup





30. Votre ressenti vis-à-vis de l'expérience en cours :


Pas du tout Beaucoup





31. L'environnement virtuel :


Pas du tout Beaucoup




32. Les modalités d'interaction (utilisation de la manette et du jeu Minecraft) :



Pas du tout Beaucoup




33. D'autres sujets non corrélés à l'expérience en cours (possibilité de remplir les champs ci-dessous) :


.....

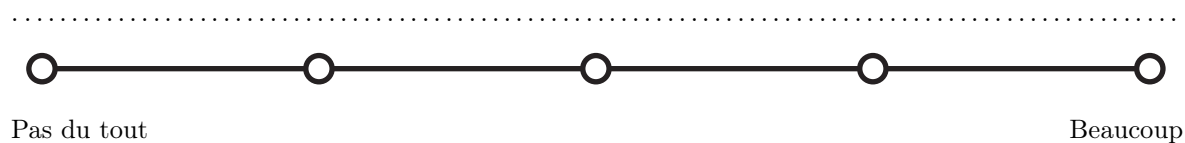
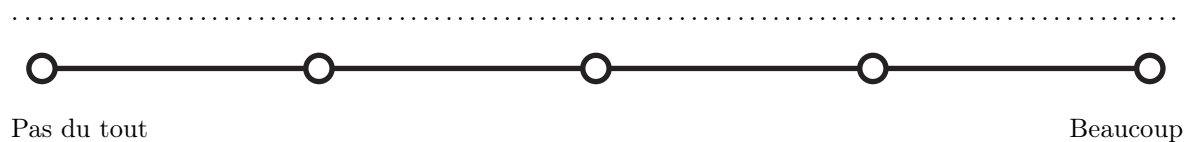
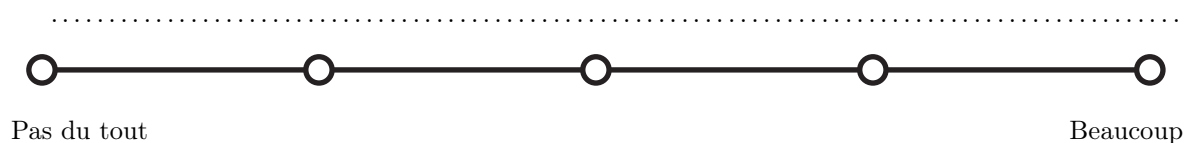

Pas du tout Beaucoup



.....


Pas du tout Beaucoup





34 . Merci d'inscrire ici toutes remarques que vous désiriez faire vis-à-vis de cette expérimentation :

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



**Fin du questionnaire,
merci de votre participation.**

B.6 Questionnaire a posteriori de l'expérimentation

Ce questionnaire a été adressé aux participants de l'expérimentation par courriel via l'outil « *Google Forms* ».

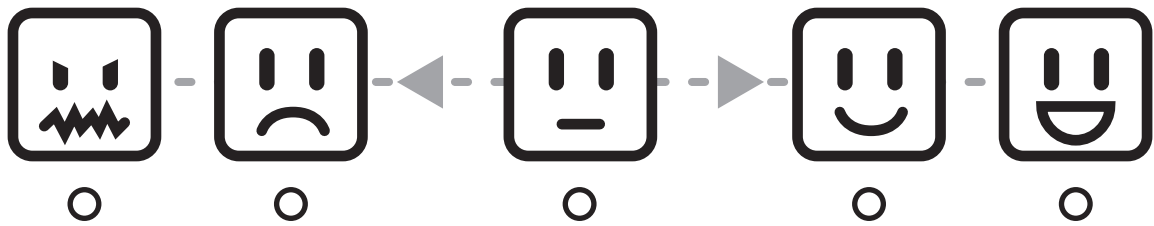
Expérimentation Minecraft

Questions supplémentaires

Nom :

Prénom :

1. Laquelle de ces illustrations correspond le mieux à l'état émotionnel dans lequel vous étiez lors de l'expérimentation ?



Frustration

Joie

2. Quel fut votre ressenti vis-à-vis de cette expérience :



Ennuyeuse

Captivante

3. Auriez vous envie dans un cadre personnel d'utiliser à nouveau la plateforme qui vous a été proposée ?

Plateforme = Minecraft + Oculus Rift + Manette



Pas du tout

Oui beaucoup

4. Recommanderiez vous l'utilisation de cette plateforme à d'autres personnes ?

Plateforme = Minecraft + Oculus Rift + Manette



Non pas du tout

Oui tout à fait

XX

Annexe C

Annexes à l'expérimentation 2

C.1 Configuration Minecraft

Logiciels externes à *Minecraft* nécessaires à la plateforme d'expérimentation :

Les mêmes logiciels externes que pour la configuration de l'expérimentation 1 sont requis, voir en section B.2.

Ressources pour la configuration de *Minecraft* — *Client* :

Suivre la procédure présentée en section B.2 en occultant les points f. et h. Nous ne verrons ici que les *mods* supplémentaires nécessaires à cette seconde configuration.

- a) Faire fonctionner *Minecraft* avec les lunettes 3D *Nvidia*. Tout d'abord télécharger et installer le *Shader Mod* permettant l'emploi de *shaders* comme son nom l'indique. Placer le fichier dans ...*.minecraft\mods*.
↓ **Mod :** <http://goo.gl/CGukdp>
- b) Télécharger le *3D Wrapper* qui va forcer *Java* à rendre deux images. Attention, il est nécessaire de choisir la version à télécharger en fonction de la version de *Java* installée (32 ou 64 bits.)
↓ **3D Wrapper :** <http://3dsurroundgaming.com/OpenGL3DVisionGames.html>
- c) Installer le *3D Wrapper*. Après avoir décompressé le fichier *Minecraft_3DVisionOGL_x...zip*, déplacer le fichier *Mike Shaders.zip* dans ...*.minecraft\shaderpacks*. Ce *shader* permet à *Minecraft* d'afficher les deux images nécessaires à la stéréoscopie.
- d) Déplacer les six fichiers contenus dans le dossier *Minecraft_3DVisionOGL_x..* et comportant les extensions suivantes **.dll*, **.ini*, **.nlp* et **.bat* dans *C:\Program Files\Java\jre1.8.0_73\bin*. Si le *3D Wrapper* ne fonctionne pas après cela, c'est que *Minecraft* fonctionne sur une version de *Java* embarquée, les fichiers devront alors être placés dans ...*Minecraft\runtime\jre-x64\version\bin*, le dossier ...*Minecraft* désignant le dossier dans lequel est placé l'exécutable du jeu, *Minecraft.exe*.

- e) Tester le fonctionnement du *3D Wrapper*. Au lancement de *Minecraft* une *pop-up* relative au *3D Wrapper* doit s'ouvrir, cliquer alors sur «OK». Une fois *Minecraft* lancé, activer le *shader* nécessaire au fonctionnement des lunettes. Aller dans *Options* → *Options graphiques* → *Shaders* et cliquer sur *Mike Shader.zip*. *Minecraft* fonctionnera à présent avec les lunettes 3D *Nvidia*.
- f) Mise en place de la vidéo d'introduction. Télécharger et installer les *mods MCEF* et *WebDisplay*. Le premier permet la prise en charge par *Minecraft* du moteur *Chromium* issu du navigateur *Chrome* de *Google*. Le second permet l'affichage de pages internet dans *Minecraft* par le biais de l'écran placé dans l'environnement. Placer les deux fichiers téléchargés dans le dossier `... \.minecraft\mods\`.
- g) Le *mod WebDisplay* permet d'afficher des pages internet comme nous venons de l'expliquer, dans le cas de notre configuration, nous étions en réseau local sans connexion internet. Nous avons donc créé un ensemble de pages HTML que nous avons placées à la racine du lecteur *C*. Ne pouvant entrer de chemins locaux par les fonctionnalités de base de *WebDisplay*, nous avons utilisé les scripts des ordinateurs du *mod ComputerCraft* afin de forcer l'injection des chemins dans l'affichage. Quatre vidéos sont disponibles en fonction de l'environnement choisi et correspondent chacune à un contexte SCV-SC, SCV-AC, ACV-SC et ACV-AC. C'est donc le choix du monde installé sur le *serveur* qui détermine quelle vidéo est jouée.
 - ↓ **Mod MCEF** : <http://goo.gl/suPE1C>
 - ↓ **Mod WebDisplay** : <http://goo.gl/UWz37A>
 - ↓ **Vidéos (placer le dossier dans C:\)** :
<http://remy-eynard.net/minecraft-exp/Ressources/minecraft-display.zip>
- h) Nous ne traiterons pas la configuration relative au cube émotionnel, le *auraMod* n'étant pas aussi facilement réutilisable que le reste de la configuration. Néanmoins, les sources peuvent être transmises à qui le désirerait. Dans ce cas, merci d'adresser un courriel à l'auteur de ce document.
- i) Lancer *Minecraft* en prenant soin de bien sélectionner le bon profil. Cliquer sur «*Multijoueur*» puis «*Connexion rapide*». Entrer l'adresse du serveur dont nous allons à présent voir la configuration. Les postes participants sont correctement configurés pour notre première expérimentation.

Ressources pour la configuration de *Minecraft* — Serveur :

Suivre la procédure présentée en section B.2. Seul le point e. relatif à l'installation de l'environnement est différent et est traité ici.

Installer un environnement virtuel sur le serveur. Le dossier à télécharger contient ici quatre environnements différents en fonction du contexte, la vidéo de présentation étant différente. Copier le contenu du dossier choisi dans le dossier `... \world\` du serveur.

↓ **Dossier** : <http://remy-eynard.net/minecraft-exp/Ressources/EV-2.zip>

C.2 Livret d'informations aux participants

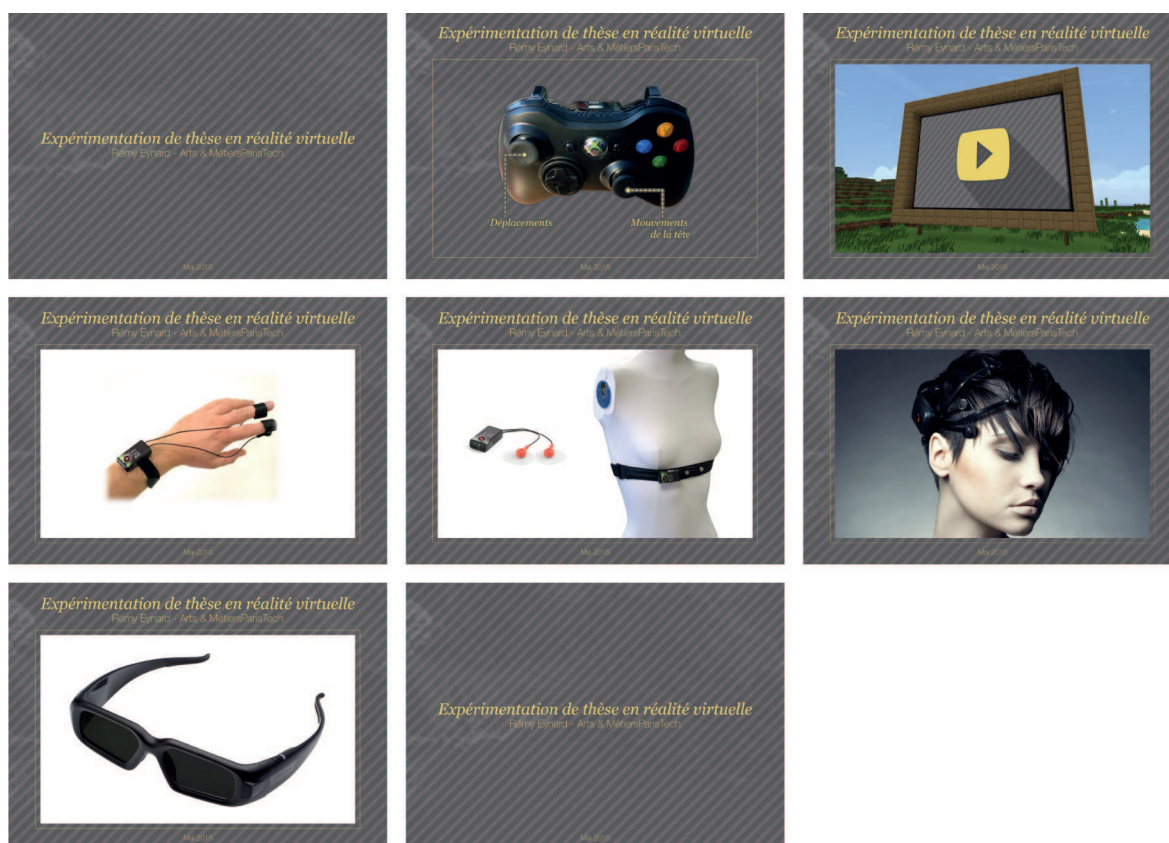


FIGURE C.1 – Livret d'informations aux participants - Expérimentation 2.

C.3 Questionnaire Post-Expérimentation

Les questions 23 à 35 n'étaient présentes que dans la version du questionnaire remplie pour les participants ayant passé l'expérimentation dans le contexte ACV. Pour faciliter la lecture, les numéros des questions concernées sont en **orange et gras**. La question 19 n'était présente que pour les participants ayant été immergés avec le **Cube émotionnel**, son numéro est en **vert et gras**.

Date & Numéro de session d'expérimentation : &

Numéro de la salle d'expérimentation : ☐ 01 ☐ 02

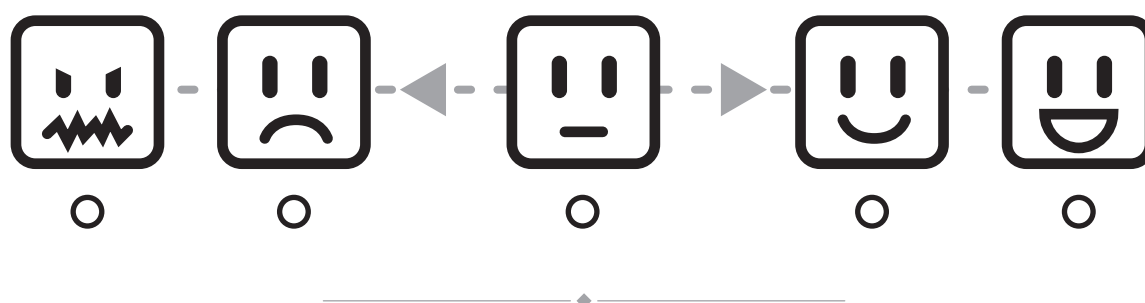
Nom :

Prénom :

Questionnaire

Post-expérimentation

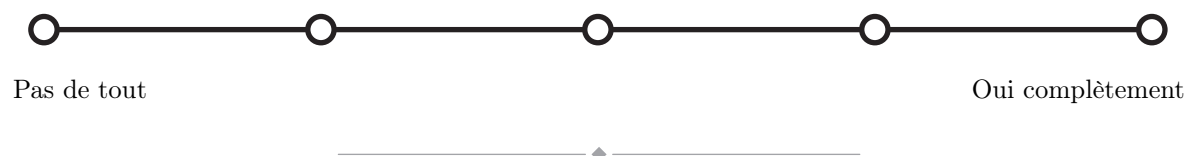
1. Votre ressenti global vis-à-vis de cette expérience :



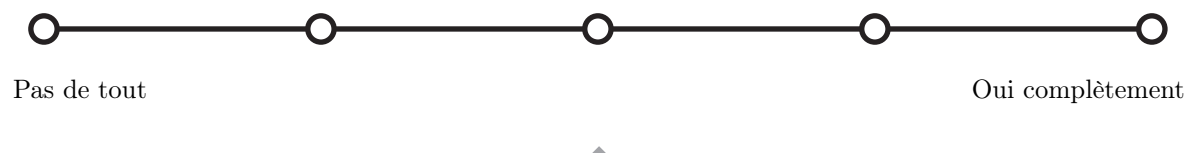
2. Vous avez ressenti une forme de challenge durant cette expérience :



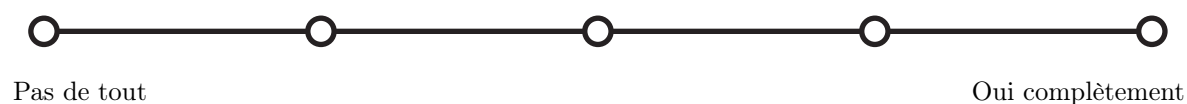
3. Vous étiez focalisé sur le tâche à accomplir :



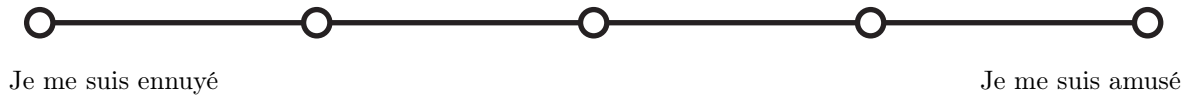
4. Vous vous êtes sentis maître de vos mouvements, de vos actions dans l'environnement virtuel :



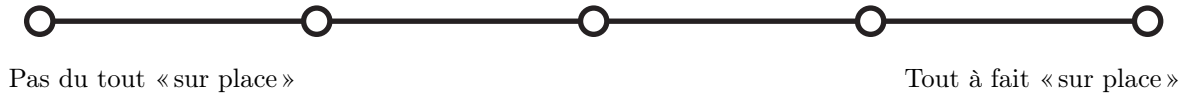
5. Votre avatar répondait bien aux actions que vous désiriez faire :



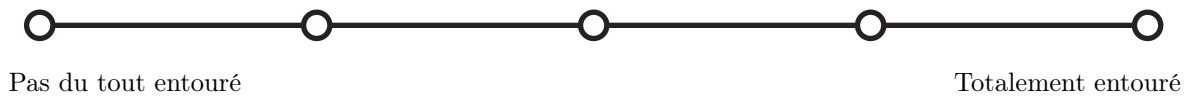
6. Je me suis amusé durant cette expérience :



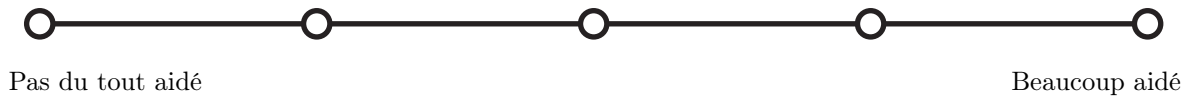
7. Dans l'environnement virtuel, vous aviez le sentiment « d'être sur place, présent » :



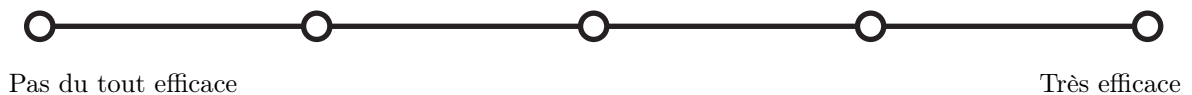
8. D'une certaine façon, vous aviez l'impression que le monde virtuel vous entourait :



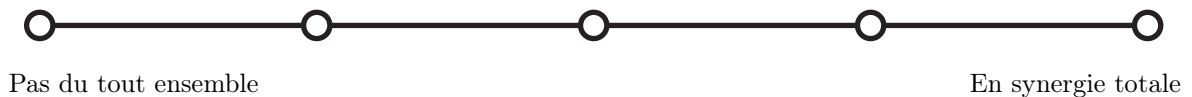
9. Votre partenaire vous a aidé dans l'accomplissement de la tâche :



10. Vous avez été efficace vis-à-vis de l'accomplissement de la tâche :

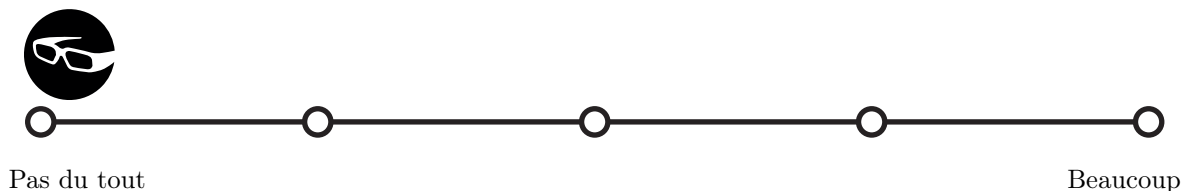


11. Vous avez eu le sentiment de travailler ensemble sur la tâche :

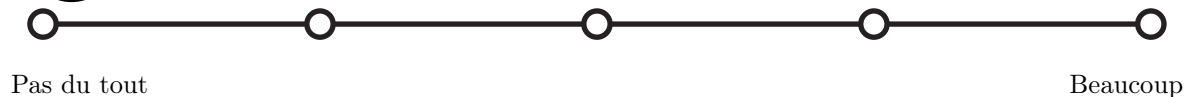


12. Votre performance a-t-elle été altérée par le matériel dont vous étiez équipé :

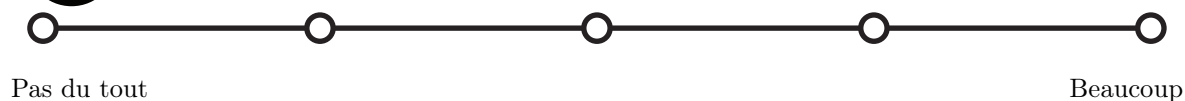
12.1 Les lunettes 3D :



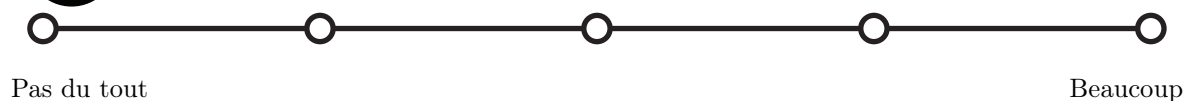
12.2 Le casque à électroencéphalogramme :



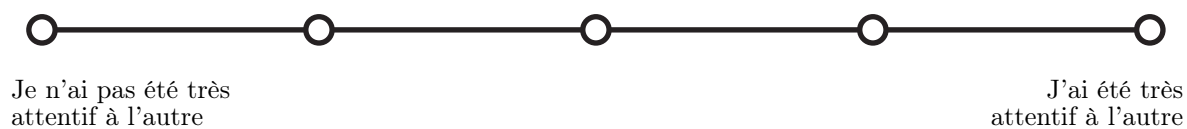
12.3 Le capteur de rythme cardiaque :



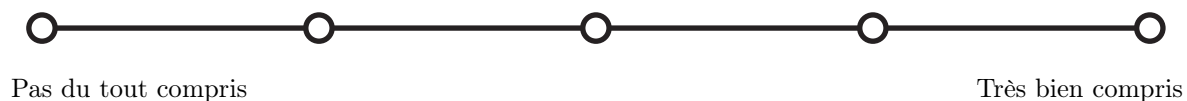
12.4 Le capteur de conductance de la peau :



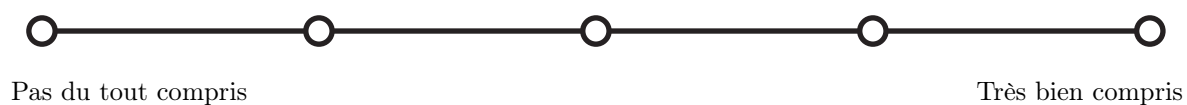
13. Vous avez fait preuve d'une attention particulièrement vive vis-à-vis de l'autre personne :



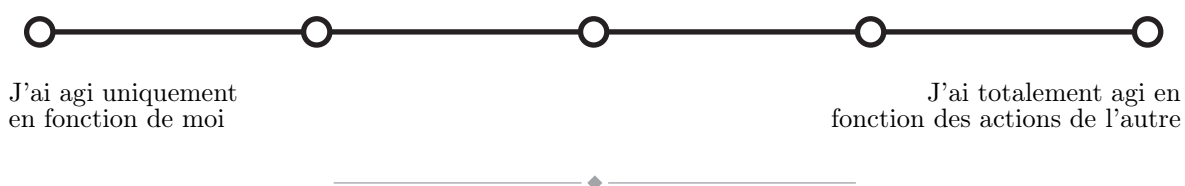
14. L'autre personne vous a paru comprendre ce que vous désiriez faire (actions) :



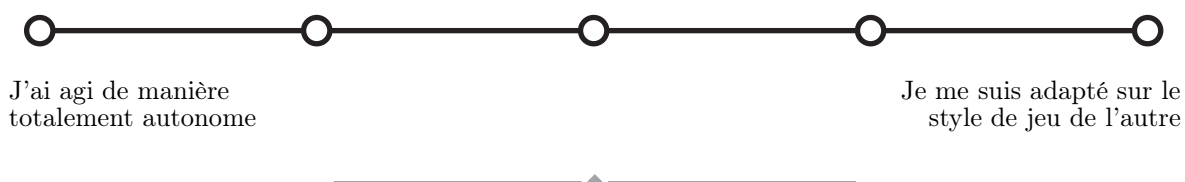
15. J'ai compris ce que désirait l'autre personne (actions) :



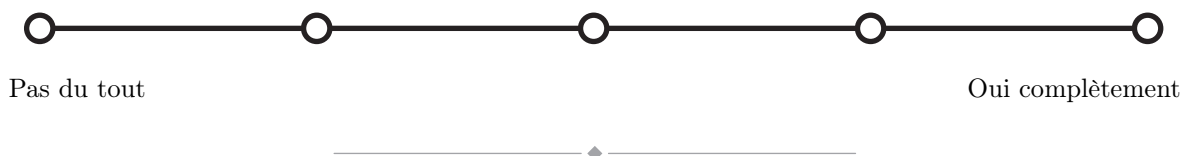
16. Vos actions étaient dépendantes de celles de l'autre personne :



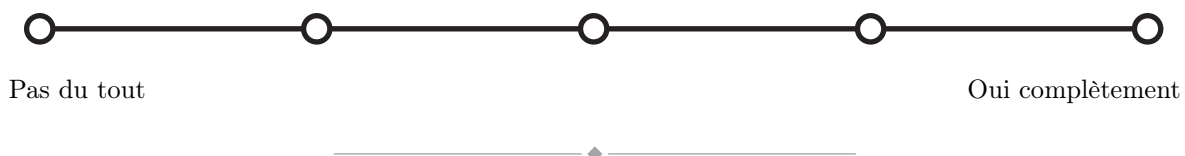
17. Votre style de jeu (par exemple : manière de construire la statue, récupération de matériaux) était une réponse directe au style de jeu de l'autre personne :



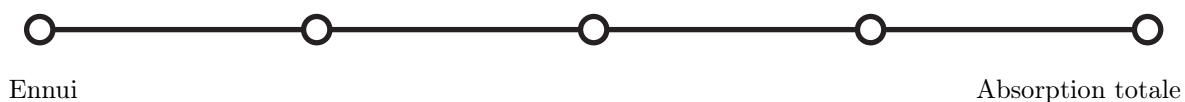
18. Vous avez ressenti l'état émotionnel de votre partenaire :



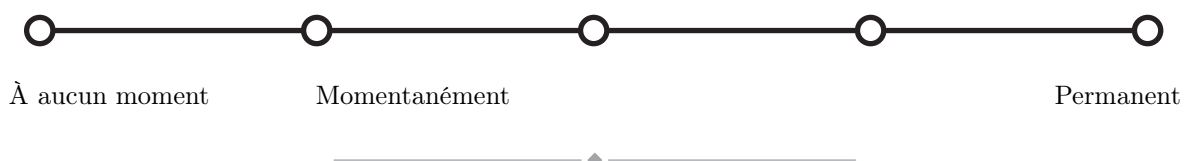
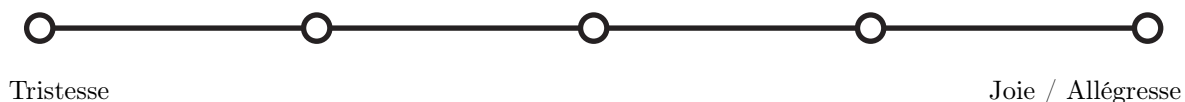
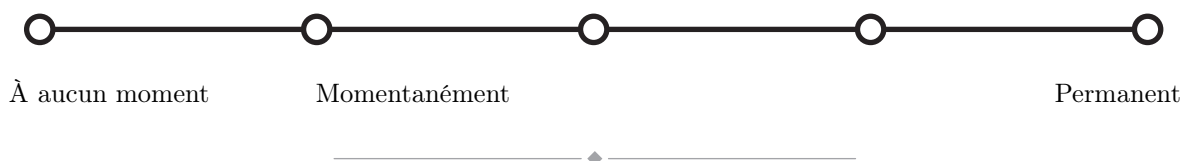
19. Le « cube émotionnel » vous a aidé à comprendre l'état émotionnel de votre partenaire :



20. Décrivez votre état émotionnel global vis-à-vis de cette expérience ::



Cet état émotionnel fut-il momentané ou durable pendant l'expérimentation :



Colère Satisfaction

À aucun moment Momentanément Permanent

Peur Confiance

À aucun moment Momentanément Permanent

21. Quel était votre sentiment au sujet de cette expérience ?

Ennuyeux Captivant

22. Vous avez interagi avec votre partenaire (conversation) : ☐ Oui ☐ Non

Si oui :

23. Vous avez interagi à la fréquence suivante :

Rarement De nombreuses fois

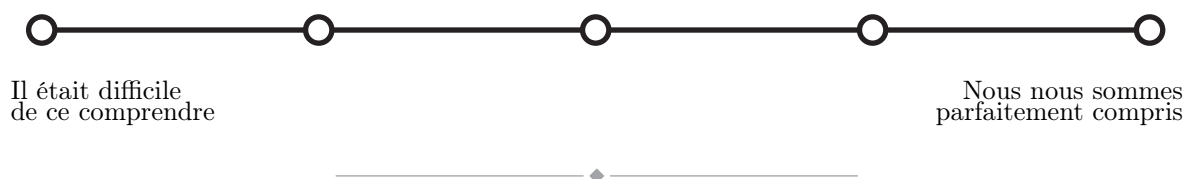
24. Votre partenaire vous a appris des choses durant l'expérience :

Rien appris Beaucoup appris

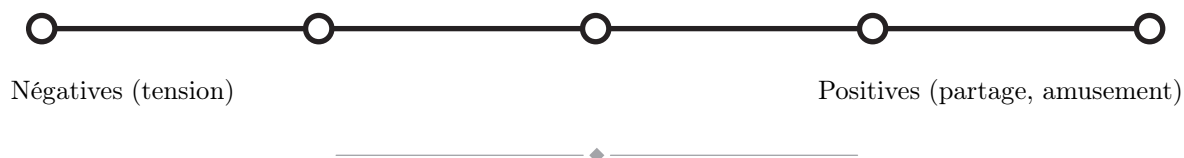
25. Vous avez, au travers de vos conversations, ressenti l'état émotionnel de votre partenaire :

Pas du tout Complètement

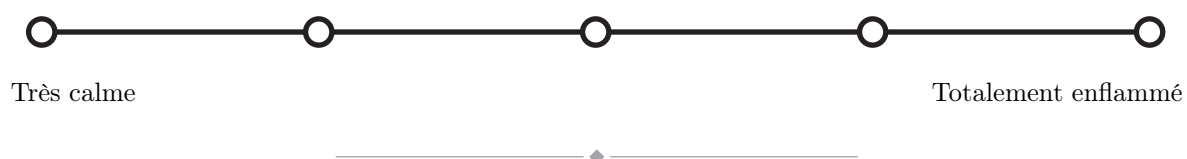
26. Vous avez ressenti un manque de compréhension entre vous et votre partenaire :



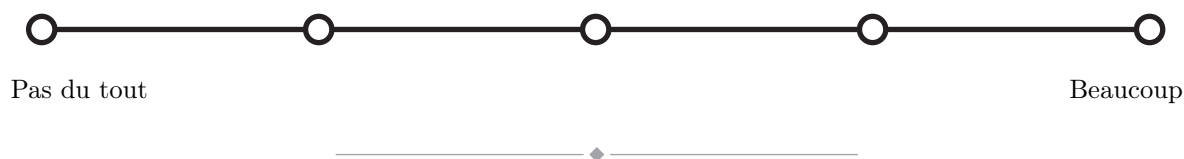
27. Vos discussions étaient plutôt :



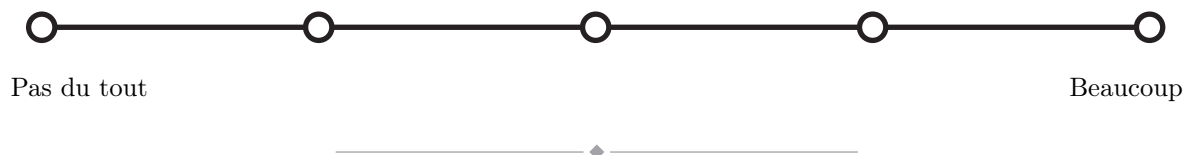
28. Le degré d'intensité de vos conversations était :



29. Vous vous êtes mutuellement encouragés durant l'expérience :

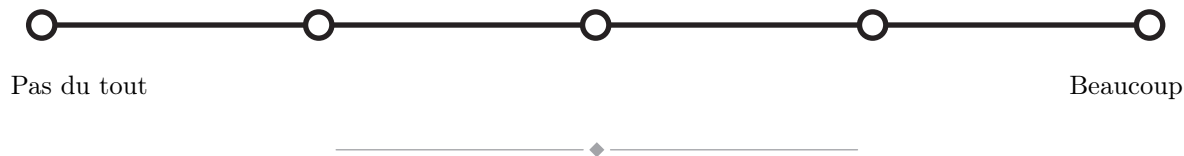


30. Vous vous êtes mutuellement conseillés vis-à-vis de la tâche à accomplir :

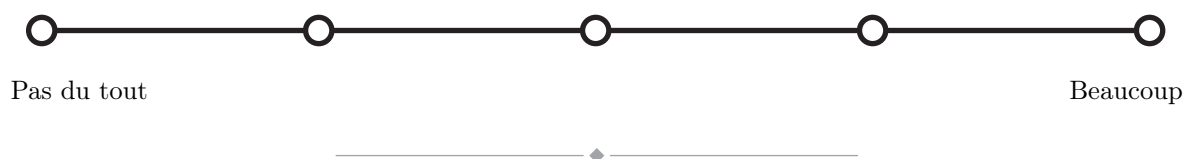


Sur quels thèmes avez-vous échangés ?

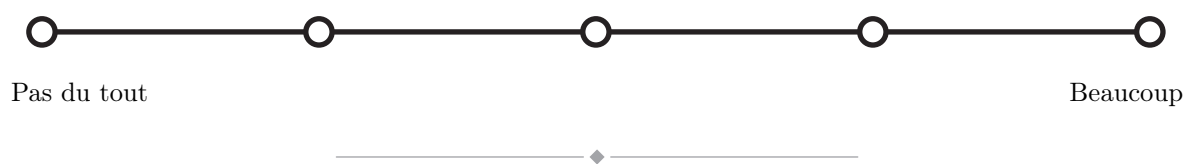
31. La réalisation de la tâche demandée :



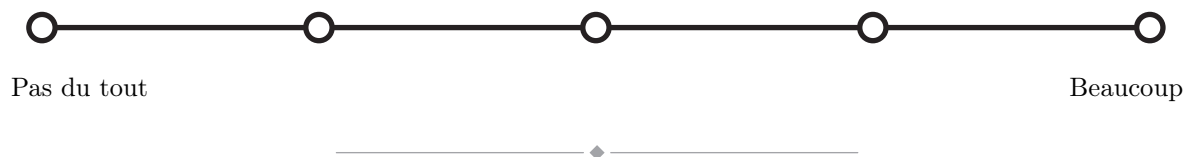
32. Votre ressenti vis-à-vis de l'expérience en cours :



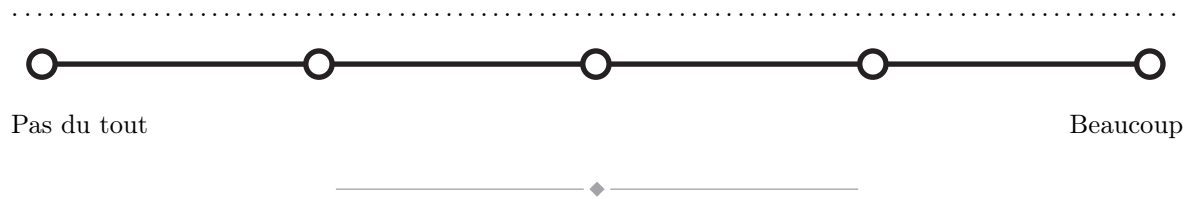
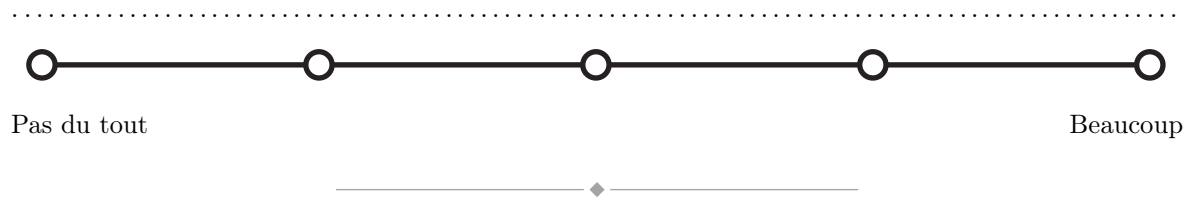
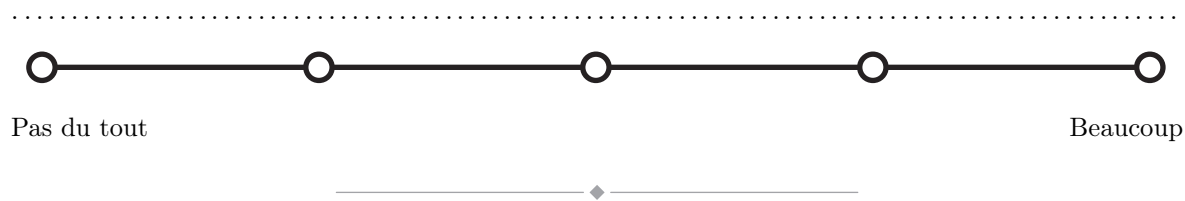
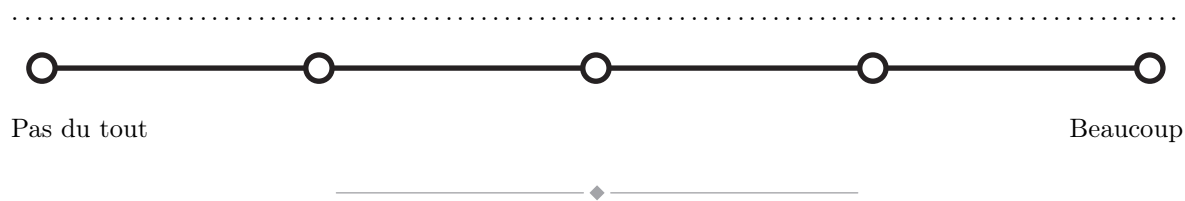
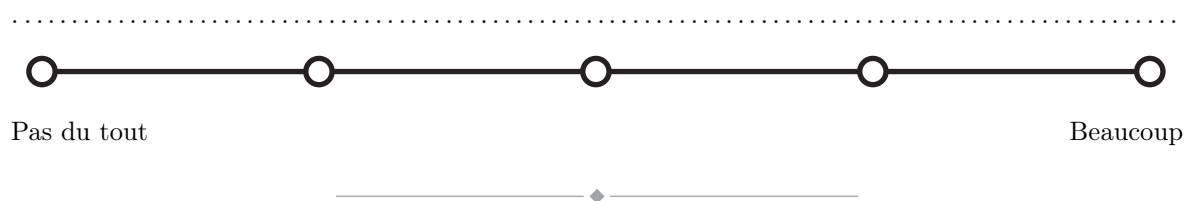
33. L'environnement virtuel :



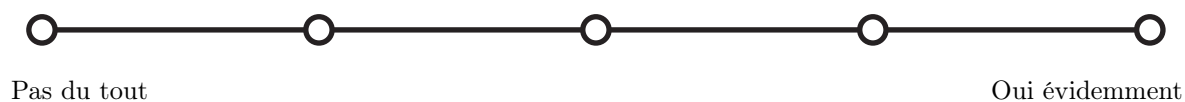
34. Les modalités d'interaction (utilisation de la manette et du jeu Minecraft) :



35. D'autres sujets non corrélés à l'expérience en cours (possibilité de remplir les champs ci-dessous) :



36. Souhaiteriez-vous utiliser à nouveau cette application (tâche comprise) dans un contexte personnel ?



37. Recommanderiez-vous l'utilisation de cette application à d'autres personnes ?



Pas du tout

Oui évidemment

38 . Merci d'inscrire ici toutes remarques que vous désireriez faire vis-à-vis de cette expérimentation :

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



**Fin du questionnaire,
merci de votre participation.**

Annexe D

Annexes à l'expérimentation complémentaire

Dans le cadre de cette expérimentation, chaque échantillon se trouvait immergé dans les deux contextes (ACV et SCV). C'est pourquoi, pour chaque question, deux échelles de Likert spécifiquement dédiées à chacun des contextes sont présentes. Pour l'échantillon 3, quatre échelles de Likert sont présentes pour chaque question, car les participants ont été immergés dans les deux environnements virtuels et, pour chaque environnement, sont passés sur les deux contextes ACV et SCV.

D.1 Questionnaire Post-Expérimentation - Echantillons 1 & 2

Date :
 Heure :
 Expérimentation : ☐ 01 ☐ 02
 Poste d'expérimentation : ☐ 01 ☐ 02
 Nom :
 Prénom :
 Adresse mail :

Questionnaire

Post-expérimentation

1. Dans l'environnement virtuel, vous aviez le sentiment « d'être sur place, présent » :

Avec Comm. : ☐ — ☐ — ☐ — ☐ — ☐

Sans Comm. : ☐ — ☐ — ☐ — ☐ — ☐

Pas du tout « sur place » Tout à fait « sur place »

————— ◆ —————

2. Avez-vous eu le sentiment d'être incarné dans votre avatar ?

Avec Comm. : ☐ — ☐ — ☐ — ☐ — ☐

Sans Comm. : ☐ — ☐ — ☐ — ☐ — ☐

Non, pas du tout Oui, je me suis totalement transposé dans l'avatar

————— ◆ —————

3. Vous avez été efficace vis-à-vis de l'accomplissement de la tâche :

Avec Comm. : ☐ — ☐ — ☐ — ☐ — ☐

Sans Comm. : ☐ — ☐ — ☐ — ☐ — ☐

Pas du tout efficace Très efficace

————— ◆ —————

4. Vous avez eu le sentiment de travailler ensemble sur la tâche ?

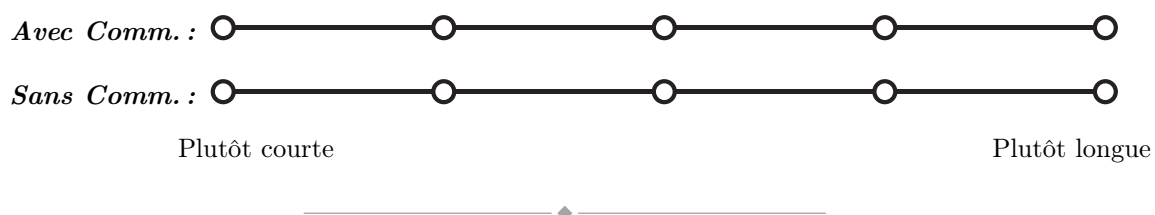
Avec Comm. : ☐ — ☐ — ☐ — ☐ — ☐

Sans Comm. : ☐ — ☐ — ☐ — ☐ — ☐

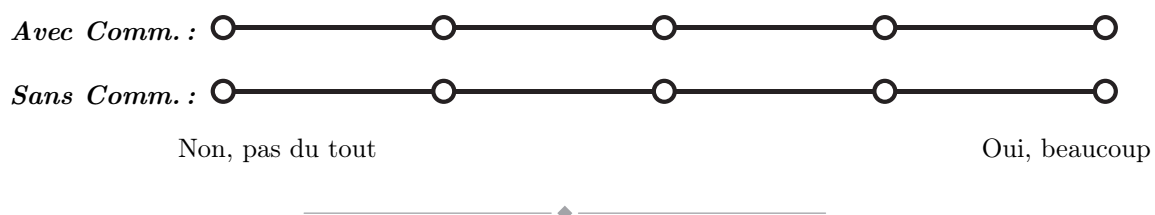
Pas du tout ensemble En synergie totale

————— ◆ —————

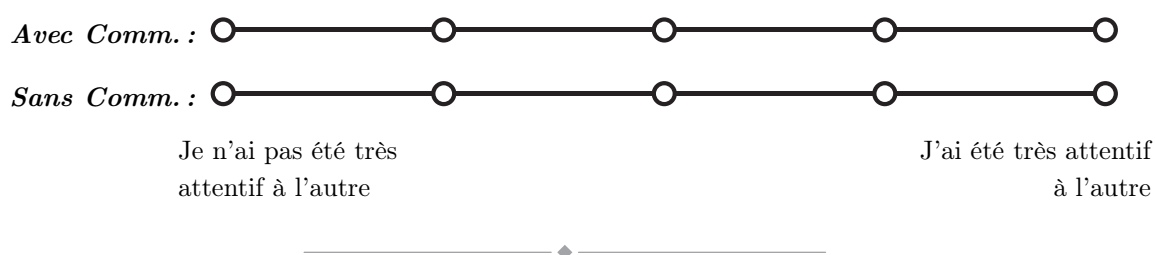
5. Subjectivement, comment estimeriez-vous la durée de l'expérience ?



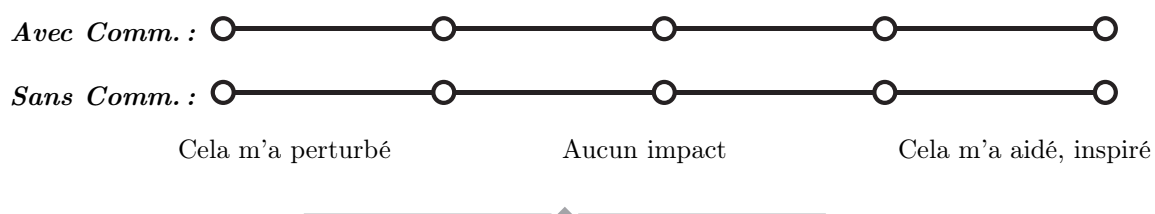
6. Vous êtes-vous amusés durant cette expérience ?



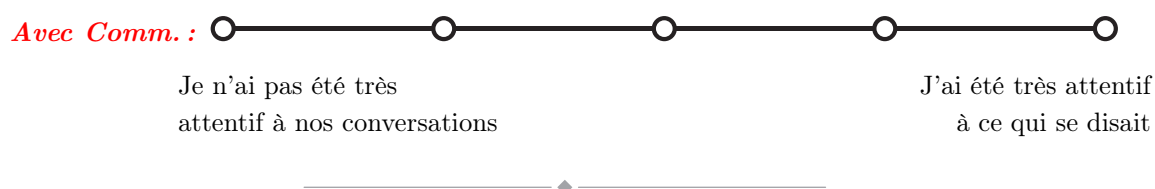
7. Avez vous fait preuve d'une attention particulière vis-à-vis du comportement de l'autre personne (déplacement, méthode utilisée pour poser les blocs, etc...) ?



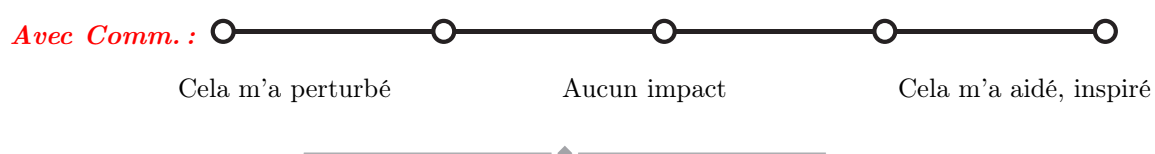
8. Comment estimeriez vous l'impact de ce comportement sur vos actions :



9. Avez vous été attentif à ce qu'a pu vous dire l'autre personne ?



10. Comment estimeriez vous l'impact de ces échanges sur vos actions :



D.2 Questionnaire Post-Expérimentation - Echantillon 3

Date :
 Heure :
 Poste d'expérimentation : ☐ 01 ☐ 02
 Nom :
 Prénom :
 Adresse mail :

Questionnaire

Post-expérimentation

1. Dans l'environnement virtuel, vous aviez le sentiment « d'être sur place, présent » :

<i>Constr. Vert.</i>	
<i>Avec Comm. :</i>	○ ———— ○ ———— ○ ———— ○ ———— ○
<i>Constr. Vert.</i>	
<i>Sans Comm. :</i>	○ ———— ○ ———— ○ ———— ○ ———— ○
<i>Constr. Horiz.</i>	
<i>Avec Comm. :</i>	○ ———— ○ ———— ○ ———— ○ ———— ○
<i>Constr. Horiz.</i>	
<i>Sans Comm. :</i>	○ ———— ○ ———— ○ ———— ○ ———— ○
Pas du tout « sur place »	Tout à fait « sur place »
<div style="display: flex; align-items: center; justify-content: center; margin-top: 10px;"> <div style="flex-grow: 1; border-top: 1px solid #ccc; position: relative;"> <div style="position: absolute; left: 50%; top: -5px;">◆</div> </div> </div>	

2. Avez-vous eu le sentiment d'être incarné dans votre avatar ?

<i>Constr. Vert.</i>	
<i>Avec Comm. :</i>	○ ———— ○ ———— ○ ———— ○ ———— ○
<i>Constr. Vert.</i>	
<i>Sans Comm. :</i>	○ ———— ○ ———— ○ ———— ○ ———— ○
<i>Constr. Horiz.</i>	
<i>Avec Comm. :</i>	○ ———— ○ ———— ○ ———— ○ ———— ○
<i>Constr. Horiz.</i>	
<i>Sans Comm. :</i>	○ ———— ○ ———— ○ ———— ○ ———— ○
Non, pas du tout	Oui, je me suis totalement transposé dans l'avatar
<div style="display: flex; align-items: center; justify-content: center; margin-top: 10px;"> <div style="flex-grow: 1; border-top: 1px solid #ccc; position: relative;"> <div style="position: absolute; left: 50%; top: -5px;">◆</div> </div> </div>	

3. Vous avez été efficace vis-à-vis de l'accomplissement de la tâche :

Constr. Vert.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Vert.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Pas du tout efficace Très efficace

————— ◆ —————

4. Vous avez eu le sentiment de travailler ensemble sur la tâche ?

Constr. Vert.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Vert.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Pas du tout ensemble En synergie totale

————— ◆ —————

5. Subjectivement, comment estimeriez-vous la durée de l'expérience ?

Constr. Vert.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Vert.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Plutôt courte Plutôt longue

————— ◆ —————

6. Vous êtes-vous amusés durant cette expérience ?

Constr. Vert.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Vert.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Non, pas du tout Oui, beaucoup

————— ◆ —————

7. Avez vous fait preuve d'une attention particulière vis-à-vis du comportement de l'autre personne (déplacement, méthode utilisée pour poser les blocs, etc...) ?

Constr. Vert.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Vert.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Je n'ai pas été très attentif à l'autre J'ai été très attentif à l'autre

————— ◆ —————

8. Comment estimeriez vous l'impact de ce comportement sur vos actions :

Constr. Vert.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Vert.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Avec Comm. : ○ — ○ — ○ — ○ — ○

Constr. Horiz.
Sans Comm. : ○ — ○ — ○ — ○ — ○

Cela m'a perturbé Aucun impact Cela m'a aidé, inspiré

————— ◆ —————

9. Avez vous été attentif à ce qu'a pu vous dire l'autre personne ?

Constr. Vert.

Avec Comm. :



Constr. Horiz.

Avec Comm. :



Je n'ai pas été très
attentif à nos conversations

J'ai été très attentif
à ce qui se disait



10. Comment estimeriez vous l'impact de ces échanges sur vos actions :

Constr. Vert.

Avec Comm. :



Constr. Horiz.

Avec Comm. :



Cela m'a perturbé

Aucun impact

Cela m'a aidé, inspiré



11 . Merci d'inscrire ici toutes remarques que vous désireriez faire vis-à-vis de cette expérimentation :

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



Fin du questionnaire,
merci de votre participation.

D.3 Guide pour groupe de discussion

Les termes en **orange** ont été ajoutés aux questions posées à l'échantillon 3, ce dernier ayant expérimenté les deux environnements virtuels, contrairement aux échantillons 1 & 2.

1. Le phénomène de translucence : environ 15 minutes.

- Quel impact le comportement, les actions de votre partenaire ont-ils eu sur le vôtre ?
- Selon le contexte **ou l'environnement virtuel**, quelles différences avez-vous observées vis-à-vis de cet impact ?

2. L'efficacité : environ 15 minutes.

- À votre avis, pourquoi pourriez-vous être plus efficace dans l'un des deux contextes **ou environnements virtuels** (mentionner lequel) ?
- Quelle fut la nature de vos interactions verbales ? Et en quoi vous furent-elles utiles ?
- L'un de vous s'est-il positionné en tant que leader de la tâche à accomplir ? Quel est votre ressenti vis-à-vis de cela ?
- Qu'est-ce qui fait que vous pourriez plus vous amuser dans l'un ou dans l'autre des contextes **ou environnements virtuels** (mentionner lequel) ?

3. Le sentiment de présence virtuelle : environ 15 minutes.

- Comment décririez-vous votre sentiment de présence dans l'environnement virtuel et la manière dont vous avez interagi avec ce dernier ?
- De quelle manière avez-vous perçu l'autre utilisateur dans l'environnement virtuel ?
- Le fait de communiquer verbalement avec votre partenaire a-t-il eu un impact sur votre sentiment de présence dans l'environnement virtuel ?
- Lors de vos communications verbales, avez-vous été gêné par quelque chose, qui vous aurait semblé étrange, irréel (incarnation par l'avatar) ?

4. La qualité de service : environ 5 minutes.

Que conseillerez-vous aux concepteurs de ces expériences en vue de les améliorer ? Que ces améliorations soient technologiques, de l'ordre de l'interaction, de la narration ou du gameplay.

5. L'acceptation de la technologie : environ 5 minutes.

- Si l'on vous proposait de refaire ces expériences, quel serait votre ressenti ?
- Que diriez-vous à une personne qui vous demande si faire ces expériences en vaut la peine ?

INVESTIGATION DE L'IMPACT DE L'INTERACTION SOCIALE VOCALE SUR L'EXPERIENCE UTILISATEUR DANS LES ENVIRONNEMENTS 3D TEMPS-RÉEL IMMERSIFS.

RESUME :

L'objectif de ce travail de thèse était d'apporter une meilleure compréhension de l'impact des interactions sociales vocales (le fait de se parler) sur l'expérience utilisateur (UX) en environnement virtuel immersif. L'expérience utilisateur englobe les émotions, les états d'esprit, mais aussi le fait de s'amuser, d'être impliqué dans l'activité en cours et permet surtout une évaluation de la satisfaction résultant de l'usage.

Afin d'évaluer l'impact de la communication vocale sur l'UX, nous avons mis en place deux expérimentations dans lesquelles nous avons séparé nos participants en deux échantillons. Ces échantillons furent immergés dans le même environnement virtuel, pour participer à la même activité, la seule différence résidait dans la possibilité ou non de communiquer vocalement. Nos résultats ont montré que la communication vocale a un impact sur plusieurs facteurs de l'UX comme la performance ou la présence virtuelle. Il faut néanmoins être conscient de la nature variable de l'UX qui fait que les constats émis ici pourraient être différents dans d'autres contextes expérimentaux.

Mots clés : Réalité Virtuelle ; Environnements 3D Immersifs ; Expérience Utilisateur ; Interactions Sociales ; Communication Verbale ; Présence Sociale.

INVESTIGATION OF THE IMPACT OF VOCAL SOCIAL INTERACTION ON USER EXPERIENCE IN 3D REAL-TIME IMMERSIVE VIRTUAL ENVIRONMENTS.

ABSTRACT:

The objective of this doctoral thesis was to provide a better understanding of the impact of vocal social interactions (the act of speaking) on the user experience (UX) in immersive virtual environment. User experience includes emotions, states of mind, but also entertainment, involvement in the current activity, and provides above all an assessment of the satisfaction resulting from the use.

To assess the impact of vocal communication on UX, we set up two experiments in which we have separated our participants in two samples. These samples were immersed in the same virtual environment, participated at the same activity, the only difference was the capability to communicate verbally or not. Our results showed that the vocal communication has an impact on several UX factors like performance or virtual presence. We should nevertheless be aware of the varying nature of UX, that makes that our observations could be different in other experimental contexts.

Keywords : Virtual Reality ; 3D Immersive Environment ; User Experience ; Social Interactions ; Verbal Communication ; Social Presence.

