

Doctorat de l'Université de Toulouse

préparé à l'Université Toulouse - Jean Jaurès

Faire comprendre les mécanismes de l'autisme en situation
d'apprentissage grâce à la Réalité Virtuelle Incarnée

Thèse présentée et soutenue, le 26 février 2026 par

Aurélien MOLINER

École doctorale

CLESCO - Comportement, Langage, Éducation, Socialisation, Cognition

Spécialité

Psychologie

Unité de recherche

CLLE - Unité Cognition, Langues, Langage, Ergonomie

Thèse dirigée par

Franck AMADIEU et David PANZOLI

Composition du jury

M. Jean-Michel BOUCHEIX, Président, Université de Bourgogne

M. André TRICOT, Rapporteur, Université de Montpellier Paul-Valéry

Mme Valérie GOURANTON, Rapporteur, Université de Rennes

Mme Lucie BOUVET, Examinatrice, Université Toulouse - Jean Jaurès

M. Franck AMADIEU, Directeur de thèse, Université Toulouse - Jean Jaurès

M. David PANZOLI, Co-directeur de thèse, INU Champollion

Remerciements

Ce travail de thèse n'aurait pu aboutir sans le soutien et l'accompagnement de nombreuses personnes que je tiens à remercier chaleureusement.

Je souhaite tout d'abord remercier la région Occitanie pour avoir financé ce travail.

Je souhaite ensuite exprimer ma profonde gratitude à mes directeurs de thèse, Franck Amadiou et David Panzoli, pour leur encadrement rigoureux, leur disponibilité et leurs précieux conseils tout au long de ces années de recherche. Leur expertise et leur confiance m'ont permis de mener à bien ce projet.

Je remercie également Thomas Tanzi, ami inénarrable et relecteur attentif, qui a égaré tant d'heures de sa vie sur cette thèse. Ton aide et tes remarques acerbes ont été inestimables pour faire aboutir ce qui se trouve sous vos yeux.

À Judith Coddens, ma compagne, merci d'être là, tout simplement. Ta présence à mes côtés a été un pilier essentiel dans cette aventure.

Je pense à ma famille, mon troupeau de frères et sœurs, Théo, Maëlle, Énora, Thalie, Jérémy et David, qui ont toujours cru en moi. À mon père, merci pour ta fierté et ton soutien indéfectible, ainsi que pour tout le reste. À toi Éric, qui m'as vu grandir et qui est comme un père pour moi. À Laurence, pour cette belle personne que tu es et qui a su être la juste quand il le fallait. À mon Bro, merci pour tes illustrations qui m'ont sorti de l'embarras à de nombreuses reprises, pour tes conseils souvent curieux et pour ce caractère unique qui est le tien.

Enfin, je remercie tous ceux qui, de près ou de loin, m'ont apporté leur soutien, leurs encouragements et leur aide durant ces années. Votre contribution, quelle qu'elle soit, a compté.

À ma mère, partie trop tôt, qui aurait été si fière de voir l'aboutissement de ce travail : cette thèse t'est dédiée.

Résumé

Mots-clés : Réalité virtuelle incarnée, Trouble du spectre de l'autisme, Apprentissage, Sensibilisation, Empathie, Changement d'attitude, Formation des enseignants, Modèle CAMIL

Cette thèse interdisciplinaire, au croisement de la psychologie cognitive et de l'informatique, examine l'efficacité de la réalité virtuelle incarnée (RVI) comme outil de sensibilisation au trouble du spectre de l'autisme (TSA) auprès des enseignants. S'appuyant sur le modèle CAMIL (Cognitive-Affective Model of Immersive Learning) de Makransky et Petersen (2021), cette recherche interroge les mécanismes psychologiques par lesquels l'immersion virtuelle peut favoriser la compréhension du fonctionnement cognitif des personnes avec TSA et modifier les attitudes à leur égard.

Le cadre théorique mobilise les modèles explicatifs du fonctionnement cognitif dans le TSA, comme le modèle Enhanced Perceptual Functioning (EPF) de Mottron et al. (2006), et les théories de l'apprentissage en environnement immersif, avec une attention portée aux concepts de présence, d'agentivité et d'empathie. L'objectif central consiste à évaluer si l'incarnation d'un avatar représentant une personne avec TSA peut générer une expérience empathique suffisamment significative pour induire une meilleure compréhension des particularités sensorielles et cognitives du TSA.

Un apport majeur réside dans la conception et le développement d'un dispositif de RVI fonctionnel, déployé sur casque occlusif (HMD). L'environnement virtuel simule une salle de classe du point de vue d'une personne avec TSA, intégrant des modifications perceptives (hypersensibilités auditives et visuelles) et des mécaniques d'interaction permettant d'expérimenter les difficultés caractéristiques du TSA. Cette dimension technique constitue une contribution méthodologique importante, articulant contraintes ergonomiques, objectifs pédagogiques et validité écologique.

La méthodologie repose sur deux expérimentations menées auprès d'enseignants. La première étude (N = 112) compare quatre conditions : RVI, réalité virtuelle non incarnée sur écran plat, diaporama audio et diaporama. La seconde (N = 67) affine le protocole avec une refonte de l'outil. Les variables mesurées incluent la présence, l'agentivité, l'empathie envers l'avatar, les connaissances factuelles sur le TSA et les stéréotypes.

Les résultats révèlent une validation partielle du modèle CAMIL. Les variables présence, agentivité et empathie sont significativement plus élevées en condition RVI qu'en condition vidéo, validant la voie affective du modèle. Cependant, aucun avantage significatif de la RVI n'est observé pour l'acquisition de connaissances factuelles. Paradoxalement, l'introduction de pauses pédagogiques structurées améliore significativement l'efficacité de la RVI, démontrant que l'immersion seule ne garantit pas l'apprentissage. Concernant le changement d'attitude, si une réduction des stéréotypes est observée, l'effet reste modeste et ne se traduit pas par une modification substantielle des attitudes explicites.

Cette recherche apporte plusieurs contributions. Sur le plan théorique, elle révèle les limites du modèle CAMIL dans son application au domaine socio-affectif et identifie les variables modératrices manquantes. Sur le plan méthodologique, elle démontre l'importance d'une approche intégrative combinant immersion, guidage pédagogique et temps de réflexion. Sur le plan technique, elle propose un dispositif de RVI opérationnel et reproductible.

Les implications pratiques soulignent que la RVI doit s'inscrire dans une approche pédagogique plus large, combinant expériences immersives, rencontres avec des personnes avec TSA, formations théoriques et accompagnement réflexif. L'optimisation de l'efficacité pédagogique passe moins par l'amélioration de la fidélité technologique que par la sophistication du design pédagogique.

En conclusion, cette recherche contribue à une vision scientifiquement fondée de l'intégration de la réalité virtuelle dans les pratiques de sensibilisation, démontrant que son efficacité dépend moins de ses propriétés technologiques que de son articulation avec une ingénierie pédagogique adaptée.

Abstract

Keywords: Embodied virtual reality, Autism spectrum disorder, Awareness training, Learning, Empathy, Attitude change, Teacher training, CAMIL model

This interdisciplinary thesis, at the intersection of cognitive psychology and computer science, examines the effectiveness of embodied virtual reality (EVR) as an awareness tool for autism spectrum disorder (ASD) among teachers. Based on Makransky and Petersen's (2021) CAMIL model (Cognitive-Affective Model of Immersive Learning), this research investigates the psychological mechanisms through which virtual immersion can foster understanding of the cognitive functioning of people with ASD and modify attitudes toward them.

The theoretical framework mobilizes explanatory models of cognitive functioning in ASD, particularly Mottron et al.'s (2006) Enhanced Perceptual Functioning (EPF) model, and theories of learning in immersive environments, with attention to the concepts of presence, agency, and empathy. The central objective is to assess whether virtual embodiment of an avatar representing a person with ASD can generate a sufficiently significant empathic experience to induce better understanding of the sensory and cognitive particularities of ASD.

A major contribution lies in the design and development of a functional EVR system, deployed on a head-mounted display (HMD). The virtual environment simulates a classroom from the perspective of a person with ASD, integrating perceptual modifications (auditory and visual hypersensitivities) and interaction mechanics allowing users to experience characteristic difficulties of ASD. This technical dimension constitutes an important methodological contribution, articulating ergonomic constraints, pedagogical objectives, and ecological validity.

The methodology relies on two experiments conducted with teachers. The first study (N = 112) compares four conditions: EVR, non-embodied virtual reality on flat screen, audio slideshow, and slideshow. The second (N = 67) refines the protocol with a redesigned tool. Measured variables include presence, agency, empathy toward the avatar, factual knowledge about ASD, and stereotypes.

Results reveal partial validation of the CAMIL model. Proximal variables (presence, agency, and empathy) are significantly higher in the EVR condition than in the video condition,

validating the affective pathway of the model. However, no significant advantage of EVR is observed for factual knowledge acquisition. Paradoxically, the introduction of structured pedagogical pauses significantly improves EVR effectiveness, demonstrating that immersion alone does not guarantee learning. Regarding attitude change, while stereotype reduction is observed, the effect remains modest and does not translate into substantial modification of explicit attitudes.

This research provides several contributions. Theoretically, it reveals the limits of the CAMIL model in its application to the socio-affective domain and identifies missing moderating variables. Methodologically, it demonstrates the importance of an integrative approach combining immersion, pedagogical guidance, and reflection time. Technically, it proposes an operational and reproducible EVR system.

Practical implications emphasize that EVR must be part of a broader pedagogical approach, combining immersive experiences, encounters with people with ASD, theoretical training, and reflective support. Optimizing pedagogical effectiveness depends less on improving technological fidelity than on sophisticating pedagogical design.

In conclusion, this research contributes to a scientifically grounded vision of integrating virtual reality into awareness practices, demonstrating that its effectiveness depends less on its technological properties than on its articulation with appropriate pedagogical engineering.

Table des matières

Remerciements.....	3
Résumé	4
Abstract.....	6
Glossaire.....	12
Introduction	16
Partie 1 - Cadre théorique	19
Chapitre 1 - Trouble du Spectre de l'Autisme et fonctionnement cognitif	20
1. Principes généraux	20
1.1 Définitions du TSA.....	20
1.2 Fonctionnement cognitif chez les personnes avec TSA sans déficience intellectuelle	22
1.3 Modèle Enhanced Perceptual Functioning (EPF) de Mottron et al. (2006).....	27
2. Formations des professionnels sur le TSA et idée reçues.....	29
2.1 Les stéréotypes sur le TSA : changer les perspectives de conception	30
2.2 Notion d'attitude et de comportement sous le regard de la psychologie sociale et de l'ergonomie des IHM.....	33
2.3 Quelques modèles à connaître : De persuader à influencer, l'évolution de la vision du changement.....	36
2.4 Situations d'apprentissage et professionnels concernés	37
3. Synthèse.....	39
Chapitre 2 - La réalité virtuelle (incarnée)	41
1. Caractéristiques de la réalité virtuelle.....	42
1.1 Quelles définitions de la réalité virtuelle ?	42
1.2 Les concepts et modèles principaux en RV	45
1.3 Le modèle présence - incarnation	49
1.4 L'immersion technologique.....	58
1.5 La conception en réalité virtuelle sous CCU et les modèles précurseurs.	62
2. La réalité virtuelle pour l'apprentissage	64
2.1 Les intérêts possibles de la RVI pour l'apprentissage	64
2.2 Les limites de la RVI à prendre en compte pour les tâches d'apprentissages.....	67
2.3 L'acquisition de connaissances et compétences en RV - le facteur humain.....	76
2.4 Le CAMIL – présence et agentivité au cœur de l'apprentissage en RV	82
3. Synthèse.....	89
Chapitre 3 - Problématique et objectifs.....	91
1. Relations causales principales :	96
1.1 L'immersion et le contrôle augmentent la présence et l'agentivité.....	96
1.2 La fidélité de représentation facilite l'incarnation de l'avatar avec TSA.....	96
1.3 L'incarnation et la présence génèrent de l'empathie envers les personnes avec TSA	97
1.4 L'empathie et la motivation médiatisent l'effet sur la compréhension	97
1.5 La compréhension améliore les attitudes et réduit les stéréotypes	98
1.6 Une charge cognitive trop élevée peut diminuer l'efficacité de l'intervention	98
2. Synthèse des mécanismes théoriques	99
3. Questions de recherches et hypothèses	100
Partie 2 – Ingénierie d'un outil RVI pour la sensibilisation au TSA	102
Chapitre 1 - Motivations.....	103

1. Les ateliers itératifs et le scénario pédagogique.....	104
1.1 Cadre et finalités.....	105
1.2 Objectifs pédagogiques généraux.....	105
Chapitre 2 - La conception de la maquette	109
1. Scène 1 - Le tutoriel.....	110
2. Scène 2 – l’expérience immersive.....	111
2.1 La phase de présentation de l’environnement.....	111
2.2 L’expérience immersive TSA.....	112
2.3 Détails des tâches pédagogiques et objectifs	115
3. Scène 3 - La résolution du problème.....	130
3.1 Phase de résolution	130
3.2 La phase de renforcement.....	131
3.3 La phase de sortie	132
Chapitre 3 - Développement du prototype et du MVP	133
1. Cadre technologique.....	134
1.1 Choix du moteur de jeu	134
1.2 Langage de programmation et IDE.....	134
1.3 Périphérique cible : Meta Quest 3 HMD stand-alone.....	134
1.4 Chronologie de développement	134
2. Architecture fonctionnelle du prototype	135
2.1 Organisation du projet Unity	135
2.2 Scripts clés et logique comportementale	135
2.3 Gestion des assets et des performances	137
3. Conception des environnements et personnages	138
3.1 Salle de classe virtuelle.....	138
3.2 Avatars et agents/PNJ.....	138
4. Dispositifs d’interaction utilisateur.....	140
4.1 Contrôleurs et mapping des entrées	140
4.2 Objets interactifs.....	140
4.3 Feedbacks sensoriels	140
5. Séquences d’activité	141
6. Les modifications de l’outil en cours de thèse – les évolutions du prototype vers le MVP .	143
Partie 3 - Partie empirique	146
Chapitre 1 - Préambule.....	146
1. Rappel des hypothèses générales.....	148
2. Les variables	152
Chapitre 2 - Expérimentation 1 : test du prototype de sensibilisation au TSA - les effets de l’immersion sur la compréhension du TSA.....	153
1. Présentation/introduction	153
2. Hypothèses	156
3. Matériel et méthode.....	157
3.1 Participants	157
3.2 Application	158
3.3 Équipement.....	158
3.4 Procédures	158
3.5 Mesures	160
4. Résultats attendus et contribution	164

4.1	Cinétose	165
4.2	Présence	166
4.3	Agentivité	167
4.4	Empathie.....	168
4.5	Compréhension et stéréotypisation du TSA	168
5.	Discussion.....	173
5.1	Synthèse de la validation des hypothèses	173
5.2	Interprétation des résultats.....	174
5.3	Limites.....	177
5.4	Perspectives d'amélioration et de recherche.....	177
5.5	Conclusion.....	178
Chapitre 3 - Expérimentation 2 : test du MVP de sensibilisation au TSA - les effets de l'immersion sur la compréhension du TSA – V2		179
1.	Présentation / Introduction.....	179
2.	Hypothèses	181
3.	Matériel et Méthodes.....	182
3.1	Participants	182
3.2	Application	183
3.3	Équipement.....	183
3.4	Procédure (≈ 30 min).....	183
3.5	Mesures	184
4.	Résultats attendus et contribution	188
4.1	Cinétose	188
4.2	Charge cognitive.....	189
4.3	Présence.....	190
4.4	Agentivité	191
4.5	Empathie.....	192
4.6	Performances d'apprentissage	192
4.7	Attitude.....	193
5.	Discussion.....	195
5.1	Limites.....	198
5.2	Perspectives d'améliorations et de recherches.....	198
5.3	Conclusion.....	199
Discussion générale.....		200
1.	Rappel des objectifs et de la démarche scientifique	200
1.1	Rappel des hypothèses générales	201
2.	Apports empiriques et théoriques : une validation partielle du modèle CAMIL	204
2.1	Confirmation des mécanismes de présence, d'agentivité et d'empathie.....	204
2.2	Résultats contrastés sur les résultats d'apprentissages : limites du modèle CAMIL	205
3.	Limites et perspectives de recherche	207
3.1	Limites méthodologiques liées à la conception expérimentale.....	207
3.2	Limites liées aux participants et au contexte	207
3.3	Limites méthodologiques et instrumentales	208
3.4	Limites des analyses	209
3.5	Perspectives de recherche	210
4.	Conclusion générale	212
Références		214
Bibliographie.....		217
Annexes		245
ANNEXE A - CADRE THÉORIQUE : TROUBLE DU SPECTRE DE L'AUTISME.....		246

A.1	Définitions et classifications diagnostiques	246
A.2	Modèles théoriques explicatifs du TSA	250
A.3	Modèle Enhanced Perceptual Functioning (EPF)	258
ANNEXE B -	CADRE THÉORIQUE : RÉALITÉ VIRTUELLE IMMERSIVE	264
B.1	Embodiment et sentiment d'incarnation	265
B.2	Cinétose et confort en RV	280
B.3	Principes d'interfaçage en RV	301
B.4	Le tableau récapitulatif des avantages et inconvénients/limites de la RVI	315
ANNEXE C -	CONCEPTION DU DISPOSITIF DE RVI	321
C.1	Documentation technique pour les développeurs	321
C.2	Matériaux et assets du dispositif	328
C.3	Système de résolution de l'énigme	335
C.4	Captures d'écran du moteur Unity et du prototype	338
ANNEXE D -	INSTRUMENTS DE MESURE.....	347
D.1	Questionnaires de référence issus de la littérature	347
D.2	Batterie de mesures - expérimentation 1	353
D.3	Batterie de mesures expérimentation 2	359
ANNEXE E -	RÉSULTATS STATISTIQUES	368
E.1	Résultats expérimentation 1	368
E.2	Résultats expérimentation 2	374

Glossaire

BES : *Basic Empaty Scale* : Échelle de mesure de l'empathie

BES - A : *Basic Empaty Scale - Adult* : Échelle de mesure de l'empathie version adulte

CAO : conception assistée par ordinateur

CCU : Conception Centrée Utilisateur : Paradigme de conception de projet qui met l'utilisateur au cœur du processus. Norme ISO 9241-210 (2010 : Ergonomie de l'interaction homme-système - Partie 210 : Conception centrée sur l'opérateur humain pour les systèmes interactifs)

CIM : Classification Internationale des Maladies. Pour les besoins de ce travail, seule la codification du DSM a été prise en compte

Cinétose : Mal du simulateur, cybermalaise ou *Virtual sickness* (utilisé aussi pour le mal des transports). Situation de discordance entre la perception visuelle et le système vestibulaire provoquant nausée, vertige, vomissement...

CL : *Cognitive Load* : charge cognitive

CNDP : Centre National de Documentation Pédagogique

CLT : *Cognitive Load Theory*, de Sweller (1994)

CTML : *Cognitive Theory of Multimedia Learning*, de Mayer (2005)

CVTAE : *Control-Value Theory of Achievement Emotions* : théorie de la valeur de contrôle des émotions liées à l'accomplissement

DMM : *Distance Matters Measurement* : rapport d'échelle pour adapter la taille des objets en réalité virtuelle afin qu'ils restent cohérents et perceptibles

DSM : *Diagnostic and Statistical Manual of Mental Disorders*

DSM - IV - V - TR : *Diagnostic and Statistical Manual of Mental Disorders* : Manuel Diagnostique et Statistique des Troubles Mentaux, IV^{ème} et V^{ème} éditions, Texte Révisé

ECG : Électrocardiographie

EDA : *Electrodermic Activity* : Activité électrodermique

EEG : Electroencéphalographie

ERVE : Emotional Responses in Virtual Environments

EV : Environnement Virtuel

FH : Facteur Humain

FOV : *Field Of View* : champs de vision

FPS : *First Person Shoot* : Vue à la première personne

GEM : Groupe Entraide Mutuelle

HAS : Haute Autorité de la Santé

HMD : *Head Mounted Display* ou visiocasque. Casque obstructif/occlusif qui permet l'immersion dans les environnements virtuels.

HR : *Heart Rate* : battements cardiaques

IHM : Interaction Homme Machine

IPQ : Igroup Presence Questionary

IRMfc : Imagerie par Résonance Magnétique de Connectivité Fonctionnelle

JSPE : Jefferson Scale of Physician Empathy

NTIC : Nouvelles Technologies de l'Information et de la Communication

PD : *Pupillar Diameter* : diamètre pupillaire

RA : Réalité Augmentée

RGPD : Règlement Général sur la Protection des Données

RM : Réalité Mixte

Roomscale : Paradigme de conception qui permet aux utilisateurs de se déplacer librement en RV dans un espace défini par le dispositif

RV : Réalité Virtuelle

RVI : Réalité Virtuelle Incarnée

SAMI : Système d'Apprentissages Multimédia Interactif

SCR : *Skin Conductive Response* : réponse électrodermale

SDI : Sans Déficience Intellectuelle

SDT : *Self Determination Theory* : théorie de l'autodétermination

SNC : Système Nerveux Central

SOO : *Sense of Ownership* : sentiment de possession

SSQ : *Simulator Sickness Questionary* : questionnaire mesurant le niveau de cinétose dans un simulateur.

SUS : *System Usability Scale* : Échelle d'utilisabilité d'une interface

TAR : Théorie de l'Action Raisonnée

TCC : Thérapies Cognitivo-Comportementales

TCP : Théorie du Comportement Planifié

TED : Trouble Envahissants du Développement

TERV : Thérapies par Exposition à la Réalité Virtuelle

TND : Troubles Neuro-Développementaux

TPI : *Temple Presence Inventory*

TSA : Trouble du Spectre de l'Autisme

UI : *User Interface* : Interface Utilisateur

UX : *User experience* : Expérience Utilisateur

XR : Réalité étendue

Introduction

J'aimerais bien que mon professeur puisse vivre ma vie juste un instant, pour voir ce que je vois.

Propos recueillis lors des ateliers GEM la Bulle de Rodez, septembre 2022.

L'objectif de ce travail de thèse est de concevoir, développer et évaluer un dispositif innovant de sensibilisation aux troubles du spectre de l'autisme (TSA), fondé sur les principes de la réalité virtuelle incarnée (RVI). S'appuyant sur les spécificités immersives et perceptives des dispositifs de type casques occlusifs ou *Head Mounted Display* (HMD), la RVI permet d'immerger l'utilisateur dans un environnement virtuel en adoptant une perspective en première personne et en induisant un sentiment d'incarnation, à travers l'agentivité, la localisation de soi et la possession du corps virtuel. Ce type d'expérience ouvre la possibilité d'un changement de perspective susceptible d'impacter significativement les représentations, les attitudes et potentiellement les comportements des utilisateurs.

Les professionnels ciblés par ce projet sont les enseignants du secondaire et du supérieur, souvent confrontés à une méconnaissance du TSA, en particulier des profils dits « sans déficience intellectuelle » (TSA-SDI) tels que ceux décrits sous les terminologies dites d'« autisme de haut niveau » ou de « syndrome d'Asperger » chez Attwood (2022). Ces profils, bien que fréquemment rencontrés dans les établissements scolaires, restent mal compris, en partie en raison des stéréotypes persistants associés au TSA (Willay, 2020 ; Bennett & Goodall, 2016). Ces représentations inexactes peuvent générer des situations pédagogiques inadaptées, voire délétères pour les élèves concernés.

Afin de répondre à cette problématique, ce projet s'appuie sur les apports combinés de la psychologie cognitive, de la psychologie sociale, mais aussi de l'informatique et de l'ergonomie des interfaces immersives. L'outil développé vise à favoriser une meilleure compréhension des particularités cognitives et perceptives propres au TSA, telles qu'exposées dans le modèle EPF (*Enhanced Perceptual Functioning*) de Mottron et al. (2006), qui met en évidence un fonctionnement perceptif accru, notamment dans les modalités visuelles et auditives, ainsi qu'un traitement local renforcé des informations. La reconnaissance de ces spécificités permet de concevoir un contenu pédagogique pertinent, en cohérence avec les processus sous-jacents au traitement de l'information chez les personnes avec TSA.

L'intérêt d'utiliser la réalité virtuelle incarnée dans ce contexte trouve un écho dans plusieurs expériences récentes. Des projets immersifs tels que « *The Displaced* » de Salomon et Ismail (2015), ou encore « *Dans la peau d'une personne autiste* » de Lanovaz (2023), illustrent les tentatives de mobilisation de la RVI pour provoquer un changement de perspective chez les utilisateurs. Ces travaux suggèrent que la RVI, via des mécanismes d'immersion et d'incarnation, permet de transmettre de manière efficace un certain nombre d'éléments sensoriels, émotionnels et cognitifs qui peuvent influencer la compréhension de l'autre de manière plus ou moins profonde. La technologie ne se limite donc pas à une simple reproduction de stimuli visuels ou auditifs, mais favorise une intégration affective et sensorimotrice susceptible de modifier les schémas mentaux existants.

Dans cette perspective, la RVI est ici envisagée comme un outil de formation à fort potentiel, en particulier lorsqu'elle s'inscrit dans un cadre de conception pédagogique structuré. Le modèle CAMIL (*Cognitive Affective Model of Immersive Learning*) de Makransky et Petersen (2021) propose un cadre théorique robuste pour concevoir des environnements immersifs en maximisant leur efficacité pédagogique. Il identifie plusieurs facteurs clés (intérêt, auto-efficacité, motivation intrinsèque, incarnation, charge cognitive et autorégulation) ainsi que des variables modératrices (mémoire de travail, traits de personnalité, etc.) susceptibles de moduler l'impact du dispositif sur l'apprentissage. Dans cette optique, la charge cognitive, la qualité de l'interface et l'ergonomie des contenus deviennent des éléments centraux de la conception afin d'éviter toute surcharge ou distraction et de maximiser la présence et l'engagement.

Le développement de cet outil repose par ailleurs sur une approche de conception centrée utilisateur (CCU), intégrant les spécificités des enseignants en tant que public cible (Gohier et al., 2001). Ces derniers mobilisent au quotidien des compétences sociales complexes (savoir-faire sociaux, ajustements pédagogiques et gestion des interactions) qui doivent être prises en compte dans la conception de toute formation immersive visant une transformation des représentations. Comme le souligne Willay (2020), les dispositifs de formation efficaces doivent viser un « renouvellement des conceptions sur la personne dans sa globalité, et qui exprime des besoins de nature multiple ».

Cette thèse vise donc à interroger la pertinence de la RVI comme levier de compréhension et de sensibilisation au TSA, en lien avec les situations d'enseignement inclusif. À travers une approche expérimentale rigoureuse, il s'agira de tester plusieurs hypothèses relatives aux effets de la RVI sur la présence, l'agentivité, l'empathie, mais aussi sur l'évolution

des connaissances et des attitudes vis-à-vis des personnes avec TSA. L'analyse de ces effets permettra de déterminer si la RVI constitue, ou non, un outil efficace pour accompagner les enseignants dans une meilleure compréhension des besoins éducatifs spécifiques liés au TSA.

Par sa nature résolument pluridisciplinaire, cette thèse mobilise des connaissances issues de la psychologie cognitive et sociale, de l'informatique ainsi que des sciences de l'ingénierie numérique. Il ne prétend pas à l'exhaustivité mais propose une contribution originale à la réflexion sur les apports des technologies immersives dans la formation à l'inclusion des professionnels de l'éducation.

Ce travail s'inscrit dans un partenariat entre l'université Toulouse 2 Jean Jaurès et l'institut universitaire Champollion d'Albi, avec un financement de la région Occitanie.

Partie 1 - Cadre théorique

Cette première partie établit les fondements théoriques nécessaires à la compréhension de la problématique de recherche, qui articule deux champs disciplinaires distincts : le Trouble du Spectre de l'Autisme (TSA) d'une part, et la Réalité Virtuelle Immersive (RVI) comme modalité d'apprentissage d'autre part. L'enjeu consiste à examiner comment la RVI peut constituer un levier pertinent pour la sensibilisation et la formation des professionnels concernant le fonctionnement cognitif des personnes avec TSA.

Le premier chapitre est consacré au TSA et au fonctionnement cognitif spécifique des personnes concernées, particulièrement sans déficience intellectuelle. Après avoir précisé les définitions et classifications actuelles du TSA, les particularités cognitives seront explorées à travers le modèle *Enhanced Perceptual Functioning* (EPF) de Mottron et al. (2006). La question de la formation des professionnels sur le TSA sera ensuite abordée, en examinant les stéréotypes et attitudes qui influencent leurs pratiques, sous l'angle de la psychologie sociale et de l'ergonomie des interfaces homme-machine. Cette réflexion conduit à questionner les mécanismes de changement conceptuel nécessaires pour transformer durablement les représentations.

Le deuxième chapitre porte sur la RVI comme technologie et dispositif d'apprentissage. Les caractéristiques fondamentales de la RV seront présentées, ainsi que les concepts clés (présence, incarnation, immersion technologique) et les modèles théoriques structurants. Une attention particulière est accordée aux apports potentiels de la RVI pour l'apprentissage, ainsi qu'à ses limites, notamment en matière de charge cognitive et de transfert des connaissances. Le modèle CAMIL (*Cognitive Affective Model of Immersive Learning*) de Makransky et Petersen (2021), qui positionne la présence et l'agentivité au cœur des processus d'apprentissage en RV, constitue le cadre théorique principal guidant cette démarche expérimentale.

Enfin, Le troisième chapitre synthétise ces apports pour formuler la problématique, identifier les relations causales entre variables et énoncer les questions de recherche et hypothèses structurant les expérimentations présentées dans les parties suivantes.

Chapitre 1 - Trouble du Spectre de l'Autisme et fonctionnement cognitif

1. Principes généraux

Le trouble du spectre de l'autisme (TSA) constitue un domaine central de recherche en sciences humaines et cognitives en raison de ses spécificités développementales et sociales. Ce chapitre propose d'en préciser la définition, à partir des critères diagnostiques actuels, afin de mieux situer les difficultés rencontrées dans les interactions sociales, la communication et les comportements restreints. Le fonctionnement cognitif des personnes avec TSA sans déficience intellectuelle sera ensuite abordé, en particulier à travers le modèle *Enhanced Perceptual Functioning* (EPF) de Mottron et al. (2006), qui éclaire certaines particularités perceptives et cognitives. Cette analyse contribuera à identifier les spécificités des personnes avec TSA qu'il convient de traduire dans un environnement de réalité virtuelle. Sera ensuite abordées le point de vue des professionnels de l'éducation autant en termes d'idée reçues que de formations proposées à ce sujet. Ces éléments semblent, en effet, déterminants dans la qualité de l'accompagnement et dans l'identification des besoins de la population visée. Une synthèse préparera les développements ultérieurs de la thèse en dernière partie.

1.1 Définitions du TSA¹

Définitions et cadres institutionnels : Le TSA est défini par plusieurs organismes officiels. Le gouvernement français le décrit comme un ensemble de troubles neurobiologiques affectant le développement, caractérisés par des dysfonctionnements dans les interactions sociales, la communication, les comportements et les activités. L'OMS (2023) y voit des troubles comportementaux et cognitifs survenant durant le développement provoquant des difficultés importantes dans l'acquisition de fonctions intellectuelles, motrices, langagières ou sociales. La Haute Autorité de Santé (2023) le définit comme un trouble neuro-développemental selon deux dimensions symptomatiques du DSM-V : les déficits persistants de la communication et des

¹ Voir annexe A.1 : [Définitions et classifications diagnostiques](#) pour plus de détails p. 250.

interactions sociales, et le caractère restreint et répétitif des comportements, intérêts ou activités, complétés par un niveau de sévérité basé sur l'aide requise.

Évolution des classifications : Depuis Bleuler (1911), le TSA a connu plusieurs évolutions. Jusqu'en 2019, la CIM-10 classait le TSA dans les Troubles Envahissants du Développement (TED) tandis que le DSM proposait la nomenclature TSA, qui devient progressivement la référence. Le DSM-V (2013) a introduit des modifications fondamentales, rejointes par la CIM-11 (2019), bien que des débats persistent, notamment sur l'absence de référence aux anomalies sensorielles. La « triade autistique » historique a cédé la place aux deux dimensions actuelles du DSM-V.

Critères diagnostiques et manifestations : Le DSM-V établit trois critères : (1) déficits persistants dans la communication et les interactions sociales observables dans divers contextes, (2) caractère restreint et répétitif des comportements, intérêts ou activités, incluant stéréotypies motrices, adhérence aux routines, intérêts restreints et hyper- ou hypo-réactivité sensorielle, (3) manifestation précoce des symptômes, dès la petite enfance. Les difficultés relationnelles sont centrales : capacité limitée à établir des relations, initier des conversations ou prendre des initiatives sociales, accompagnées de comportements répétitifs et d'intérêts spécifiques dépassant les variations habituelles du développement.

Hétérogénéité et étiologie : La notion de spectre symbolise l'hétérogénéité phénotypique et génotypique ainsi que les variabilités interindividuelles. Cependant, Mottron et Bzdok (2020) rappellent que cette hétérogénéité relève davantage d'un consensus que d'un fait scientifique avéré. L'étiologie du TSA est multifactorielle : origine génétique (environ 50% selon Bourgeron, 2016, avec prévalence paternelle selon Iossifov, 2023), neurobiologique (Chen et al., 2015) ou facteurs environnementaux (prématurité, substances, métaux lourds). Cette compréhension permet d'appréhender la complexité du trouble et ses variations individuelles.

1.2 Fonctionnement cognitif chez les personnes avec TSA sans déficience intellectuelle

Pour représenter et faire comprendre le fonctionnement de personnes avec TSA pour opérer un changement, il est important de comprendre leur fonctionnement cognitif. Dans le cadre de ce travail, l'outil étant à destination de professionnels travaillant auprès de personnes avec TSA susceptibles d'intégrer des cursus de formation, seuls les profils sans déficience intellectuelle (SDI) sont pris en compte. Pour les profils avec déficience intellectuelle, il faudra investiguer un autre champ de recherche dédié spécifiquement.

1.2.1 *L'atypisme du TSA : notion d'intelligence et variant humain*

Les conséquences de l'étiologie du trouble sur le fonctionnement cognitif sont établies depuis plus de deux décennies. Ainsi, ces facteurs neurobiologiques vont entraîner des fonctionnements atypiques. Pour appuyer l'importance du neurobiologique, des exemples comme la théorie du « monde intense » de Markram et Markram (2010) expliquent que l'amygdale serait à l'origine d'une hyperréactivité. Selon les auteurs, le TSA serait lié à une activité excessive au sein des microcircuits neuronaux locaux, se manifestant par une sensibilité accrue aux stimuli et une capacité anormalement amplifiée à modifier les connexions synaptiques. Les travaux de Muller et al. (2011) en imagerie par résonance magnétique de connectivité fonctionnelle ont montré des profils complexes modulés par une connectivité cérébrale atypique qui dépend de paramètres développementaux et contextuels. Par ailleurs, Rubenstein et Merzenich (2003) ont montré que les personnes avec TSA ont des difficultés à différencier le « bruit » des informations. Les auteurs expliquent, via leur modèle cognitif du TSA fondé sur le déséquilibre entre excitation et inhibition, que l'excitation perturberait le traitement normal de l'information et les comportements. Les travaux de Belmonte et al. (2004) expliquent que certaines difficultés à différencier l'information pertinente de celles qui ne le sont pas (« le bruit ») pourraient être à l'origine des particularités perceptives et sensorielles des personnes avec TSA. Ces caractéristiques atypiques dues à des facteurs neurobiologiques entraîneraient donc un fonctionnement cognitif spécifique.

Le fonctionnement cognitif et la question de l'intelligence qui en découle est un débat complexe avec le TSA. Certaines idées reçues populaires dressent un portrait négatif du TSA (enquête IFOP de 2016 : 79% des français estiment qu'il existe une discrimination à l'encontre des personnes avec TSA, avec association du TSA à une insulte et la notion d'étrangeté).

Paradoxalement, sur le plan de l'intellect, certains stéréotypes positifs existent comme le préjugé entourant le syndrome du savant (cf. la comparaison populaire maladroite avec le protagoniste du film « *Rain man* », incarné par Dustin Hoffman, avec ses capacités mnésiques hors du commun). Néanmoins, que ce soit avec le syndrome du savant ou le type Asperger, le TSA concerne un pan de personnes qui montrent des habiletés particulières qui ne doivent pas être corrélées avec des contreparties obligatoirement négatives. La question de « l'intelligence » se pose donc. De quel type est-elle ? Y a-t-il des similitudes avec les neurotypiques et, si oui, quelles sont-elles ? Comment fonctionne le traitement de l'information chez les personnes avec TSA ? Ces questions ne sont que des exemples et celle de l'intelligence, trop subjective et sujette à controverse, est ramenée au fonctionnement cognitif. Pour information, cette notion de neurotypique reste un postulat pour simplifier l'appellation des personnes n'ayant pas de TSA, ce qui est bien trop vaste pour proposer de véritables points de comparaison.

Mottron (2016) explique que le TSA est un variant humain, avec ses avantages et désavantages. Selon lui, l'intelligence fluide, qui concerne la capacité à résoudre des problèmes nouveaux et abstraits, est la même que les neurotypiques. Grâce au surfonctionnement perceptif et au niveau de traitement quasi indépendant, les performances des personnes avec TSA sont accrues. Les neurotypiques, eux, voient leurs performances baisser à cause des interactions entre les systèmes de traitements. Avec ces travaux, on peut déjà voir que la question du fonctionnement cognitif des personnes avec TSA est plus complexe qu'il n'y paraît. Comme abordé dans la partie précédente, les spécificités des caractéristiques anatomiques et fonctionnelles atypiques des personnes avec TSA vont entraîner un fonctionnement qui sera différent des personnes dites neurotypiques.

Bled (2022) explique que l'expertise visuelle, qui contribue à des manifestations de l'intelligence plus élevée, serait due à une attribution particulière dans les régions corticales spécifiques. Cette attribution a pour conséquence de faciliter d'autres aspects tels que le calcul ou le raisonnement. Bled relève l'existence d'intérêts particuliers qui constitueraient une manifestation comportementale visible des processus d'acquisition de l'expertise chez les personnes avec TSA. D'après Baron-Cohen et al. (2009), les atypies sensorielles pourraient s'expliquer par une manière spécifique de traiter l'information qui se manifeste à différents niveaux du système sensoriel.

Il est important de préciser qu'un des aspects de l'expérience immersive conçue dans ce travail repose sur le traitement des informations visuelles pour résoudre une tâche. Pour cela, il

est intéressant de regarder les travaux de Stevenson et Nonack (2018). Les auteurs ont observé les mouvements oculaires chez des participants présentant divers degrés de TSA (faible, moyen et fort) dans une tâche de rotation mentale de figures cubiques tridimensionnelles. Les résultats indiquent une différence selon le degré : les participants avec degrés faible et moyen ont un comportement différent de ceux avec un degré fort, comportement renforcé également par le genre. Il semblerait donc que les stratégies soient différentes, locales pour les traits forts et plutôt heuristiques pour les traits faibles, ce qui indique que les emplacements des informations dans l'expérience doivent être disposés en fonction du niveau estimé du TSA que l'on veut simuler.

1.2.2 *Le traitement multi sensoriel*

Le traitement multi sensoriel est un élément important à aborder car il a un impact non négligeable sur le fonctionnement cognitif et également perceptif des personnes avec TSA (Mottron et al., 2009). Plus spécifiquement, c'est le traitement de l'information via le traitement multi sensoriel qui va avoir une influence sur le fonctionnement cognitif. Pour information, dans le guide « *Scolariser les élèves autistes ou présentant des troubles envahissants du développement* », des recommandations sur le fonctionnement multi sensoriel sont proposées.

Pour aborder cette notion de multi sensoriel, il est intéressant de regarder les travaux de Iarocci et McDonald (2006) qui explorent la notion de surcharges sensorielles, c'est-à-dire le fait de recevoir plus d'informations que le corps ne peut en traiter. Leurs travaux, qui se recoupent avec ceux de Robertson et Baron-Cohen (2017) ou encore ceux de Van Etten et al. (2017) sur la prévalence de l'hypo et hyper-sensibilité, montrent que ces particularités sensibles sont un processus essentiel au traitement des informations chez les personnes avec TSA. Bien que ce fait ne soit pas spécifiques à ces personnes, dans le cadre du TSA, il reste plus important que dans d'autres troubles neuro-développementaux comme l'expliquent Rogers et al. (2003). Ainsi, comme le suggère Mottron et Burack (2001), l'hypermensibilité sensorielle pourrait entraîner une focalisation accrue sur les détails, ce qui influencerait leur perception en favorisant le traitement plus localisé de l'information. Ce traitement favoriserait une attention accrue aux détails. Grandin (2009) explique que les sujets neurotypiques vont passer à côté des détails que les personnes avec TSA ne rateront pas à cause d'un meilleur traitement. Bled (2022) conclut sur cette sensibilité spécifique avec le témoignage d'une personne avec TSA qui explique que « *C'est l'approche perceptive de l'environnement qui génère notre mode de pensée et va avoir une influence sur nos représentations mentales* » *Témoignage de Stef Bonnot-Briey cité par Bled (2022)*.

1.2.3 Les pics d'habiletés ou îlots de compétences

Les pics d'habiletés ou îlots de compétences sont des notions importantes lorsque l'on aborde le TSA. Ce sont en effet ces spécificités, souvent spatiales et perceptives, mais pas uniquement, qui ont mis en lumière au niveau du grand public certains atouts du TSA avec des cas comme celui de Daniel Tammet, un écrivain anglais avec TSA possédant une synesthésie et étant capable de réciter les 22514 premières décimales du nombre Pi. Il est pertinent de s'intéresser à cet aspect du TSA car il est à la base du concept RV de ce projet. C'est par les pics d'habiletés que sera représenté une partie du traitement cognitif d'une personne avec TSA pour montrer non pas un handicap ou une difficulté comme le font souvent les expériences de découverte du TSA, mais une autre forme de pensée et de fonctionnement qui, éventuellement, peut devenir un atout dans certains cas de figure.

On retrouve, dans le cadre des matrices de Raven et les tests de QI qui y sont associés, un net avantage des personnes avec TSA. En utilisant leurs pics d'habiletés, elles performant mieux que les autres dans ce type de tâches. Bled (2022) explique que, toujours dans le même type de tâches, les cubes de Kohs, les personnes avec TSA montrent leurs avantages, perceptif et spatial, en assemblant les cubes bicolores, en modèles imposés, mieux que les neurotypiques. On retrouve cette supériorité perceptuelle dans le subtest des cubes des échelles d'intelligence de Wechsler et al. (2006) ou encore dans le test des figures emboîtées de Jolliffe et Baron-Cohen (1997). D'un autre côté, Mottron et al. (1999) montraient les importantes capacités mnésiques des personnes avec TSA ou encore leurs capacités à reproduire les « figures impossibles ». Bled (2022) cite également les aptitudes sur les tâches de recherches visuelles. Joseph et al. (2009) et Bertone et al. (2005) parlent des tâches de discrimination visuelle qui seront également très sollicitées dans l'expérience de RVI lors des tâches d'acquisition des règles de fonctionnement. Enfin, Park (2023) explique dans sa revue de littérature que l'hétérogénéité des profils des personnes avec TSA est très importante et que, malgré certains retards dans des domaines spécifiques comme le langage ou les compétences sociales, le TSA peut engendrer ces pics d'habiletés qui peuvent aller jusqu'au syndrome du savant. Ces éléments illustrent bien les particularités des personnes avec TSA et les difficultés à évaluer leur intelligence par les autres alors qu'ils sont capables de mobiliser d'importantes compétences qui leur permettent d'être plus performants dans de nombreuses situations.

1.2.4 Les modèles du fonctionnement cognitif du TSA

Les travaux de Muller et al. (2011) sur l'imagerie mentale apportent la confirmation d'une différence de fonctionnement. Les TSA présentent un style cognitif visuel spécifique, comportant de grandes capacités d'imagerie mentale et une préférence pour un apprentissage de type visuel. « Les capacités d'imagerie mentale préservées ou supérieures dans l'autisme et les particularités dans l'inspection des images mentales peuvent être liées à leur fonctionnement perceptif atypique. » (Bled, 2022). Bled continue en présentant le fonctionnement cognitif avec plusieurs théories expliquant les particularités de ce fonctionnement qui se résument sous quatre catégories : les théories perceptives, celles portant sur la cognition, les théories généralistes et computationnelles (plus récentes), et enfin les théories bayésiennes²

Pour ce travail, le choix s'est porté sur les théories perceptives et le modèle EPF de Mottron et al. (2006) (une mise à jour du modèle de Mottron & Burrack de 2001) et sa variante *veridical mapping* de Bouvet et al. (2014). Bien que d'autres cadres théoriques puissent également éclairer la conception de l'outil, le modèle EPF présente des avantages spécifiques dans le contexte de ce projet. Ce modèle rend compte du fait que certaines personnes avec TSA peuvent faire preuve de compétences exceptionnelles dans des domaines perceptifs ou techniques, sans que ces pics d'habiletés soient nécessairement associés à un déficit global de l'intelligence (Rubenstein & Merzenich, 2003 ; Müller et al., 2011). Il offre ainsi une alternative précieuse aux approches strictement déficitaires, en mettant en avant une organisation neurocognitive différente plutôt que déficiente. Par ailleurs, le modèle EPF est particulièrement adapté pour comprendre l'hétérogénéité des profils dans le TSA, en incluant les cas de compétences spécifiques marquées (jusqu'au syndrome du savant) comme l'ont montré plusieurs études dont celle d'Happé et Vital (2009) et Mottron et al. (2006). En cela, le modèle EPF constitue un cadre théorique cohérent avec les objectifs pédagogiques de ce travail, notamment en ce qu'il permet d'intégrer les forces perceptives dans la conception d'environnements interactifs tout en tenant compte de la manière singulière dont l'information est traitée par les personnes avec TSA.

² Voir annexe A.2 : [Modèles théoriques explicatifs du TSA](#) pour plus de détail p. 254.

1.3 Modèle Enhanced Perceptual Functioning (EPF) de Mottron et al. (2006)³

Le modèle EPF de Mottron et al. (2006) est un modèle majeur dans le TSA. Issu des théories perceptives et d'un long travail d'élaboration (tiré du modèle originel de Mottron & Burack de 2001), il permet d'appréhender les différences de traitements des personnes avec TSA en expliquant leurs performances supérieures. Plusieurs domaines sont concernés : les tâches cognitives dites de « bas niveau », les tâches cognitives complexes (implication atypique de la perception) et les tâches de la vie quotidienne.

Voici la liste des raisons de la performance :

- Perception visuelle et auditive orientée localement.
- Discrimination de bas niveau améliorée.
- Utilisation d'un réseau plus postérieur dans des tâches visuelles « complexes ».
- Amélioration de la perception des stimuli statiques de premier ordre.
- Diminution de la perception des mouvements complexes.
- Autonomie du traitement de l'information de bas niveau vers les opérations d'ordre supérieur.
- Relation différentielle entre la perception et l'intelligence générale.

Les auteurs englobent sous le terme « perception » les opérations supérieures chez les personnes avec TSA, perception large allant de la détection des caractéristiques à la reconnaissance des formes. Ainsi, Mottron et al. (2006) expliquent que selon le modèle EPF, la prédominance du traitement perceptif sur les fonctions cognitives de plus haut niveau entraîne un déséquilibre entre les différents niveaux de traitement de l'information. Elle rend les perceptions plus difficiles à réguler et peut interférer avec le développement d'autres compétences ou comportements. Selon eux, au-delà d'un concept réducteur de déficience que l'on pensait être, à une époque, source du TSA, ce sont des aptitudes spécifiques qui se manifestent. Elles enracinent des capacités spéciales de traitement de l'information perceptuelle, ce qui remet en cause du dogme « *la performance supérieure doit être liée à un mécanisme causal pathologique* ». Ainsi, avec l'EPF, le TSA n'est pas une déficience mais un fonctionnement alternatif qui peut s'avérer beaucoup plus performant dans certains cas.

³ Voir annexe A.3 : [Modèle Enhanced Perceptual Functioning \(EPF\)](#) pour plus de détail p. 263.

Mottron et al. (2006), avant de présenter les huit principes clés du modèle EPF, concluent en expliquant que la perception des personnes avec TSA est différente de celle des neurotypiques dans le sens qu'elle a été mesurée comme plus performante en laboratoire et en situation réelle.

Pour résumer, le modèle EPF met en lumière un traitement perceptif particulièrement développé, orienté vers les détails, et ce tant au niveau visuel qu'auditif. Il suggère une prédominance de l'activité cérébrale dans les régions sensorielles primaires et précoces, même lors de tâches engageant des informations complexes ou à forte valeur sociale. Ce modèle s'appuie sur les avancées récentes en neuroimagerie fonctionnelle et anatomique concernant la perception de bas niveau chez les personnes TSA, notamment dans l'organisation particulière du cortex visuel et dans la présence fréquente d'anomalies au niveau des régions postérieures du cerveau.

2. Formations des professionnels sur le TSA et idée reçues

Le Centre National de Documentation Pédagogique (CNDP) présente un état des lieux du trouble puis des recommandations sur le plan de la scolarisation : collaboration entre les différents partis impliqués, évaluation des potentiels d'action ou encore adaptation des démarches pédagogiques. Ces recommandations sont accompagnées d'exemples précis qui permettent de représenter des cas concrets. Il est intéressant de repérer ceux qui concernent les personnes avec TSA de haut niveau. L'objet de ce travail étant destiné à un public professionnel, les enseignants et autres professionnels susceptibles de transmettre du savoir à des personnes avec TSA SDI, il convient d'aborder l'état des lieux de leur formation. Le travail de Willay (2020) est une référence à ce sujet.

L'étude de Mac Carthaigh et Lopez (2020) sur les étudiants en soins infirmiers britanniques et sud-coréens montre que transmettre des connaissances factuelles ne suffit pas à provoquer un changement d'attitude sur le TSA. L'acquisition de connaissances factuelles sur le TSA ne permet pas de mieux le comprendre ni de le conceptualiser, ce qui ne permet pas d'adopter des comportements plus appropriés. Ce constat est un des points centraux de ce projet : la formation de professionnels susceptibles de travailler avec des personnes avec TSA.

Willay (2020) a mis en avant trois points pour appréhender la formation des professionnels du TSA :

- Approfondir la compréhension des mécanismes de la cognition sociale liés aux représentations des personnes avec TSA.
- Repérer, dans le contexte particulier du TSA, les formes d'idées reçues véhiculées par les professionnels et examiner leur influence possible sur les pratiques d'accompagnement.
- Contribuer à la recherche en ingénierie de la formation par l'exploration des obstacles cognitifs qui freinent la compréhension du fonctionnement TSA et la qualité de l'accompagnement proposé.

Il ne suffit donc pas de donner des connaissances sur le TSA pour former correctement des professionnels. Point abordé plus haut, le TSA est un fonctionnement qui possède une forte connotation sociale globalement négative. Stéréotypes et représentations sont le lot que subissent les personnes avec TSA, malgré des campagnes d'information. Il est donc essentiel d'être familiarisé avec les différents processus qui participent à leur construction, surtout quand

ces derniers participent à la création du socle de connaissances sur le TSA. Il convient donc de réfléchir à ces notions, et à la manière de les aborder, pour provoquer un changement, qu'il soit social, cognitif, perceptuel ou motivationnel.

Les travaux de Mishina et al. (2012) sur les jugements sociaux apportent un angle de réflexion intéressant. Du fait que les jugements passés influencent les perceptions futures, le passif du TSA dans les idées reçues populaires biaise encore sa perception malgré l'évolution sociétale.

2.1 Les stéréotypes sur le TSA : changer les perspectives de conception

La conception d'outils plus adaptés passe par des processus de compréhension et d'analyse des connaissances tout en tenant compte du fonctionnement cognitif et de ce qu'il privilégie par économie. La notion de changement conceptuel est un changement de paradigme basé sur une approche constructiviste des connaissances avec un rôle actif de l'individu (Sinatra & Pintrich, 2003, dans leur ouvrage généraliste). Sur les postulats du modèle EPF de Mottron et al. (2006), des auteurs ont proposé, avec un socle pluridisciplinaire, une nouvelle conception du TSA. Les travaux de Baron-Cohen (2017) proposent la notion de « *neurodiversité comme condition humaine* », travaux qui se basent eux-même sur ceux de Runswick-Cole (2013) expliquant les différences neurologiques. Selon Jaarsma et Welin (2012) le TSA est une variation naturelle de l'humain. Enfin, selon Baron-Cohen (2017), l'acceptation de ces changements au niveau sociétal est une révolution conceptuelle qui s'inscrit directement dans la lutte contre les discriminations.

Avant d'aller plus loin, il convient de présenter des définitions du stéréotype et des représentations.

- **La notion de stéréotype** : « Croyances partagées concernant les caractéristiques personnelles, généralement des traits de personnalité mais souvent aussi des comportements d'un groupe de personnes ». (Leyens et al., 1996, p. 24).
- **La notion de représentation** : « Il y a activité de représentation, lorsqu'un objet ou lorsque des éléments d'un ensemble d'objets, se trouvent exprimés, traduits, figurés sous la forme d'un nouvel ensemble d'éléments et qu'une correspondance systématique se trouve réalisée entre l'ensemble de départ et l'ensemble d'arrivée » (Denis, 1989, p. 21).

Il convient de préciser que le changement conceptuel est issu d'un croisement de disciplines et permettra d'introduire doucement les notions de changement d'attitude et de comportement.

Plusieurs modèles, comme celui de Vosniadou (1994), présentent les fondamentaux de cette notion de changement qui utilise les fonctions exécutives, la métacognition ou encore la motivation avec la notion d'utilité et de valeur de la tâche. Ils rappellent les notions de résistance et d'ancrage issues d'idées reçues, particulièrement présentes sur le plan cognitif, qui rendent difficile le changement d'attitude et de comportement. Willay (2020) ajoute que le travail de l'apprenant est long et fastidieux et qu'il ne garantit pas de résultats. Les risques pesant sur la conception des formations (complexité, évolution et variabilité du TSA qui ne sont pas suffisamment intégrés) peuvent entraîner un transfert de connaissances entravé, un traitement des informations défectueux et un raisonnement inductif généraliste. Les détails de toutes ces notions ne seront pas abordés ici, cependant, il convient d'aborder les fondamentaux pour bien comprendre que faire changer l'individu est une tâche complexe, non pas dans sa définition psychologique mais dans celle littérale. Plusieurs approches existent, comme celle de Williams et Tolmie (2000) qui montrent que la collaboration sociale peut jouer un rôle important dans le développement conceptuel. Dans leur expérience sur des connaissances en biologie, les auteurs ont montré que l'apprentissage en groupe peut faciliter le changement conceptuel, s'il est bien orchestré. Il existe également, dans l'ingénierie pédagogique, des pistes de réflexions issues des perspectives constructivistes. Duit et Treagust (2012) montrent, avec leur principe de conflit cognitif productif, que le changement conceptuel implique des ajustements cognitifs ainsi que des dimensions affectives, motivationnelles et contextuelles. Ainsi, la prise de conscience par l'apprenant de l'inadéquation de ses conceptions peut favoriser un apprentissage profond.

Comprendre les mécanismes cognitifs et motivationnels impliqués dans le changement conceptuel est indispensable pour élaborer des actions formatives qui permettent un véritable changement, particulièrement lorsqu'il s'agit de modifier des stéréotypes bien ancrés. D'après le modèle de la vraisemblance d'élaboration (ELM) de Petty et Cacioppo (1986), un changement d'attitude durable repose sur un traitement de l'information en profondeur. Bien qu'ancien, ce modèle explique que le traitement passe par une analyse critique des contenus, renforcée par une implication personnelle et une attention soutenue. Cela signifie que les approches brèves ou superficielles sont rarement suffisantes pour provoquer une transformation stable des représentations. Les travaux de Devine et al. (2012) dans le domaine des biais implicites, montrent en effet que des campagnes « *one-shot* » produisent rarement des résultats

durables, sauf si elles sont accompagnées d'un entraînement continu et d'une appropriation des nouveaux schémas cognitifs. Ces recherches soulignent l'importance d'un engagement prolongé et structuré pour assurer la pérennité du changement d'attitude. Ainsi, une session brève de 20 minutes, bien qu'elle puisse initier une prise de conscience, risque de manquer d'impact à moyen terme sans renforcement de ce processus.

Néanmoins, certains travaux, comme ceux de Lehmann-Grube et al. (2023), montrent que, dans certains cas, des interventions courtes peuvent avoir des effets concluants telles qu'une réduction des stéréotypes, une amélioration des attitudes ou encore une diminution des biais d'évaluation chez les élèves à faible statut social. Ces interventions incluent trois étapes (un feedback individuel, du contenu théorique et des stratégies pratiques) qui permettent de réduire le biais de jugement chez des élèves enseignants.

En somme, pour que la formation amorce un changement de point de vue sur le TSA, il est indispensable de combiner :

- L'identification claire des idées reçues et stéréotypes existants chez les enseignants.
- La mobilisation d'un traitement cognitif en profondeur (voie centrale du modèle ELM)
- Un engagement prolongé au-delà de la simple exposition initiale afin de consolider l'évolution cognitive et attitudinale dans la durée.

Ces éléments méthodologiques permettent de positionner l'expérience immersive de 20 minutes dans une perspective réaliste, une première impulsion vers la compréhension, mais nécessitant idéalement un suivi ultérieur pour assurer un vrai ancrage.

Pour synthétiser, le développement de la pensée complexe constitue donc un levier essentiel pour dépasser les idées reçues ou simplistes. Cela permet de mieux comprendre les objectifs de l'accompagnement individualisé et de mobiliser de manière pertinente un ensemble de savoirs nécessaires à l'adaptation des pratiques en situations réelles. Changer les attitudes pour changer les comportements sous le regard de la psychologie sociale, c'est persuader et influencer l'autre durablement et en profondeur pour faire évoluer son point de vue. Bien que dans ce travail, il ne soit abordé que la partie « compréhension » du TSA avec une tentative d'amorce de changement d'attitude, il est important de préciser ces notions pour préparer un socle de conception cohérent.

2.2 Notion d'attitude et de comportement sous le regard de la psychologie sociale et de l'ergonomie des IHM

Avant d'aller plus loin, il convient d'opérer un détour théorique indispensable sur les notions d'attitude et de comportement, qui constituent des objectifs fondamentaux dans des dispositifs de médiation tels que celui de ce projet. Si l'outil se limite à une tentative de compréhension du TSA, avec pour ambition secondaire une amorce de changement d'attitude, il s'inscrit toutefois dans une logique à plus long terme, visant à promouvoir des évolutions durables des représentations sociales. Ce lien entre représentations et TSA est notamment éclairé par les processus de catégorisation sociale de Brun et al. (2015). Les auteurs expliquent, sous l'angle de la psychiatrie, que les catégories psychiatriques ne sont pas fixes mais qu'elles évoluent selon des dynamiques sociales, institutionnelles et scientifiques. Ces catégories, qui structurent les interactions et déterminent la manière dont sont perçus les individus porteurs de différences, peuvent les stigmatiser ou les enfermer dans des identités pathologisées.

Pour qu'un dispositif puisse être persuasif, qu'il soit technologique ou non, il doit permettre, à court ou moyen terme, l'intériorisation entendue comme « *l'appropriation d'éléments de connaissance (i.e. cognitions) ou comportements nouveaux* », Fointiat et Barbier (2015). Cette perspective rejoint les apports de la captologie (*Computer As Persuasive Technologies*), concept proposé par Fogg (2003) qui croise les technologies de communication et les théories de la persuasion dans une visée d'efficacité interactionnelle. L'objectif de cette approche est de penser la conception d'interfaces non seulement fonctionnelles mais aussi capables d'engendrer des changements psychosociaux chez leurs utilisateurs.

Changer une pensée, qu'il s'agisse d'un stéréotype, d'un point de vue ou d'une simple idée, ne peut se réduire à une transmission unilatérale d'un message. En effet, selon Higgins et Walker (2012), pour parvenir à une persuasion efficace, que ce soit pour transmettre un savoir, désamorcer un conflit ou partager des idées, il est essentiel que l'orateur tienne compte à la fois de sa propre crédibilité (ethos), de la logique et de la cohérence de son discours (logos) ainsi que des émotions et de la sensibilité de son public (pathos). Cette vision, héritée de la rhétorique aristotélicienne, demeure centrale dans les théories contemporaines de la communication persuasive notamment lorsqu'il s'agit d'induire un changement profond et durable.

L'un des enjeux centraux du projet repose ainsi sur la compréhension en tant que préalable nécessaire à toute transformation significative. Il s'agit de faire comprendre, puis faire intégrer durablement, une réalité complexe telle que celle du TSA, dans le but de favoriser des

interactions plus bienveillantes et informées. L'ambition est d'observer si l'outil mobilisé, en l'occurrence un dispositif de RVI, permet effectivement une telle évolution des attitudes. Toutefois, ce processus n'est pas linéaire ni automatique. Les travaux de Sherif et Hovland (1961), bien qu'anciens, montraient déjà que cette problématique était identifiée de longue date. Selon les auteurs, il s'agit d'un processus complexe à plusieurs niveaux, impliquant des allers-retours entre l'émetteur et le récepteur de l'information, soumis à des filtres cognitifs, affectifs et parfois normatifs. L'assimilation d'une information dépend en effet de critères intrapersonnels et subjectifs mais aussi des normes sociales et institutionnelles qui en valident l'acquisition.

L'écart entre attitude déclarée et comportement effectif est illustré par une expérience célèbre menée par Bickman (1972). Lors de cette étude, les participants ont été interrogés sur leur adhésion à l'idée selon laquelle il est de la responsabilité de chacun de ramasser un papier au sol : 94 % ont répondu positivement, mais seulement 1,4 % ont effectivement ramassé le papier placé à dessein pour l'expérience. Cette dissonance souligne que nos actions ne traduisent pas toujours nos convictions. Ainsi, comme l'explique l'auteur, avoir des connaissances ou des idées sur un thème ne garantit pas leur mise en pratique.

Dans leur revue de littérature, Fointiat et Barbier (2015) proposent un état des lieux approfondi des modèles de persuasion et d'influence, à la croisée de la psychologie sociale et de la conception des IHM. Leur objectif est de dégager des axes de conception pour des outils susceptibles d'engendrer des changements durables des comportements et des attitudes. Ils présentent plusieurs modèles théoriques, allant de l'école de Yale (années 1940), à la théorie de l'action raisonnée de Fishbein et Ajzen (1980) et son extension, la théorie du comportement planifié de Ajzen (2000), jusqu'au modèle de la communication engageante de Joule et al. (2007). À l'issue de cette revue, ils introduisent la notion d'ergonomie psychosociale des IHM comme champ de recherche émergent, visant à concevoir des interfaces capables d'induire des changements pérennes.

Deux leviers fondamentaux sont ainsi identifiés pour favoriser une transformation chez l'utilisateur : la persuasion, définie comme un changement d'attitude, et l'influence sociale, qui renvoie au changement comportemental. Selon Fointiat et Barbier (2015) la persuasion est considérée comme efficace lorsqu'elle entraîne chez le récepteur un changement d'attitude conforme à celui défendu par le message. Toutefois, ils soulignent que le changement d'attitude n'implique pas nécessairement un changement de comportement, contrairement à ce que l'on pourrait supposer. Ces remarques renforcent l'intérêt d'une approche pluridisciplinaire,

combinant ergonomie cognitive, IHM et psychologie sociale, pour maximiser les chances de succès dans l'élaboration d'outils efficaces.

Enfin, il est indispensable de différencier clairement les concepts d'attitude et de comportement. Comme le précisent Fointiat et Barbier (2015), la notion d'attitude en psychologie sociale ne doit pas être confondue avec son usage courant. Elle s'inscrit dans le champ de la persuasion alors que le comportement relève de celui de l'influence. Ces deux registres interagissent, mais ne sont pas automatiquement corrélés : de la persuasion (attitude) découle l'influence (comportement).

Voici les définitions sous le regard de la psychologie sociale :

- **Attitude** : « état mental et neuropsychologique de préparation à l'action, organisée par l'expérience du sujet et exerçant une influence directrice ou dynamique sur sa réponse à tous les objets et à toutes les situations s'y rapportant ». Allport (1935), p. 6. Bien qu'ancienne, cette définition reste valable aujourd'hui.
- **Comportement** : « le comportement d'usage est un événement psychologique qui s'inscrit dans un rapport social particulier, qui en retour possède une réelle influence sur son exécution ». Terrade et al. (2009), p. 389.

2.3 Quelques modèles à connaître : De persuader à influencer, l'évolution de la vision du changement

Cette partie aborde la question de la compréhension des déterminants et processus socio-cognitifs impliqués dans les phénomènes de persuasion en partant de concepts fondateurs. Avec l'argumentation bilatérale, Hovland et al. (1949), expliquent que présenter des arguments pour ET contre en même temps serait plus efficace sur une cible instruite. Festinger et Maccoby (1964) expliquent, avec le principe d'écoute distraite, qu'un individu distrait est moins enclin à réfléchir sur le débat en cours. Selon eux, ses ressources cognitives sont occupées ailleurs, spécifiquement si cette personne ne partage pas le point de vue de l'argumentaire. Bochner et Insko (1966), avec leurs travaux sur la source digne de confiance, ont montré que le locuteur était plus convaincant avec une certaine crédibilité. Selon Brehm (1966) et sa théorie de la réactance psychologique, un individu doit penser qu'il fait son choix non contraint en étant libre de penser. Petty et Cacioppo (1986) expliquaient qu'un processus de réflexion (avec possibilité de réagir) est plus efficace qu'une simple présentation (écoute passive) pour convaincre. Ainsi, tous ces travaux ont conduit ce courant d'école de Yale à proposer trois grandes phases de la persuasion : attention, compréhension, acceptation.

Déjà abordé plus haut, le modèle de la vraisemblance d'élaboration (ELM) de Cacioppo et Petty (1986) est un modèle central à travers un concept en apparence simple : vouloir traiter le message et ensuite en avoir la capacité. Cette approche cognitive issue de la psychologie a permis, pour plusieurs approches, l'aboutissement d'un nouveau paradigme faisant le pont entre théories de l'engagement comportemental et persuasion : la communication engageante de Joule et al. (2007). Au-delà des questions « informations, médias, arguments... » s'ajoutent les actes préparatoires, point central du paradigme qui donne au récepteur de la communication un statut d'acteur. Ces travaux s'inscrivent dans la lignée de ceux de Lewin (1951) et sa théorie du champ, présentant sa célèbre équation définissant le comportement comme fonction de l'interaction entre la personnalité et l'environnement : $B = f(P, E)$. B étant le comportement, P la personnalité et E l'environnement.

Pour conclure cette partie, Fointiat et Barbier (2015) expliquent que les enjeux liés à la conception dans les IHM rejoignent ceux de la psychologie sociale dès lors qu'il est question de persuasion et de modification des comportements. Dans les deux cas, il est nécessaire de définir des indicateurs permettant d'évaluer la portée persuasive du dispositif.

2.4 Situations d'apprentissage et professionnels concernés

En situation d'apprentissage, les éléments abordés dans les chapitres précédents ont une influence déterminante. Le TSA, et la compréhension qui lui est associée, doivent être considérés comme des spécificités notables, en raison de leur impact sur les interactions entre l'apprenant et l'enseignant à différents niveaux, Willay (2020). Avant d'aller plus avant, il convient de présenter brièvement la notion de situation d'apprentissage, afin d'identifier celles qui se révèlent les plus adaptées aux mécanismes à transmettre aux utilisateurs de la future application. Champ de recherche aussi vaste que complexe, la notion de situation d'apprentissage recouvre une diversité de configurations. Dans un premier temps, il importe de rappeler brièvement les différentes typologies de connaissances. Tardif (1992) propose une classification en trois grandes catégories.

- **Les connaissances déclaratives** : théoriques, elles englobent les faits et autres connaissances précises énumérables (date, règles, etc.).
- **Les connaissances procédurales** : les procédures à suivre pour la réalisation concrète d'une action. Un des exemples les plus connus est le vélo, un ensemble de connaissance qui s'associent pour pratiquer l'activité et qui est difficile à décomposer et expliquer en détails.
- **Les connaissances conditionnelles** : les connaissances stratégiques qui concernent le moment et contexte dans lequel une connaissance procédurale peut être appliquée.

Anderson et al. (2001), sur la base des travaux de Bloom (1956), identifient de leur côté deux grandes dimensions (cognitives et type de savoir). Les quatre types de savoir (connaissances) classés du plus concret au plus abstrait sont :

- **Les faits** : connaissances de base.
- **Les concepts** : théories, classifications, etc.
- **Les procédures** : méthodes, techniques.
- **Les méta-connaissances** : conscience du savoir et des stratégies.

Willay (2020) propose de réfléchir au type de connaissances participant au développement des compétences professionnelles qui sont cohérentes avec l'accompagnement personnalisé. Pour cela, elle utilise, elle aussi, la taxonomie des objectifs éducatifs de Bloom (1956) et ses six processus cognitifs : se souvenir, comprendre, appliquer, analyser, évaluer et

créer, processus eux-mêmes décomposés en sous catégories spécifiques qui sont la mise en application des savoir-faire correspondant au processus général.

Dans ce travail, bien que l'aspect des savoir-faire sociaux de Le Boterf (2000) cité par Taupiac (2020) soit visé, ce n'est pas le seul. La notion de compétences sociales sera également abordée dans la partie RVI, ainsi que la notion de *cognitive skill* de Anderson (1982) ainsi que les types de connaissances nécessaires pour acquérir ces savoir-faire. Ces derniers permettent d'appliquer des comportements afin de communiquer, de se comporter et enfin de se conduire avec son ou ses interlocuteurs de façon à correspondre aux codes nécessaires à la situation. Dans le cadre d'un enseignement auprès de personnes avec TSA, les codes et façons de se comporter sont particulières et il convient donc d'adapter les savoir-faire sociaux. Néanmoins, pour une application correcte de ces savoir-faire sociaux dans le cadre du TSA, il convient de posséder un socle de connaissances déclaratives minimales sur le sujet. Comme le rappelle de façon plus générale Anderson (1982), la profondeur des connaissances déclaratives influence directement la compréhension.

Le présent projet s'adresse à une population cible bien définie : les enseignants. Il ne vise pas, à ce stade, une diffusion à destination du grand public. Ce choix s'explique par les spécificités de cette catégorie socio-professionnelle dont les caractéristiques sont documentées notamment par Gohier et al. (2001). Les enseignants présentent en effet une construction identitaire propre, façonnée par leur rôle, leur posture institutionnelle, ainsi que par les exigences pédagogiques et relationnelles de leur métier. Parmi ces exigences, le développement de compétences sociales constitue un enjeu central. Ces compétences, souvent complexes à acquérir, conditionnent la qualité de la relation pédagogique et facilitent la transmission des savoirs, qu'ils soient déclaratifs ou procéduraux. L'importance de ces compétences a été soulignée dans les sections précédentes et sera approfondie dans le cadre de la RVI. Compte tenu de ces éléments, la conception de l'outil pédagogique nécessite une prise en compte attentive des spécificités propres au profil enseignant. Willay (2020) souligne à ce propos que les formations doivent permettre une appropriation des connaissances centrée sur la personne formée, en s'appuyant sur ses représentations préexistantes et ses cadres de référence. L'objectif de cette approche est de contribuer à une transformation des représentations en encourageant une vision plus globale de la personne, reconnue dans la diversité et la complexité de ses besoins.

3. Synthèse

Ce travail porte sur la compréhension, par les enseignants, du fonctionnement cognitif d'une population spécifique : les personnes présentant un trouble du spectre de l'autisme (TSA). Il s'agit plus précisément d'individus suivant un parcours scolaire allant du secondaire au supérieur, correspondant au profil dit TSA sans déficience intellectuelle, incluant les formes anciennement appelées « autisme de haut niveau » ou « syndrome d'Asperger » (Attwood, 2022). Toutefois, selon Bennett et Goodall (2016), le syndrome d'Asperger ne satisfait pas l'ensemble des nouveaux critères diagnostiques bien qu'il puisse relever d'un trouble de la communication sociale.

Les travaux de Mottron et al. (2006) et ceux de Bled (2022) ont permis de préciser les particularités cognitives de ces individus, en mettant en évidence notamment les spécificités du traitement multisensoriel, l'existence de pics d'habiletés ou îlots de compétences ainsi que certains modèles explicatifs du fonctionnement cognitif propres au TSA. Parmi ces derniers, le modèle EPF de Mottron et al. (2006) s'avère central. Il propose une lecture renouvelée du fonctionnement cognitif du TSA, insistant sur une sur-fonctionnalité perceptive : focalisation accrue, sens du détail, expertise visuelle, indépendance relative des niveaux de traitement, etc. Ces caractéristiques permettent parfois des performances exceptionnelles, mais s'accompagnent aussi de difficultés, notamment dans le tri de l'information pertinente, la gestion de la surstimulation sensorielle ou encore l'articulation avec les dimensions sociales et contextuelles. Cette perspective éclaire d'un jour nouveau la compréhension du TSA : loin d'une simple déficience, il s'agit d'une organisation neurocognitive singulière, parfois déroutante pour les enseignants, mais riche de potentiels. Ces particularités justifient pleinement la nécessité de formations ciblées pour les professionnels de l'éducation.

Sur le plan pédagogique, plusieurs enjeux émergent, comme former efficacement les enseignants à mieux comprendre les TSA, mais aussi les aider à déconstruire certaines représentations erronées ou stéréotypées. Une formation pertinente devra ainsi combiner plusieurs dimensions : mise en lumière des bonnes pratiques, visualisation des surcharges sensorielles, exposition aux rapports sociaux atypiques et valorisation des besoins spécifiques, tels que le recours au calme ou à l'isolement.

Cependant, ces contenus ne peuvent être assimilés durablement qu'à travers une formation adaptée. Les travaux issus de la psychologie sociale et cognitive, notamment le

modèle de la double voie de persuasion de Petty et Cacioppo (1986), rappellent qu'un changement d'attitude durable exige un engagement cognitif profond, rendu possible par un intérêt personnel élevé et un traitement réfléchi de l'information. Par conséquent, une simple sensibilisation ponctuelle ne saurait suffire à opérer des transformations significatives. Il convient d'intégrer ces modules courts dans des parcours de formation plus longs, engageants et interactifs, susceptibles de favoriser une compréhension durable. En ce sens, la formation doit aussi servir à expliciter la diversité des profils TSA (tant sur le plan cognitif que comportemental) et à faire prendre conscience des risques de discrimination, y compris lorsque celle-ci se dissimule derrière des stéréotypes positifs. L'approche défendue ici encourage une vision plus nuancée : il ne s'agit pas de gommer les difficultés, mais d'en proposer une lecture constructive, enracinée dans une logique de bienveillance informée.

Ainsi, les travaux explorés ici convergent vers un même objectif : permettre une conception d'un dispositif pédagogique qui tienne compte à la fois des réalités neurocognitives des personnes TSA ainsi que des mécanismes d'appropriation des connaissances chez les professionnels. Pour initier un changement d'attitude, il est nécessaire d'identifier les stéréotypes initiaux, de mobiliser un traitement cognitif en profondeur et de proposer un environnement engageant. La RVI pourrait constituer un outil privilégié pour faire vivre aux professionnels une forme d'empathie cognitive, en simulant par exemple les effets de la surstimulation sensorielle ou en visualisant concrètement les spécificités perceptives décrites par le modèle EPF. De plus, par son caractère immersif et engageant, elle répond aux critères identifiés par le modèle de la double voie de persuasion, en favorisant un traitement approfondi de l'information susceptible de générer des changements d'attitudes durables. C'est dans cette perspective que vont maintenant être explorées les potentialités de la RVI pour la sensibilisation des enseignants au TSA.

Chapitre 2 - La réalité virtuelle (incarnée)

La réalité virtuelle (RV) est l'ensemble des techniques permettant de simuler un environnement en trois dimensions via un système informatique. Paradoxe technologique, la RV est relativement ancienne. Elle est imaginée sous la forme d'une vision stéréoscopique par Hugo Gernsback en 1930 avec la proposition d'un prototype théorique en 1963. Réalisée techniquement en 1968 par Ivan Sutherland, elle sera définie par Jaron Lanier en 1985. Sa démocratisation ne se fera finalement qu'en 2016 avec l'arrivée du premier dispositif grand public : l'Oculus Rift, bien que d'autres s'y soient essayés comme Morton Heilig (1962) avec son Sensorama ou encore Nintendo et Atari avec le *Virtual Boy* de 1994, l'Atari 2600 VR de 1982 ou le Jaguar VR, jamais sorti au grand public.

Aujourd'hui surtout connu pour ses applications dans le jeu vidéo, la RV est déjà étendue à de nombreux domaines. Le monde de l'industrie l'utilise au travers de dispositifs tels les CAVES de Cruz-Neira (1994). Le milieu médical se l'est également approprié avec des applications assez récentes comme par exemple *Vtopia Surgical*, dispositif d'entraînement en RV utilisée en condition réelle en France par le docteur Grégory en 2017. Au niveau thérapeutique, il existe des applications traitant différents troubles (phobies, stress post-traumatique, douleur...) et confirmant l'intérêt du développement d'une telle technologie. Le domaine de l'apprentissage et de la formation est lui aussi concerné. Un pan de la recherche s'intéresse particulièrement à la transmission et l'acquisition de savoirs en RV car l'outil, bien que prometteur, est plus difficile à appréhender de prime abord (contrairement aux espoirs des pionniers de cette technologie).

Ce projet de conception et développement, puis d'étude empirique d'un outil de compréhension des mécanismes du TSA en RVI, s'inscrit dans la lignée des médias et autres dispositifs ludiques d'apprentissages sérieux. Les attentes de ce projet elles aussi sont dans ce scope : concevoir et tester l'apport d'une technologie novatrice dans un domaine d'application concret et spécifique.

1. Caractéristiques de la réalité virtuelle

1.1 Quelles définitions de la réalité virtuelle ?

Comme l'explique Fuchs (1996), la RV n'est pas qu'un outil technologique informatique. L'auteur ajoute que la RV ne se contente plus de faire uniquement traiter l'information à l'utilisateur mais de la faire agir et percevoir dans un environnement totalement artificiel. Ces notions de « traiter, percevoir et agir » sont importantes car elles introduisent les sciences cognitives comme composantes de la RV comme l'explique Fuchs (2006) dans la réédition du premier volume du Traité de la réalité virtuelle. Le traitement de l'information et la perception sont des champs de recherche fondamentaux qui seront abordés en RV via des modèles comme le CAMIL, qui est la base de la partie conception de ce travail. Il existe ainsi plusieurs définitions et approches de la RV, en voici certaines, choisies pour leur pertinence :

La finalité, de Taupiac (2020), qui reprend les notions de Fuchs en proposant différents niveaux : « permettre à une (ou plusieurs) personne(s) des activités sensorimotrices et cognitives dans un environnement artificiel, créé numériquement et pouvant être imaginaire, symbolique ou bien une simulation de certains aspects du monde réel » (p. 37). Par « activité sensori-motrice », Fuchs et al. (2003) expliquent qu'il s'agit de percevoir et agir physiquement mais également d'avoir une activité cognitive. La finalité de RV est de simuler certains aspects du monde réel. La notion de « degrés de réalisme » est ici essentielle car il ne s'agit pas de reproduire à l'identique le monde réel mais plutôt de proposer une adaptation permettant des libertés. Ces libertés permettent par exemple de ne pas se blesser pendant une formation normalement dangereuse dans la réalité primaire, la « vraie vie » (le terme générique anglais est IRL pour In Real Life), de voir des phénomènes invisibles ou d'adapter en modifiant ou ajoutant des éléments visant à aider la compréhension. Les possibilités sont vastes et celles citées ne sont que des exemples parmi tant d'autres. Fuchs parle également de monde symbolique et monde imaginaire, montrant les possibilités quasi infinies que peut proposer la RV à ses utilisateurs.

La définition technique, établie avec Fuchs et al. (2003) : La RV est un « domaine scientifique et technique exploitant l'informatique et des interfaces comportementales en vue de simuler dans un monde virtuel le comportement d'entités 3D, qui sont en interaction en temps réel entre elles et avec une personne en immersion pseudo-naturelle par l'intermédiaire de canaux sensori-moteurs ». Fuchs décompose ensuite les différents éléments de sa définition (en sachant que la

notion d'immersion a évolué avec le temps pour être progressivement remplacé par « immersion technologique ». Voici la proposition de Fuchs et al. (2003) :

L'aspect informatique : exploitation des potentiels matériels et logiciels pour réaliser techniquement un environnement virtuel interfacé qui est dynamique (animation en temps réel, lois physique et comportementales).

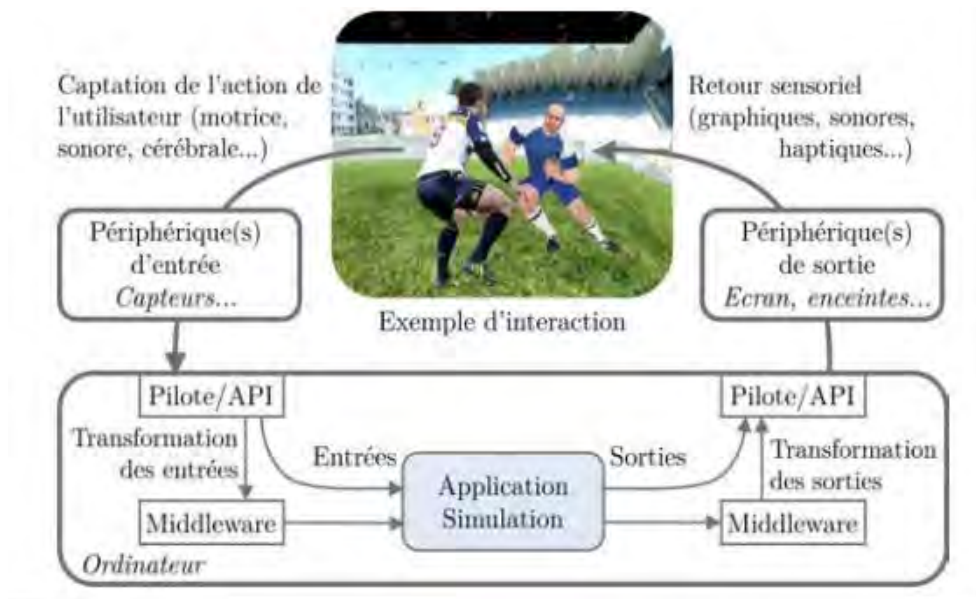
Les interfaces comportementales : elles sont composées des interfaces sensorielles (qui informent l'utilisateur des composants du monde virtuel, et ce, par l'intermédiaire de ses sens, vue et ouïe) et des interfaces motrices (qui informent l'ordinateur des actions motrices de l'utilisateur)

Le monde virtuel : un environnement virtuel en temps réel, interactif avec les contraintes qu'il pose (modélisation, numérisation et traitement, puis interfaçage utilisateur / environnement virtuel

L'interaction en temps réel : La notion de latence issue de la boucle d'interaction et qui explique qu'il ne faut pas de décalage entre action et perception, qui sera vu plus en détail dans la section concept et modèles principaux.

L'immersion pseudo-naturelle : se rapprocher de ce que pourrait vivre l'utilisateur afin de lui faciliter son immersion en RV, une notion subjective car le « naturel » ne peut pas encore être atteint mais qui se doit d'être le plus performant possible pour faciliter l'utilisabilité.

Une grande partie de cette définition technique s'appuie sur le concept de boucle d'interaction qui sera vu dans la section concernée. À noter : s'il n'y a pas d'interaction humaine, alors on parle d'animation et plus de RV. Voici un aperçu de la boucle dans son aspect technique issue d'Arnaldi et al. (2018) en fig. 1.

Figure 1
Boucle d'interaction technique - issue de Arnaldi, Guitton & Moreau (2018)


La notion d'aspect fonctionnel : La RV est une technologie offrant la possibilité de s'affranchir de l'environnement physique pour se projeter virtuellement dans un autre espace, un autre moment et/ou un autre mode d'interaction. Cette notion de fonctionnel va introduire indirectement la notion d'utilisateur comme le montrera Taupiac (2020) ci-après. Selon Fuchs et al. (2003) le fonctionnel est important car il permet d'appréhender la nouvelle réalité proposée par la RV qui questionne les trois unités de temps, de lieu et d'action et ce, du point de vue de l'humain.

Ainsi, dans la continuité de cette notion de « fonctionnel », Taupiac continue en proposant une définition anthropocentrée, c'est-à-dire centrée sur l'utilisateur, impliquant le rejet des visions techno centrée. Fuchs explique qu'elles sont trop axées sur l'aspect technique et restreignent les usages. Selon la définition de la RV par Taupiac (2020), qui reprend une partie de la définition de Fuchs : la RV est une « *technologie amenant ses utilisateurs à s'extraire de la réalité physique pour changer virtuellement de temps et/ou de type d'interaction, en leur permettant des activités sensorimotrices et cognitives dans un environnement artificiel, créé numériquement et avec lequel il va interagir.* » p. 37. Ces différentes définitions montrent que le domaine est encore en pleine essor du point de vue de la recherche comme l'expliquent Makransky et Petersen (2021) avec le nombre d'articles publiés sur la RV depuis 2016, et qui ne cesse d'augmenter.

Pour finir cette présentation générale il est important de préciser, comme le rappelle Fuchs (2006) dans la dernière édition du *Traité de la réalité virtuelle*, un ouvrage majeur composé de cinq volumes qui fait office de référence aujourd'hui, que la RV est avant tout un domaine pluridisciplinaire dans lequel se côtoie de nombreuses disciplines. Ainsi, informatique, télé-opération, robotique, mécanique, optique et acoustique collaborent étroitement avec la psychologie expérimentale, les sciences cognitives, l'ergonomie, la neurobiologie, pour ne citer que les principales.

La notion d'incarnation dans la réalité virtuelle : le I de RVI. Ce « I » renvoie au terme « Incarnée ». La RVI se distingue de la RV que l'on appelle « générique », comme dans les modalités de *flatscreen* (projection de l'environnement sur un écran plat), de dispositifs CAVE, ou l'utilisateur n'incarne pas via un avatar mais demeure lui-même ou encore via autres dispositifs non inclusifs. La RVI correspond donc tout simplement à la RV dans des casques occlusifs qui permettent de bloquer l'accès aux informations du monde réel en isolant totalement l'utilisateur comme décrit par Loomis et al. (1999). Cela offre un sentiment d'immersion globale dans un environnement virtuel totalement coupé de toute interférence extérieure tout en incarnant un avatar. Un autre élément important est le niveau d'interaction possible, questionné Makransky et al. (2020a), c'est-à-dire la liberté de contrôler en partie l'expérience (notion relative aux possibilités laissées par le logiciel). Pour illustrer cela, la RV sous *flatscreen*, bien qu'interactive, n'est pas de la RVI mais la RV uniquement.

1.2 Les concepts et modèles principaux en RV

Si l'on souhaite que les éléments cognitifs, émotionnels et comportementaux abordés précédemment soient efficacement intégrés par l'utilisateur dans une expérience en RV, il est impératif de concevoir cette dernière en tenant compte de plusieurs fondements théoriques et ergonomiques. Ces fondements constituent les conditions préalables nécessaires à une expérience à la fois fluide, engageante et cognitivement efficace. Autrement dit, l'outil ne peut remplir ses objectifs pédagogiques que s'il respecte les exigences d'une bonne ergonomie immersive : éviter les perturbations sensorielles (telles que la cinétose), favoriser une interaction naturelle avec l'environnement virtuel, et, surtout, induire un engagement actif de l'utilisateur. Voici les concepts fondamentaux qui président à la conception d'une expérience de RV véritablement réussie.

1.2.1 *L'illusion de plausibilité et les modèles attentionnels*

Pour qu'une expérience en RV atteigne son plein potentiel, elle doit générer ce que Slater (2009) nomme l'illusion de plausibilité. Il s'agit d'un « leurre cognitif » au cœur duquel l'utilisateur en vient à croire, à un niveau implicite, qu'il est réellement présent dans l'environnement simulé. Ce phénomène, loin d'être purement perceptif, repose sur un ensemble complexe de mécanismes attentionnels et cognitifs. Les modèles du traitement de l'information permettent d'éclairer ce processus.

Les modèles classiques du filtre attentionnel (Broadbent, 1958 ; Deutsch & Deutsch, 1963 ; Treisman, 1969) ont tenté de décrire la manière dont l'attention humaine trie, hiérarchise et sélectionne les informations sensorielles pertinentes dans un environnement saturé de stimuli. Si ces modèles ont longtemps été perçus comme rigides, les travaux de Johnston et Wilson (1980) ont introduit une flexibilité contextuelle précieuse, soulignant la variabilité des seuils attentionnels selon l'environnement ou la tâche.

Dans le cadre de la RV, cette logique de filtrage prend un sens nouveau. Comme le souligne Pereira da Silva (2010) dans son modèle computationnel de l'attention, la force de la RV réside précisément dans sa capacité à réduire le bruit informationnel externe en concentrant la charge perceptive sur des stimuli intentionnellement choisis et cohérents. L'environnement virtuel devient ainsi un canal privilégié de focalisation attentionnelle, renforçant la vraisemblance de la simulation et, par conséquent, l'efficacité cognitive de l'expérience.

1.2.2 *Facteurs favorisant l'immersion*

Un point de vigilance s'impose néanmoins dans l'approche théorique : la littérature en RV présente souvent des chevauchements conceptuels entre différentes notions. Ainsi, le *Sense of Embodiment* (SoE), ou sentiment d'incarnation, tel que défini par Kilteni et al. (2012), intègre en son sein des composantes telles que l'agentivité, la localisation corporelle et l'appropriation du corps virtuel. Ce même SoE est également mobilisé dans le modèle CAMIL pour expliquer à la fois la présence (*Sense of Presence*) et l'agentivité (*Sense of Agency*), illustrant la porosité des frontières conceptuelles dans les modèles immersifs.

Cette complexité n'est pas un obstacle mais un indice de la richesse phénoménologique de la RV. Pour clarifier ces interrelations et préparer à leur mobilisation dans la conception d'outils immersifs, il convient de présenter de manière structurée les principaux concepts

impliqués dans une expérience de RV efficace. Voici les concepts-clés à la base d'une expérience immersive réussie selon Arnaldi et al. (2018) :

- **La boucle perception-décision-action** : Ce mécanisme représente la dynamique fondamentale de toute expérience interactive en RV. L'utilisateur perçoit une stimulation dans l'environnement, prend une décision en fonction de celle-ci, puis agit. La boucle se referme sur une rétroaction environnementale qui nourrit une nouvelle perception. Plus cette boucle est fluide et naturelle, plus l'illusion immersive est renforcée.
- **La présence** : Il s'agit du sentiment d'« être là » dans l'environnement simulé. C'est une construction perceptive et cognitive dépendant de nombreux facteurs : cohérence sensorielle, interactions en temps réel, crédibilité de la narration, etc. La présence est essentielle pour l'engagement de l'utilisateur et constitue une condition sine qua non pour l'apprentissage immersif.
- **L'incarnation** : Ce concept désigne la capacité de l'utilisateur à se sentir corporellement présent dans l'environnement virtuel en adoptant un corps virtuel (avatar). L'incarnation repose sur la congruence entre les mouvements physiques de l'utilisateur et leur représentations virtuelles. Elle est un levier puissant pour l'agentivité et l'implication émotionnelle.
- **L'agentivité** : L'agentivité renvoie à la sensation d'être à l'origine de ses propres actions dans l'environnement virtuel. Elle est renforcée par la précision des interactions, la latence minimale et le sentiment de contrôle sur les événements. L'agentivité est cruciale pour l'apprentissage actif et l'autodétermination.
- **L'immersion technologique** : Enfin, il est important de distinguer l'immersion subjective (cognitive et émotionnelle) de l'immersion technologique, laquelle dépend des caractéristiques matérielles du système (HMD, champs visuels, taux de rafraîchissement, *tracking*, etc.). Bien qu'insuffisante à elle seule, l'immersion technologique est un facilitateur fondamental des autres dimensions.

1.2.3 La boucle « perception, décision, action » ou boucle d'interaction

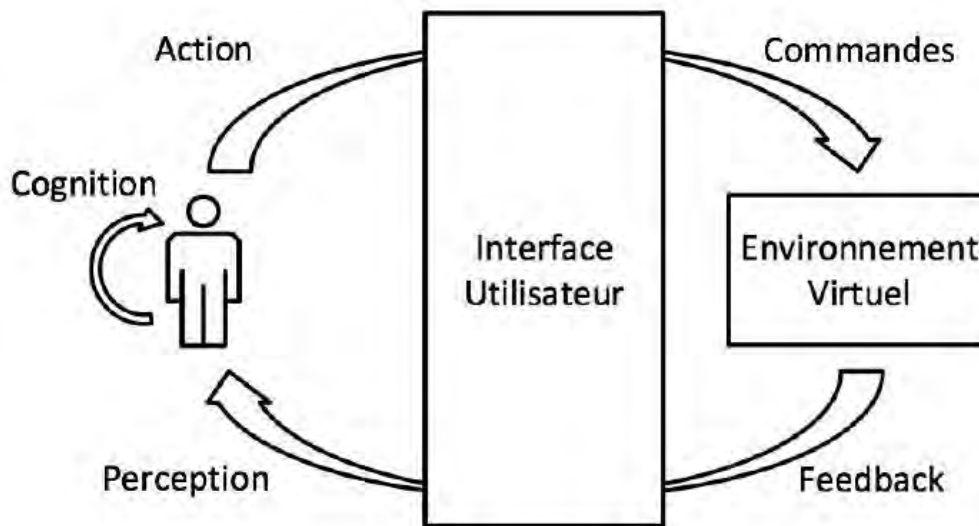
La boucle « perception, décision, action » est un concept clé de RV basé sur les principes d'interaction homme machine (IHM). Ces concepts de boucle peuvent être appelé « boucle d'interaction » en simplification. Cette boucle est la base du concept informatique qui permet la RV. C'est une limite technique majeure qui est repoussé à chaque génération de casques

occlusifs (HMD) et qui permet aux autres concept clé de RV de prendre forme. Ainsi, si la boucle n'est pas optimale (avec un défaut de conception par exemple par une surestimation des capacités matérielle d'un logiciel, trop de ressources nécessaires pour afficher les modèles...), la présence, le *Sense of Embodiment*, l'incarnation ou encore l'agentivité seront grandement impactés, rendant l'expérience très désagréable pour l'utilisateur par l'apparition quasi instantanée de cinétose comme le rappellent Rebenitsch et Owen (2016).

Avec cette boucle (fig. 2), Arnaldi et al. (2018) apportent un aspect technique aux définitions de la RV par la « boucle de perception-action » et la « boucle d'interaction ».

Figure 2

La boucle d'interaction issue de Arnaldi, Guitton & Moreau (2018)



Au cœur du concept technique qui permet de produire le leurre cognitif de la RV, ce sont ces boucles qui, en plus de montrer comment fonctionne la RV, vont donner le temps de latence de l'information entre sa captation et son affichage. Plus précisément, il s'agit du temps de calcul entre l'action de l'utilisateur (un mouvement de tête par exemple), sa perception par le dispositif (acquisition via la centrale gyroskopique de l'information), son calcul par les processeurs et le retour (qui vont calculer la nouvelle spatialisation) puis le rendu visuel correspondant au mouvement initial dans le casque occlusif (une image qui suit le mouvement faisant croire à l'utilisateur que l'environnement virtuel est stable autour de lui). Cette notion de latence est importante notamment au niveau de la cinétose.

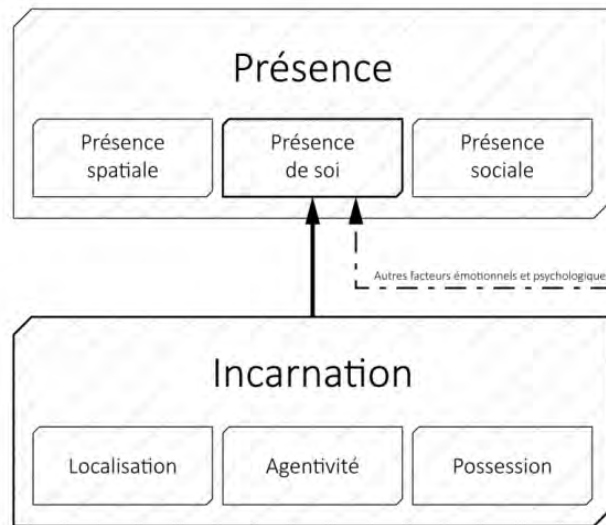
Il est intéressant de préciser que la RV se situe dans le spectre des technologies immersives incluant RV, RA (réalité augmentée) et RM (réalité mixte ou MR en anglais pour *mixed reality*) bien que différente sur un point central, l'isolement. La RVI, par la conception

occlusive des casques, isole l'utilisateur dans l'environnement virtuel comme par Loomis et al. (1999).

Fuchs et al. (2003) ont proposé dans leur ouvrage de référence trois type d'interfaces de RV, dont l'interface sensorielle, utilisée dans ce projet. Il s'agit d'ensembles matériels destinés à restituer, sur le plan sensoriel, les composantes de l'environnement virtuel (visuel, auditif, kinesthésique, etc.). Ils permettent une vision stéréoscopique d'un environnement virtuel projeté sur deux écrans, dans un casque (HMD) qui bloque toute perception visuelle de l'environnement extérieur dit réel, tout en étendant le champ de vision (*Field of View* - FOV) à des degrés proches de celui de l'homme (110° en moyenne, et 130° pour les HMD les plus performants). Via une centrale gyroscopique composée de trois éléments (gyroscope, accéléromètre et boussole magnétique), le HMD capte chaque mouvement de l'utilisateur et lui restitue une image (via la boucle d'interaction, Arnaldi et al., 2018) qui correspond de manière cohérente à ce qu'il verrait si la situation était réelle. Makransky et Petersen (2021) expliquent qu'un dispositif RV permet de suivre les mouvements de la tête tout en restituant une image différente dans chaque lentille, recréant ainsi l'illusion de profondeur via de faux repère visuel.

1.3 Le modèle présence - incarnation

En 2018, Gorisse et al., toujours dans une optique pluridisciplinaire, ont proposé un nouveau méta-modèle nommé « présence - incarnation » afin de synthétiser l'ensemble des travaux sur ces thèmes. Ce modèle lie ces deux composantes de la RV pour proposer une explication au fonctionnement du leurre cognitif, analysant pourquoi la RV fonctionne-elle sur le cerveau alors que l'utilisateur sait pertinemment qu'il ne se trouve pas réellement là où il est en RV. Voici un aperçu schématisé du modèle « présence - incarnation » (fig. 3) :

Figure 3
Le modèle présence incarnation, Gorisse, Christmann & Richir (2018)


1.3.1 La présence

Concept clé de la RV mais qui s'utilise dans d'autres domaines, la présence renvoie au fait, dans une définition large, de se trouver physiquement et matériellement dans un lieu déterminé. Cette notion oscille entre la « réalité primaire » (IRL) et la « réalité médiatisée » (la RV dans ce travail) d'après Gorisse et al. (2018). Quand Sheridan (1992) parlait pour la première fois de présence virtuelle pour les espaces numériques, Slater et al. (1994) définissaient ce concept par le sentiment « d'être là ». Liée à la perception, la présence comporte une composante individuelle. Toujours inspiré de Sheridan (1992), Riva et al. (2003) suggèrent trois types de déterminants de la présence : l'étendue des informations sensorielles présentées, le degré de contrôle que l'on a et enfin le degré auquel on peut modifier l'environnement.

Avant d'aller plus loin, il est important de faire un bref aparté sur l'immersion car il y a encore débat entre présence et immersion, bien que Slater proposait déjà en 1999 une extension à sa propre définition de l'immersion en incluant l'aspect « immersion technologique ». Le terme « immersion » est donc mis de côté dans ce travail afin d'éviter toute confusion.

Les travaux de référence de Lee (2004) subdivisent la présence en trois sous concepts :

- La présence physique/spatiale, qui concerne la relation entre l'utilisateur et l'environnement virtuel, composé des trois éléments définis par Slater (1999) :
 - Le sentiment « d'être là » dans l'environnement virtuel

- Le degré selon lequel cet environnement prend le pas sur la réalité, amenant l'utilisateur à réagir prioritairement aux événements virtuels plutôt qu'aux stimuli du monde réel.
- La tendance des utilisateurs, après l'expérience, à se remémorer l'environnement comme un lieu réellement visité, plutôt que comme la simple observation d'images générées par ordinateur.
- La présence de soi, qui explique comment s'élabore la version alternative virtuelle de l'utilisateur, physiquement et psychologiquement. Ce concept est introduit par Biocca (1997) qui expliquait que la présence de soi est la représentation mentale qu'un utilisateur construit de lui-même au sein d'un environnement virtuel, intégrant son corps, ses états physiologiques et émotionnels ainsi que les traits perçus et son identité.
- La présence sociale désigne l'état psychologique dans lequel les utilisateurs perçoivent les acteurs sociaux virtuels comme de véritables interlocuteurs, qu'ils soient appréhendés par des canaux sensoriels ou à travers d'autres formes d'interaction.

Lee (2004) explique aussi qu'une expérience de RV peut être qualifiée de « para-authentique » si elle conserve un lien avec le monde réel (visioconférence, visite virtuelle et concept de métaverse) en opposition à « artificielle » dans le cadre d'environnement virtuel totalement fictif (jeux vidéo, environnement onirique...).

Sur la base des travaux de Lee (2004), Gorisse et al. (2018) ont donc proposé un modèle de présence pour la RV qui définit la présence de la sorte : « *Un état psychologique où la virtualité de l'expérience n'est pas décelée. Ainsi, si nous considérons la présence comme un continuum (intensité du sentiment de présence sujette à des variations continues), alors, à son paroxysme, l'utilisateur ne serait pas en mesure d'identifier la « réalité » dans laquelle il évolue.* » p. 2 (notre traduction).

A ce stade, il est intéressant de citer les travaux de Riva et al. (2011, p. 7, notre traduction), qui ont proposé le concept de « niveau de présence » décomposé en trois éléments distincts définis de la sorte

- La protoprésence : capacité à associer perception et mouvements et à discerner l'environnement interne et externe
- La présence centrale : capacité à concentrer les processus sensoriels sur la tâche en cours
- La présence étendue : capacité à sélectionner des expériences antérieures comme pertinentes pour soi

Les auteurs ajoutent qu'une vue FPS (*first person shoot* - vue à la première personne) avec un avatar intégral, renforce la présence. Dans cette continuité, il est donc pertinent de s'intéresser aux autres éléments qui impactent la présence ou qui sont impactés par elle. Cadet et Chainay (2021), qui ont étudié l'impact des corps incarnés sur les processus cognitifs (mémoire, émotion et sentiment de présence), ont constaté que la présence était liée à l'émotion et que cela impactait grandement la mémorisation. Makowski et al. (2017) faisaient le même constat sur la mémoire avec un état conscient de présence dont les variations influençaient la mémorisation. Sur un autre registre, Reiners et al. (2014), cités par Jensen et Konradson (2018), expliquent que le fait d'être debout, plutôt qu'assis, renforce le sentiment de présence. Cet élément est surtout représentatif de la cohérence de la posture car il est difficile de se représenter une position debout et haute dans l'environnement virtuel alors que l'utilisateur est assis IRL.

Taupiac (2020) propose une revue plutôt complète des différentes méthodes pour mesurer la présence comme le questionnaire *Igroup Presence Questionnaire* de Shubert (2003). Le *Temple Presence Inventory* (TPI) de Lombard et al. (2009, 2011), est un des plus complets validé par la littérature (8 thèmes décomposés en 39 items). Le TPI a pour vocation de mesurer avec précision la présence, ce qui n'est pas le but ici. Il est cité pour montrer qu'il est possible d'affiner le recueil de données sur la présence. La présence est un composant important de l'illusion de plausibilité en RV.

1.3.2 *L'incarnation ou SoE (Sense of Embodiment)*

L'incarnation ou SoE, est la seconde composante du modèle « présence - incarnation » de Gorisse et al. (2018) et donc, par extension, du leurre cognitif. Sur la base des travaux De Vignemont (2011) et Blanke et Metzinger (2009), cités par Kilteni et al. (2012), une définition du SoE a été proposée par Kilteni et al. : « *Le SoE à l'égard d'un corps B est le sentiment qui émerge lorsque les propriétés de B sont traitées comme si elles étaient les propriétés de son propre corps biologique* » (p. 375, notre traduction). Cette définition explique que la notion d'incarnation ou SoE est une sorte de transfert puis d'investissement du corps virtuel jusqu'à la perception de ce corps virtuel comme le nôtre.

Gorisse et al. (2018) expliquent qu'il n'est pas possible pour l'utilisateur de réellement dissocier totalement son corps réel et bloquer les stimuli externes, lors d'une exposition RV. Il est néanmoins possible, comme l'ont montré Botvinick et Cohen (1998), avec leur fameuse expérience de la « *Rubber hand* », ou main en caoutchouc, qu'il est possible d'avoir malgré tout le sentiment, pour le participant, d'être en possession de cette main factice. Le plus

surprenant est que cette incarnation de la main est possible malgré l'erreur de perception induite par le décalage créé lors de l'expérimentation (décalage de plusieurs centimètres).

Le processus d'incarnation implique donc la notion du corps et de sa perception. L'importance du corps est illustrée par la théorie du Dilemme du cyborg de Biocca (1997). Cette théorie explique que le processus « *d'identification à soi* » en RV passe par le corps de l'utilisateur. Biocca propose trois éléments importants dans ce processus qui permet de se sentir être soi-même dans un corps qui n'est pas le sien et qui n'est pas réellement là :

- Le corps objectif (corps réel de l'utilisateur)
- Le corps virtuel (corps virtuel représenté ou non dans l'univers simulé)
- Le schéma corporel (représentation mentale du corps)

Le concept sous-jacent est celui du transfert du schéma corporel propre à l'individu, depuis son corps réel vers son corps virtuel. C'est une capacité qui dépend de la qualité de la simulation de RV mais surtout des capacités de l'utilisateur à pouvoir transférer son schéma corporel. Bien entendu, la qualité des environnements virtuels et les autres composants (présence, agentivité, etc.) vont faciliter et renforcer ce transfert. Des variations et perturbations sont néanmoins possibles, surtout avec le décalage des perceptions du corps réel dont certaines peuvent induire de la cinétose ou rompre l'illusion de plausibilité si elles sont trop brutales ou trop denses.

Les travaux de Kilteni et al. (2012) repris et confirmés par le modèle « présence - incarnation » de Gorisse et al. (2018) décomposent l'incarnation ou SoE et font ressortir trois concepts principaux :

- **La localisation de soi** (*Sense of self-location*) de Arzy et al. (2006), c'est à dire « être dans le corps », se sentir « dedans ». Cette notion est en lien avec la vue en FPS et la notion d'égoцентризм dû à une prédisposition sensorielle de cette vue pour se situer dans le corps, Lee (2004). Les auteurs rappellent également l'importance des signaux vestibulaires pour cette localisation.
- **Le sentiment d'agentivité** (*Sense of agency*) de Newport et al. (2010) qui correspond au contrôle moteur global avec coordination des actions et feedback visuel.
- **Le sentiment de possession** (*sense of body ownership*) de Lopez et al. (2008) qui correspond à sentir que ce corps (celui de l'avatar) est le sien avec une influence *top down*

(visuel, tactile... qui prouve que le corps virtuel est le sien) et une influence *bottom up* (le processus cognitif qui permet de se convaincre que le corps virtuel est le nôtre).

Deux types de facteurs impactent l'incarnation et ses composants ainsi que la relation entre eux : les facteurs externes et internes.

Pour les facteurs externes, Kim et al. (2020) ont montré que le fait de posséder un avatar complet (avatar en condition humaine et vue FPS) est plus influent sur le SoE qu'un avatar en nuage de points (14 points en vue FPS) et que si les mouvements réels sont synchronisés avec leurs représentations virtuelles, alors le SoE augmente. Confirmant ces résultats, Dewez et al. (2019) expliquent que la vue à la troisième personne entraîne une baisse du sentiment du SoE via le sentiment de localisation de soi. Les auteurs nuancent néanmoins leur propos en précisant qu'avec la stimulation adaptée (visuotactile synchrone), il est possible de déplacer progressivement la perception vers la vue troisième personne. Selon Kokkinara et Slater (2014), la synchronisation est un élément essentiel, qu'elle soit visuotactile (retour toucher au niveau des capteurs correspondant au mouvement réel) ou visuomoteur (retour visuel dans l'environnement virtuel correspondant au mouvement réel), avec une influence plus importante du visuomoteur. L'agentivité est également impactée par la synchronisation, avec une possibilité de décaler légèrement l'action et le retour d'information (jusqu'à une certaine limite après laquelle l'utilisateur perd toute agentivité). Kim et al. (2020) ont également travaillé sur la similitude entre le corps réel et le corps virtuel. Le SoE est renforcé si les deux entités (réelle et virtuelle) sont de morphologies équivalentes via le sentiment de possession. Ces résultats sont confirmés par les travaux de Kokkinara et Slater (2014) qui expliquent que l'apparence de l'avatar est importante pour maximiser l'illusion d'appropriation. Avec un teint et des vêtements similaires à ceux de l'utilisateur, le sentiment de possession est plus fort. Les auteurs confirment par ailleurs le fait que de simples mains et pieds permettent déjà de s'approprier le corps virtuel (en cas d'autoreprésentation, le sentiment est encore plus dense). Il existe néanmoins des nuances dans la littérature. Dans leurs travaux, Kim et al. (2020) expérimentent trois types d'avatar dont un nuage de point en vue troisième personne influant, contre toute attente, sur le SoE. Kokkinara et Slater (2014) avaient déjà eu des résultats similaires. Enfin, ces derniers rappellent que la RV permet d'incarner des avatars ne correspondant pas aux caractéristiques de l'utilisateur (taille, morphologie, genre, etc.) mais qui génèrent néanmoins de l'incarnation (même avec des avatars et positions non réalistes) malgré la baisse du sentiment de localisation de soi.

Pour les facteurs internes, Dewez et al. (2019), avec leurs travaux d'identification des facteurs suscitant et perturbant l'incarnation (résistance ou prédisposition), ont proposé une expérience particulièrement intéressante. Dans cette expérience, le participant se fait poignarder, par surprise, sa main virtuelle alors qu'il réalise une tâche basique. Un certain nombre de facteurs internes ont donc été identifiés (l'article porte sur l'influence de la personnalité et la conscience du corps sur le SoE). Toujours selon Dewez et al. (2019) le locus de contrôle est le facteur interne qui impacte le plus de composantes de l'incarnation. Pour rappel, le principe du locus de contrôle est issu de la théorie de l'attribution causale de Weiner (1986), qui permet d'attribuer à un acte une origine, plaçant l'individu selon différentes postures devant l'acte/situation en question. Ainsi, la causalité attribue la responsabilité de la réussite ou de l'échec à des facteurs internes ou externes et contrôlable ou incontrôlable par exemple. Ainsi, parmi ses composantes, on trouve l'agentivité (positivement corrélé avec le locus de contrôle interne) et le sentiment de possession (positivement corrélé avec un locus de contrôle externe).

En ce qui concerne les composantes du SoE, dans les autres disciplines comme la neurorobotique, Tsakiris et Critchley (2016) et Ehrsson (2020), cité par Segil et al. (2022), expliquent que différentes combinaisons de modalités (vision, toucher, proprioception, audition et intéroception) influencent le sentiment de possession. La revue de littérature de Segil et al. (2022) traite de la question des prothèses avec d'étonnantes similitudes avec la RVI. La comparaison entre un membre virtuel et une prothèse doit être faite néanmoins avec prudence. La littérature partage certains auteurs (De Vignemont, Blanke, etc.) et les sciences cognitives sont clairement citées (Segil et al., 2022) dans les fondamentaux du SoE en neurorobotique. Cette comparaison est donc intéressante car dans les deux cas, le membre « factice » est perçu comme sien par l'utilisateur avec le même processus d'incarnation comportant les mêmes composants. Il est intéressant de noter l'ajout d'un concept de processus subconscient neurophysiologique (Arzy et al., 2006 ; Blanke, 2012 ; Makin et al., 2017 ; Braun et al., 2018, cités par Segil et al., 2022) qui sert d'ancrage à l'incarnation. Il est ici question d'une approche à deux niveaux avec, d'un côté, le conscient, avec les trois composants du SoE vu précédemment (localisation de soi, agentivité et possession, considérés comme séparables), et, d'un autre côté, le subconscient, avec des composants comme le schéma corporel de Biocca (1997) ou l'intégration multi sensorielle, qui interagissent entre eux pour moduler les composants conscients. Selon Gouzien et al. (2017), cités par Segil et al. (2022), l'incarnation survient lorsque l'outil, le dispositif ou le membre artificiel, est intégré et reconnu par les

systèmes neurobiologiques de façon similaire à un membre naturel. Il est à noter les extensions « outil » et « dispositif » sont des appellations applicables à des membres virtuels.

La revue de littérature de Segil et al. (2022) fait le point sur le SoE. Ils présentent une infographie qui permet d'appréhender le SoE de façon claire⁴. Il est intéressant de noter que l'agentivité et le sentiment de possession ne sont pas influencés par les mêmes traits, suggérant qu'elles peuvent apparaître indépendamment. Segil et al. (2022) présentent enfin différentes méthodes de mesure du SoE : comportementales, qui correspondent aux mesures explicites et physiologiques qui correspondent aux mesures implicites.

Les travaux de Falcone (2022) proposent une approche similaire mais simplifiée pour obtenir une bonne perception du SoE avec des mesures explicites qui correspondent aux entretiens et questionnaires comme chez Peck et Gonzalez-Franco (2021), ainsi que des mesures implicites comme le diamètre pupillaire (*pupillar diameter* - PD), les battements cardiaques (*heart rate* - HR) et enfin la réponse électrodermale (SCR) de Ehrsson et al. (2007) avec une expérience sur une *rubber hand* qui confirment les résultats précédents.

1.3.3 L'agentivité

Bien qu'elle soit un composant du SoE dans les modèles de RV comme celui de Kilteni et al. (2012), l'agentivité est traitée dans le CAMIL comme un élément à part entière à cause de son importance sur l'expérience de RVI. De plus, des travaux récents tels que ceux de Jeunet et al. (2018) ont décomposé l'agentivité en plusieurs sous parties de façon très pertinente, ce qui incite à considérer cet élément comme plus important que la littérature ne le supposait de prime abord. L'agentivité sera traitée ici de façon spécifique bien que le concept apparaisse dans la littérature comme composant du SoE dans l'ensemble des articles.

Haggard (2017) explique que l'agentivité correspond à « *l'expérience subjective d'être la source de l'action* », l'utilisateur étant celui qui produit l'action via son avatar virtuel. Caspar et al. (2015) expliquent que l'agentivité correspond également aux conséquences des actions de l'utilisateur dans l'environnement, c'est à dire les *feedback* visuels et auditifs qu'il va recevoir après avoir effectué une action et qui correspondent à la conséquence logique de cette action. La notion de cohérence est ici importante et il faut que le retour soit cohérent avec l'action, que ce soit de manière « réelle » ou bien que cela corresponde à la logique de la situation avec les

⁴ Voir annexe B.1.1 : [Embodiment framework de Segil et al. \(2022\)](#) p. 271.

règles établies spécifiquement pour la situation (cas d'un environnement virtuel onirique par exemple). Enfin, comme l'explique De Vignemont (2011), cité par Brugada-Ramentola et al. (2019), l'agentivité est étroitement lié à la conscience de l'action ainsi qu'à sa planification.

Blakemore et al. (2002) ont travaillé sur les éléments qui affectent l'agentivité, comme par exemple le fait de créer un décalage (visuel ou temporel) qui va abolir l'agentivité. C'est le cas dans l'expérience de la *rubber hand*, où un seuil maximum qui va rompre l'illusion de plausibilité (le leurre cognitif proposé par Slater en 2009). Brugada-Ramentola et al. (2019) citent également les travaux d'Haggard (2017) qui soulignent l'importance du visuomoteur pour montrer le mouvement afin de générer de l'agentivité. Brugada-Ramentola et al. (2019) présentent également plusieurs travaux qui nuancent néanmoins l'agentivité avec la possibilité d'avoir de l'agentivité malgré une absence de mouvement réel et ce, uniquement grâce aux mouvements de l'avatar, comme ceux de Tieri et al. (2015b) ou encore Wegner et al. (2004), cités par Brugada-Ramentola et al. (2019). Ces résultats sont confirmés par les travaux de Wegner et al. (2004) ; Tieri et al. (2015b) et Pezzetta et al. (2018), cité par Brugada-Ramentola et al. (2019) qui expliquent que même sans implication motrice directe, le simple fait d'observer passivement un membre en mouvement peut induire un sentiment d'agentivité par procuration sur ce mouvement.

Gorisse (2019) approfondit la notion d'agentivité en présentant lui aussi les travaux de Jeunet et al. (2018). Ces travaux introduisent deux composants : le sentiment d'agentivité et le jugement de l'agentivité. Ces travaux divisent donc cette notion (le sentiment d'agentivité devient un sous composant et non plus une autre façon d'appeler l'agentivité) et reposent sur trois principes clés :

- **La priorité** : « l'intention consciente d'accomplir un acte doit précéder immédiatement l'action, qui à son tour doit précéder immédiatement le résultat ». Une boucle intention, action, résultat, qui reprend les fondamentaux des boucles d'interactions.
- **La cohérence** : « *le résultat sensoriel doit correspondre au résultat prédit* ». Cette notion de cohérence s'applique quelle que soit la situation (réelle ou non). Ce qui se passe doit correspondre à ce que l'utilisateur a envisagé.
- **L'exclusivité** : « *la pensée doit être la seule cause apparente du résultat (absence d'influence extérieure)* ». L'utilisateur doit penser que la source d'une action est la seule cause de cette dernière. Ce principe est en lien avec celui de la cohérence.

Pour illustrer ceci, Gorisse (2019) explique, en citant Slater (2009), qu'il existe une production importante de projets (ludiques ou professionnels) non réalistes mais qui n'en demeurent pas moins efficaces. Outre la notion de subjectivité, le plus important reste de garantir la cohérence et la plausibilité des événements dans l'univers dédié. Il est question ici de l'illustration de l'illusion de plausibilité proposée par Slater (2009). L'agentivité serait liée à une modélisation du système nerveux central (SNC) comme chez Farrer et al. (2008). Elle représente les conséquences sensorielles prévues du mouvement, qui sont ensuite comparées aux signaux de rétroaction sensorielle réels résultant de ce mouvement. Farrer et al. (2008) expliquent que si ces signaux sont corrélés, ils donnent naissance à l'agentivité. La représentation du corps et la capacité de contrôler cette représentation sont donc importantes pour l'agentivité dans les environnements immersifs.

Pour conclure, l'impact de l'agentivité est important car, étroitement lié au sentiment de possession (*sense of ownership* - SOO), il affecte l'incarnation directement par les actions de l'utilisateur et peut même détruire l'illusion de plausibilité si aucun retour du résultat de l'action n'est proposé à l'utilisateur, Caspar et al. (2015).

1.4 L'immersion technologique

Slater (1999), introduisait déjà une extension du concept « d'immersion » vers « immersion technologique » afin de limiter le recouvrement et l'ambiguïté avec la présence. En partant de l'immersion et ses nombreuses problématiques de similitude avec la présence, pour aller vers l'immersion technologique, Gorisse et al. (2018) ont repris le concept de Slater, l'idée étant de proposer une évolution d'un concept flou qui a posé de nombreux problèmes à plusieurs auteurs. En s'appuyant sur la méta-analyse par Bailenson et Cummings (2016), ne reprenant que des points techniques, voici les facteurs technologiques de l'immersion décrits par Gorisse et al. (2018), p. 99 :

- Capture de mouvement : nombre et type de degrés de liberté avec lesquels les mouvements d'un utilisateur sont capturés dans un système.
- Vision : comparaison de la vision stéréoscopique et monoscopique.
- Champ de vision : ouverture du champ de vision.
- Taux de rafraîchissement : nombre d'images par seconde rendues par le système.
- Qualité de l'image : éléments influençant la qualité générale tels que le réalisme, la fidélité visuelle, la résolution, l'éclairage, les textures, etc..
- Qualité du son : présence ou absence de son, spatialisation, nombre de canaux, etc..

- Point de vue : comparaison des points de vue (première et troisième personne).

De cette méta analyse ressort un certain nombre d'éléments qui vont impacter tout particulièrement l'immersion. Ainsi, la capture du mouvement de l'utilisateur et sa restitution dans le casque occlusif est importante. Pour rappel, cela demande une synchronisation visuo-motrice et un taux de rafraîchissement qui respectent le temps minimum requis par la boucle d'interactions pour ne pas créer de latence. Enfin, la vision stéréoscopique et le champ de vision sont les derniers éléments importants qui vont permettre un affichage et un retour visuel qui vont immerger l'utilisateur dans l'environnement virtuel. Une vision en 2D seulement n'a pas d'effet immersif aussi puissant qu'une vision en 3D et ce, de loin, selon Gorisse (2019). Gorisse et al. (2018) terminent en expliquant que la qualité visuelle et sonore de l'environnement semble avoir une incidence nettement plus modérée que les autres facteurs. Cela semble être confirmé par les travaux d'Arnaldi et al. (2018) ainsi que ceux de Bailenson (2018) qui proposent comme concept de conception la création d'environnements virtuels et contenus allégés visuellement mais aux performances accrues (sans que cela impacte l'expérience des utilisateurs). Les auteurs rappellent également que le niveau de réalisme est considéré comme un facteur d'efficacité de la RV, souvent mis en avant alors que ce sont les autres facteurs abordés qui sont les véritables garants d'une expérience de RV réussie. Pour finir, Moreau et al. (2006) rappelaient que la RV avait pour intérêt de représenter la « meilleure copie » du monde réel et non pas sa parfaite reproduction.

1.4.1 *Les avatars et agents - UX et PNJ*

Figure représentant l'utilisateur au sein d'un espace virtuel, l'avatar (et son équivalent PNJ - les agents ou personnages non joueurs, ou encore NPC pour *Non Personal Character* en anglais) est au cœur des concepts de la RV. Sans avatar ou agent, on ne parle pas de RV mais de film d'animation, d'expérience télévisuelle ou l'utilisateur n'est qu'un spectateur sans action possible.

Les avatars et agents PNJ ont un impact majeur sur l'expérience de l'utilisateur. La cohérence de leur comportement passe par leur expression et cette notion est un défi de taille en RV. Mais tout d'abord, il faut faire les distinguer l'un de l'autre. Gorisse (2019) explique qu'un « avatar » est contrôlé par l'humain alors qu'un « agent » est contrôlé par IA (les PNJ). Il parle également d'avatar incarné (vue FPS notamment), les plus pertinents en RVI. Bien que les travaux de Gorisse (2019) portent principalement sur les avatars (donc humain), l'auteur fait

des parallèles avec les agents (PNJ - IA) en expliquant que ce qui s'applique aux avatars s'applique aussi aux agents.

Gorisse (2019) rappelle que pour le moment, la problématique de réalisme des avatars avant tout une limite technique car ne pouvant atteindre de véritable réalisme. Le résultat actuel, bien que d'excellente facture et atteignant un degré de réalisme très élevé comme l'explique Seymour et al. (2017), reste contre-productif dans la recherche de présence et d'incarnation ; allant jusqu'à provoquer des rejets de la part des utilisateurs (cas de certains jeux vidéo où les agents étaient jugés dérangeants, comme le cas de « *Mass Effect Andromeda* »). La littérature obtient de meilleurs résultats avec des avatars et agents moins réalistes (mais parfaitement animés et réactifs) que leurs homologues photoréalistes par exemple. Gorisse (2019) le confirme avec les études de Zell et al. (2015) et de Fleming et al. (2017). Selon ces auteurs, le degré de réalisme n'est pas l'élément qui rendra les agents ou avatars plus attractifs. Les travaux de Zell et al. (2015) parlent des incohérences des modèles et de l'impact négatif que cela a sur l'attractivité, tandis que Fleming et al. (2017) démontrent que les avatars stylisés sont évalués plus positivement que les modèles issus de reconstruction 3D. Gorisse (2019) explique également qu'un haut degré de réalisme induit un horizon d'attente plus élevé et donc plus sensible aux incohérences. Gorisse (2019) nomme le phénomène opposé qui en découle comme la « suspension consentie de l'incrédulité » qui consiste à inhiber temporairement le jugement vis-à-vis de la vraisemblance de la situation (inversement proportionnel au degré de réalisme).

1.4.2 *La fidélité (visuelle et comportementale)*

Fidélité visuelle : Malgré ces limites techniques, qui seront un jour franchies, il convient d'aborder la question de fidélité visuelle qui dépend, d'après les travaux de Garau (2003) et Mansour et al. (2006), cités par Gorisse (2019), de trois critères :

- **L'anthropomorphisme :** L'aspect non-humanoïde en opposition à l'aspect humanoïde de l'avatar. Comme le rappelle Biocca (1997) avec le Dilemme du cyborg : il est possible de choisir l'incarnation technologique pour amplifier le corps, mais il faut prendre garde que le schéma corporel et l'identité puissent s'adapter à cette forme non humanoïde (cyborg). Dans cette logique, il est intéressant d'évoquer Steptoe et al. (2013) qui précisent qu'il est malgré tout possible d'éprouver un sentiment de possession et d'agentivité envers des avatars n'appartenant pas au schéma corporel humanoïde.

- **Le réalisme** : Le fait d'opposer avatars stylisés d'un côté (comme ceux utilisés dans ce travail) et photo réalisme de l'autre (avec les limites et problématiques actuelles).
- **La véracité de l'avatar** : Le fait de ressembler ou non à l'utilisateur et qui peut impacter l'expérience de RVI de plusieurs manières, comme le montre l'expérience « *Virtually being Einstein* » de Banakou et al. (2018) altérant les performances cognitives des utilisateurs en modifiant simplement l'apparence de l'avatar des utilisateurs et en la remplaçant par celle du célèbre physicien.

Fidélité comportementale : Elle dépend des propriétés dynamiques, de l'animation et des expressions faciales selon Gorisse (2019). Ces éléments doivent répondre aux principes de cohérence, suivre et répondre aux *inputs* de l'utilisateur avec précision et proposer des réponses expressives représentatives. Cela qui demeure toujours à l'heure actuelle un souci dû aux limites techniques, malgré les progrès considérables et les bibliothèques de plus en plus fournies (ex : la bibliothèque Omniverse Audio2face générée par IA de Nvidia). Gorisse (2019) explique que les animations, composantes essentielles de la communication non verbale, permettent de s'exprimer via le langage corporel.

Une question se pose alors : pourquoi les avatars et agents ont-ils, malgré un réalisme relatif, un tel impact ? Pour cela, il faut se pencher sur un processus d'identification en RV. Dans l'étude de Taylot (2002) cité par Gorisse (2019) portant sur l'expérience « *The dreamscape* », les utilisateurs s'identifient plus à leur avatar qu'à leur image réelle. Selon eux, les avatars retranscrivent une meilleure version d'eux-mêmes. Dans cette logique, la personnalisation (la possibilité de modifier son avatar à volonté) revêt un rôle important, faisant de l'avatar un « individu » et du corps virtuel une « réalité ». Le but : atteindre le stade du « c'est moi ». Ce constat est confirmé par Messinger et al. (2008), sur l'analyse du jeu vidéo « *Second life* » avec la création d'avatar comme étant des meilleures versions des utilisateurs, et ce, pour les rendre plus attractifs aux yeux de la communauté.

Pour rappel, tout ce qui s'applique aux avatars s'applique aussi aux agents (PNJ) car dans ce projet la personnalisation de l'avatar ne sera pas possible. Tous les principes abordés vont donc s'appliquer à tous les agents.

1.5 La conception en réalité virtuelle sous CCU et les modèles précurseurs.

1.5.1 La CCU en RVI

Avant d'aborder le modèle CAMIL, il est nécessaire d'aborder la conception en RV en conception centrée utilisateur (CCU, un concept de Nielsen, 1994, qui met au centre de la conception l'utilisabilité et l'expérience de l'utilisateur). La conception centrée utilisateur en RVI possède un certain nombre de spécificités. Pour rappel, la norme ISO 9241-210, définit l'expérience utilisateur comme « *les perceptions et les réponses résultantes de l'utilisation d'un produit, un système ou un service* ». La norme ISO 924-11 quant à elle, définit l'utilisabilité par 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* ».

Selon Cook et al. (2019) la conception en RV est un challenge de design centré sur l'humain car la RV est liée directement à l'humain, avec l'ergonomie notamment, mais aussi à cause des contraintes physiques et mentales qui sont inhérentes à cette technologie. La RV, via l'illusion de plausibilité, s'adresse directement au cerveau de l'utilisateur grâce à un appareillage occlusif. Concevoir un outil dans de telles conditions implique une structure particulière et la conception centrée utilisateur permet d'offrir un cadrage adéquat. Les allers-retours avec les utilisateurs vont permettre de corriger, lors des phases de maquettage et prototypage, de nombreux problèmes qui seraient passés à travers les mailles de la conception.

Hassenzahl (2001) avec ses travaux en psychologie cognitive sur la qualité hédonique d'un produit, complète cette qualité hédonique avec la dimension ergonomique (en sachant qu'un produit peut être évalué sur une échelle opposant ergonomie et hédonisme aux deux extrêmes, l'équilibre étant la meilleure des options) qui va impacter l'attractivité du produit, ce qui détermine ses comportement et émotion vis-à-vis du produit. Les travaux de Tcha-Tokey et al. (2015) proposent une revue de littérature sur l'expérience utilisateur en environnement virtuel immersif qui présente différentes mesures ainsi qu'un modèle holistique. Les auteurs définissent la notion de *flow* (importante en conception centrée utilisateur et qui permet d'évaluer le degré de divertissement de l'utilisateur) de la sorte : « *état psychologique de sentiment de contrôle, entre le stress et l'ennui* ». Le *flow* est influencé par l'interactivité, les compétences, l'attention portée et la téléprésence. Les auteurs présentent le modèle de Cheng

et al. (2014), basé une approche de modélisation par équation structurelle, le modèle TAM (modèle d'acceptation technologique) de Shin et al. (2013) inspiré des modèles de l'acceptabilité, et enfin le modèle de Lin et Parker (2007) basé sur l'influence des caractéristiques de l'environnement virtuel. Les différents éléments composant ces modèles seront tous repris par le modèle heuristique de Tcha-Tokey et al. (2015)⁵. Ce modèle, théorique, n'a pas été validé et constitue une piste de réflexion pour comprendre sur quelles bases les modèles ultérieurs ont été pensés. Il reste néanmoins très intéressant car il complète les trois autres modèles en ajoutant à chacun d'entre eux les composants manquants.

1.5.2 La notion de métaphores et Schème Comportementaux Importés (SCI)

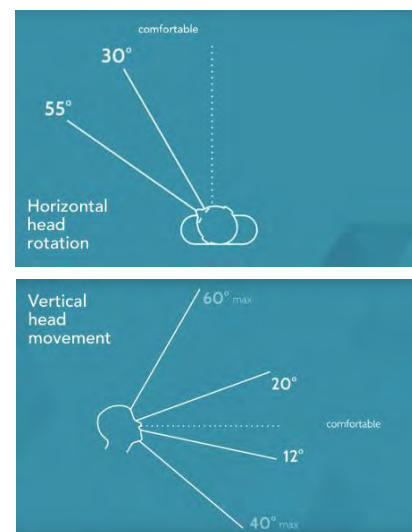
Abordée par Fuchs en 2003 dans le *Traité de la réalité virtuelle* vol. 1, cette notion est importante en conception. Il convient de créer des situations qui vont correspondre au réel ou bien créer des métaphores qui vont simplifier par représentation une action ou situation. Un exemple concret illustre cette notion : le cas du « piratage informatique » dans les jeux vidéo (ou l'on « clic » sur un objet représentant un écran d'ordinateur mais sans clavier, tour...) pour lancer la séquence complète au lieu de le faire comme en vrai, ce qui serait trop long et très dur à réaliser. Il n'est pas possible de réaliser une séquence totale et il est nécessaire de simplifier les tâches pour les rendre abordables avec des « schémas » représentant une illustration possible et concise de la tâche.

1.5.3 L'interface et l'UX design

Les interfaces : Cook et al. (2019), rappellent que, malgré ses atouts indéniables, la RV reste une technologie relativement immature et que la courbe de progression de ses contenus, comme les interfaces, est grande. D'après eux, ces interfaces sont encore complexes tant pour les apprenants que pour les enseignants qui vont utiliser la RV. Les principes fondamentaux de l'UX design adapté de la CCU de Nielsen (1994) sont capitaux est encore aujourd'hui trop sommaire. Voici l'interfaçage optimisé actuel. La zone de confort horizontale doit être située entre 30° et -30° avec un maximal

Figure 4

La zone de confort visuel



⁵ Voir annexe B.1.3 : [Modèle heuristique de Tcha-Tokey et al. \(2015\)](#) p. 280.

de 50° à -50°. La zone de confort verticale doit être située entre 20° et -12° avec un maximal de 60° à -40° avec un centre visuel standard 6° en dessous de la ligne d'horizon (voir fig. 4). L'interface doit être de 120° horizontal et de 60° vertical. La taille des éléments en fonction de leur distance se mesure en DMM (*Distance Matters Measurement* - qui est un rapport d'échelle de taille, 1DMM : 1mm = 1m) avec une distance de confort située entre 5m et 10m. Il est important de noter qu'il existe une zone interdite située entre 0 m et 2m dans laquelle aucun élément ne doit jamais être inséré⁶.

Limite due à la conception : Une des principales limites au niveau pédagogique se situe en amont de la RV du côté de la conception. Cette limite est clairement abordée dans le modèle CAMIL de Makransky et Petersen (2021). La conception en RV pour un usage pédagogique demande une connaissance des caractéristiques qui l'influencent, avec les méthodes d'instruction adaptées comme le rappelle Lewis et al. (2021). D'ailleurs, Moreno expliquait déjà en 2006 la complémentarité de la méthode et du média, d'où l'importance de concevoir une méthode en fonction du média cible en prenant en compte les spécificités de ce dernier. Enfin, Lewis et al. (2021) expliquent qu'il existe un déficit dans la conception des applications pédagogiques dans le domaine de la RV. Ce problème aurait pour origine, selon ces auteurs, la complexité de développement des applications éducatives qui demande une interdisciplinarité qu'il est difficile de mettre en œuvre.

2. La réalité virtuelle pour l'apprentissage

2.1 Les intérêts possibles de la RVI pour l'apprentissage

La RVI est une technologie relativement récente qui a bénéficié d'un engouement croissant auprès du grand public, Bailenson (2018). Bien qu'elle présente de nombreuses limites (Poupard et al., 2024), plusieurs études ont mis en évidence ses atouts potentiels (Lewis et al., 2021 ; Makransky & Petersen, 2021). Examinons ses principaux avantages dans le cadre de l'apprentissage et de la compréhension.

Comme l'explique le premier chapitre de l'ouvrage dirigé par Arnaldi et al. (2018), la RV constitue une véritable révolution dans le domaine de l'ingénierie pédagogique. Parmi les

⁶ Voir annexe B.3 : [Principes d'interfaçage en RV](#) pour plus de détail p. 308.

avantages mis en avant figurent la possibilité de scénarisation des apprentissages, la réduction des risques liés à certaines pratiques professionnelles (« dérisquage »), l'économie de moyens dans certaines configurations de formation ainsi que l'attrait intrinsèque de la technologie elle-même. Ces caractéristiques font de la RV un outil particulièrement pertinent dans des domaines à fort enjeu sécuritaire ou technique, comme le secteur médical ou l'industrie, où la simulation permet d'optimiser la formation sans exposition au danger : former un chirurgien constitue une tâche complexe dans laquelle la RV peut jouer un rôle significatif.

Carruth (2017) propose une présentation globale des principaux outils et modèles utilisés dans ce domaine, tels que la conception assistée par ordinateur (CAO) ou la photogrammétrie. Cette analyse met en lumière les atouts des dispositifs de RV tout en soulevant certaines limites liées à leur utilisation dans un cadre pédagogique. La diversité des possibilités offertes par ces technologies y est soulignée, ainsi que plusieurs pistes potentielles d'exploitation à des fins éducatives. D'un point de vue méthodologique, Dewez et al. (2019) soulignent que la RV offre un contrôle précis sur de nombreux paramètres expérimentaux, allant des conditions de stimulation à l'apparence ou la configuration de l'avatar utilisé. Ce contrôle fait de la RV un dispositif expérimental particulièrement pertinent, capable de s'adapter à une grande diversité de contextes et d'objectifs de recherche. Schutte et Stilinović (2017) confirment, par leurs travaux sur l'empathie, la puissance de l'outil RV qui entraîne un engagement supérieur des utilisateurs. De même, Bouchard et al. (2013) soulignent l'importance de la présence tout en rappelant la notion de Vallée de l'étrange de Mori (1970). Ce concept décrit le malaise ressenti face à des robots ou personnages quasi-humains et révèle un mécanisme cognitif où le système de reconnaissance des visages et corps humains détecte une incongruence : les signaux familiers entrent en conflit avec des indices subtils d'anormalité, Kätsyri et al. (2015). Néanmoins, ce concept est encore débattu et sujet à controverses, il convient donc de rester prudent. Dans une revue de littérature consacrée à la RV et à ses effets pédagogiques, Lewis et al. (2021) établissent un état des lieux relativement exhaustif des bénéfices associés à cette technologie. En premier lieu, ils évoquent les effets positifs sur la motivation des apprenants dans les environnements d'apprentissage. Les auteurs abordent également des aspects liés à la conception pédagogique ainsi qu'à la collaboration entre apprenants, axe qui, ne constitue pas un point d'intérêt dans le cadre du projet étudié. Voici un résumé de cette revue de littérature de Lewis et al. (2021) :

La motivation : Ce facteur a une influence sur une fonction cognitive importante dans l'acquisition de connaissances : l'attention, comme chez Long et al. (2011), cités par Lewis et

al. (2021). Selon Mildner et al. (2015) et sous les principes de la gamification, cette amélioration de la motivation serait due aux axes du partage et de la conservation des acquis en jeu. Il faut néanmoins rester prudent avec la gamification (abordé dans la partie des limites de la RV). D'après Cook et al. (2019) si un artefact est bien conçu, alors il sera motivant et efficace. Selon Cooper et al. (2019) la possibilité de pouvoir explorer ou encore modéliser des phénomènes abstraits est motivant. Par exemple, les expériences « *in the cell* » ou encore « *the body VR - Journey inside a cell* » illustrent parfaitement le fait de pouvoir visiter des lieux impossibles dans la réalité (IRL). Dans ces expériences, les utilisateurs sont miniaturisés et projetés dans un organisme. Avec de tels processus, l'engagement est accru, ce qui est motivant et amène à dépenser plus de temps sur la tâche éducative. Enfin, en sus de la motivation, c'est de la persévérance que la RV peut entraîner selon Chen et al. (2020) avec leurs travaux sur *Google Earth VR*. Il convient de rappeler que la RVI reste sujette à controverses sur le domaine de la motivation comme l'expliquent Poupard et al. (2024) dans leur revue systématique.

L'immersion : Pour rappel, à cause du flou de la frontière entre présence et immersion, le terme « d'immersion technologique » est favorisé dans la littérature, avec une focalisation sur les aspects techniques permettant d'immerger l'utilisateur en RV. Néanmoins, certains auteurs comme Lewis et al. (2021), p. 18, utilisent encore cette définition, en incluant les aspects psychologique et physique, avec la définition suivante : « *ce qui permet l'introduction chez la personne de la croyance qu'elle a quitté le monde réel et qu'elle est maintenant présente dans l'environnement virtuel* ». Leur revue de littérature aide à appréhender ce concept « d'être immergé ». Jensen et Konradsen (2018) expliquent que les technologies type casque oclusif permettent la mise en place d'approches pédagogiques constructivistes, ce qui permet des apprentissages basés sur la simulation. Également cités par Lewis et al. (2021), Edwards et al. (2019), expliquent qu'avec le multi sensoriel notamment, la RV améliore la compréhension des concepts abstraits par exemple. Enfin, d'après Sol Roo (2017) ; Ucar et al. (2017) ; et Yoon et al. (2017), cités par Lewis et al. (2021), sur la base des principes de la théorie du développement cognitif de Piaget (1964), la RV permettrait de pallier à certains déficits comme celui de la pensée opérationnelle concrète en permettant de créer des modèles mentaux corrects avec des expériences de type « conte et histoire ».

Lewis et al. (2021) font ensuite un listing des autres avantages qui recoupe certains points d'Arnaldi, Guiton et Moreau (2018). Selon eux, la RVI présente de nombreux avantages dans les contextes éducatifs et de formation professionnelle. Sur le plan de la sécurité et de l'accessibilité, cette technologie permet notamment la manipulation d'objets dans des

environnements potentiellement dangereux de manière entièrement sécurisée comme l'expliquent Jensen et Konradsen (2018) et Makransky et al. (2019). Elle contribue également à une réduction significative des coûts de formation, particulièrement dans les domaines dont les équipements requièrent des investissements importants et comportent des risques tant pour les apprenants que pour l'environnement selon Cook et al. (2019) ainsi que Cooper et al. (2019). En matière d'autonomie et de contrôle de l'apprentissage, la RVI augmente substantiellement les possibilités pour les apprenants de gérer leur propre processus d'acquisition de connaissances selon Altinpulluk (2019). Hite (2022) explique que cette technologie offre un engagement sensoriel comparable au processus naturel d'apprentissage observé chez l'enfant, favorisant ainsi une appropriation plus intuitive des contenus pédagogiques. Sur le plan du développement des compétences cognitives, la RVI démontre des effets bénéfiques multiples. Elle améliore l'acuité spatiale, stimule la créativité et enrichit les processus métacognitifs des apprenants, Lee et Shea (2020). Enfin, en termes de communication et de collaboration, la RVI présente des avantages substantiels. Cette technologie facilite l'interaction en créant des environnements d'apprentissages plus dynamiques et interactifs, Demitriadou et al. (2020).

Cette liste, non exhaustive, des atouts de la RV permet de mieux comprendre pourquoi cette technologie a suscité un tel engouement lors de son grand retour en 2016. Néanmoins elle n'est pas exempte de limites, spécifiquement dans les tâches d'apprentissage, certaines entrant en contradiction avec certains résultats présentés plus haut (motivation, charge cognitive, etc.).

2.2 Les limites de la RVI à prendre en compte pour les tâches d'apprentissages

Les premières limites de la RV n'ont pas été mise en lumière immédiatement mais sont aujourd'hui connu de la littérature. Certains éléments demeurent encore incertains par manque de recul et de données. Le principe de précaution est donc appliqué sur certaines recommandations d'usage en l'attente de recherches plus complètes.

Le rapport de l'Agence nationale de sécurité sanitaire (ANSES) de 2021 référence actuellement l'ensemble des limites des technologies immersives, RV incluse. Ce rapport, élaboré par les experts du domaine, est un travail conséquent qui ne sera abordée ici que partiellement. Il est intéressant de noter que parmi les experts mandatés pour rédiger le rapport, figure Philippe Fuchs et Guillaume Moreau, pour ne citer qu'eux, deux des plus éminents experts de la RV. Plus récemment, Poupard et al. (2024) ont mené une revue systématique sous méthodologie PRISMA pour clarifier comment et dans quelle mesure les technologies

immersives (RV et RA) influencent réellement l'apprentissage, en examinant spécifiquement leurs effets sur la charge cognitive et la motivation intrinsèque. Ce qui est intéressant avec ces travaux, c'est qu'ils évaluent également le niveau de preuve des concepts et les méthodologies des articles concernés. Cela permet de prendre du recul sur certaines données présentées.

2.2.1 *Les limites issues du rapport ANSES 2021*

Le manque d'articles : Un des premiers soucis mis en avant par l'ANSES est le manque de données. Technologie ancienne et récente, la RV n'a été étudiée que récemment en. Malgré le graphique de Markransky et Petersen (2021) qui montre que le nombre d'articles est en pleine expansion, la littérature a besoin de s'étoffer. De nombreux pans de la RV sont encore trop peu documentés et l'ANSES recommande de mener davantage de travaux. Par exemple, l'effet des rayonnements bleu, dont l'impact a été prouvé sur le rythme circadien (voir ANSES 2019) lors d'utilisation d'écran *flatscreen*, est encore à déterminer pour les écrans des HMD. Ergonomie, psychologie avec l'aspect psycho social, neurologie, perception de soi et développement sont des domaines qui manquent de travaux pour conclure à l'existence ou non d'effets de la RV. Le principe de vigilance est donc appliqué à l'ensemble de ces domaines bien que l'ANSES les traite dans des parties distinctes avec des recommandations spécifiques. Des études longitudinales sont aujourd'hui nécessaires afin de déterminer l'impact réel de la RV sur l'humain.

L'ergonomie : Au niveau ergonomique, c'est le poids des casques occlusifs qui pose aujourd'hui problème, bien que ce poids ait une chance de se réduire avec les futurs modèles et les évolutions de la technologie. A l'heure actuelle le poids d'un Meta Quest 3 (casque de référence utilisé dans ce travail) est de 515g contre 503g pour le modèle antérieur (Meta Quest 2). Ce poids pèse sur la tête des utilisateurs et entraîne un inconfort qui freine l'utilisabilité. Le rapport ANSES cite également la question de la fatigue visuelle, évoquée mais avec des résultats contradictoires. Kim et al. (2019) constatent de la fatigue visuelle tandis que Gao et al. (2019) ont des résultats plus mitigés. Cette fatigue serait due à la proximité des écrans.

Les effets persistants / consécutifs à l'exposition : Les perturbations de l'habileté manuelle et celle de la perturbation de l'orientation du corps constituent un autre élément à analyser. Champney et al. (2007) ont proposé une expérience avec une tâche d'insertion d'une cheville dans un trou et une tâche de marche le long d'une ligne en post exposition RV. Bien qu'ancienne, cette étude est intéressante. Elle démontre l'effet d'un post effet avec, pour 60% des participants, un taux doublé d'erreurs. Il s'ensuit une amélioration mais, au bout d'une

heure, les mesures ne sont toujours pas revenues aux valeurs initiales. Ces résultats sont confirmés par Harm et al. (2008) qui ajoutent un ordre de grandeur de l'erreur spatiale de 15% et un retour à la normale au bout de six heures. Enfin, plus récemment, Day et al. (2018) ont montré dans leur expérience avec pointage sur cible dans un environnement virtuel puis en réel, qu'il fallait environ soixante-cinq essais pour revenir à la normale. Il en va de même pour l'orientation du corps avec l'expérience de Viaud-Delmon et al. (1999). Ces travaux expliquent qu'après une rotation en RV, les participants sous évaluent la rotation réelle qu'ils ont fait dans l'environnement réel. Bien qu'anciens également, ces travaux sont intéressants pour montrer le décalage que la RV peut produire sur la perception de l'environnement réel.

Les recommandations pour les troubles mentaux - les utilisateurs à proscrire : Sur la base des réflexions de Viaud-Delmon (2007) sur le rapport RV et réel, certains auteurs se sont penchés sur la confusion que pourrait engendrer la RV sur des populations à risques (personnes sujettes à l'idéation paranoïaque et aux troubles dissociatifs). Les résultats des études Freeman et al. 2007, Valmaggia et al. (2007) et plus récemment, Hesse et al. (2017), cités dans le rapport ANSES 2021, confirment que les populations à risques sont à exclure de la RV car elles sont plus susceptibles de développer des effets post exposition. Ces travaux suggèrent une certaine relativité mais propose de rester prudent en excluant ces populations dans l'attente de plus de résultats. Lewis et al. (2021) confirment, avec les travaux de Kenwright (2018), la possible apparition de traumatismes physiologiques et changements comportementaux qui pourrait être causée par des modifications psychologiques ou neurologiques. Ce point est confirmé également par Ramirez et LaBarge (2018), bien qu'ils expliquent eux aussi qu'il faille rester prudent avec ces résultats. Par ailleurs, Ramirez et LaBarge (2018) recommandent l'application du principe d'équivalence qui est d'éviter de faire vivre aux utilisateurs une expérience en RV qui serait mauvaise pour eux IRL.

La cyber sécurité et l'éthique : Le rapport ANSES propose un cadrage éthique et un rappel des droits, avec les problématiques d'atteinte à la vie privée issue de la cyber sécurité. La question de la législation est abordée. La notion d'écologie est également évoquée avec les conséquences de la production d'unités technologies comme les casques occlusifs et l'impact environnemental. Il aborde enfin des questions sociétales à la fois pour l'utilisateur mais aussi pour son entourage direct ou indirect. Ramirez et LaBarge (2018) expliquent qu'il convient d'être prudent avec la RV à cause d'un risque moral unique et qu'il convient de concevoir avec une analyse accrue des risques afin de se prémunir d'éventuels écueils.

2.2.2 *Les limites issues de la revue systématique de Poupard et al. (2024)*

Le niveau d'expertise : Cette revue récente sous méthodologie PRISMA révèle un résultat important : l'efficacité de la RV dépend fondamentalement du niveau d'expertise préalable des apprenants. Cette observation est conforme à l'effet de réversion de l'expertise documenté dans la littérature sur la charge cognitive. Les études rapportant des effets négatifs impliquaient majoritairement des apprenants novices, sans connaissances préalables sur le sujet enseigné. Pour ces débutants, la RV entrave significativement l'apprentissage en raison de la surcharge cognitive extrinsèque qu'elle impose. En revanche, toutes les études montrant des effets positifs concernaient des apprenants de niveau intermédiaire, possédant déjà une base de connaissances dans le domaine. Cette base préexistante leur permet de gérer la charge supplémentaire imposée par la RV et de prévenir la surcharge cognitive, transformant un obstacle de la technologie en un facilitateur d'apprentissage.

Cette découverte a des implications pratiques majeures. La majorité des situations d'apprentissage formel concernent des apprenants novices découvrant un nouveau domaine de connaissances. Or, c'est précisément pour cette population que la RV se révèle la plus problématique. L'utilisation de la RV pour l'enseignement de contenus nouveaux à des débutants risque donc d'entraver l'apprentissage plutôt que de le faciliter. La RV semble davantage appropriée pour le perfectionnement ou l'approfondissement de compétences chez des apprenants qui possèdent déjà les bases conceptuelles nécessaires. Cette restriction considérable du champ d'application efficace remet en question l'utilité de la RV pour la majorité des contextes éducatifs traditionnels et suggère que son rapport coût-bénéfice, déjà questionnable sur le plan financier et logistique, l'est également sur le plan pédagogique.

La motivation : L'impact de la RV sur la motivation intrinsèque révèle des résultats préoccupants. Seules 20% des études démontrent un effet positif concomitant sur la motivation et la performance d'apprentissage, tandis que 67% rapportent des effets incohérents entre ces deux dimensions. Cette incohérence révèle un problème conceptuel majeur. La revue met ainsi en évidence la confusion terminologique entre motivation intrinsèque, motivation à apprendre, plaisir perçu et intérêt situationnel. Poupard et al. (2024) proposent une interprétation intéressante, expliquant que la majorité des études rapportent des effets positifs de la RV sur la motivation (combinés avec des effets nuls ou négatifs sur l'apprentissage effectif). Ce pattern suggère que la motivation induite par la RV est dirigée vers l'exploration de la technologie plutôt que vers l'apprentissage du contenu. Cette dissociation entre motivation d'usage et

motivation d'apprentissage représente une limite fondamentale et remet en question l'un des arguments majeurs justifiant l'adoption de la RV en éducation. La revue systématique de Poupard et al. (2024) dresse ainsi un portrait critique. Contrairement aux prédictions optimistes, la RV tend à surcharger cognitivement les apprenants novices et la motivation induite est mal orientée pédagogiquement. L'effet de réversion de l'expertise suggère que la RV pourrait être bénéfique pour des apprenants de niveau intermédiaire mais ce champ d'application restreint limite son utilité pour l'éducation de masse.

2.2.3 *La charge cognitive en réalité virtuelle : entre limites et potentialités*

La revue systématique de Poupard et al. (2024) soulève également un autre enjeu fondamental avec la notion de charge cognitive en RV. L'analyse de plusieurs études portant sur la RV révèle des résultats préoccupants. Parmi celles ayant examiné l'impact de la RV sur l'apprentissage en relation avec la charge cognitive, 38% rapportent des effets négatifs contre seulement 15% d'effets positifs. Ce déséquilibre contredit les prédictions des modèles théoriques tels que le CATLM (*Cognitive-Affective Theory of Learning with Media*) ou le CAMIL, qui postulent que la richesse sensorielle et l'interactivité élevées de la RV devraient optimiser la charge cognitive et favoriser l'apprentissage.

Fondements théoriques de la charge cognitive, les travaux issus de la théorie de la charge cognitive ou *Cognitive Load Theory* (CLT) de Sweller (1994) et de la *Cognitive Theory of Multimedia Learning* (CTML) de Mayer (2005) permettent d'appréhender ce concept clé en RVI. Les derniers travaux de Sweller (2023) sur la CLT confirment, malgré certains effets non répliqués (effet de doubles modalités par exemple), l'évolution et l'intérêt de la CLT ainsi que, plus généralement, son ouverture vers la cognition. La charge cognitive fait référence au stockage de l'information en mémoire de travail et à son traitement. La CLT explique notamment les limites du traitement cognitif des informations. Elle s'appuie sur un modèle architectural cognitif caractérisé par un système de mémoire de travail aux capacités limitées, comprenant des unités de traitement semi-autonomes dédiées respectivement au traitement des informations visuelles et auditives selon Paas et al. (2011). Ces dernières entretiennent des interactions dynamiques avec un système de mémoire à long terme disposant d'une capacité de stockage théoriquement illimitée (Paas et al., 2011).

Au niveau des effets délétères documentés, les recherches démontrent que la RV, particulièrement lorsqu'un casque immersif est utilisé, tend à augmenter la charge cognitive

extrinsèque des apprenants, résultant en une performance d'apprentissage diminuée avec des tailles d'effet moyennes à larges. Trois mécanismes explicatifs convergents ont été identifiés.

- **Premièrement** : La complexité excessive de l'environnement virtuel conduit à une surcharge cognitive. Le nombre accru d'éléments avec lesquels interagir et les distractions présentes dans l'environnement immersif imposent une charge cognitive extrinsèque substantielle qui entre en compétition avec les ressources nécessaires à l'acquisition des connaissances. Jensen et Konradsen (2018) expliquent, en citant les travaux de Moesgaard et al. (2015), que les participants ont indiqué que l'environnement virtuel a « trop capté » leur attention et qu'ils n'ont pas fait attention aux informations qu'ils devaient remarquer normalement. Jensen et Konradsen (2018) ajoutent que la RV peut être distrayante voire contre-productive, indiquant que les méthodes traditionnelles sont plus performantes dans certains cas.
- **Deuxièmement** : L'environnement de RV agit comme un « détail séduisant » qui entrave l'assimilation cognitive. La richesse perceptive capte l'attention sur des éléments visuellement attractifs mais pédagogiquement non pertinents, créant un paradoxe où les apprenants sont émotionnellement engagés mais cognitivement désengagés du contenu à apprendre. Albus et al. (2021) expliquent qu'une surcharge d'éléments visuels peut gêner la sélection et la structuration d'informations en surchargeant le traitement cognitif de l'apprenant.
- **Troisièmement** : L'effet de nouveauté de la technologie impose une double pénalité. Les étudiants doivent apprendre à utiliser le système (charge extrinsèque supplémentaire) et perçoivent la RV comme un outil de divertissement plutôt que d'apprentissage, détournant l'attention de l'objectif pédagogique vers l'exploration ludique de la technologie.

Ces constats ont été confirmés dans divers contextes d'apprentissage. Chang et al. (2019) mentionnent que l'utilisation d'un artefact pédagogique en RV peut augmenter la charge de travail. Ces résultats sont également confirmés au niveau de la RVI pour l'apprentissage en chirurgie avec les travaux de Frederiksen et al. (2020). Par ailleurs, Rettinger et al. (2022) ont constaté que des spectateurs avaient une influence sur la charge cognitive en augmentant cette dernière.

Malgré cela, des pistes d'optimisation et potentialités existent. Ainsi, en parallèle de ces limites documentées, plusieurs travaux suggèrent que la charge cognitive en RV ne constitue

pas uniquement un obstacle mais peut, sous certaines conditions, être optimisée voire exploitée comme un atout pédagogique. D'abord, une conception appropriée de l'environnement virtuel apparaît déterminante. Il convient de générer un environnement virtuel avec le bon nombre d'informations, afin de ne pas surcharger l'utilisateur d'informations parasites. Albus et al. (2021) précisent qu'avec la conception d'une signalisation adéquate, il serait éventuellement possible de réduire le problème de surcharge cognitive (bien que pas en profondeur).

Ensuite, l'intégration d'éléments spécifiques peut alléger la charge cognitive. D'après Steed et al. (2016), un avatar actif peut alléger le problème de charge cognitive sur une tâche de rappel, appuyant l'importance d'un avatar approprié en RVI. Néanmoins, il convient de traiter les agents PNJ avec prudence, malgré le fait que la présence d'agents PNJ dans l'environnement virtuel augmente la présence de l'utilisateur.

Enfin, des approches innovantes émergent pour mesurer et adapter dynamiquement la charge cognitive. Collins et al. (2019) ont proposé une évaluation de la charge cognitive, l'ERVE (*Emotional Responses in Virtual Environments*), principalement centrée sur les états physiologiques des utilisateurs et qui utilise des mesures d'électrocardiogrammes (ECG), d'électroencéphalogrammes (EEG) et d'activités électrodermales (EDA). Dans cette optique, Gerry et al. (2018) proposent un système adaptatif qui répond à la charge cognitive pour s'adapter à l'utilisateur. Ce système, baptisé « *Levity* », permet, grâce à des EEG, de mesurer et d'ajuster de manière interactive l'affichage d'un tableau visuel pendant une tâche de recherche visuelle.

La charge cognitive en RV constitue ainsi un phénomène ambivalent. En effet, si elle représente indéniablement une limite majeure documentée par la littérature, avec des effets négatifs prédominants sur l'apprentissage, elle ne constitue pas pour autant une barrière insurmontable. Certains travaux comme ceux de Wenk et al., 2023, montrent que la charge cognitive ne diffère pas significativement entre la RVI et les écrans 2D traditionnels. Il convient donc d'adopter une posture prudente. Les travaux récents suggèrent qu'une conception pédagogique rigoureuse, tenant compte des principes de la CLT et intégrant des stratégies d'optimisation appropriées, pourrait permettre de transformer cette contrainte en opportunité.

2.2.4 *La cinétose ou cybermalaise*

Le rapport ANSES (2021), tout comme la revue de Poupart et al. (2024) mettent l'accent sur un autre point noir de la RV : la cinétose ou cybermalaise (*virtual sickness*). Ce point doit être traité spécifiquement et demeure l'un des plus problématique parmi les limites de la RV.

Ce trouble proche de celui du mal des transports au niveau des symptômes serait issu, entre autres, de l'incohérence sensori-motrice (ou visuo-vestibulaire) lors d'effet de type «vection» (c'est-à-dire, par exemple, des déplacements de l'avatar en RV tandis que le corps de l'utilisateur reste immobile dans le monde réel). Certains utilisateurs y sont également sensibles par le simple fait de porter un casque occlusif.

Proposé en 1975 par Reason et Brand, la théorie du conflit sensoriel est la première théorie sur la cinétose. Selon les auteurs, l'origine serait un conflit multisensoriel ou intra-sensoriel (par exemple entre le conflit des organes canalaire et otolithique du système vestibulaire comme l'explique Griffin, 1990, dans le rapport de l'ANSES, 2021).

Kennedy et al. (1993) développent en expliquant que l'incohérence sensori-motrice est issue d'un conflit entre trois systèmes sensoriels : le visuel, le vestibulaire et le proprioceptif. Dans leur test *Simulator Sickness Questionnaire* (SSQ), qui reste aujourd'hui un des tests les plus utilisés pour mesurer la cinétose, les auteurs regroupent les symptômes en trois groupes (seize symptômes au total) :

- Les symptômes nauséux : nausée, maux de ventre, salivation accrue, etc.
- Les symptômes oculomoteurs : fatigue visuelle, vision floue, migraines, etc.
- Les symptômes de désorientation : vertiges, étourdissements, etc.

Deux scores sont ensuite calculés : Un score de sévérité globale et un score de sévérité pour chaque symptôme grâce auquel on peut établir le niveau de cinétose de l'utilisateur. Ce test SSQ sera utilisé dans ce travail pour surveiller le degré de cinétose afin de pouvoir corriger l'expérience immersive lors de son développement.

L'origine de l'incohérence sensori-motrice entre les différents systèmes se situe, selon Arnaldi et al. (2018), sur le plan technique notamment. Les auteurs expliquent que la latence de la boucle d'interaction est en lien avec la cinétose et que ce sont les variations de latence des casques occlusifs, d'une amplitude comprise entre 10 et 100 ms (due aux fréquences d'acquisition et de rafraîchissements de capteurs) qui provoqueraient la cinétose. C'est un problème qui était fréquent sur les casques occlusifs grand public des années 90. Le *Virtual boy* de Nintendo notamment n'était pas assez puissants pour effectuer le calcul en dessous de cette limite et ce, avec des fluctuations importantes de la latence, provoquant chez les utilisateurs une puissante cinétose et expliquant l'échec commercial.

Cette cinétose ou cybermalaise est un frein majeur à la RV d'autant plus qu'il est difficile de prévoir quel utilisateur y sera sensible ou non. Pour le moment la littérature ne propose pas de solution pour repérer ceux qui y seront sensibles bien que le rapport de l'ANSES (2021) propose une liste de personnes à surveiller (femmes enceintes, personnes avec troubles oculomoteurs, migraines, troubles de la proprioception, anxiété et mal des transports). Certains auteurs proposent néanmoins des solutions techniques pour atténuer la cinétose⁷.

⁷ Voir annexe B.2 : [Cinétose et confort en RV](#) p. 287.

2.3 L'acquisition de connaissances et compétences en RV - le facteur humain

Avant d'aller plus loin, il est nécessaire d'aborder l'acquisition de connaissances et de compétences en RV. Ce détour conceptuel permet d'introduire la notion de compréhension, centrale dans ce travail. L'objectif est de permettre à des novices de comprendre un concept précis. Si les usages ludiques de la RV demeurent les plus visibles, elle s'impose désormais comme un outil d'apprentissage prometteur, comme en témoignent de nombreuses applications éducatives (*That dragon cancer, Dans la peau d'une personne autiste, Journey inside the cell*).

Cette section abordera successivement les compétences sociales et capacités cognitives mobilisées en RV, les types de connaissances transmissibles (avec leurs atouts et limites), puis la question de la motivation et de l'engagement, qui introduira la partie suivante sur la compréhension d'autrui.

2.3.1 Compétences et habilités sociales, apprentissage cognitif

Pour comprendre ce que la RV peut apporter à l'apprentissage, il convient d'abord de clarifier la nature des compétences visées. Le schéma proposé par Taupiac (2020) concernant la compétence sociale de l'orateur⁸ offre un cadre éclairant. Il illustre la manière dont les éléments constitutifs de toute compétence s'articulent : connaissances déclaratives (le « savoir »), connaissances conditionnelles (savoir quand et pourquoi agir), et connaissances procédurales (le « savoir-faire »). Ces trois types de connaissances se déploient simultanément à travers des dimensions cognitives, comportementales et relationnelles.

Dans l'exemple de l'orateur, les savoirs déclaratifs concernent les conventions de présentation, la posture, l'utilisation du regard ou la gestion de l'espace. Les savoir-faire relèvent de la structuration du discours. Les savoir-être englobent l'adaptabilité, la gestion du stress et la réponse aux imprévus (Gorisse, 2019). Cette architecture tripartite correspond à la définition que Wiemann (1977) donne des compétences sociales : « *la capacité d'un acteur social à choisir, parmi les comportements communicatifs disponibles, ceux qui lui permettront d'atteindre efficacement ses objectifs interpersonnels tout en maintenant une relation équilibrée avec autrui* » (p. 198, notre traduction). Taupiac (2020) utilise d'ailleurs cette expression pour

⁸ Voir annexe B.1.2 : [Schéma des compétences sociales de l'orateur \(Taupiac, 2020\)](#) p. 276.

désigner l'ensemble des aptitudes mobilisées dans les interactions interpersonnelles, particulièrement pertinentes dans un cadre éducatif impliquant des personnes avec TSA.

Au sein de ces compétences sociales, les habiletés sociales constituent un sous-ensemble opérationnel central. Selon Favrod et al. (1998), elles renvoient à l'ensemble des compétences cognitives et comportementales mobilisées pour exprimer de façon adéquate émotions et besoins, et pour atteindre des objectifs interpersonnels. Ces habiletés se déclinent en trois catégories : les habiletés réceptives (décoder les indices contextuels et sociaux), les habiletés décisionnelles (analyser et planifier la réponse adaptée), et les habiletés émettrices (traduire la réponse en comportements verbaux et non-verbaux observables).

Comme le souligne Taupiac (2020), ces compétences s'acquièrent par l'expérience, ce qui en fait une cible naturelle pour la RV. Toutefois, Favrod et al. (1998) introduisent également la notion d'« ensemble de capacités cognitives », renvoyant à une dimension plus fondamentale : celle de l'apprentissage cognitif et de l'acquisition des connaissances élémentaires. Cette dimension cognitive préalable pose question dans le contexte de la RV.

En effet, si l'on revient au schéma de Taupiac (2020), on constate que les connaissances déclaratives constituent la base sur laquelle s'appuient les deux autres types de connaissances. Or, plusieurs études, comme celle de Parong et Mayer (2018), indiquent que la RV, malgré son efficacité remarquable sur l'engagement et la motivation, serait parfois moins performante que des supports traditionnels pour la transmission de ces connaissances déclaratives. Ce paradoxe soulève une interrogation centrale dans ce travail : comment concevoir des environnements virtuels qui, tout en exploitant le potentiel immersif de la RV pour l'acquisition d'habiletés sociales et de compétences expérientielles, ne négligent pas la transmission efficace des savoirs théoriques fondamentaux nécessaires à la compréhension des mécanismes du TSA ?

2.3.2 L'acquisition de connaissances déclaratives et factuelles en réalité virtuelle

Cette question des types de connaissances que l'on peut acquérir en RV soulève une problématique majeure dans le cadre d'un projet de RV. Les connaissances déclaratives (correspondant aux faits, règles ou normes sociales) constituent une base essentielle dans le développement des compétences sociales, notamment chez les personnes avec TSA. Comme l'ont montré Tardif (1992) et Anderson et al. (2001), ces savoirs s'inscrivent dans une typologie plus large comprenant également les connaissances procédurales (savoir-faire) et conditionnelles (stratégies d'application contextuelle).

Bien que les connaissances déclaratives englobent un vaste ensemble de savoirs théoriques, le présent travail se concentrera spécifiquement sur leur dimension factuelle, c'est-à-dire les faits concrets et données empiriques relatives au fonctionnement cognitif des personnes avec TSA. Ce choix permet de cibler un périmètre plus circonscrit et mesurable, tout en reconnaissant la porosité conceptuelle entre connaissances factuelles et déclaratives. Cette focalisation sur les connaissances factuelles apparaît plus adaptée aux objectifs de cette recherche visant la compréhension de mécanismes cognitifs spécifiques.

Il est donc essentiel de pouvoir transmettre des connaissances factuelles dans le cadre de ce projet. Pourtant, les travaux actuels en RV montrent que la RV n'est pas le média le plus adapté pour la transmission de connaissances factuelles (il est bien plus adapté aux connaissances procédurales). De plus, sans réactivation, le niveau de connaissances diminue avec le temps dans la période post acquisition.

Néanmoins, certains travaux comme ceux de Buttussi et Chittaro (2023) montrent que l'acquisition et la rétention de connaissances spatiales est meilleure en RVI par rapport à de la RV non immersives (*flatscreen*). Plus récemment, les travaux de Lindner et al. (2025) montrent que la formation immersive en RV (mais nécessairement combinée à un retour d'information automatisé), permettent une rétention légèrement accrue des savoirs à moyen terme (par rapport à des supports plus passifs comme la vidéo). Ces travaux montrent qu'il est donc possible, dans certains cas particuliers, que la RVI permette une meilleure rétention de certaines informations selon certains types de connaissances. Pour répondre à cette problématique, et, bien que non testé dans ce protocole, le principe des pauses pédagogiques de Parong et Mayer (2018) sera mis en œuvre afin de permettre une transmission et une acquisition convenable de connaissances de type factuelles en RVI.

2.3.3 *La motivation et l'engagement*

La motivation est une question centrale en apprentissage et dépend d'un certain nombre de facteurs. Les travaux de Bandura (2001) et sa théorie sociocognitive (issue de sa théorie de 1986, mise à jour en 2001 avec la perceptive d'agentivité) apporte un éclairage intéressant sur l'importance de la motivation avec une causalité impliquant les caractéristiques individuelles, l'environnement et enfin les actions et comportements mis en place par l'apprenant. De leur côté, Ryan et Deci (2000) avec leur théorie de la motivation (*Self-Determination Theory* - SDT), proposent deux définitions qui subdivisent la motivation en deux aspects distincts :

- Motivation intrinsèque : « Intérêt et plaisir qu'un individu trouve à effectuer une action, sans attente de récompense externe ». Ryan et Deci (2000, p. 70, notre traduction).
- Motivation extrinsèque : « Inspiration qu'un individu trouve à effectuer une action, provoquée par une circonstance extérieure (récompense, sanction, approbation, etc.) ». Ryan et Deci (2000, p.71, notre traduction).

Stone (2001) explique que la RV augmente la motivation intrinsèque grâce au degré de liberté qu'elle apporte, renforcé par l'aspect attirant de la haute technologie qu'elle représente (variante de l'effet « *wow* »). Pour finir, Bouffard et Vezeau (2010) rappellent, en citant Piaget (1950), les facteurs (environnement, cognitif et émotionnel) de l'apprentissage en mettant l'accent sur le rôle central de la perception de sa propre compétence, par la personne elle-même. D'après les auteurs, être un agent actif (autorégulation, contrôle de ses pensées, auto-évaluation, etc.) de son apprentissage est une des clés de la réussite de ce dernier. Toujours selon ces auteurs, il convient donc de prendre en compte ces dimensions complexes lors de la conception de contenus et outils pédagogiques.

Dans le prolongement des mécanismes de motivation, la notion d'engagement occupe une place centrale dans les expériences immersives visant la sensibilisation. En psychologie sociale, l'engagement constitue un levier puissant de transformation des attitudes et des comportements. Il est particulièrement mobilisé dans les dispositifs de persuasion engageante, dans lesquels l'individu, par ses propres actions, s'implique progressivement dans un processus de changement. À cet égard, Girandola et Roussiau (2003) rappellent que l'engagement permet d'apporter une source de modification durable, en ce qu'il lie l'individu à ses propres actes.

La définition de l'engagement, telle que formulée par Beauvois et Joule (1998), cités par Girandola et Roussiau (2003), souligne que l'engagement correspond à l'ensemble des conditions qui font que l'accomplissement d'un acte ne peut être attribué qu'à la personne qui l'a effectué. Ainsi, plus une action semble librement choisie, plus elle engage l'individu et augmente les chances qu'il adapte ses croyances ou comportements en conséquence. Ce principe est central dans les approches de communication engageante et dans les dispositifs de formation visant des transformations profondes.

La RVI constitue un terrain particulièrement favorable à ce type de dynamique. En effet, en plongeant l'utilisateur dans une expérience incarnée et interactive, elle favorise un engagement cognitif et émotionnel élevé. Les travaux d'Allcoat et von Mühlénen (2018) montrent notamment que l'utilisation de la RVI dans le cadre de leçons pédagogiques induit un

niveau d'engagement supérieur à celui observé avec des supports classiques tels que les textes ou les vidéos. L'effet est attribué à l'expérience de présence renforcée et à la stimulation multisensorielle qui renforcent la concentration, l'implication affective et la mémorisation.

Plus largement, des études récentes ont mis en lumière le potentiel de la RVI à favoriser l'engagement prosocial à travers des mises en situation immersives. Par exemple, Herrera et al. (2018) ont démontré que les environnements virtuels favorisent l'adoption de perspectives empathiques, notamment lorsqu'ils sont conçus pour simuler l'expérience subjective de personnes en situation de vulnérabilité. Ces expériences, en facilitant la prise de rôle, induisent une implication émotionnelle forte, condition préalable à un engagement moral durable.

En outre, Slater et Wilbur (1997) posaient déjà les fondements du lien entre présence virtuelle et engagement comportemental, en soulignant que plus un individu se sent « présent » dans un environnement virtuel, plus il est susceptible d'agir comme s'il se trouvait réellement dans cette situation, ce qui renforce la dimension engageante de l'expérience.

Enfin, l'engagement en RVI est aussi corrélé à l'interactivité perçue et au sentiment de contrôle. Makransky et Lilleholt (2018) ont mis en évidence que les environnements virtuels interactifs favorisent un état d'engagement profond, appelé « *flow* », propice à l'apprentissage, à la réflexion personnelle et à la révision des représentations initiales.

Ces éléments montrent que la réalité virtuelle, en tant que média immersif, n'est pas simplement un outil de présentation d'informations, mais un catalyseur de processus psychosociaux complexes où l'engagement agit comme une charnière entre émotion, compréhension et transformation. Cet engagement constitue une étape préalable essentielle à la compréhension par l'autre, sujet de la section suivante, dans laquelle seront explorées plus en détails les conditions d'apparition et d'efficacité de l'empathie induite en contexte immersif.

2.3.4 La compréhension d'autrui - la notion d'empathie en réalité virtuelle

Récemment, Leveau et Camus (2022) ont montré que la RV produit un effet positif sur les intentions comportementales et que les réactions d'ordre émotionnel ont une incidence directe sur la motivation à agir, réagir ou s'abstenir. Selon eux, le sentiment d'incarnation joue un rôle déterminant sur l'induction émotionnelle de l'individu. Gall et al. (2021), en s'inspirant du modèle de Kilteni et al. (2012), ont démontré que l'illusion d'incarnation intensifie le traitement émotionnel de l'environnement virtuel et que l'excitation, la dominance et la valence sont plus intenses en condition de forte incarnation.

A partir du modèle de Davis (1980) et des travaux qui ont suivis, la notion d'empathie, un élément très important en RV, prend une place particulière en RVI. La RV est un outil pour stimuler l'empathie qui semble être assez prometteur comme le montre les travaux récents, tels que « *the displaced* ». Cette expérience en vidéo 360° démontre qu'à travers la RV, l'empathie des utilisateurs était renforcée après immersion. Todo et Binkley (2016), avec le concept d'auto-empathie virtuelle, expliquent le processus qui permet de ressentir par empathie ce qui est vécu par l'avatar de l'utilisateur, renforcé par la présence et l'incarnation. En effet l'empathie est liée à ces concepts clé de la RV. Shin (2017) explique que l'empathie a un effet positif sur l'engagement des utilisateurs de la RV. Il explique ainsi que les utilisateurs ont un contrôle intentionnel et volontaire sur l'histoire qu'ils vivent en RV tout en ajoutant que les processus cognitifs, par lesquels ils font l'expérience de qualité, de présence et de flux, déterminent la manière dont ils vont compatir avec les histoires et ainsi les incarner.

L'aspect ludique est également une notion intéressante pour apporter de l'empathie, qui entre en compte dans une expérience immersive ayant pour but de changer les comportements. Ainsi, dans l'expérience vidéoludique « *That Dragon, Cancer* », Chen et al. (2018) ont mesuré l'impact de l'aspect vidéoludique sur l'empathie. Dans cette expérience, 84 étudiants incarnaient le rôle des parents d'un enfant étant atteint du cancer. Via diverses tâches, ils devaient s'en occuper. Les auteurs ont constaté une augmentation du score d'empathie expliquant que l'aspect vidéoludique était préférable pour comprendre l'empathie (en comparaison avec des procédés classiques). Ventura et al. (2020) rappellent néanmoins d'être prudent avec l'empathie en RV car le manque d'articles dans le domaine et certains résultats non significatifs semblent indiquer que la RV peut parfois échouer à provoquer l'effet escompté dans certains cas de figure.

2.4 Le CAMIL - présence et agentivité au cœur de l'apprentissage en RV

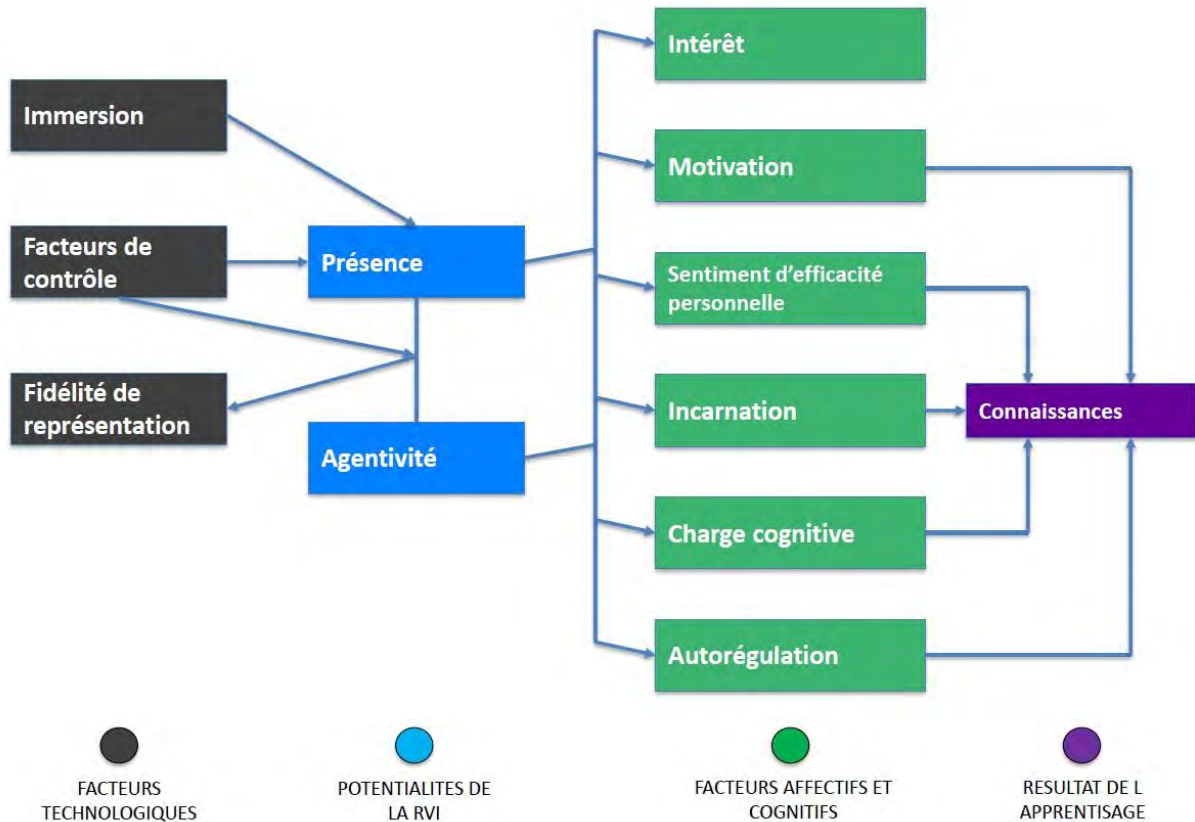
Synthèse des recherches sur les processus d'apprentissage et des travaux abordés, le modèle cognitivo-affectif de l'apprentissage immersif (CAMIL) de Makransky et Petersen (2021) est le modèle clé de ce travail. Ils expliquent que l'efficacité pédagogique ne réside pas intrinsèquement dans la RVI en tant que telle, mais plutôt dans l'adéquation entre l'approche méthodologique déployée (au sein de la séquence d'apprentissage) et l'exploitation optimale des spécificités inhérentes à ce dispositif technologique.

Le CAMIL utilise les facteurs de l'immersion technologique, du contrôle et de la fidélité de représentation vu plus haut pour déterminer quels types d'apprentissages peuvent bénéficier des atouts de la RVI et comment concevoir ces contenus. De ces facteurs découlent des concepts comme la présence et l'agentivité. Ces derniers, qui sont au cœur des médias immersifs, vont faire qu'une méthode pédagogique enrichissant l'apprentissage par présence et agentivité augmentera donc l'apprentissage via la RVI. Les auteurs ajoutent qu'en cas de connaissances préalables, ces concepts renforceront l'apprentissage en RVI. A l'inverse, n'avoir aucune connaissance ne permettra pas d'organiser et d'intégrer les informations à long terme (malgré une expérience ludique positive). Autre concept clé, l'incarnation (avec des agents actifs) stimule la présence sociale et la motivation de l'apprenant et ce, particulièrement en RVI (en opposition avec du *flatscreen* qui ne provoque pas d'incarnation). Adoptant le point de vue selon lequel le média interagit avec la méthode, le CAMIL utilise les théories de la motivation et de l'apprentissage en prédisant une interaction si l'instruction facilite la présence et l'agentivité. Via l'immersion technologique, le CAMIL prédit un apprentissage par le biais de processus affectifs et cognitifs. Six facteurs sont décrits : l'intérêt, la motivation intrinsèque, l'auto-efficacité, l'incarnation, la charge cognitive et l'autorégulation. Ils conduisent à l'acquisition de connaissances factuelles, conceptuelles et procédurales ainsi qu'au transfert de connaissances.

Voici le schéma du CAMIL - traduit et issu de Makransky et Petersen (2021) (fig. 5) :

Figure 5

Vue générale du CAMIL – traduit en français et adapté de l’original de Makransky et Petersen (2021)



Ce schéma présente l’ensemble des relations entre les différents éléments qui composent le CAMIL.

Le modèle s’appuie sur des recherches antérieures comme celles de Makransky et Lilleholt (2018) et Makransky et Petersen (2019) sur le rôle des facteurs affectifs et cognitifs dans l’apprentissage immersif, des études qui ont testé des théories dont celle de Lee et al. (2010) et d’autres anciens modèles comme ceux de Salzman et al. (1999) ; Alavi et Leidner (2001) ; Piccoli et al. (2001) ; Wan et al. (2007), cités par Makransky et Petersen (2021).

Les auteurs expliquent, en citant Witmer et Singer (1998), que les facteurs de contrôle englobent le degré, l’immédiateté et le mode de contrôle. La fidélité de la représentation comprend quant à elle le réalisme et la fluidité de l’affichage, ainsi que la cohérence du comportement de l’objet selon Dalgarno et Lee (2010). Pour information, il existe un contre-courant issu des travaux de Clark (1994) expliquant que la méthode est le seul facteur pertinent

pour l'apprentissage, le média étant un simple vecteur d'enseignement. Ces travaux étaient appuyés par le manque de preuves de l'époque. Les auteurs rappellent également des travaux montrant que la RVI, dans certains cas, n'est pas plus efficace qu'un support vidéo traditionnel. Enfin, les auteurs utilisent le terme « immersion » en se référant à Cummings et Bailenson (2016) qui, eux, utilisent plutôt le terme complet « immersion technologique ». Il s'agit de la mesure de la vivacité qu'un système peut proposer et à quel point celui-ci peut isoler du monde extérieur.

Ci-dessous se trouvent les six facteurs affectifs et cognitifs qui peuvent conduire à des résultats d'apprentissage basés sur la RVI via la présence et l'agentivité :

Intérêt : Cette notion correspond à la construction psychologique définie par Krapp (1999) comme représentant l'articulation entre un individu et un objet d'étude ou un champ disciplinaire particulier. Cette construction se manifeste à travers des dimensions affectives et cognitives. Cette conceptualisation permet de distinguer deux modalités complémentaires : d'une part, l'intérêt individuel, ancré dans les dispositions personnelles du sujet, et d'autre part (dans le contexte spécifique de la RVI), l'intérêt situationnel que Hidi et Renninger (2006) caractérisent comme l'orientation attentionnelle focalisée accompagnée d'une activation affective immédiate, en réponse à des stimuli environnementaux déterminés. L'intérêt situationnel correspond à l'envie d'en savoir plus, un comportement de recherche de connaissances situationnelles à court terme selon Knogler et al. (2015). Ce comportement peut néanmoins déclencher de l'intérêt individuel qui permettra de réintéresser au contenu ultérieurement. La présence permet de le déclencher via les aspects « nouveauté et intensité », mais aussi via l'autonomie de l'expérience, des déclencheurs de l'intérêt situationnel selon Schraw et al. (2001).

Motivation intrinsèque : Deci et Ryan ont proposé en 2015 la théorie de l'autodétermination (SDT). Elle consiste à s'engager dans une activité pour la satisfaction intrinsèque associée à l'activité elle-même, plutôt que pour une conséquence distincte. Trois besoins clés doivent être satisfaits : l'autonomie, la compétence et la relation. RVI et motivation intrinsèque sont corrélées positivement avec la présence et l'agentivité en médiateur. Par ailleurs, la théorie de l'agentivité sociale de Mayer (2014), cité par Makransky et Petersen (2021), explique que les interactions sociales constituent un facteur motivationnel déterminant. Elles déclenchent chez l'apprenant une activation de processus sociocognitifs, engendrant la mobilisation d'activités cognitives orientées vers la construction de sens et ce, à partir du matériel pédagogique proposé. Enfin, selon la théorie de la valeur de contrôle des émotions liées à l'accomplissement (CVTAE) de Pekrun (2006), les activités d'accomplissement, dont la valeur perçue et la contrôlabilité sont

élevées, déclenchent du plaisir. Ainsi, une autonomie élevée entraîne de la motivation intrinsèque.

Auto-efficacité : Schunk et DiBenedetto (2016) expliquent qu'il s'agit de la perception des capacités d'une personne à apprendre ou à réaliser des actions. S'appuyant sur les travaux de Bandura (1977), Makransky et Petersen (2021) expliquent que quatre éléments permettent d'accroître l'efficacité personnelle : la performance, l'expérience vicariante, la persuasion verbale et les états physiologiques. En RVI, la présence et l'agentivité permettent, grâce à l'aspect « réel », de vivre l'activité comme un accomplissement de performance tout en ayant le sentiment de contrôler les actions. La méta analyse de Gegenfurtner et al. (2014) montrent que des niveaux plus élevés d'interaction et de contrôle se traduisent par des estimations plus élevées de l'auto-efficacité. Le feedback immédiat et la fidélité le renforce également selon Makransky et al. (2020b). Enfin, Makransky et Petersen (2019) ont montré, via la modélisation par équations structurelles, un chemin positif entre la présence et l'auto-efficacité passant par la motivation intrinsèque.

Incarnation : L'incarnation est un élément central pour décrire le fait de posséder un avatar et les sensations associées. Présence et agentivité sont liées directement à l'incarnation (l'agentivité étant un des composants). Pour rappel, un des composant de la présence (l'auto-présence) est supposée être associée à des niveaux accrus d'incarnation expérimentés par les apprenants dans la RVI selon Biocca (1997).

Charge cognitive : La charge cognitive est une composante spécifiquement importante de la compréhension du processus d'apprentissage lors de l'apprentissage dans le cadre de la RVI selon Makransky et al. (2019b) ainsi que Meyer et al. (2019). Pour rappel, elle est plus importante en RVI. Elle est liée à la présence et peut résulter d'un haut niveau d'agentivité. Ainsi, Makransky et al. (2020) rappellent que la RVI, malgré son agentivité élevée et ses possibilités d'autonomie, n'est pas optimale et propice dans certains cas à l'apprentissage.

Autorégulation : Boyd et al. (2005) expliquent qu'elle correspond à l'aptitude à contrôler ses actions en inhibant les impulsions, en maintenant son attention et en s'engageant dans des activités nécessaires, y compris lorsque des alternatives plus séduisantes se présentent. Elle génère des pensées qui permettent d'atteindre les objectifs fixés, un processus facilité en RVI grâce à la présence (sociale via les interactions agent/avatar). Néanmoins, une corrélation négative est possible si les contenus n'apportent de réflexion naturelle selon Makransky et al. (2019). Les auteurs mettent également en garde les concepteurs à cause de l'aspect hédonique

potentiel qui engendre des stratégies plus superficielles et donc un apprentissage plus faible malgré des opportunités d'apprentissage autorégulé. L'introduction d'activité de réflexion est importante pour y palier.

Le CAMIL aborde ensuite les relations entre connaissances factuelles, conceptuelles et procédurales, ainsi que le transfert de connaissances d'un côté et la RVI de l'autre, basée sur les travaux de Radianti et al., 2020 et la taxonomie d'Anderson et al., 2001. Le CAMIL rappelle que la RVI n'est pas le média le plus adapté pour transmettre des connaissances factuelles comme dit chez Parong et Mayer (2018). C'est possible, mais il convient de concevoir le contenu pédagogique spécifiquement pour la RVI. A contrario, la RVI offre des conditions optimales pour l'acquisition de connaissances procédurales selon Radianti et al. (2020). Le contrôle et les modifications possibles des environnements virtuels permettent de proposer des conditions pédagogiques surclassant celles de la réalité. Markransky et al. (2019) expliquent enfin que le transfert de connaissances peut être amélioré en RVI de par son aspect procédural ou conceptuel.

Voici comment les six facteurs impactent l'apprentissage :

L'intérêt situationnel : Augmentation de l'attention et de l'engagement, permet d'obtenir une sensation d'apprentissage sans effort, Harackiewicz et al. (2016).

La motivation intrinsèque : Persistance et curiosité entraînant un plaisir issu d'activités motivantes ayant un impact positif, grâce à l'utilisation de stratégies flexibles et créatives, Pekrun (2006). A un impact sur l'ensemble des connaissances.

L'auto-efficacité : Conviction que l'on peut exécuter efficacement les comportements nécessaires pour produire un résultat. Permet la fixation d'objectifs et choix d'activités, ainsi que volonté et persévérance selon Eccles et Wigfield (2002).

L'incarnation : Lien entre les processus moteurs et visuels avec efficacité accrue quand la fidélité est forte. Amélioration des performances de transferts, très efficace pour les connaissances procédurales, mais permet également le factuel et conceptuel, via le renforcement des voies neuronales par renfort des traces mnésiques par ajouts de modalités motrices, Goldin-Meadow (2011). Par exemple : manipuler un objet physique en RVI est mieux pour le comprendre qu'une simple explication de sa manipulation.

La charge cognitive : Facteur avec une influence négative, la charge cognitive doit être réduite au maximum. Les « *bells and whistles* » (détails séduisants) de Moreno et Mayer (2002), ainsi

que la charge cognitive étrangère, Makransky et al. (2020a), doivent être mimées. Il s'agit d'un équilibre entre fidélité de représentation, niveau de détails des environnements virtuels et choix des facteurs pertinents.

L'autorégulation : Technique du résumé via l'apprentissage génératif. Cette théorie de Wittrock (1974) reprise par Fiorella et Mayer (2016), suggère que le processus par lequel une personne génère et transfère la signification des stimuli et des événements (à partir de ses antécédents, de ses attitudes, de ses compétences et de ses expériences) demeure opérant en RVI, malgré les contraintes liées à la charge cognitive et à la présence de distracteurs. Makransky et al. (2020) ont constaté qu'elle améliorerait les connaissances procédurales et le transfert (résumé fait après séance de RVI) mais est inutile avec une séance vidéo. Un constat similaire est fait par Klingenberg et al. (2020) sur les connaissances déclaratives.

Un certain nombre de facteurs externes importants n'apparaissent pas dans le CAMIL mais influencent néanmoins le modèle. Listés par Makransky et Petersen (2021), ils englobent la facilité d'utilisation, les facteurs sociaux et une série de variables individuelles qui sont les suivantes : l'âge, le genre, la cinétose, la mémoire de travail, la personnalité, la prédisposition à l'absorption et la capacité spatiale. Les capacités cognitives propres à chacun sont aussi un facteur à prendre en compte en RVI selon Li et al. (2020), tout comme la préexposition à la technologie de RVI et l'impact du facteur nouveauté (effet « *wow* »). Attention cependant, ces facteurs, qui vont impacter l'apprentissage en RVI, demandent pour certains plus de recherches (comme la personnalité). Pour information, dans le présent travail, les facteurs d'auto-efficacité et d'autorégulation seront mis de côté car l'outil développé ne s'appuie pas sur ces mécanismes actuellement.

2.4.1 Concevoir sous le CAMIL

Les recommandations de conception du CAMIL sont les suivantes : un affichage réaliste, avec un degré de contrôle élevé et immédiat (avec prise en compte des limites dues à la charge cognitive), pour favoriser la présence et l'agentivité mais aussi l'autonomie. Les principes de cohérence et de signalisation doivent également être respectés comme précisé par Mayer et Fiorella (2014). Il convient ensuite de limiter les traitements parasites, soit l'équilibre avec l'hédonisme et de nouveau la charge cognitive, c'est-à-dire surveiller le matériel non pertinent qui pourrait distraire et conduire à des activités uniquement hédoniques. Pour finir, l'ajout d'activités d'autorégulation et de réflexion (explication des contenus via des pairs

avatars, Klingenberg et al., 2020 et résumé des segments, Parong & Mayer, 2018) est essentiel et ne doit pas être mis de côté.

Il est également conseillé d'appliquer certains principes de conceptions pédagogiques comme les principes de modalité de Mayer et Pilegard (2014), la segmentation de Parong et Mayer (2018) et la préformation, Meyer et al. (2019). Il convient ensuite de prendre en compte l'utilisabilité. En cas de facilité d'utilisation faible (ISO 2018), les utilisateurs n'utiliseront pas le dispositif, du fait de défaut de conception qui gêne durant la séance, de saccades dues à un mauvais calcul du ratio performance/affichage ou encore niveau de contrôle insuffisant réduisant l'agentivité.

Pour clôturer cette partie et avant de présenter la conception de l'outil, voici un rappel global des recommandations de Makransky et Petersen (2021). Les auteurs expliquent que pour concevoir efficacement des outils pédagogiques en RVI, il est essentiel de tirer parti de ses spécificités. Concrètement, cela suppose une conception optimisée pour les casques occlusifs, une restitution réaliste et fluide des environnements et une navigation réactive. En intégrant ces principes, il est possible de favoriser l'émergence des sentiments de présence et d'agentivité chez l'apprenant, rendant possibles des expériences formatives proches du réel, parfois inaccessibles ou irréalisables autrement.

3. Synthèse

La réalité virtuelle incarnée (RVI), qui place spécifiquement l'utilisateur en situation d'incarnation constitue une technologie à la fois ancienne, dans ses fondements conceptuels, et moderne dans ses applications, offrant des potentialités multiples. Les principaux atouts de la RVI ont été examinés, tant sur le plan émotionnel, à travers les travaux portant sur l'empathie, facteur clé de la compréhension d'autrui, que sur le plan cognitif, en lien avec l'acquisition de connaissances et la motivation intrinsèque. Le recensement des bénéfices de la réalité virtuelle (RV) établi par Lewis et al. (2021), en cohérence avec les observations d'Arnaldi et al. (2018), a permis d'identifier une série d'avantages non exhaustifs, parmi lesquels certains demeurent sans doute encore inexplorés⁹.

Les limites de la RV ont également été discutées. Le rapport de l'ANSES (2021) souligne notamment l'insuffisance d'études, en particulier longitudinales, malgré l'essor rapide de cette technologie. Les contraintes ergonomiques liées aux casques occlusifs, encore trop volumineux et peu accessibles, ont été mises en évidence. Les effets résiduels ou persistants post-exposition, bien que généralement bénins et transitoires, ont également été mentionnés. D'autres points critiques ont été abordés, tels que les contre-indications relatives à certains troubles mentaux, les questions de cybersécurité et les enjeux éthiques. Un point central a concerné la charge cognitive, à la fois contrainte et opportunité en RVI : une surcharge informationnelle peut nuire à l'apprentissage, tandis qu'un dosage optimal permet de capter et de maintenir l'attention. La question des environnements virtuels distrayants et des interfaces mal conçues a ainsi été évoquée, en lien avec l'utilisabilité globale du dispositif. Cette discussion a été enrichie des recommandations du modèle CAMIL (*Cognitive Affective Model of Immersive Learning*) de Makransky et Petersen (2021), notamment en matière de conception. La problématique de la cinétose, encore aujourd'hui l'un des principaux freins à l'usage de la RV, a fait l'objet d'une attention particulière. Ses manifestations et les stratégies d'atténuation actuelles ont été présentées, tout en soulignant la variabilité intra individuelle des réactions. Des recommandations pratiques pour l'usage de la RVI ont conclu cette section.

Les boucles d'interaction « perception, décision, action » ont été détaillées pour expliciter le traitement des actions de l'utilisateur à travers le dispositif, générant ainsi une

⁹ Voir annexe B.4 : [Le tableau récapitulatif des avantages et inconvénients/limites de la RVI](#) p. 322.

dynamique d'engagement itératif. Plusieurs variantes de ces boucles impliquent différentes interfaces (sensorielles, motrices et sensorimotrices) selon leur rôle dans la transmission et le retour d'information. Le modèle de la « présence-incarnation » proposé par Gorisse et al. (2022) permet de conceptualiser ces deux piliers fondamentaux de la RVI.

L'immersion technologique, entendue comme l'ensemble des caractéristiques techniques (résolution, champ de vision, qualité du *tracking*, etc.), a également été abordée. Des concepts secondaires tels que la transmission de l'empathie en RVI, la conception des avatars et agents (dont la cohérence prime parfois sur le réalisme), ou encore les modalités d'acquisition des connaissances en RV ont été intégrés à l'analyse, en articulation avec les éléments développés dans le chapitre sur le TSA. Ces considérations sont venues nourrir les réflexions autour de la motivation et des compétences sociales, en lien avec les conditions d'apprentissage optimales.

Enfin, les principes fondamentaux de la conception centrée utilisateur (CCU) ont permis d'introduire le modèle CAMIL, qui constitue l'ossature théorique de cette recherche. Le CAMIL postule que l'efficacité pédagogique d'un dispositif immersif repose sur l'exploitation optimale des spécificités du média. Le modèle identifie six facteurs clés favorisant l'apprentissage : l'intérêt, la motivation intrinsèque, l'auto-efficacité, l'incarnation, la charge cognitive et l'autorégulation. Il offre par ailleurs des recommandations de conception issues de la littérature, visant à adapter le contenu pédagogique aux propriétés uniques de la RVI. Trois facteurs généraux (l'immersion, le contrôle utilisateur et la fidélité de la représentation) sont considérés comme centraux dans la génération du sentiment de présence, chacun étant structuré autour de sous-composants spécifiques. Bien que non exhaustif, le CAMIL prend également en compte plusieurs variables modératrices, telles que la mémoire de travail, les prédispositions individuelles ou encore les traits de personnalité. Le modèle insiste enfin sur l'importance de l'utilisabilité du dispositif, tant sur le plan technique que pédagogique, condition sine qua non pour atteindre les objectifs de formation visés.

Chapitre 3 - Problématique et objectifs

L'ambition de cette recherche est de concevoir et de développer un outil pédagogique autonome de compréhension du TSA en RVI puis de tester de manière empirique cet outil. Le but est de tester sa capacité à favoriser la mémorisation et l'assimilation de connaissances factuelles sur le TSA, en vue de soutenir l'acquisition d'un savoir-faire social pertinent à moyen terme, soit une amorce de changement d'attitude. L'objectif est d'aider les enseignants à prendre conscience du fonctionnement des étudiants avec TSA en situation d'apprentissage.

La compréhension du fonctionnement des apprenants avec TSA sans déficiences intellectuelles par les personnes du corps enseignants s'inscrit dans une perspective de changement d'attitude au sens des théories de la psychologie sociale (Fointiat & Barbier, 2015). Il vise une évolution du regard porté par les utilisateurs sur la population des personnes avec TSA. Pour atteindre cet objectif, l'approche retenue repose sur l'utilisation de la RVI, technologie choisie pour ses spécificités en matière d'apprentissage et de transmission de connaissances, qu'elles soient factuelles, conceptuelles, procédurales ou de transfert via des mécanismes d'immersion et d'empathie. Le modèle CAMIL proposé par Makransky et Petersen (2021) constitue un cadre théorique central dans cette approche. Ce modèle postule que la RVI peut améliorer l'apprentissage à condition que le contenu pédagogique soit spécifiquement conçu pour tirer parti des caractéristiques propres au dispositif immersif. Le contenu pédagogique est, à cette fin, conçu de manière à exploiter les spécificités de la RVI selon le modèle CAMIL. Cela est rendu possible grâce à la facilitation de l'apprentissage par des processus affectifs et cognitifs, déclenchés par la présence et l'agentivité propres aux environnements immersifs. Ces processus interagissent pour favoriser l'acquisition de connaissances factuelles, conceptuelles et procédurales, ainsi que leur transfert.

Ce travail s'articule donc autour de deux objectifs principaux :

- a) La conception d'un environnement virtuel de RVI simulant la perception et le traitement cognitif d'un apprenant avec TSA sans déficience intellectuelles (processus de conception avec les principaux critères à respecter pour une certaine efficacité : ateliers itératifs de maquettage puis prototypage et enfin de développement sous Unity 3D et visual studio).
- b) L'examen empirique des effets de cet environnement virtuel de RVI sur des utilisateurs non TSA (les effets de la RVI, et de la manipulation de l'immersion et contrôle, sur les

principales VD : compréhension, changement d'attitude et empathie) afin de favoriser, au sein d'une population cible spécifique, en l'occurrence les enseignants et formateurs, une meilleure compréhension, via la mémorisation des connaissances et l'amorce d'un changement d'attitude, des particularités de fonctionnement d'un apprenant avec TSA sans déficiences intellectuelles. L'objectif était d'évaluer l'efficacité de ce dispositif de RVI dans la transmission de connaissances afin d'obtenir une meilleure compréhension du fonctionnement en cours des étudiants présentant un TSA sans déficiences intellectuelles et une amorce de changement d'attitude envers cette population, auprès du public cible d'enseignants et de formateurs.

L'objectif de conception (a) a été traité sous le modèle CAMIL en suivant un protocole de conception spécifique et rigoureux, utilisant les concepts de conception centrée utilisateur. Cette phase de conception est appuyée par la mise en œuvre d'ateliers de conception composés d'experts du TSA et d'apprenants TSA (Cellule GEM de Rodez) se déroulant par itération sur les deux premières années de thèse. Le dispositif informatique pédagogique complet issu de ces ateliers a été conçu sous Unity 3D. L'objectif était d'exploiter pleinement les possibilités offertes par les différents outils de conception (informatiques et humains) pour favoriser la conception d'un outil performant permettant une meilleure compréhension du TSA auprès d'un public neurotypique composé d'enseignants et de formateurs. Un certain nombre de critères technologiques, sont à prendre en compte pour concevoir un outil de RVI pertinent, tels que l'adjonction de présence et agentivité, la manipulation de l'immersion technologique et du contrôle ou encore les limites dues à la charge cognitive et la cinétose. A l'heure actuelle, il n'existe qu'une seule application en RVI traitant de la perception du TSA par des utilisateurs neurotypiques : « Dans la peau d'une personne autiste » de Lanovaz (2023). Elle est disponible sur le marché (plateforme Meta). Cette application, entièrement modélisée en 3D et accessible gratuitement, a fait l'objet d'un travail de thèse par Koniou (2023).

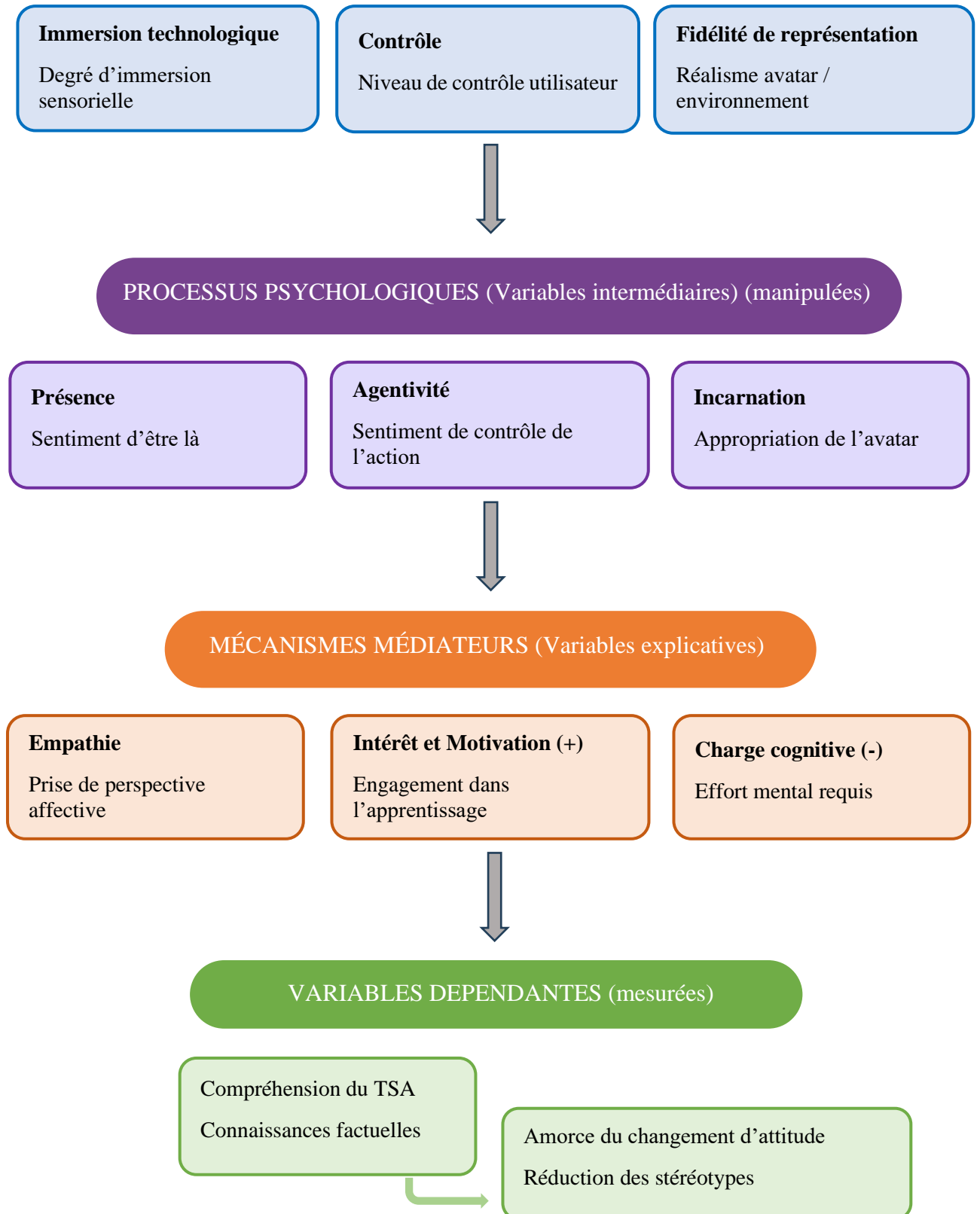
L'objectif empirique (b) était abordé à travers un protocole expérimental afin de tester les effets de l'outil de RVI conçu dans le cadre de cette thèse sur la compréhension et l'acquisition de connaissances sur le fonctionnement de personnes avec TSA conformément aux prédictions du modèle CAMIL qui postule que la présence et l'agentivité induites par la RVI faciliteront l'apprentissage via des processus affectifs et cognitifs, particulièrement en présence de connaissances préalables. Également, les effets sur une amorce de changement d'attitudes vis-à-vis des personnes avec TSA était examiné.

Les objectifs visés étaient d'évaluer les effets de l'outil RVI en comparaison avec d'autres outils ne proposant pas les mêmes caractéristiques que la RVI comme l'absence d'agentivité ou encore l'absence d'environnement dynamique 3D. Il était attendu que l'outil de RVI soutienne :

1. Une meilleure compréhension et mémorisation du fonctionnement cognitif propre aux personnes avec TSA en situation de cours, en s'appuyant sur le modèle EPF (Mottron et al., 2006). L'immersion technologique et l'agentivité offertes par la RVI favorisent l'engagement cognitif et émotionnel de l'utilisateur, renforçant ainsi l'encodage mnésique des informations. La présence ressentie dans l'environnement virtuel active des processus attentionnels soutenus et une expérience incarnée qui facilitent l'ancrage des connaissances en mémoire à long terme. Ces mécanismes, conformes au modèle CAMIL, permettent un apprentissage plus profond et durable que les supports pédagogiques traditionnels.
2. Une réduction des stéréotypes et idées reçues associés au TSA, souvent renforcés par des approches traditionnelles (Willay, 2020) et l'amorçage d'un changement d'attitude sociale, à travers une expérience immersive empathique. Comme le souligne Willay (2020), les outils traditionnels de transmission de connaissances, bien qu'efficaces pour diffuser une information, tendent paradoxalement à renforcer certains stéréotypes négatifs à propos du TSA. Dans cette optique, la présente recherche propose de concevoir un outil de RVI puis d'expérimenter l'usage de la RVI afin d'évaluer sa capacité à limiter ou éviter les biais renforcés par les médias traditionnels. Pour ce faire, cette recherche repose sur une mesure comparative du niveau de stéréotypisation sur le TSA, avant et après exposition à l'outil RVI, ainsi que plus de deux semaines après l'expérience, afin de mesurer la pérennité des effets observés.
3. En outre, il était attendu une validation expérimentale de certaines dimensions du modèle CAMIL, notamment la relation entre présence, agentivité et apprentissage.

Figure 6

Schéma de modélisation des variables et interactions



IMPORTANT : une absence de cinétose est nécessaire pour le bon déroulement d'une expérience RV et la mise en place des relations de cette séquence

Légende du schéma :

En italique : les facteurs affectifs et cognitifs du CAMIL qui influencent l'apprentissage immersif mais qui ne sont mesurés directement. (+) : *affecte positivement*. (-) : *affecte négativement*

Variables indépendantes : Caractéristiques techniques de la RVI manipulées expérimentalement (facteurs technologiques du CAMIL)

Processus psychologiques : États mentaux induits par l'expérience de RVI (potentialités de la RVI du CAMIL)

Mécanismes médiateurs : Facteurs cognitifs et affectifs (du CAMIL) qui expliquent les effets (excepté le facteur empathie, issu d'autres pans de la littérature)

Variables dépendantes : *Outcomes* mesurés pour évaluer l'efficacité de l'intervention (résultat de l'apprentissage du CAMIL)

1. Relations causales principales :

1.1 L'immersion et le contrôle augmentent la présence et l'agentivité

L'immersion sensorielle (via les technologies RV) et le contrôle utilisateur sont les antécédents technologiques de la présence. L'interactivité, l'agentivité et le sens de contrôle sur l'expérience virtuelle contribuent au sentiment de présence. Witmer et Singer (1998) expliquent que la présence résulte de l'interaction entre les facteurs d'immersion (stimuli sensoriels) et les facteurs individuels. Le contrôle est identifié comme un facteur central de la présence. De son côté, Slater (2009) postule que l'agentivité renforce l'illusion de plausibilité (sentiment que les événements se produisent réellement). Enfin, Gutiérrez-Maldonado et al. (2011) expliquent que le contrôle sur l'environnement virtuel influence la présence (et donc l'efficacité d'une intervention de RV).

Mécanisme psychologique : L'immersion technologique active les systèmes perceptuels qui créent l'illusion de « *being there* ». Le contrôle, via le *feedback* sensori-moteur, renforce l'agentivité et la présence.

1.2 La fidélité de représentation facilite l'incarnation de l'avatar avec TSA

La fidélité visuelle et comportementale de l'avatar influence l'appropriation corporelle virtuelle. Plus l'avatar est réaliste (Ce point est pondéré par l'utilisation du *low poly* et la caricature des représentations émotionnelles afin d'éviter l'effet Vallée de l'étrange), plus l'utilisateur peut s'y projeter et l'incarner. Avec leurs travaux sur le *Sense of Embodiment* (SoE ou incarnation), Kilteni et al. (2012) expliquent que la fidélité visuelle-motrice influence l'incarnation. Peck et al. (2013) suggèrent que la fidélité de l'avatar influence l'identification et les processus de changement d'attitude. Ainsi, plus l'avatar est détaillé, plus l'incarnation est forte (toujours avec la limite de la Vallée de l'étrange). Enfin, Gonzalez-Franco et Peck (2018) abordent quant à eux le fait que la ressemblance visuelle et la synchronisation comportementale sont cruciales.

Mécanisme psychologique : Le réalisme visuel permet la reconnaissance de patterns humains ce qui facilite la projection de soi. Enfin, la cohérence comportementale, via le *feedback* visuo-moteur, renforce l'appropriation corporelle.

1.3 L'incarnation et la présence génèrent de l'empathie envers les personnes avec TSA

L'incarnation d'un avatar représentant une personne avec TSA active les mécanismes de prise de perspective, un processus cognitif central de l'empathie. Avec l'effet Proteus, Yee et Bailenson (2007) montrent que l'incarnation d'avatars influence les attitudes et comportements. Selon eux, l'internalisation des caractéristiques de l'avatar permet un changement de soi. Les travaux de Herrera et al. (2018) expliquent que l'incarnation d'avatars via les illusions de propriété corporelle permet d'apprendre l'empathie (mécanisme de « *perspective-taking* » renforcé par l'incarnation). Maister et al. (2015) montrent que l'incarnation d'avatars différents influence les attitudes implicites et que l'incarnation active les réseaux neuronaux de l'empathie. Enfin, Banakou et al. (2013) expliquent que l'incarnation influence la cognition et la perception (base neurologique : activation des aires somato-sensorielles et pariétales).

Mécanisme psychologique : L'incarnation d'une personne avec TSA via la simulation incarnée permet une meilleure compréhension des expériences sensorielles du TSA. La présence permet un engagement émotionnel qui va activer les mécanismes empathiques.

1.4 L'empathie et la motivation médiatisent l'effet sur la compréhension

L'empathie facilite l'attention aux détails pertinents et la motivation soutient l'effort cognitif nécessaire à l'apprentissage conceptuel. Les travaux en psychologie sociale de Batson et al. (1997) montrent que l'empathie induite améliore l'apprentissage sur les groupes stigmatisés et la motivation à comprendre. Au niveau de la motivation intrinsèque, Deci et Ryan (2000) expliquent qu'elle facilite l'apprentissage approfondi et que l'engagement émotionnel soutient la persévérance cognitive. De leur côté, Decety et Jackson (2004) décrivent l'architecture fonctionnelle de l'empathie, processus qui, selon Batson et al. (1997), peut motiver l'acquisition de connaissances sur autrui, l'émotion empathique activant la curiosité qui facilite l'apprentissage (Hidi & Renninger, 2006).

Mécanisme psychologique : L'empathie active une attention focalisée sur les expériences TSA ce qui permet un encodage privilégié. Ainsi, la motivation, qui demande un effort cognitif soutenu, permet une consolidation des apprentissages.

1.5 La compréhension améliore les attitudes et réduit les stéréotypes

L'acquisition de connaissances factuelles sur les TSA remplace les stéréotypes par des représentations plus nuancées et précises. Avec sa théorie du contact, Allport (1954) explique que la connaissance de l'autre réduit les préjugés. Il ajoute des conditions qui sont une connaissance approfondie, une égalité de statut et un but commun. Dans leur méta analyse de 2008, Pettigrew et Tropp montrent que la connaissance médiatise la réduction des préjugés. Selon eux, la connaissance permet une réduction de l'anxiété ce qui améliore les attitudes. Crisp et Turner (2009) montrent que l'apprentissage conceptuel sur les groupes minoritaires améliore les attitudes (plus efficace que le simple contact). Enfin, Campbell et al. (2019) expliquent que l'éducation sur les TSA réduit significativement les stéréotypes (médiation par l'augmentation des connaissances factuelles).

Mécanisme psychologique : Les connaissances factuelles remplacent progressivement les stéréotypes ce qui réduit les biais cognitifs. Ainsi, la compréhension nuance les représentations et permet une attitude plus positive envers la population cible.

1.6 Une charge cognitive trop élevée peut diminuer l'efficacité de l'intervention

Selon la théorie de la charge cognitive, les ressources mentales limitées peuvent être saturées par la complexité technique, réduisant les ressources disponibles pour l'apprentissage empathique. La théorie de la charge cognitive de Sweller (1988) montre que la surcharge cognitive entrave l'apprentissage et la mémorisation. Plus récemment, Makransky et Lilleholt (2018) ont montré que la complexité de la RV peut créer une surcharge cognitive. Cela a un impact négatif sur l'apprentissage et l'engagement émotionnel. Paas et Van Merriënboer (2020) ont montré que, en RV, la charge extrinsèque (interface, navigation) peut limiter l'apprentissage. Selon eux, il est donc nécessaire d'optimiser l'ergonomie cognitive. Dans leur revue de 2020, Radianti et al. suggèrent que la cinétose et la complexité d'interface réduisent l'efficacité pédagogique. Selon eux, le confort et la sécurité sont essentiels (la cinétose peut ruiner le sentiment de présence).

Mécanisme psychologique : Une surcharge cognitive limite les ressources disponibles ce qui entraîne une réduction de l'attention empathique. Avec une complexité technique inutile, l'effort cognitif est détourné ce qui entraîne un engagement réduit dans l'apprentissage.

2. Synthèse des mécanismes théoriques

En plus du CAMIL, ce travail s'appuie en partie sur plusieurs théories convergentes :

1. **Théorie de la Présence** (Witmer & Singer, 1998).
2. **Théorie de l'Embodiment** (Kilteni et al., 2012).
3. **Théorie de la Charge Cognitive** (Sweller, 1988).

Ces théories forment un cadre cohérent expliquant comment la technologie RV peut induire des changements d'attitude via des mécanismes psychologiques bien établis. Sur cette base, plusieurs modalités expérimentales sont ainsi mises en place :

- **Modalité à forte immersion technologique** : Expérience RVI via un casque occlusif, avec agentivité et présence accrue (contrôle direct de mains virtuelles).
- **Modalité à moyenne immersion technologique** : Logiciel interactif en *flatscreen*, avec un contenu identique, sur écran plat, mais permettant une certaine agentivité (contrôle via clavier /souris, avec présence d'un avatar non incarné à l'écran, en vue subjective).
- **Modalité à faible immersion technologique** : Présentation PowerPoint, PowerPoint Audio et vidéo sur écran plat avec défilement de diapositives / vidéo contenant les mêmes informations factuelles, mais sans agentivité ni présence (aucune interaction ou représentation de l'utilisateur).

3. Questions de recherches et hypothèses

Ce projet de recherche vise à concevoir un dispositif de RVI sous un processus de conception centrée utilisateur via des outils de conception puis de tester l'efficacité de ce dispositif de manière empirique. Voici les questions de recherche :

1. Dans quelle mesure le niveau d'immersion, d'agentivité et de présence influence-t-il l'utilisateur dans une tâche de sensibilisation au TSA ?
2. La RVI immersive améliore-t-elle la compréhension factuelle du TSA par rapport à des modalités non immersives ?
3. L'immersion permet-elle une rétention des connaissances factuelles à moyen terme ?
4. La RVI immersive induit-elle un accroissement mesurable de l'empathie ?
5. Peut-elle contribuer à réduire les stéréotypes et changer les attitudes liées au TSA ?
6. Les effets de la RVI sur les stéréotypes et les attitudes sont-ils durables ?

La RVI est comparée à d'autres outils qui n'offrent pas les mêmes caractéristiques d'immersion, d'interaction et de dynamique afin de mesurer les effets différentiels sur la compréhension, la mémoire, l'empathie et la stéréotypisation. Voici les hypothèses :

- 1. Présence et agentivité :** L'exposition en RVI immersive produira un niveau significativement plus élevé de présence et d'agentivité que les modalités non immersives. Le réalisme perceptif, l'incarnation corporelle et le contrôle direct de l'avatar sont connus pour renforcer l'engagement et la sensation de présence (Makransky & Petersen, 2021).
- 2. Rétention des connaissances :** La rétention des connaissances factuelles sera supérieure dans la condition RVI immersive. L'expérience immersive augmente l'encodage en mémoire épisodique, ce qui favorise le rappel différé (Makransky et al., 2019).
- 3. Empathie :** La condition RVI immersive induira un score d'empathie plus élevé que les autres modalités. L'incarnation virtuelle facilite la *perspective-taking* et l'activation de réponses affectives pro-sociales (Herrera et al., 2018).
- 4. Stéréotypisation :** La RVI immersive réduira davantage les stéréotypes liés au TSA que les modalités non immersives, avec un effet maintenu à moyen terme. Les

expériences immersives vécues à la première personne peuvent modifier durablement certaines représentations sociales grâce à l'empathie et la projection « à la place d'autrui » qui permet de prendre en compte un point de vue auquel qu'il est difficile de se représenter autrement (Banakou et al., 2016).

5. Changement d'attitude : L'amorce de changement d'attitude sociale sera plus marquée et plus durable dans la condition immersive. La combinaison de présence, d'engagement émotionnel et de vécu expérientiel favorise la reconfiguration cognitive et affective des attitudes (Slater & Sanchez-Vives, 2016).

La suite de ce travail s'organise comme suit :

Un chapitre est consacré au travail de conception qui a consisté à conduire des ateliers itératifs puis de concevoir et de développer en laboratoire l'outil de RVI de sensibilisation au TSA sous le model CAMIL.

Un autre chapitre présente les deux études empiriques conduites pour tester les hypothèses relatives aux effets de l'outil RVI.

Partie 2 - Ingénierie d'un outil RVI pour la sensibilisation au TSA

Cette seconde partie présente le processus d'ingénierie pédagogique et technologique ayant conduit à la conception et au développement de l'outil de RVI dédié à la sensibilisation au TSA. La démarche adoptée s'inscrit dans une approche itérative associant réflexion pédagogique, contraintes techniques et retours d'expérience utilisateurs.

Elle s'articule autour de trois temps structurants. Dans un premier temps, les motivations guidant la conception de l'outil, ainsi que la méthodologie d'ateliers itératifs mobilisée pour élaborer le scénario pédagogique seront abordées. Cette phase exploratoire a permis de définir les objectifs d'apprentissage et les mécanismes d'immersion à privilégier. Dans un second temps, la conception de la maquette sera détaillée à travers ses trois scènes principales : le tutoriel d'initiation, l'expérience immersive centrale et la résolution du problème. Chacune de ces séquences répond à des intentions pédagogiques spécifiques tout en s'articulant dans une progression cohérente. Enfin, le troisième temps porte sur le développement du prototype puis du produit minimum viable (MVP), abordant les choix technologiques, l'architecture fonctionnelle, la conception des environnements et personnages ainsi que les dispositifs d'interaction. Les évolutions successives de l'outil au cours de la thèse témoignent d'un processus de raffinement continu guidé par les retours des expérimentations.

Cette démarche de conception incarnée illustre les défis et les choix inhérents au développement d'un dispositif de RVI à visée pédagogique, articulant exigences scientifiques, contraintes techniques et impératifs d'utilisabilité.

Chapitre 1 - Motivations

Pour ce projet, un outil spécifique a été imaginé : une simulation en réalité virtuelle d'une situation d'apprentissage fictive durant laquelle l'utilisateur se retrouve à la place d'un élève dans une salle de classe durant un cours fictif. L'utilisateur incarne une personne avec TSA. Cette idée se base sur le concept d'incarnation de Banakou et al. (2008) « *Virtually being Einstein* ». Pendant la durée de l'exposition, l'utilisateur perçoit les stimuli du monde extérieur à la manière d'une personne avec TSA (amplification de certains stimuli, focalisation accrue sur certains éléments via des mini jeux de type *serious game* ou encore du bruit blanc).

L'outil a été maqueté en conception centrée utilisateur (CCU) suivant un scénario pédagogique spécifique puis prototypé et développé sous le moteur Unity3D afin d'être proposé en application de RV utilisable sur un HMD.

Dans un premier temps, les ateliers de *design thinking* organisés entre novembre 2022 et octobre 2024 dans les locaux de l'association GEM La Bulle Rodez ont permis de poser les bases des idées de ce projet. Toutes les phases de conception ont été menées en collaboration directe avec les membres de l'association lors de ces ateliers. Les nombreuses itérations ont permis de faire évoluer progressivement le projet jusqu'à l'outil actuellement développé. Il est essentiel de souligner que les étapes de conception détaillées dans les sections suivantes, à savoir la phase de maquetage puis celle de prototypage et enfin le développement du MVP ont toutes été soumises, à un moment donné, à une validation en atelier.

La phase de maquetage a ensuite été amorcée. À chaque étape, les éléments conceptuels en cours de développement ont été présentés aux membres lors d'ateliers de vérification des contenus. Une fois validés, les contenus ont été intégrés à la feuille de route du projet. Ce processus a été reconduit au cours de la phase de développement afin de garantir que chaque étape scénaristique soit conforme à la réalité vécue par les personnes concernées. Ainsi, chaque étape a fait l'objet d'une validation explicite par les membres du GEM.

À l'issue du développement du prototype, l'outil finalisé a été de nouveau soumis aux membres du GEM pour validation, consolidant ainsi une démarche centrée sur la co-conception, le retour d'expérience, et l'adéquation au vécu des utilisateurs finaux.

Résumé des phases principales de l'outil issu du concept initial :

- Un tutoriel immersif : Faire rentrer l'utilisateur dans la VR en lui fournissant les bases d'utilisabilité pour qu'il puisse vivre la séance de façon optimale.
- Une phase d'exposition : Une fois l'utilisateur préparé, il rejoint la salle de classe virtuelle et se familiarise quelque instant avec son environnement avant que l'exposition ne commence.
- Une expérience immersive : Après la phase de familiarisation, l'activité commence. Il s'agit d'un enchaînement de tâches générales qui vont immerger l'utilisateur dans la peau d'une personne avec TSA en lui faisant vivre des expériences proches de celles qu'une personne avec TSA peut vivre et ressentir.
- Une phase de sortie : L'activité se clôture avec une tâche de résolution avec explication et retour à un état « initial ». L'utilisateur est ensuite invité à enlever le casque pour reprendre le cours de sa vie. Cette phase permet à l'utilisateur de retourner dans le monde réel lors d'une transition « douce ».

Les parties suivantes proposent la trame de conception lors des ateliers itératifs, du développement de la maquette, puis du prototype et enfin du MVP (étape 2 du développement).

1. Les ateliers itératifs et le scénario pédagogique

En réponse à la situation particulière de la reconnaissance sociale du TSA et aux conséquences que cela implique dans la formation et l'accompagnement des personnes concernées, une initiative a été lancée pour constituer un groupe de travail composé de personnes avec TSA. Ce groupe de volontaires, issu de l'association GEM La Bulle (parrainée par les PEP 12 de Rodez), avait pour objectif principal de produire un socle de témoignages permettant d'élaborer une base conceptuelle pertinente, centrée sur l'environnement de la salle de classe.

À partir de ce socle, des ateliers itératifs ont été mis en place avec les volontaires du GEM (des personnes avec TSA sans déficience intellectuelle) en vue de générer les premières idées susceptibles de représenter au mieux leur réalité en contexte d'apprentissage. Plusieurs situations concrètes ont émergé de ces échanges, servant de base à la construction des premières tâches de l'outil. Parmi les éléments directement issus de ces ateliers figurent notamment l'idée d'une énigme centrale, l'intégration d'interactions sociales, la présence d'indices dissimulés ainsi que la séquence d'*aftermath*. Ces éléments correspondent à des demandes explicites

formulées par les membres du GEM concernant leur propre vécu avec le TSA et les modalités jugées les plus adéquates pour le faire comprendre à autrui. Voici le scénario pédagogique issue de ces itérations : conception d'un outil de RVI pour la compréhension du TSA (sans déficience intellectuelle)

1.1 Cadre et finalités

- Public cible : Enseignants et formateurs (université, enseignement secondaire, formation professionnelle, éducation spécialisée).
- Finalité : favoriser une meilleure compréhension des particularités cognitives et sociales des personnes avec TSA, afin de soutenir l'adaptation des pratiques pédagogiques et éducatives.
- Cadre théorique : modèle CAMIL pour l'exploitation optimale de la RVI ; modèle EPF pour la description des particularités cognitives du TSA ; ancrage en psychologie sociale (changement d'attitude et réduction des stéréotypes).

1.2 Objectifs pédagogiques généraux

À l'issue de l'expérience, les participants devront être capables de :

1. Identifier les principales particularités cognitives liées au TSA sans déficience intellectuelle (ex. hyperfocalisation perceptive, difficultés de traitement social implicite, etc.).
2. Comprendre les conséquences de ces particularités dans un contexte scolaire (ex. interprétation littérale des consignes, rigidité cognitive).
3. Reconnaître des situations pédagogiques à risque (ex. surcharge sensorielle en classe) et envisager des stratégies d'adaptation.
4. Savoir dans quel contexte appliquer ces stratégies pour soutenir la réussite et l'inclusion d'un élève TSA.
5. Développer une posture empathique en percevant la situation « depuis la perspective » d'un élève TSA, réduisant ainsi les stéréotypes et favorisant l'amorce d'un changement d'attitude.

La structure pédagogique (progression en 3 actes immersifs)

L'utilisateur est placé dans une salle de classe virtuelle et se voit attribuer l'avatar d'un étudiant avec TSA. Après une brève introduction guidée (voix off), sa perception est modifiée

afin de refléter certaines caractéristiques de l'expérience du TSA : hypersensibilités sensorielles (intensité des sons et couleurs), surcharge cognitive lors de la réception d'instructions et difficultés d'interprétation des interactions sociales.

Acte 1 - Introduction et contextualisation (1 minute)

Phase d'introduction (immersion et tutoriel) : L'utilisateur apprend à manipuler l'environnement en RV. Il est ensuite « incarné » dans l'avatar étudiant TSA et perçoit la salle de classe de manière altérée (sons amplifiés, couleurs saturées, ambiance déroutante).

- Avant immersion dans la salle de classe, l'utilisateur suit un tutoriel d'assimilation des bases de la RV pour le familiariser avec l'environnement virtuel et réduire la cinétose.
- L'utilisateur est déplacé dans une salle de classe virtuelle afin d'incarner un avatar.
- Une voix off présente les objectifs pédagogiques et explique que l'expérience se fera « à la première personne », dans la peau d'un élève TSA.
- L'environnement s'altère et l'utilisateur le perçoit comme avec un TSA :
 - Surcharge sensorielle : amplification des bruits ambiants, scintillements lumineux, voix multiples qui se superposent.
- **But** : créer un cadre de sécurité cognitive, clarifier le rôle, préparer l'utilisateur à l'immersion.

Acte 2 - Expérience immersive (10-15 minutes)

- L'utilisateur perçoit toujours l'environnement comme un élève TSA :
 - Surcharge sensorielle.
 - Difficultés sociales : difficultés lors d'échanges avec des pairs/enseignants, perception des visages moins expressive (atténuation des indices sociaux).
 - Rigidité cognitive : perturbation provoquée par un changement imprévu.
- Interactivité :
 - L'utilisateur doit accomplir des tâches pédagogiques spécifiques.
 - Le système génère des difficultés spécifiques au TSA (distracteurs sensoriels, surcharge cognitive, ambiguïtés sociales).
- **But** : provoquer une compréhension expérientielle et affective des difficultés rencontrées par un élève TSA.

Déroulement pédagogique détaillé :

1. **Présentation de la tâche cognitive (l'énigme au tableau) :** L'enseignant explique une énigme de labyrinthe. Pour l'avatar étudiant TSA, la consigne devient complexe, car la surcharge perceptive et l'activation d'une « carte mentale » envahissante brouillent la compréhension.
2. **Première tâche - L'horloge comme repère :** L'utilisateur doit focaliser son attention sur l'horloge. En observant ses symboles et le passage de l'aiguille, un premier indice de résolution de l'énigme est intégré à la carte mentale.
3. **Phase de stress social (interaction avec l'enseignant) :** L'enseignant s'approche et demande si l'étudiant va bien. La voix et l'expression faciale deviennent incohérentes, renforçant la confusion. La solution consiste à détourner le regard pour rétablir la compréhension verbale. Cette étape illustre la stratégie d'évitement visuel souvent adoptée par les personnes avec TSA.
4. **Seconde tâche - Le tableau de consignes :** Grâce à une régulation sensorielle (casque audio + bruit blanc), le stress diminue. L'utilisateur peut alors détecter un indice sur le tableau, renforçant la carte mentale.
5. **Troisième tâche - Les plantes (routine et persévérance) :** L'utilisateur doit maintenir son attention sur une séquence de plantes pour compléter un indice supplémentaire malgré les interruptions de l'enseignant. Cela illustre la difficulté à revenir dans une routine interrompue et la nécessité de persévérance malgré les aléas sociaux.
6. **Résolution de l'énigme :** L'utilisateur est invité par l'enseignant à venir résoudre l'énigme. Avec l'aide de la carte mentale, l'utilisateur trace la solution au tableau. Il expérimente le mélange entre surcharge cognitive et stratégie de compensation.
7. **Phase finale - L'exposé oral :** L'enseignant félicite l'étudiant et lui demande d'expliquer sa démarche à ses pairs. Cette situation sociale réactive le stress (pulsations, bruit de fond, murmures), permettant à l'utilisateur de ressentir la difficulté d'un contexte d'exposition sociale.

Acte 3 - Débriefing réflexif guidé (2 minutes)

Conclusion pédagogique : La simulation se termine par une voix off rappelant les principaux phénomènes vécus (surcharge sensorielle, stratégies compensatoires, gestion du stress social). L'utilisateur retire son casque après avoir acquis une expérience incarnée de certains aspects du TSA.

- Retour à la perspective d'un observateur « extérieur ».

- Analyse des difficultés vécues : messages explicatifs présentant les liens avec le modèle EPF et les caractéristiques cognitives du TSA.
- Mise en avant de stratégies pédagogiques concrètes (ex. réduire les distracteurs sensoriels, prévoir des routines stables).

Sur la base de ce scénario pédagogique, une maquette est ensuite conçue et développée.

Chapitre 2 - La conception de la maquette

La maquette est la première étape de conception concrète du dispositif. C'est la première mise en forme des idées issues des ateliers de *design thinking* et de mise en relation avec la littérature. On y décrit l'expérience utilisateur telle qu'elle est pensée avant de partir en production. Chaque partie technique de la maquette fait référence à une des futures scènes (qui seront développées sous Unity - 3 scènes distinctes - partie prototype). Cette partie reprend donc la trame complète du projet, incluant les 3 scènes, qui seront détaillées dans les parties suivantes :

- a) Le tutoriel d'acquisition des fondamentaux d'utilisabilité en RV
- b) L'expérience immersive (incluant phase exposition + tâches générales)
- c) La résolution du problème (incluant tâche de résolution + sortie)

Il est important de noter que cette maquette ne représente pas la forme définitive du logiciel mais les premières itérations qui ont évolué au cours du développement (notamment lors des tests en CCU qui ont permis d'affiner les contenus). Le contenu de la maquette ne correspond pas au contenu du logiciel et cette version est présentée pour comprendre le processus de conception dans son ensemble.

1. Scène 1 - Le tutoriel

Un tutoriel d'acquisition des fondamentaux d'utilisabilité ainsi que de réduction de cinétose a été conçu au préalable. La nécessité d'un tutoriel à destination des novices avant de les plonger dans un environnement virtuel est fondamentale. En effet, ce tutoriel vise à améliorer l'expérience initiale des utilisateurs, souvent source de rejet ou de cinétose, en particulier chez les novices. En s'appuyant sur les recommandations de la littérature, il propose une introduction progressive dans un environnement neutre et dépourvu de profondeur visuelle, inspiré des standards de l'industrie. Trois tâches fondamentales (contrôle des manettes, interaction avec objets, exploration visuelle à 360°) permettent de développer la présence et l'agentivité. Ce dispositif facilite l'adaptation cognitive et sensorielle à la RV, en minimisant les effets indésirables tout en préparant l'utilisateur à l'expérience immersive principale. Une transition visuelle discrète permet enfin de basculer sans heurts vers la scène suivante.

Le principe de base de ce tutoriel fait suite à une problématique récurrente en RV : celle de l'immersion des nouveaux utilisateurs. Si des structures telle que Meta, ou encore certains titres vidéoludique (« *Lies Beneath* » - où l'environnement, blanc au départ, se crée à la manière d'un dessin en ajoutant progressivement quelques lignes lors de la progression de l'utilisateur) ont fait le choix d'immerger de manière douce leurs utilisateurs afin de réduire les risques de rejet, ce n'est pas encore le cas de toutes les expériences de RV. Les salles de jeu et leur système qui rend malade beaucoup d'utilisateurs, telles que les expériences types montagne russe, issue du rapport de l'Anses (2021) sont de bons exemples relatifs aux effets sanitaires liés à une exposition aux technologies de RV. Cette problématique est aujourd'hui essentielle car de nombreux utilisateurs se plaignent de leur première expérience de RV suite à une mauvaise immersion initiale. Ainsi, selon le rapport de l'Anses (2021), 30 à 50 % des utilisateurs seraient sujet à la cinétose.

Le but de ce tutoriel : accueillir l'utilisateur en RV dans un environnement le moins anxiogène et provoquant le moins de cinétose possible afin de la préparer à l'environnement de RV dans les conditions optimales préconisées par la littérature (Arnaldi et al., 2018). Le leurre cognitif démarrant dès la première seconde d'utilisation du casque occlusif, il convient d'immerger l'utilisateur le plus efficacement possible.

2. Scène 2 - l'expérience immersive

Pour rappel, la finalité est de favoriser une meilleure compréhension des particularités cognitives et sociales des personnes avec TSA, afin de soutenir l'adaptation des pratiques pédagogiques et éducatives.

2.1 La phase de présentation de l'environnement

La première étape est une présentation **Figure 7**
de la salle de classe en vue externe pour

familiariser l'utilisateur (fig. 7). Une fois la porte située à la fin du tutoriel franchie, l'environnement se transforme en une salle de classe simplifiée (moins d'informations à traiter).

La salle de classe sans avatars ni agents



Descriptif de la tâche : L'utilisateur apparaît dans cette salle au troisième rang, sur le côté droit du bureau du milieu. Pour le moment, il n'a pas encore d'avatar le représentant dans l'environnement. Cette phase va durer 20 secondes. L'objectif est que l'utilisateur se familiarise avec l'environnement. La logique est la même que pour le tutoriel, accueillir l'utilisateur dans un environnement peu anxiogène qui ne provoquera pas de cinétose et qui ne créera pas de rupture dans le leurre cognitif. Cette tâche n'est pas en lien direct avec la compréhension du TSA mais avec l'utilisabilité de l'outil.

Choix de conception : L'objectif est d'alléger la charge cognitive de l'utilisateur. En 1993 déjà, Hedberg et al. conseillaient aux développeurs de faire attention à la charge cognitive lors du développement d'interfaces interactives, ainsi ne sont présent que les éléments minimums nécessaires à la cohérence de l'environnement. Tous les éléments sont à l'échelle 1 : 1.

La vue externe (sans corps virtuel) permet d'assister en tant que spectateur à la scène, adoptant ainsi une posture d'observateur détaché. Elle induit une forme d'immersion technologique distincte de celle permise par la vue en première personne, en facilitant une perception globale de l'environnement grâce à une meilleure conscience spatiale comme chez Gorisse et al. (2017). Cette configuration permet à l'utilisateur de se concentrer davantage sur

les éléments à découvrir dans l'environnement, plutôt que sur sa propre présence virtuelle. Ce recentrage s'explique par l'articulation entre facteurs endogènes (tels que l'état émotionnel, l'attention ou les attentes individuelles) et facteurs exogènes (comme la perspective adoptée ou le niveau de contrôle exercé sur l'environnement), qui influencent conjointement le positionnement cognitif et perceptif de l'utilisateur au sein de la simulation immersive (Slater & Wilbur, 1997 ; Rothbaum & Hodges, 1999).

Une voix *off* guide également l'utilisateur en lui expliquant la situation : il se trouve dans une salle de classe et il va vivre l'expérience de vie commune d'un étudiant avec TSA dans une situation d'apprentissage. Pour cela, il va intégrer le corps virtuel de l'étudiant situé à côté de lui. La transition se fera dans un fondu noir (pour ne pas provoquer de mouvement et réduire les risques de cinétose, en utilisant le principe de *l'eyelide*, le clignement des yeux). Cette voix *off*, sous le principe du CAMIL, va donner les premières informations nécessaires au bon déroulement de l'expérience. Le reste du matériel audio de cette partie se compose d'une bande son d'ambiance globale. Une fois cette phase d'immersion terminée, l'utilisateur incarne l'étudiant avec TSA et la phase suivante commence.

2.2 L'expérience immersive TSA

2.2.1 Principe

Cette phase est la plus longue de l'immersion. D'une durée d'environ 10 à 15 minutes, elle est composée de plusieurs étapes et doit amener l'utilisateur à sentir qu'il incarne une autre personne que lui-même. Une fois familiarisé avec l'environnement virtuel et les agents, l'utilisateur est amené dans le corps qu'il doit incarner. C'est à partir de ce moment que l'immersion commence réellement.

Le principe est de faire incarner à l'utilisateur un étudiant avec TSA pour ressentir ce qu'il vit dans une situation d'apprentissage. Ce principe d'incarnation est aujourd'hui utilisé par quelques applications (RV sous casque occlusif, mais également RV sous smartphone, laptop, etc.). La RVI est une des solutions que l'on retrouve dans la littérature comme système de prise de conscience ou de compréhension d'une situation par un individu habituellement non concerné par la situation. Les travaux de Schutte et Stilinovi (2017) avec leur expérience de RVI (*the displaced* - incarner une réfugiée de guerre) montrent que l'empathie et l'engagement sont supérieurs en condition de RV. Banakou et al. (2018) avec leur expérience « *Virtually being Einstein* » ont également montré qu'incarner quelqu'un d'autre avait une influence

positive sur les performances de celui qui incarne. Ce principe a été repris dans d'autres expériences comme les travaux de Dyer et al. (2018) dans un projet de RVI basé sur l'empathie (la dégénérescence maculaire et la perte d'audition, vivre à la place d'une personne âgée), ce qui a pour effet d'augmenter l'empathie des soignants envers les personnes âgées. Il en existe d'autres, certaines plus vidéoludiques, qui montrent que ce champ qui est actuellement en pleine exploration n'est pas nouveau comme cekke de Peck et al. (2013) « *Putting yourself in the skin of a black avatar reduces implicit racial bias* », une expérience en RV pour vivre à la place d'une personne d'ethnie différente.

Un des éléments essentiels du dispositif est le recours à des pauses pédagogiques durant l'immersion. Chaque phase sera entrecoupée d'explications (hors *gameplay*) pour permettre à l'utilisateur de comprendre et d'acquérir les connaissances présentées. Les travaux de Parong et Mayer (2018) ont montré que faire une pause durant une séance de RVI et de demander aux utilisateurs de faire un résumé permettait de pallier aux lacunes de la RV au niveau de l'acquisition des connaissances factuelles sans rompre l'immersion. Via le principe de cohérence, la segmentation et les stratégies d'apprentissage génératives, les auteurs ont trouvé une solution pour permettre aux utilisateurs d'apprendre en RV aussi bien qu'avec un support plus traditionnel. Pour rappel, en RV, les utilisateurs étaient sous l'effet « *wow* » donc moins concentré sur l'acquisition des connaissances, ce qui faisait qu'ils performaient moins bien à la restitution de connaissances que sur d'autres supports. Cette solution permet d'acquérir les connaissances dont les utilisateurs auront besoin pour mieux comprendre les personnes avec TSA.

Le matériel audio¹⁰ comprend la bande son « ambiance salle de classe », déjà utilisée et qui continue jusqu'à la phase de reprise de contrôle de l'environnement. Le reste du matériel audio concerne la voix de l'enseignant lorsque ce dernier va s'adresser aux étudiants et à l'utilisateur. La voix de l'enseignant est associée au genre de l'enseignant (féminin).

La trame scénaristique qui permet de mettre en place les tâches est issue du scénario « conséquences du comportement de l'utilisateur ». C'est un déroulé qui commence avec le premier comportement de type TSA de l'utilisateur, qui est la première réaction à la demande initiale de l'enseignant (la tâche initiale et la création de la carte mentale). Ce type de comportement est issu des entretiens avec les membres du collectif GEM. Il représente une des

¹⁰ Voir annexe C.2.2 : [Matériaux audio : Voix off et ambiances sonores](#) p. 336.

réactions possibles à une confrontation directe avec l'environnement telle que le vivrait une personne avec TSA. Cette trame est construite sous forme d'histoire pour rendre le déroulé des événements le plus réaliste possible et proche d'une véritable expérience TSA. L'ensemble de l'expérience a été construit sur cette base scénaristique qui correspond à la retranscription des diverses expériences recueillies lors des ateliers de design thinking avec le collectif GEM.

Cette trame de scénario¹¹ représente une base de construction qui a permis de concevoir les différentes phases de ce projet en gardant une trame de cohérence globale. L'expérience immersive s'articule comme une histoire dans laquelle le protagoniste est entraîné malgré lui vers l'étape suivante, bien qu'il tente de reprendre le contrôle de ce qui lui arrive, comme le ferait une personne avec TSA.

Abordées dans la partie théorique, les pauses pédagogiques constituent un principe de conception directement issu des travaux de Parong et Mayer (2018). Ces auteurs ont mis en évidence certaines limites de la RV dans la transmission de connaissances factuelles (objectif pédagogique général : identifier les principales particularités cognitives liées au TSA) en lien avec la charge cognitive induite par l'environnement immersif. Pour répondre à cette problématique, ils ont proposé un dispositif spécifique : l'insertion de pauses pédagogiques au sein des expériences immersives.

Ces pauses se caractérisent par de courtes interruptions de l'action, conçues de manière à ne pas provoquer de rupture du leurre immersif, autrement dit à préserver l'illusion de plausibilité telle que proposée par Slater (2009). Concrètement, ces pauses prennent la forme d'un ralentissement temporaire de l'activité : les agents virtuels poursuivent leurs tâches mais sans progression réelle, adoptant une posture d'attente ou de réflexion apparente. Ce procédé permet à l'utilisateur de percevoir une continuité temporelle crédible sans que le déroulement de la scène semble figé de manière artificielle, ce qui risquerait de compromettre l'expérience.

Dans cet intervalle, un message audio sous forme de voix off est diffusé, apportant une clarification ou une information contextuelle pertinente. Cette modalité de transmission permet d'introduire des éléments factuels sans générer de surcharge cognitive excessive, en s'appuyant sur un canal sensoriel distinct de celui mobilisé par l'interaction directe. Ce principe de conception a été intégré dans le dispositif développé, bien qu'aucune évaluation empirique

¹¹ Voir annexe C.2.3 : [Matériaux scénaristiques : Narration de l'expérience](#) p. 337.

indépendante de son efficacité n'ait encore été conduite. Une analyse plus approfondie de ces mécanismes serait souhaitable.

2.3 Détails des tâches pédagogiques et objectifs

Sur la base de ces concepts et idées, une structure pédagogique décomposée en six phases a été conçue, avec les correspondances du scénario pédagogique initial. Chacune de ces grandes phases comportent également des sous-parties et certaines phases sont simultanées avec d'autres comme la confrontation avec l'environnement (partie 3) et la reprise de contrôle de l'environnement (partie 4), qui sont le cœur de la structure. Voici ces parties :

- 1) L'incarnation de l'étudiant TSA
- 2) La présentation du problème initial : correspondant à la phase de présentation de la tâche cognitive du scénario pédagogique (l'énigme au tableau)
- 3) La confrontation avec l'environnement virtuel et les agents PNJ (simultanée phase 2)
 - a) Première tâche - L'horloge comme repère
 - b) Phase de stress social (interaction avec l'enseignant)
 - c) Seconde tâche - Le tableau de consignes
 - d) Troisième tâche - Les plantes (routine et persévérance)
- 4) La reprise du contrôle de l'environnement (simultanée phase 3)
- 5) La résolution du problème (l'énigme)
- 6) La phase de confrontation : la phase finale - l'exposé oral

2.3.1 *L'incarnation de l'étudiant avec TSA*

La prise de possession de l'avatar étudiant avec TSA est le premier élément qui va permettre à l'utilisateur de percevoir qu'il y a une différence entre une perception dite neurotypique et une perception sous le filtre des TSA. Ce principe est possible grâce à une modification des éléments visuels pour qu'ils correspondent à une représentation se rapprochant de ce que peuvent voir les personnes avec TSA.

L'objectif pédagogique (et les moyens de l'atteindre) :

- **Objectif principal** : Identifier les principales particularités cognitives liées au TSA sans déficience intellectuelle
 - Faire comprendre à l'utilisateur les différences de perception sensorielle entre un individu neurotypique et un individu avec TSA.

- Amorcer l'approche perceptive du traitement de l'information (modèle EPF de Mottron)
- **Objectif externe** : Faire un rappel en situation des fondamentaux d'utilisabilité de la RV (vision et cinétose)

Descriptif de la tâche : Une fois « à l'intérieur » du corps de l'étudiant TSA (l'avatar), la perception par l'utilisateur de l'environnement virtuel est modifiée pour correspondre à la perception d'une personne atteinte de TSA. Les couleurs sont saturées et entourées d'un effet de halo. Les sons sont également amplifiés et modifiés (cf. les expériences d'incarnation de personnes atteintes de TSA). Cette différence de perception aurait également un impact sur le mode de pensée des personnes avec TSA comme chez Bled (2022), citant Stef Bonnot-Briey (personne avec TSA, consultante et formatrice). La perception de l'environnement est différente selon les filtres perceptifs propres aux TSA (Willay, 2020) comparativement avec une perception neurotypique (fig. 8 et fig. 9).

Figure 8

Salle de classe avant perception TSA



Figure 9

Salle de classe après perception TSA



Choix de conception : Au niveau de la littérature, cette technique a déjà été pratiquée en jouant sur d'autres aspects de la saturation (gros plan, son amplifié, cadrage spécifique, etc. comme on peut le voir dans les vidéos 360°). Ces types de vidéo sont conçues pour un support écran (pas de RV) et présentent la perception d'un enfant TSA dans différentes situations. Les auteurs ont choisi certains aspects de la saturation et pour ce travail, d'autres aspects, plus pertinents pour de la RV (les cadrages et grand plan risquant de rendre malade l'utilisateur), ont été sélectionnés (saturation des couleurs et effet de halo).

Toujours d'après Bled (2022), citant Stef Bonnot-Briey, la sensibilité importante dans les TSA entraînerait l'hyperperceptivité et un mode de pensée particulier : « *C'est l'approche perceptive de l'environnement qui génère notre mode de pensée et va avoir une influence sur nos représentations mentales* ».

Un des risques de telles modifications des informations sensorielles est l'augmentation de la cinétose. Dans leur revue systématique, Radianti et al. (2022) ont montré que pour réduire les risques de cinétose en RV, il faut être prudent sur les jeux de caméra comme les changements rapides avec accélération et décélération, le FOV, *field of view* ou champs de vision, sur lesquels jouent la latence et le frame rate ainsi que le *flicker* (une fluctuation de la brillance de surface) pour ne pas provoquer de cinétose. Pour rappel, les écrans du HMD sont situés à moins de 2 cm des yeux de l'utilisateur. Divers éléments représentatifs ont été conçus spécifiquement pour réduire au maximum les risques de cinétose chez les utilisateurs ne présentant pas de troubles spécifiques comme le suggère Radianti et al. (2022). Cette altération constitue la première interaction avec l'environnement virtuel que l'utilisateur sera amené à parcourir dans cette expérience. Avec ce changement radical de l'environnement, l'utilisateur se confronte donc pour la première fois à une perception de type TSA. Cette partie de l'expérience va durer environ 15 secondes, puis le PNJ enseignant (situé au tableau) va prendre la parole et désigner un élément sur le tableau principal.

2.3.2 *La présentation du problème initial : La consigne initiale et la carte mentale*

Certaines personnes présentant un TSA, lorsqu'elles sont confrontées à des situations d'évaluation impliquant une consigne explicite, peuvent produire une représentation mentale du problème plus complexe que celle attendue, appelée carte mentale selon Boldsen (2018). Ce phénomène rend la tâche initiale plus difficile à résoudre, car des éléments non pertinents au regard de la résolution sont intégrés au traitement cognitif. Ces informations supplémentaires doivent alors être traitées, ce qui alourdit la charge cognitive.

Les objectifs pédagogiques (*et les moyens de les atteindre*) :

- **Objectif principal** : Comprendre les conséquences des particularités du TSA dans un contexte scolaire
 - Montrer la création d'un système de pensée spécifique au TSA à travers l'élaboration d'une carte mentale lors de la présentation d'une consigne scolaire
- **Objectif externe** : Dans une logique de gamification, susciter l'engagement de l'utilisateur par le biais d'une expérience ludique et immersive.

Descriptif de la tâche : Dans l'expérience, la consigne est formulée par l'agent enseignant PNJ. Il s'agit de la seule instruction explicitement donnée à l'utilisateur. Le reste de l'environnement doit être compris de manière autonome, par observation et analyse. Des aides à la résolution

sont rendues disponibles afin que le gameplay ne se substitue pas à l'objectif principal. Toutefois, cette dimension ludique est essentielle car elle structure l'expérience : c'est par la confrontation à la difficulté et aux mécanismes de résolution que la phase de renforcement prendra tout son sens. Cette orientation répond à des attentes exprimées lors des entretiens avec les cellules du GEM.

L'énigme proposée est une tâche de fond construite autour d'un problème de logique de type « sens commun », ne nécessitant aucune connaissance spécifique (fig. 10). Elle s'inspire du jeu vidéo *The Witness*, fondé sur un gameplay d'apprentissage implicite et de découverte par essai-erreur. Cette énigme est affichée au tableau. L'enseignant en lit la consigne, inscrite sous forme de sous-titre. Pour la résoudre,

Figure 10

L'enseignant présentant l'énigme



l'utilisateur doit d'abord en comprendre les règles, qui seront explicitées ultérieurement.

Cependant, au moment où l'utilisateur concentre son attention sur l'énigme, le tableau se brouille (fig. 11). Il se transforme en une carte mentale complexe, illustrant le processus cognitif observé chez certaines personnes avec TSA. Chaque élément perçu est interprété, puis schématisé, produisant une arborescence qui démultiplie les possibilités de résolution et finit par rendre la tâche initiale presque insoluble. Cette carte mentale va progressivement devenir incompréhensible par surcharge d'informations inutiles pour l'utilisateur. Il sera alors impossible pour l'utilisateur de résoudre l'énigme, même s'il se concentre dessus. Les informations sont volontairement dissimulées.

Figure 11

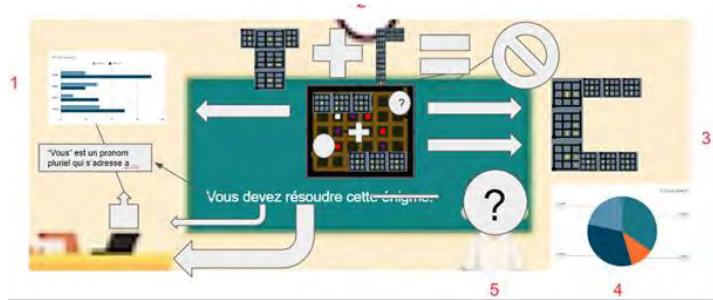
Début de la carte mentale



La carte mentale devient alors **Figure 12**

illisible en raison d'une surcharge d'informations non pertinentes (fig. 12). L'utilisateur ne peut plus résoudre l'énigme, même en y portant toute son attention, car les données nécessaires sont dissimulées. Toutefois, certaines zones restent

Carte mentale avec les zones



accessibles : elles correspondent à des sous-parties que l'utilisateur doit identifier et résoudre à l'aide d'indices dissimulés dans l'environnement virtuel. L'utilisateur ne peut progresser tant qu'il n'a pas trouvé un moyen de réduire la densité d'informations affichées au tableau. Une fois les indices découverts, les sous-zones s'éclaircissent progressivement, offrant une interprétation explicite de leur fonctionnement.

La seule manière de reprendre le contrôle de la tâche et de résoudre l'énigme consiste à gérer l'environnement via des tâches annexes, cherchant un équilibre entre résolution logique et maîtrise de l'environnement virtuel.

Choix de conception : La dimension ludique est justifiée par les travaux de Gorisse et al. (2018) qui soulignent que lorsqu'une expérience est purement contemplative, l'engagement de l'utilisateur peut diminuer, réduisant ainsi le sentiment d'agentivité, lui-même lié à la sensation de présence. Comme le soulignent Muletier et al. (2014), la gamification présente des avantages à court terme, notamment en termes de motivation et d'intérêt. Néanmoins, son usage doit être nuancé : Dicheva et Dichev (2015) montrent que seuls 24 % des études analysées rapportent un effet positif sur l'apprentissage, contre 49 % de résultats non concluants, appelant à une prudence méthodologique quant à son apport éducatif. Le but est de proposer un système de compréhension de règles simples sans explications qui, une fois acquises, vont permettre de résoudre la tâche avec un degré de satisfaction qui aide à renforcer l'immersion via le fait d'être plongé dans la tâche.

Le principe de carte mentale est documenté par Boldsen (2018) à partir du cas de Tito Rajarshi Mukhopadhyay (et corroboré par les entretiens menés avec les cellules du GEM). La carte mentale vise à représenter la surcharge cognitive générée par un traitement excessif de l'information et les stratégies de représentation utilisées par les personnes avec TSA. La fragmentation de la carte mentale illustre la perte de contrôle perceptif et cognitif qu'une personne avec TSA peut expérimenter dans certaines situations selon Boldsen (2018).

2.3.3 L'énigme à résoudre

Cette énigme repose sur deux niveaux de règles :

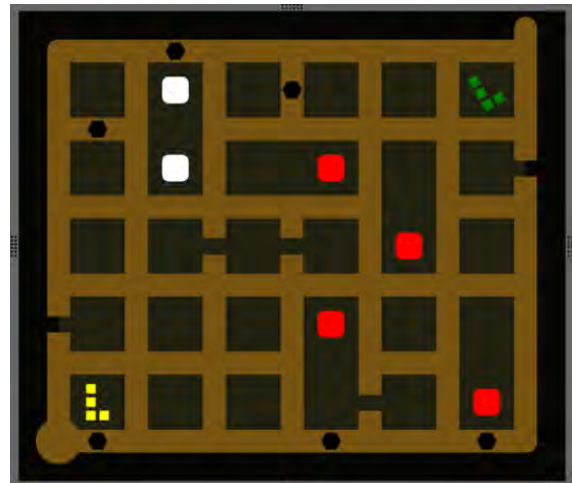
- Des règles explicites, clairement données à l'apprenant (par le PNJ enseignant : « *des indices sont dissimulés, à vous de les trouver* »).
- Des règles implicites, que l'apprenant doit inférer par exploration et raisonnement inductif (les indices trouvés doivent être liés à l'énigme par déduction).

Ces règles ne sont pas expliquées verbalement mais doivent être déduites de l'environnement par l'utilisateur. Une fois identifiées, ces règles permettent de résoudre l'énigme en moins de cinq minutes¹².

Le design du labyrinthe a été conçu à **Figure 13**

l'issue de plusieurs séries de tests lors des ateliers *La forme non résolue de la tâche*

itératifs, afin d'offrir une expérience à la fois stimulante et non frustrante (fig. 13). L'architecture de l'énigme réduit progressivement les options possibles (passages bloqués, règles contraignantes), facilitant la progression par essais/erreurs. Le système de correction intégré en RV soutient ce processus, en évitant toute perte de temps inutile.



L'objectif est ici de produire une expérience gratifiante, où l'utilisateur, initialement confronté à une tâche perçue comme difficile, parvient à la résoudre grâce à l'acquisition progressive des règles. Ce renforcement par la réussite renvoie au concept de satisfaction ludique décrit par McGonigal, cité par Muletier et al. (2014). Toutefois, il convient de rappeler que les *serious games* ne peuvent être directement assimilés aux jeux vidéo classiques, et présentent des contraintes spécifiques comme chez Hocine et al. (2011).

¹² Leur nature est détaillée en annexe C.3 : [Système de résolution de l'énigme](#) p. 335.

Enfin, la progression dans l'énigme déclenche le déverrouillage de zones de l'environnement, accompagnées d'éléments d'aide et de rappels des règles. La carte mentale de l'utilisateur se complète progressivement : les règles deviennent explicites, les éléments parasites disparaissent, et l'énigme devient résoluble (fig. 14). Cette dynamique illustre une agentivité progressive, dans le sens de Bandura (2008/2019), selon laquelle les actions de l'individu modifient activement son environnement. La mobilisation de cette agentivité est essentielle pour maintenir attention, motivation et engagement.

Figure 14

L'énigme avec les zones de règle



Mise à jour méthodologique :

Lors des tests utilisateurs internes de vérification, il est apparu que pour les utilisateurs neurotypiques, la simple juxtaposition d'une énigme et de règles symbolisées par des points d'interrogation suffisait à produire un effet comparable à celui d'une carte mentale. Cet élément est un point d'amélioration qui pourrait nécessiter un protocole plus approfondi qu'il n'a pas été possible de mettre en place pour des raisons de délais.

2.3.4 La confrontation avec l'environnement virtuel et la reprise de contrôle

Cette phase marque le véritable début de l'expérience immersive pour l'utilisateur. L'ensemble des éléments scénaristiques et des stratégies pédagogiques présentés en amont sont ici mis en œuvre dans un environnement hostile et socialement complexe. À l'issue de la présentation de la tâche par l'enseignant, l'utilisateur engage deux séquences d'activités parallèles pour commencer à préparer la résolution de l'énigme : la confrontation à l'environnement (et la recherche d'indices) et la reprise de contrôle de cet environnement. Ces deux dimensions se déroulent de manière simultanée, participant conjointement à la progression de l'expérience et à l'acquisition des connaissances.

2.3.4.1 La confrontation avec l'environnement et la récupération d'indices

La trame narrative repose sur le scénario des « conséquences du comportement de l'utilisateur » et s'articule autour de tâches successives, combinant interactions avec les agents PNJ, mécanismes de focalisation attentionnelle et régulation émotionnelle. Voici l'ensemble des objectifs pédagogiques traités dans cette partie :

Les objectifs pédagogiques (*et les moyens de les atteindre*) :

- Identifier les principales particularités cognitives liées au TSA sans déficience intellectuelle
 - Introduire les systèmes de routines spécifiques au TSA (Willay, 2020).
 - Introduire la notion de focalisation.
- Comprendre les conséquences de ces particularités dans un contexte scolaire.
 - Montrer les conséquences des routines et de la focalisation sur le comportement.
- Reconnaître des situations pédagogiques à risque.
 - Présenter les effets d'un environnement socialement menaçant pour une personne avec TSA.
- Savoir dans quel contexte appliquer ces stratégies pour soutenir la réussite et l'inclusion d'un élève avec TSA.
 - Exposer les stratégies alternatives mobilisées par les personnes avec TSA pour faire face à ce type de situation.
 - Illustrer des techniques d'apaisement spécifiques comme le recours au bruit blanc, souvent rapporté lors des entretiens menés avec les cellules du GEM.
- Développer une posture empathique en percevant la situation « depuis la perspective » d'un élève TSA, réduisant ainsi les stéréotypes et favorisant l'amorce d'un changement d'attitude.
 - Faire vivre la situation à l'utilisateur via l'incarnation de l'avatar avec TSA.

2.3.4.2 *La focalisation du détail - l'horloge et la règle des hexagones*

Les objectifs pédagogiques (*et les moyens de les atteindre*) :

- **Objectifs principaux** : Identifier les principales particularités cognitives liées au TSA sans déficience intellectuelle et comprendre les conséquences de ces particularités dans un contexte scolaire.
 - Introduire la notion de focalisation.
 - Montrer les conséquences de la focalisation sur le comportement
- **Objectif externe** : mise en place des pauses pédagogiques.

Descriptif de la tâche : Figure 15

L'utilisateur doit localiser la seule zone nette de l'environnement. Un effet de flou progressif (*shader*) restreint le champ visuel (voir fig. 15), contraignant l'attention sur un point unique : l'horloge (voir fig. 16 ci-dessous). Le maintien du regard durant trois secondes

L'environnement virtuel en mode flou avec zone horloge



Figure 16

déclenche l'apparition de la première règle, qui vient remplacer une zone parasite de la carte mentale. Simultanément, l'enseignant commence à interagir avec l'utilisateur, amorçant ainsi la mise en scène des difficultés sociales typiques d'un TSA.

L'horloge avec tracé



Choix de conception : Le *shader* de flou progressif permet d'illustrer le principe de focalisation via un effet visuel simple. En occultant artificiellement une partie du champ de vision de l'utilisateur, on impose à celui-ci la focalisation sur un point telle que le perçoit une personne avec TSA, c'est-à-dire pratiquement « forcée » cognitivement par son fonctionnement inhérent. Une pause pédagogique comme le suggère Parong et Mayer (2018) est, dès lors, activée pour expliciter l'expérience vécue, favorisant ainsi l'acquisition de connaissances factuelles sur le TSA.

2.3.4.3 Détourner le regard pour comprendre le professeur

Les objectifs pédagogiques (et les moyens de les atteindre) :

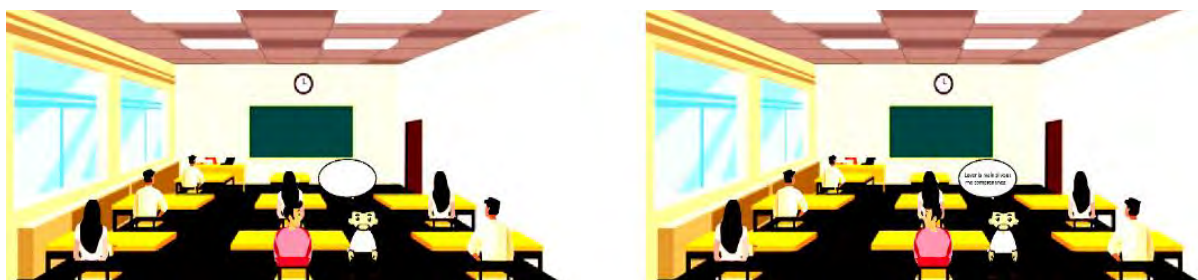
- **Objectifs principaux :** Reconnaître des situations pédagogiques à risque et savoir dans quel contexte appliquer ces stratégies pour soutenir la réussite et l'inclusion d'un élève TSA.
 - Présenter les effets d'un environnement socialement menaçant pour une personne avec TSA.
 - Exposer les stratégies alternatives mobilisées par les personnes avec TSA pour faire face à ce type de situation.

Descriptif de la tâche : Lorsque l'utilisateur regarde l'enseignant, les propos du PNJ deviennent incompréhensibles (floutage sonore, saturation). Seul le détournement du regard de l'utilisateur permet de rétablir la clarté sonore. Ce mécanisme met en scène une particularité d'interaction souvent observée dans le TSA.

Si l'utilisateur ne détourne pas le regard dans les 15 secondes, une aide contextuelle (voix *off*) intervient. Le regard dirigé vers le sol déclenche la bande son « clair » et amorce la suite de tâche (fig. 17).

Figure 17

Le détournement du regard pour comprendre l'enseignant



Voici ce que dit l'enseignant à l'utilisateur : « Tout va bien ? Si vous avez un souci, n'hésitez pas à me le dire. Levez la main si tout va bien ».

Si l'utilisateur manifeste un signe d'approbation (levée de main ; mouvements détectés via le dispositif de réalité virtuelle HMD) l'enseignant retournera au tableau, signalant ainsi que l'utilisateur est prêt à poursuivre vers la tâche suivante. À l'issue de cette étape, une première représentation des manifestations physiologiques du stress et de l'anxiété sera induite par le biais du casque occlusif.

Choix de conception : la technique de floutage sonore et saturation permet d'illustrer avec un procédé simple la difficulté de compréhension lors d'une interaction directe. Ce simulacre est une représentation figurée qui oblige l'utilisateur à détourner son regard (dans un cas réel, les personnes avec TSA comprennent les propos, mais peuvent détourner le regard pour se concentrer sur les paroles).

La représentation des **Figure 18**

manifestations physiologiques du *Overlay effet « blessure critique » de jeu vidéo*

stress et de l'anxiété repose sur l'activation d'un *overlay* visuel discret composé de pulsations lumineuses rouges, couplées à des vibrations synchronisées imitant les battements cardiaques (fig. 18).



2.3.4.4 *L'hyper concentration : le bruit blanc et la focalisation du détail via la règle cachée*

L'objectif pédagogique (et les moyens de l'atteindre) :

- **Objectif principal :** Identifier les principales particularités cognitives liées au TSA sans déficience intellectuelle et savoir dans quel contexte appliquer des stratégies pour soutenir la réussite et l'inclusion d'un élève avec TSA.
 - Illustrer des techniques d'apaisement spécifiques comme le recours au bruit blanc.
 - Montrer la capacité de focalisation malgré la contrainte.

Descriptif de la tâche : Le stress simulé **Figure 19**

général (vibrations, pulsations visuelles) *Bureau avec écouteur et main virtuelle*

traduit les réactions physiologiques observées dans le TSA. Pour y faire face, l'utilisateur doit activer un dispositif d'apaisement : il saisit des écouteurs en surbrillance, les place à ses oreilles virtuelles (fig. 19), déclenchant ainsi une pause sensorielle de 15 secondes sous bruit blanc (fig. 20). Durant cette période, une tâche s'active, matérialisant la troisième règle/indice (une règle visible dès le départ mais dissimulée dans un texte global (fig. 21)



Figure 20

Le bruit blanc et sa représentation figurée



ainsi que la structure séquentielle de résolution. Ces éléments viennent enrichir la carte mentale. Parallèlement à l'acquisition de la règle visuelle, le personnage de l'enseignant, intrigué par le comportement inhabituel de l'utilisateur (ce qui est souligné par la narration en voix *off*) recommence à manifester de l'intérêt pour lui.

Figure 21

La règle cachée (L jaune)



Toutefois, à ce stade, l'utilisateur ne perçoit pas encore cette interaction, ce qui contribue à souligner le décalage communicationnel souvent observé chez les personnes avec TSA.

Choix de conception : Le bruit blanc est apaisant pour une personne avec TSA (propos recueillis lors des entretiens avec les cellules du GEM). Il ne provoque pas d'état anxieux et n'est pas source de stress. Ce type de comportement, bien qu'apaisant pour les personnes présentant un TSA, est fréquemment source d'incompréhension de la part des observateurs extérieurs. Cette réaction sociale est souvent liée à une stéréotypisation du TSA fondée sur des représentations erronées ou simplistes, Willay (2020).

2.3.4.5 La focalisation du détail 2 : la concentration sur la constellation de point - la règle des plantes

L'objectif pédagogique (et les moyens de les atteindre) :

- **Objectif principal :** Identifier les principales particularités cognitives liées au TSA sans déficience intellectuelle.
 - Introduire les systèmes de routines spécifiques au TSA (Willay, 2020).
 - Renforcer la notion de focalisation.

Descriptif de la tâche : Une série d'objets sur la gauche de l'écran (des plantes) s'illumine. L'utilisateur doit conserver son regard sur une série de quatre points dans un ordre précis, guidé par un système de clignotement. Toute déviation réinitialise le compteur (fig. 22). À partir de la dixième seconde, l'enseignant (en continuité avec la phase précédente) tente d'interrompre la concentration par un signal d'alerte phasique (Posner & Cohen, 1984), rendant la tâche particulièrement exigeante sur le plan attentionnel. L'utilisateur doit résister à cette interruption et rester concentré sur les points.

Figure 22

Les L verts penchés et le timer



En cas de réussite (4×3 secondes), la quatrième règle s'intègre à la carte mentale. Après trois échecs, une aide est automatiquement déclenchée. Une pause pédagogique explicative s'ensuivra détaillant différemment la situation en fonction de la réussite ou de l'échec.

Choix de conception : ce système de focalisation diffère du premier système de l'horloge. Cette fois-ci, l'utilisateur est contraint, grâce au système de timer et de jauge, de focaliser avec une obligation de résistance aux stimuli externes. Cet artifice permet de simuler les effets de la perturbation sur les personnes avec TSA lors d'une tâche de focalisation avec un processus automatisé de décompte temporel.

2.3.5 La reprise du contrôle de l'environnement (simultanée avec 2.3.4.)

Cette phase est une dynamique parallèle, décrivant les processus de récupération des ressources indispensables à la résolution de l'énigme principale, définie comme tâche complexe selon Gauthier et al. (2016). Ces auteurs précisent qu'une tâche complexe exige la mobilisation conjointe de savoirs, savoir-faire et savoir-être, s'éloignant des tâches purement procédurales. Dans le cadre de cette expérience en RV, la reprise du contrôle de l'environnement correspond à une démarche au cours de laquelle l'utilisateur mobilise successivement des ressources internes et externes.

Les objectifs pédagogiques (et les moyens de les atteindre) :

- **Objectif principal** : Développer une posture empathique en percevant la situation « depuis la perspective » d'un élève avec TSA, réduisant ainsi les stéréotypes et favorisant l'amorce d'un changement d'attitude.
 - Faire vivre la situation à l'utilisateur via l'incarnation de l'avatar avec TSA.
- **Objectif secondaire** : Mettre en évidence l'équilibre entre acquisition de compétences et gestion de l'environnement, à la fois physique et social.

2.3.5.1 Mobilisation des ressources internes, Fagnant et al. (2014)

La compréhension des règles de résolution de l'énigme (cf. phase précédente) permet à l'utilisateur de transformer des informations initialement perçues comme ambiguës ou parasites en savoirs opérants. Il s'agit ici d'un transfert de connaissances implicites en connaissances explicites, mobilisables dans la résolution d'une tâche complexe. Cette acquisition passe par :

- Une assimilation de connaissances factuelles relatives aux TSA (via les pauses pédagogiques).
- Une mise en œuvre de connaissances procédurales (ex. appliquer les règles de résolution, manipuler les écouteurs, détourner le regard pour mieux entendre).
- Un développement de connaissances conatives, liées à l'attitude à adopter face à une surcharge cognitive, une interaction sociale ou un besoin d'autorégulation émotionnelle.

Ce processus vise à faire expérimenter à l'utilisateur, de manière immersive, certaines compétences socio-cognitives caractéristiques de stratégies d'adaptation chez les personnes avec TSA.

2.3.5.2 Mobilisation des ressources externes

L'environnement virtuel propose à l'utilisateur une série de supports extérieurs (indices visuels, rétroactions, écouteurs, PNJ, voix *off*) qui font office de ressources externes. Ces éléments sont mis en œuvre tout au long de l'expérience : Ils agissent comme des guides structurants (ex : flèches, points jaunes, consignes explicites).

2.3.5.3 Une coordination multi-niveau

La réussite de la tâche principale (résolution de l'énigme finale) nécessite une coordination harmonieuse entre plusieurs niveaux de ressources :

- Cognitifs : compréhension des règles, traitement de l'information, abstraction.
- Attentionnels : maintien de l'attention, inhibition des distracteurs, focalisation visuelle.
- Émotionnels : gestion de l'anxiété, stratégies d'apaisement (bruit blanc), autorégulation.

Ce triptyque est au cœur de la reprise du contrôle de l'environnement et illustre la nécessité, pour l'utilisateur, d'endosser un rôle actif et réflexif dans la démarche. Il devient alors agent de son propre apprentissage selon Bandura (2008/2019) et ses travaux sur l'agentivité. Cette dernière intervient dans un environnement dynamique où ses actions modifient directement les conditions de résolution de la tâche. L'utilisateur doit progressivement apprendre à mobiliser ces ressources à bon escient.

2.3.5.4 Préparation à la tâche finale

À l'issue de cette phase, l'utilisateur a acquis l'ensemble des règles nécessaires à la résolution de l'énigme et a expérimenté différentes modalités de régulation cognitive et émotionnelle. Cette phase fait office de transition vers la tâche finale, qui s'effectuera dans une autre scène. Dans celle-ci, l'autonomie et l'intégration des apprentissages seront évaluées de manière implicite par la réussite de la tâche.

3. Scène 3 - La résolution du problème

3.1 Phase de résolution

Les objectifs pédagogiques (et les moyens de les atteindre) :

- **Objectif principal** : Identifier les principales particularités cognitives liées au TSA sans déficience intellectuelle.
 - Montrer les habiletés de raisonnement des personnes avec TSA sans déficience intellectuelle.
 - Montrer l'émergence de raisonnements alternatifs viables.

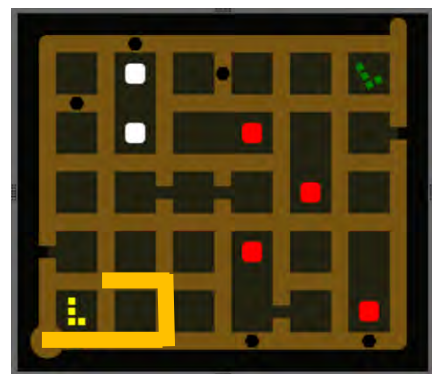
Objectif externe : Intégrer une phase ludique favorisant la satisfaction de l'utilisateur.

Descriptif de la tâche : À l'issue de la phase *d'aftermath* et de la découverte complète des règles, l'utilisateur entre dans la phase de résolution de l'énigme. Le tableau devient alors l'élément visible principal de l'environnement virtuel. Il se rapproche visuellement et gagne en taille afin de faciliter l'interaction. Cette transition visuelle est effectuée à l'aide d'un effet de type *eyelid*, conçu pour limiter l'apparition de cinétose (cf. section dédiée à ce phénomène). L'utilisateur peut alors résoudre l'énigme en saisissant la craie virtuelle et en dessinant le tracé sur le tableau virtuel.

Choix de conception : Devant le tableau, l'utilisateur se voit proposer un système d'action. Il doit engager, via l'utilisation des contrôleurs du casque oculatif, une tentative de résolution à partir des règles précédemment intégrées. Muni d'une craie virtuelle, l'utilisateur active un système de dessin en deux dimensions qui permet de tracer ses gestes. La craie, une fois déposée, reste fixée sur le tableau, obligeant l'utilisateur à la reprendre pour continuer.

L'utilisateur doit réaliser le parcours correctement et atteindre la zone de sortie (fig. 23). Lorsque l'énigme est correctement résolue, un retour visuel immédiat est proposé : le curseur ainsi que le tracé s'illuminent par clignotement de couleur verte (rouge en cas d'erreur). Ce code couleur, hérité des pratiques de gamification et de marketing, est toutefois à manier avec précaution, les perceptions chromatiques pouvant différer selon les cultures comme l'expliquent Madden et al. (2000). L'activation de la solution

Figure 23
Début de résolution



entraîne un changement de comportement des PNJ. Une nouvelle bande sonore est introduite, marquant l'entrée dans la phase de renforcement, initiée par l'attitude et la voix de l'enseignant.

3.2 La phase de renforcement

Les objectifs pédagogiques (et les moyens de les atteindre) :

- **Objectif principal** : Reconnaître des situations pédagogiques à risque.
 - Présenter les effets d'un environnement socialement menaçant pour une personne avec TSA.
- Objectifs secondaires.
 - Valoriser la reconnaissance sociale.
 - Identifier les obstacles au développement des interactions sociales.
 - Sensibiliser aux effets anxiogènes caractéristiques du TSA.

Descriptif de la tâche : Cette phase débute dès que les agents PNJ réagissent à la réussite de l'utilisateur. Leur réaction, exprimant la surprise ou l'admiration peut être perçue de manière déstabilisante. Bien que socialement valorisante, ce type de réaction ambiguë est souvent difficile à décoder par des personnes avec TSA.

Choix de conception : Le renforcement positif s'appuie alors sur les *feedbacks* explicites de l'enseignant amorcées dans la phase précédente, valorisant clairement la performance de l'utilisateur. Cette demande d'explication amorce ainsi la phase d'*aftermath*.

Notion d'*aftermath* : Bien que la résolution soit acquise, cette nouvelle phase introduit une tension anxiogène, illustrant le coût cognitif et émotionnel de la centralisation de l'attention sociale pour une personne avec TSA. Ce contraste est matérialisé par une série d'effets sensoriels : vision périphérique teintée de rouge, pulsations lumineuses synchronisées avec simulation des vibrations cardiaques croissantes, inspirés des représentations de stress dans les jeux vidéo. Cette mise en scène immersive vise à illustrer les conséquences souvent méconnues d'une sollicitation pourtant bienveillante. Ce qui serait anodin pour une personne neurotypique peut représenter une surcharge insurmontable pour une personne avec TSA. La visualisation progressive d'un *overlay* saturé de stimuli symbolise la création désorganisée d'une carte mentale globale face à une demande sociale intense. Cette dimension a été intégrée sur recommandation de Mme Bouvet, spécialiste du TSA.

3.3 La phase de sortie

L'objectif pédagogique (et les moyens de l'atteindre) :

- **Objectif principal** : Développer une posture empathique en percevant la situation « depuis la perspective » d'un élève avec TSA, réduisant ainsi les stéréotypes et favorisant l'amorce d'un changement d'attitude.
 - Permettre une dernière comparaison entre la perception TSA et neurotypique de l'environnement afin de rappeler les différences.
 - Proposer une sortie en douceur de l'expérience immersive.

Descriptif de la tâche : À l'issue de l'*aftermath*, **Figure 24**

la voix *off* annonce la fin de l'expérience et l'extraction prochaine de l'utilisateur du corps de l'étudiant. Une transition très progressive ramène alors l'utilisateur vers une perception neurotypique : les couleurs se normalisent, les sons se stabilisent et l'environnement virtuel reviens à la norme initiale (fig. 24).

Perception neurotypique initiale



La voix *off* invite alors l'utilisateur à retirer le casque, appuyée par un texte explicatif accompagnant cette sortie (fig. 25). L'utilisateur est ensuite accueilli par un accompagnant (expérimentateur dans le cadre d'un protocole de recherche ou membre du personnel dans un contexte de prévention ou de médiation pédagogique).

Figure 25

Écran de fin

« Vous avez fini – Vous pouvez enlever le casque »

Choix de conception : par la réinitialisation des paramètres initiaux, le comparatif perception TSA / perception neurotypique est rappelé à l'utilisateur.

Une option de *Game Over* pourrait être développée pour anticiper les situations d'échec, bien qu'en l'absence d'un tel module, une supervision humaine est prévue durant les expérimentations. Une prise en charge spécifique sera assurée pour les cas concernés.

Chapitre 3 - Développement du prototype et du MVP

Cette partie est consacrée à l'aspect informatique de l'outil. Les différentes versions ont été élaborées avec le moteur Unity 3D, le langage de programmation C# et le logiciel Visual Studio. Il s'agit de la dernière étape méthodologique présentant le développement effectif¹³. Le prototype a été élaboré entre février et octobre 2023. Il est le fruit d'une collaboration avec un développeur indépendant qui a été recruté pour les besoins du projet. Voici le socle technique (moteur, langage, assets...), les choix de conception justifiés par la littérature scientifique et les contraintes pédagogiques. Le tableau 1 est un récapitulatif des étapes principales qui seront abordé ici.

Tableau 1

Résumé des étapes (tâches de l'activité par ordre chronologique)

Étape	Rôle technique
<i>Tutoriel :</i>	Calibration et familiarisation.
<i>Téléportation dans la classe :</i>	Fade-in <i>eyelid</i> pour éviter vection.
<i>Séquences d'activité (indices, PNJ, bruit blanc...):</i>	Gestion par un Sequence Manager : déclencheurs, timers, voix-off, enregistrement <i>logs</i> .
<i>Phase de résolution :</i>	Tableau agrandi, choix d'outils (craie, pointeur, regard).
<i>Aftermath & Renforcement :</i>	Activation <i>overlay</i> stress, scripts PNJ, contrôle de charge GPU.
<i>Sortie & Désimmersion :</i>	Transition inverse, zone refuge neurotypique, message <i>off-boarding</i> .

¹³ Voir annexe C.1 : *Documentation technique pour les développeurs* p. 328.

1. Cadre technologique

1.1 Choix du moteur de jeu

La maquette a été développée avec le moteur Unity 3D (version 2022.3 LTS), choisi pour la richesse de son écosystème VR et la compatibilité multi-plateforme (communauté, plug-ins VR, pipelines URP/HDRP, etc.).

1.2 Langage de programmation et IDE

Le langage C# - édité sous Visual Studio 2022 - permet une intégration native des scripts et un débogage unifié (productivité, débogage, intégration directe dans Unity).

1.3 Périphérique cible : Meta Quest 3 HMD stand-alone

- Deux écrans LCD de 2,48 pouces avec résolution de 2064 x 2208 pixels par œil Champ de vision : 110° avec lentilles Pancake.
- CPU/GPU embarqués : Qualcomm Snapdragon XR2 Gen 2 et Adreno 740.
- RAM de 8 Go et Mémoire interne de 512 Go.
- Système de *tracking* : 6 DoF avec 6 caméras intégrées.
- Optimisation *low poly*, *shaders* légers.

1.4 Chronologie de développement

Développement d'un prototype (fév.-oct. 2023 - fév.-oct. 2024) en collaboration avec un développeur indépendant.

2. Architecture fonctionnelle du prototype¹⁴

2.1 Organisation du projet Unity

Le prototype a été développé dans l'environnement Unity en s'appuyant sur une organisation modulaire visant à assurer une certaine clarté au projet. Il s'articule autour de quatre scènes principales représentant les temps narratifs de l'expérience : *tutoriel*, *salle de classe*, *résolution* et *aftermath*. Ces scènes sont accompagnées de dossiers (Ressources, Scripts et Prefabs) permettant une séparation claire entre logique, contenu dynamique et objets instanciables.

Chaque scène a été pensée pour représenter un moment distinct de l'expérience utilisateur (UX), tant sur le plan narratif que technique. Par exemple, la scène de tutoriel initie l'utilisateur aux interactions spécifiques à la RV tandis que la scène de résolution met en œuvre les mécaniques de compréhension liées à l'objectif pédagogique.

Cette segmentation permet une gestion optimisée des ressources, une meilleure lisibilité du code, ainsi qu'un déploiement plus simple sur des casques de réalité virtuelle autonomes, pour lesquels les contraintes en termes de performance et de poids mémoire sont importantes.

2.2 Scripts clés et logique comportementale

La logique fonctionnelle repose sur un ensemble de scripts organisés autour de managers centralisant les comportements clefs, couplés à des scripts secondaires assurant l'interaction utilisateur et le traitement des états.

Classroom Manager : Le Classroom Manager est le cœur de l'expérience. Il centralise la progression du scénario en s'appuyant sur une variable *Progress*, incrémentée soit manuellement (*Progress++*), soit via une fonction d'attente temporisée (*Wait()*). Ce système d'étapes conditionnelles permet une synchronisation fine entre l'audio, les animations des agents PNJ et les sous-titres.

¹⁴ Partie en collaboration avec M. Poulet – développeur full stack indépendant

La fonction *soustitre.Soustitre()* permet, par ailleurs, l'affichage séquencé de segments de texte, chacun accompagné d'une durée d'apparition personnalisée, assurant une expérience textuelle fluide et alignée avec le contenu sonore.

Élèves Manager : L'Élèves Manager est responsable de la synchronisation des animations des PNJ. Il expose des références à chaque élève animé, ce qui permet au *Classroom Manager* de déclencher ou interrompre dynamiquement leurs comportements selon la progression.

Audio Manager : L'Audio Manager regroupe l'ensemble des *audio sources* utilisées dans le prototype, comprenant les voix off, les dialogues contextualisés de l'enseignant, les variations sonores floues et les effets sonores comme les battements de cœur. Il permet une spatialisation fine, indispensable à l'induction d'un sentiment de présence en VR.

Input et interaction utilisateur : Le système d'interaction repose sur une abstraction des entrées utilisateur à travers les scripts *InputVR* associés à la main gauche et à la main droite. Ces scripts encapsulent les appels aux API XR Unity, en renvoyant des booléens simples selon l'action détectée (pression de bouton, saisie, relâchement, etc.). L'ensemble est regroupé dans un *Input Manager* accessible depuis les managers centraux, ce qui permet une gestion unifiée des contrôles.

Les objets interactifs utilisent le script *Interactable*, qui permet de saisir ou relâcher des objets selon leur *collider* et leur *rigidbody*. Si la variable *returnback* est activée, l'objet revient automatiquement à sa position initiale. Ce système permet d'assurer la stabilité des scènes et d'éviter des pertes d'objets en dehors du champ de vision utilisateur.

Les autres scripts importants :

- **LookSpy :** vérifie si l'utilisateur maintient son regard sur une cible pendant une durée définie.
- **MainLeve :** détecte le geste de la main levée à partir de la position relative de la tête de l'utilisateur.
- **EnigmeTableau :** gère la résolution de l'énigme interactive, via des objets « Balise » devant être activés avec la craie virtuelle.

2.3 Gestion des assets et des performances

Un effort particulier a été porté à la gestion des performances, critère essentiel pour assurer un bon confort d'utilisation sur les casques occlusifs. Les actifs 3D sont globalement en *low poly*, permettant un rendu visuel acceptable tout en maintenant un budget polygone réduit.

Trois techniques principales ont été mises en œuvre :

- **Batching statique et dynamique** : réduction du nombre de *draw calls* via la combinaison des meshes statiques, notamment pour les décors.
- **Niveaux de détail (LODs)** : application de niveaux de détail différenciés pour certains objets, notamment les personnages, en fonction de leur distance à la caméra.
- **Occlusion culling** : les objets non visibles par la caméra principale sont automatiquement exclus du rendu, réduisant la charge GPU.

Ces optimisations permettent de garantir un *framerate* stable au-dessus de 72 FPS sur des casques de type Meta Quest 3, condition indispensable pour limiter la cinétose et garantir une immersion crédible.

Un autre aspect technique notable réside dans l'héritage de composants issus d'anciennes versions du projet. En raison de la complexité croissante du prototype, certains éléments comme le système *Parent Constraint Custom* (reliant les modèles de mains à la caméra) ou des scripts non utilisés subsistent dans l'arborescence. Une documentation rigoureuse a permis de contenir ces résidus sans impact significatif sur les performances ou la stabilité.

Enfin, le système de caméra principale intègre un objet *ConeVision* utilisé pour détecter le regard de l'utilisateur. Dans certaines scènes, une caméra secondaire (*CameraFocusHorloge*) est utilisée pour isoler des éléments de l'environnement, comme une horloge, tout en rendant floue la caméra principale, ce qui participe à la mise en scène sensorielle et narrative.

3. Conception des environnements et personnages

3.1 Salle de classe virtuelle

La salle de classe virtuelle repose sur des modèles prédéfinis, conçus à partir des standards de modélisation proposés par Unity¹⁵. Tous les éléments sont entièrement paramétrables et dotés d'une physique, permettant des interactions les plus réalistes possibles. L'aménagement comprend un mobilier scolaire classique, avec huit bureaux étudiants (accompagnés de cahiers et de stylos) ainsi qu'un bureau enseignant enrichi d'un léger décorum.

Afin de limiter la surcharge cognitive, conformément aux recommandations en RVI, les éléments décoratifs ont été volontairement restreints. Ceux qui ont été conservés proviennent d'une bibliothèque interne du SGRL, alimentée par des ressources de *l'Asset Store Unity*. Une étagère, placée à gauche du bureau de l'enseignant, contient les objets nécessaires à la tâche 4. Sur le mur droit, un tableau dédié à la tâche 2 est positionné à proximité de la porte d'entrée des participants.

L'environnement a été conçu pour maintenir une cohérence visuelle : deux fenêtres assurent un éclairage naturel, et les textures appliquées au sol, aux murs et au plafond (parquet, murs blancs, plafond en bois) évoquent un cadre scolaire traditionnel. Quelques éléments symboliques, tels qu'un globe terrestre et une carte, viennent compléter cette ambiance sans compromettre la lisibilité ou la sobriété de l'espace virtuel.

3.2 Avatars et agents/PNJ

Répartition des agents/PNJ : 1 enseignant (script interactif), 7 étudiants passifs, 1 avatar contrôlé/téléporté.

L'avatar et les agents non-joueurs (PNJ) ont été modélisés en *low poly* afin de limiter la charge de calcul, en cohérence avec les capacités techniques des casques de RV

Figure 26

Avatar en photogrammétrie



¹⁵ Voir annexe C.2 : [Matériaux et assets du dispositif](#) p. 335.

autonomes. Ces choix visent également à éviter les effets de « Vallée de l'étrange », Mori (1970), susceptibles de générer un malaise chez l'utilisateur lorsque les avatars sont trop réalistes (exemple des avatars en photogrammétrie, fig. 26, Schwind et al., 2018 ; Miltcheva, 2021).

Le style graphique de type « *cartoon* », **Figure 27**

associé à une expressivité simplifiée, permet *Agent enseignant low poly*

ainsi d'assurer un résultat satisfaisant tout en réduisant les risques de rejet émotionnel (fig. 27).



En outre, les expressions faciales stylisées facilitent la lisibilité des émotions, ce qui est essentiel dans un contexte de sensibilisation aux TSA. Les expressions trop réalistes étant difficiles à modéliser avec précision selon Jourabloo et al. (2022). Ce

même style est appliqué aux PNJ et à l'avatar de l'utilisateur afin d'assurer une cohérence esthétique dans l'environnement.

Les agents PNJ sont organisés en trois groupes : un enseignant (PNJ actif), un ensemble d'étudiants (PNJ passifs) et l'étudiant avec TSA que l'utilisateur incarne après la phase d'introduction. Ces personnages disposent de scripts comportementaux qui simulent des routines scolaires. L'enseignant interagit avec l'utilisateur selon un scénario précis, tandis que les autres PNJ adoptent des comportements neutres et répétitifs, limités à la tâche en cours. Le choix de neutralité expressive vise à refléter les difficultés d'interprétation des émotions chez les personnes avec TSA, tout en minimisant les distractions dans un environnement scolaire modélisé comme celui d'un contexte d'examen.

4. Dispositifs d'interaction utilisateur

4.1 Contrôleurs et mapping des entrées

L'utilisateur interagit avec l'environnement grâce à des mains virtuelles contrôlées via les manettes du HMD ainsi que par un système de pointeur *raycast* invisible partant du regard. Les mains permettent une manipulation simple et directe des objets, tandis que le *raycast* est utilisé pour les sélections ou les actions à distance ainsi que le déclenchement contextuel d'événements (ex. : activation d'une animation après fixation prolongée).

4.2 Objets interactifs

Certains objets ont été conçus pour être manipulables ou déclencher des événements spécifiques, contribuant à la progression des tâches expérimentales. Parmi eux, on trouve la sphère de référence utilisée lors du tutoriel pour l'apprentissage des gestes de base, une craie virtuelle servant à écrire sur le tableau dans certaines phases, ou encore un casque audio activant un bruit blanc simulé lors de situations sensoriellement chargées (ex. tâche de surcharge auditive). Ces objets, modélisés en *low poly*, sont dotés de *colliders* et de comportements interactifs simples, assurant une prise en main fluide.

4.3 Feedbacks sensoriels

Pour enrichir l'expérience et améliorer la compréhension des stimulus sensoriels propres au TSA, plusieurs *feedbacks* ont été implémentés. Des *overlays* visuels dynamiques (pulsations rouges en surimpression pour simuler le stress ou la surcharge sensorielle) sont utilisés de façon ponctuelle pour illustrer l'état émotionnel ou physiologique de l'avatar.

5. Séquences d'activité

Voici à présent le détail des séquences d'activités (tableau 2).

Tableau 2

Le tableau récapitulatif des séquences d'activités

Séquence d'activité	Objectifs pédagogiques	Conception UX À travers interaction UX. ex : forcer UX à regarder un élément	Contenu - comment cela se représente en développement
<i>L'immersion TSA</i>	Percevoir l'environnement sous les filtres du TSA (via saturation du son et de la lumière ambiante)	Saturation des contrastes et application d'un filtre de bordure Augmentation du volume de la bande son et saturation sonore	Saturation sonore et visuelle. Prise de parole du PNJ enseignant pour présenter l'énigme. Apparition de l'énigme sur le tableau
<i>La carte mentale</i>	Comprendre le concept de « carte mentale » qui se présente comme une complexification des informations sous le filtre TSA	Apparition des éléments de complication après timer - recouvrement des zones actuelles	Apparition des 4 cases
<i>Première tâche - L'horloge</i>	Comprendre le processus de recherche d'indices (avec mise en exergue des indices saillants - le filtre cognitif du TSA permet de les mettre en évidence en occultant le reste de la scène)	Forcer l'UX à fixer une zone définie Via le <i>hit</i> du <i>raycast</i> de l'utilisateur, activation d'un <i>line renderer</i> jaune de 3 secondes sur l'extrémité de l'objet « aiguille trotteuse ». Intégration dans la zone 1 de la carte mentale	Flou gaussien progressif avec 1 point non affecté : l'horloge centrale (composé de symboles similaires à ceux de l'énigme). Modification du tracé de l'aiguille qui laisse un sillon jaune. Mise en place d'un timer
<i>Phase de stress - le rapport social</i>	Comprendre les difficultés du rapport social quand on a un TSA (avec les spécificités du filtre TSA : interaction directe entraînant du stress - besoin de détourner le regard pour comprendre ses interlocuteurs...)	Activation de la <i>voice line</i> brouillée tant que l'utilisateur a le regard (<i>raycast</i>) sur le visage de l'enseignant. Activation <i>overlay</i> « stress » (pulsation rougeâtre/cardiaque)	L'enseignant s'approche de l'utilisateur. Voix off pour consigne (détourner le regard vers le sol). Dès que le regard est détourné, la voix devient claire (transmission de la consigne de résolution). Le suivi de la consigne entraîne le départ de l'enseignant
<i>Seconde tâche - Le tableau de consignes</i>	Comprendre les stratégies de régulation du stress chez les personnes avec TSA Comprendre le processus de recherche d'indices	Arrêt des <i>overlays</i> « stress ». Le son et la lumière reviennent à leur état initial. Intégration en zone 2 de la carte mentale. Activation signal lumineux (première plante étagère gauche) en fin de la tâche	Le casque sur le bureau clignote. L'utilisateur doit le saisir et les mettre sur sa tête. Le tableau sur la droite clignote. 10 consignes sont écrites (La septième consigne est un des indices « le début est en bas à gauche »). La consigne passe en surbrillance et clignote
<i>Troisième tâche - Les plantes</i>	Comprendre la notion de « besoin de focalisation » et les risques d'une perturbation par un tiers lors de ce type de tâche (ce qui entraîne un abandon de	Activation curseurs en remplissages si <i>raycast</i> utilisateur on hit.	L'utilisateur doit maintenir son regards trois secondes sur chaque plante (3) pour activer 3 jauges. L'enseignant perturbe la tâche. Si l'utilisateur détourne le regards les curseurs redescendent. Une voix <i>off</i> guident l'utilisateur a la première

	la tâche avec besoin de recommencer)	<p>Fin curseur 1 : voix <i>off</i> explicative + <i>voice line</i> enseignant « perturbation ».</p> <p>Tous les curseurs <i>full</i> : Activation animation de compréhension.</p> <p>Intégration en zone 3 de la carte mentale + activation zone 4 (rappel global)</p>	plante, en cas de temps trop long ou de non résolution
<i>La résolution de l'énigme</i>	Tâche secondaire hors TSA (Effet de récence pour augmenter ludiquement l'agentivité sur la dernière tâche)	<i>Eyelide</i> pour téléporter l'utilisateur vers la zone « tableau » Activation de la craie et du tableau. <i>Line renderer</i> vert pour représenter le tracé. Craie mobile sur 2 dimensions	L'enseignant invite l'utilisateur à réaliser la solution au tableau. Il doit saisir la craie virtuelle fixée en bas à gauche et déplacer cette craie pour faire le tracé (complété en rejoignant le point en haut à droite)
<i>L'aftermath</i>	<p>Comprendre la difficulté de la prise de parole sans préparation</p> <p>Ramener l'utilisateur a un état neurotypique avant de clôturer la simulation</p>	<p>Activation <i>overlay</i> de stress</p> <p>Activation animation de fin</p> <p>Retour état initial son et lumière</p>	L'enseignant demande à l'utilisateur de se retourner pour expliquer comment il a fait. Observation des PNJ pendant 5 secondes. La voix <i>off</i> fait un rappel de l'expérience et invite l'utilisateur à enlever le casque

6. Les modifications de l'outil en cours de thèse - les évolutions du prototype vers le MVP

À la suite des retours obtenus lors du premier protocole expérimental, de nombreuses modifications ont été apportées au prototype initial. Ces ajustements visaient à corriger des bugs techniques et des défauts de conception ayant altéré la qualité de la première passation. Le résultat de ces améliorations est une nouvelle version du programme, conçue comme un MVP (Minimum Viable Product), considérée à ce stade comme la version la plus aboutie du dispositif. Cette version a fait l'objet de plusieurs tests utilisateurs selon une démarche de CCU, tous évalués comme satisfaisants, et a ensuite été utilisée comme support pour le second protocole expérimental. Une galerie visuelle d'extraits du dispositif est disponible¹⁶.

Les modifications :

- **Upscale de l'enseignant (effet perturbateur social)** : Dans la tâche des plantes, l'effet de perturbation sociale induit par la présence de l'agent PNJ enseignant a été renforcé. À la suite d'un retour du GEM dans le cadre des tests en conception centrée utilisateur, un *upscale* de la tête de l'enseignant a été intégré afin d'augmenter la perception de pression sociale et de rendre la présence de cette figure d'autorité plus saillante dans l'espace virtuel.
- **Alignement des plantes** : Lors de la première expérimentation, le chevauchement des zones de détection (zones de *hit*) des objets « plantes » a généré des difficultés d'interaction, certains objets se superposant de manière imprécise. Pour faciliter l'action de sélection, les plantes ont été réalignées spatialement de manière plus nette.
- **Affinage du raycast de sélection** : Toujours sur la tâche des plantes, le rayon de sélection (*raycast*) entraînait en collision avec plusieurs objets simultanément, ce qui rendait certaines activations aléatoires ou difficiles. Ce rayon a été affiné pour qu'il corresponde plus fidèlement à la direction du regard de l'utilisateur, améliorant ainsi la précision et la fluidité des interactions.
- **Ajout d'animations aux PNJ élèves** : Les agents PNJ représentant les élèves ont été enrichis d'animations comportementales plus naturelles et contextualisées, notamment

¹⁶ Voir annexe C.4 : [Captures d'écran du moteur Unity et du prototype](#) p. 345.

dans la séquence d'*aftermath*. Cela contribue à renforcer la crédibilité de la scène et l'immersion de l'utilisateur dans un environnement scolaire crédible.

- **Suppression de la tâche au sol** : L'élément interactif « sol » situé au sol, qui faisait partie d'une des tâches initiales, a été retiré. En effet, cet objet n'avait pas été compris ni identifié comme pertinent par la quasi-totalité des participants. Sa suppression vise à alléger la charge cognitive et à éviter les confusions.
- **Réorganisation des règles du tableau** : La règle n°3, initialement positionnée par rapport à l'élément « sol » (supprimé) a été déplacée et reclassée comme règle n°2. Elle figure désormais sur le tableau des consignes à droite de l'utilisateur, afin d'en améliorer la visibilité et la logique d'enchaînement.
- **Réduction de l'échelle du labyrinthe (*downscale*)** : L'environnement du labyrinthe, jugé trop vaste par plusieurs participants, a été redimensionné. Cette réduction visait à faciliter la navigation et l'atteinte des objectifs, tout en prévenant la fatigue ou la désorientation.
- **Correction de l'animation de l'enseignant** : L'animation de marche de l'enseignant a été retravaillée afin de produire des déplacements plus fluides et cohérents, réduisant l'effet de décalage perceptif et renforçant la vraisemblance des mouvements.
- **Ajout d'un texte récapitulatif en fin de session** : Afin de soutenir l'ancrage des apprentissages et de permettre à l'utilisateur de synthétiser ce qu'il vient de vivre, un texte récapitulatif a été intégré à la fin de l'expérience. Ce résumé reprend les messages clés transmis durant le parcours immersif.
- **Ajustement des bandes-son** : Certaines pistes audio ont été modifiées pour mieux s'aligner avec les nouveaux éléments visuels et interactifs. Ces ajustements visaient à assurer la cohérence de la narration et la fluidité de l'expérience.
- **Réduction générale de l'ambiance sonore** : Les sons d'ambiance ont été globalement diminués, notamment à partir de la tâche du bruit blanc, pour éviter qu'ils ne masquent les voix off ou ne fatiguent l'utilisateur. Cette optimisation sonore participe également à la gestion de la charge sensorielle.
- **Ajout de décors interactifs (agentivité)** : Dans un souci d'augmenter l'agentivité de l'utilisateur, des objets interactifs supplémentaires (un cahier et un stylo) ont été placés sur le bureau. Ils permettent une manipulation libre et renforcent le sentiment de contrôle et d'engagement dans l'environnement via l'agentivité.

L'ensemble de ces modifications s'inscrit dans une logique d'amélioration continue, articulant rigueur expérimentale, retours d'utilisateurs et contraintes techniques, pour aboutir à une version stable, réaliste et engageante du dispositif, prête à être évaluée de manière comparative dans un protocole rigoureusement contrôlé.

Conclusion

Le développement du prototype, puis du MVP, a permis d'aboutir à un dispositif fonctionnel et cohérent, intégrant les dimensions techniques, ergonomiques et pédagogiques définies en amont. Le recours à Unity 3D et C#, couplé à une optimisation ciblée pour le Meta Quest 3, a assuré la faisabilité technique tout en maintenant un niveau de fluidité compatible avec les standards de la RV éducative. Les ajustements successifs, allant du raffinement des interactions à l'amélioration des animations et des feedbacks sensoriels, témoignent d'une démarche itérative et centrée utilisateur, indispensable dans un projet de conception en RV.

Toutefois, plusieurs limites doivent être soulignées. Sur le plan méthodologique, le développement s'est appuyé sur un seul moteur et un ensemble restreint de périphériques, ce qui réduit la généricité du dispositif. De plus, certains héritages de versions antérieures subsistent dans l'architecture logicielle, créant une complexité qui pourrait entraver la maintenance ou l'extension future du programme.

Les perspectives d'amélioration sont multiples. D'un point de vue technique, une modularisation plus poussée du code, accompagnée d'une documentation consolidée, faciliterait l'évolutivité du système. Sur le plan pédagogique, des scénarios supplémentaires pourraient être développés afin d'explorer la variabilité des contextes scolaires ou des profils utilisateurs. Enfin, l'intégration de mesures physiologiques (par ex. suivi oculaire, biométrie) offrirait une voie prometteuse pour enrichir l'évaluation de l'expérience et renforcer la pertinence scientifique du dispositif.

En somme, ce travail constitue une première étape structurée vers un outil immersif robuste, tout en appelant à un élargissement des expérimentations et à une consolidation technique en vue d'une utilisation plus large et durable.

Partie 3 - Partie empirique

Chapitre 1 - Préambule

L'objectif de ce travail est d'examiner les effets de l'environnement en RVI sur la mémorisation et l'assimilation de connaissances factuelles sur le TSA pour acquérir un savoir-faire social à moyen terme avec une attitude adaptée aux personnes avec TSA. Plus concrètement, l'objectif est de mesurer l'impact du facteur technologique : soit à quels degrés d'immersion et de contrôle, générant la présence et l'agentivité, la compréhension du TSA est-elle impactée ?

Plusieurs modalités de comparaison ont été conçues pour étudier les effets de l'immersion technologique afin de faire varier celui-ci sur plusieurs niveaux (de l'absence totale d'immersion à une immersion maximum). La première expérimentation traitera de quatre modalités et la seconde de deux modalités de cette liste.

- **Modalité à forte immersion technologique (RV)** : RVI (environnement virtuel projeté dans un casque occlusif avec présence et agentivité accrue via le contrôle de mains virtuelles).
- **Modalité à moyenne immersion technologique (FS)** : Logiciel en *flatscreen* avec un contenu identique permettant d'agir sur l'environnement (réalité virtuelle non immersive et non occlusive avec agentivité maintenue via des contrôleurs de type clavier/souris et présence réduite sans avatar incarné : avatar non incarné à l'écran en vue FPS).
- **Modalité à faible immersion technologique :**
 - Présentation diaporama (DP), soit des informations et images présentées en diapositives (sans aucune bande son) qui défilent et dont le contenu est identique aux expériences RVI et *flatscreen* en termes de contenu factuel mais sans agentivité ni présence (aucun contrôle possible sur le contenu et pas de représentation de l'utilisateur).
 - Présentation diaporama audio (DP-A), identique à la modalité DP mais avec ajout d'une bande son composée de voix off reprenant le contenu des modalités RV et FS.

- Présentation en vidéo (FS) : une capture vidéo de la modalité RVI mais sans aucune présence et agentivité, ainsi qu'une interactivité ou contrôle par le participant.

Un premier protocole expérimental a été mis en place : il avait pour objectif de tester la première version du logiciel (le MVP) sur le plan de la présence et de l'agentivité. Une première série de relevé sur la compréhension du TSA a été effectuée pour tester les séquences d'activités et les mesures du niveau de compréhension du TSA (questions).

Un second protocole découlant du premier a ensuite été mis en place : il visait à tester les améliorations du logiciel (post MVP) et le nouveau protocole de recueil de données, plus abouti et ciblé, afin de mesurer l'impact de l'empathie générée en RVI sur les représentations du TSA.

1. Rappel des hypothèses générales

Ce projet de recherche vise à concevoir un dispositif de RVI sous un processus de conception centrée utilisateur via des outils de conception puis de tester l'efficacité de ce dispositif de manière empirique.

- 1. Présence et agentivité :** L'exposition en RVI immersive produira un niveau significativement plus élevé de présence et d'agentivité que les modalités non immersives. Raisonement : Le réalisme perceptif, l'incarnation corporelle et le contrôle direct de l'avatar sont connus pour renforcer l'engagement et la sensation de présence, Makransky et Petersen (2021).
- 2. Apprentissage factuel :** Une différence significative est attendue entre les modalités sur l'acquisition immédiate de connaissances factuelles (même si on observera qu'il n'y a pas d'effet). Raisonement : les connaissances factuelles peuvent être transmises efficacement quel que soit le support, si le contenu est constant, Mayer (2014).
- 3. Rétention des connaissances :** La rétention des connaissances factuelles à deux semaines sera supérieure dans la condition RVI immersive. Raisonement : l'expérience immersive augmente l'encodage en mémoire épisodique, ce qui favorise le rappel différé, Makransky et al. (2019).
- 4. Empathie :** La condition RVI immersive induira un score d'empathie plus élevé que les autres modalités. Raisonement : l'incarnation virtuelle facilite la prise de perspective et l'activation de réponses affectives pro-sociales, Herrera et al. (2018).
- 5. Stéréotypisation :** La RVI immersive réduira davantage les stéréotypes liés au TSA que les modalités non immersives avec un effet maintenu à moyen terme. Raisonement : les expériences immersives vécues à la première personne peuvent modifier durablement certaines représentations sociales, Banakou et al. (2016).
- 6. Changement d'attitude :** L'amorce de changement d'attitude sociale sera plus marquée et plus durable dans la condition immersive. Raisonement : la combinaison de présence, d'engagement émotionnel et de vécu expérientiel favorise la reconfiguration cognitive et affective des attitudes selon Slater et Sanchez-Vives (2016).

Le plan de travail est le suivant :

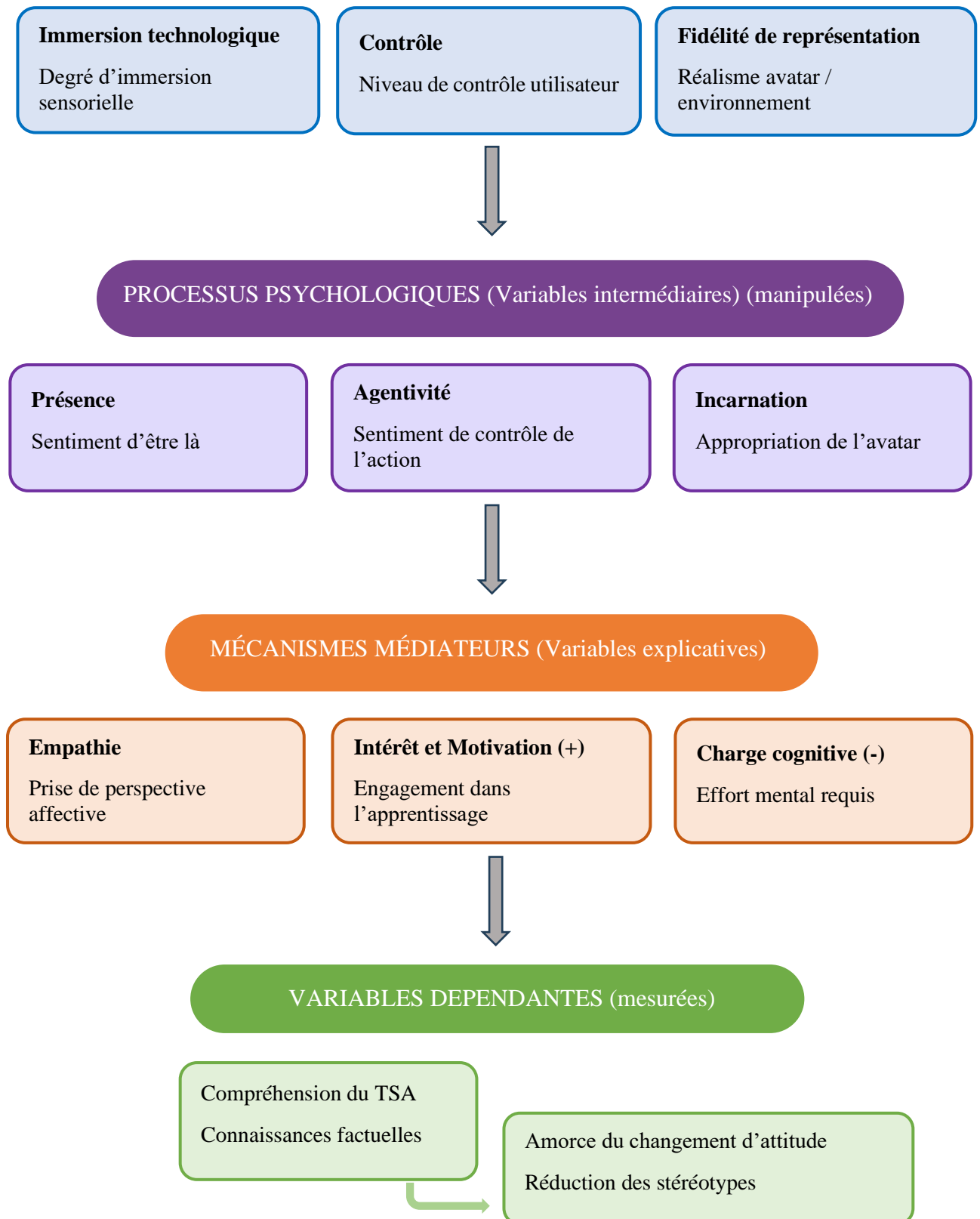
1. Volet conception : Après constitution d'un recueil de données des ateliers itératifs, conception en laboratoire de l'outil de RVI de sensibilisation au TSA sous le model CAMIL.
2. Expérimentation 1 : étude empirique du prototype de l'outil de RVI de sensibilisation au TSA.
3. Expérimentation 2 : étude empirique de la version revue et corrigée de l'outil et des modalités de recueil de données (questionnaire).

Pour synthétiser :

Le modèle illustre comment un environnement de RV bien conçu (grâce au contrôle et la fidélité de représentation) peut générer des effets de présence et d'agentivité. Ces effets activent des leviers cognitifs et affectifs (motivation, incarnation, etc.) qui facilitent l'intégration des connaissances factuelles. Cela permet, via une médiation pédagogique, d'améliorer la compréhension, de favoriser des changements conceptuels et de susciter l'empathie, aboutissant à un changement d'attitude.

Figure 6 (bis)

Schéma de modélisation des variables et interactions



IMPORTANT : une absence de cinétose est nécessaire pour le bon déroulement d'une expérience RV et la mise en place des relations de cette séquence.

Légende du schéma :

En italique : les facteurs affectifs et cognitifs du CAMIL qui influencent l'apprentissage immersif mais qui ne sont mesurés directement. (+) : *affecte positivement*. (-) : *affecte négativement*.

Variables indépendantes : caractéristiques techniques de la RVI manipulées expérimentalement (facteurs technologiques du CAMIL).

Processus psychologiques : états mentaux induits par l'expérience de RVI (potentialités de la RVI du CAMIL).

Mécanismes médiateurs : facteurs cognitifs et affectifs (du CAMIL) qui expliquent les effets (excepté le facteur empathie, issu d'autres pans de la littérature).

Variables dépendantes : *outcomes* mesurés pour évaluer l'efficacité de l'intervention (résultat de l'apprentissage du CAMIL).

2. *Les variables*

Variables indépendantes : Trois variables de RVI seront manipulées, toutes autour du facteur technologique.

- L'immersion technologique.
- Le degré de contrôle.
- La fidélité de représentation.

Processus psychologique (Variables intermédiaires) : Ces variables font la médiation entre les facteurs technologiques et les mécanismes médiateurs.

- Présence.
- Agentivité.
- Incarnation (pas mesurée directement).

Mécanismes médiateurs (Variables explicatives) :

- Empathie.
- Intérêt et motivation (pas mesurée directement).
- Charge cognitive (mesurée uniquement dans l'expérimentation 2).

Variables dépendantes (effets attendus) : Les effets de la RVI sur le rapport des participants vis-à-vis des personnes avec TSA, qui est l'objectif final.

- Compréhension du TSA.
 - Connaissances factuelles.
- Amorce de changement d'attitude.
 - Réduction des stéréotypes.
 - Score d'attitude vis-à-vis du TSA.

Chapitre 2 - Expérimentation 1 : test du prototype de sensibilisation au TSA - les effets de l'immersion sur la compréhension du TSA

1. Présentation/introduction

Cette première expérimentation a été menée à partir du prototype de sensibilisation au TSA en RVI, développé au sein du laboratoire SGRL au cours de la première année du projet. Il y avait deux objectifs principaux dans cette phase exploratoire, les enjeux étant de vérifier leur conformité avec les prédictions issues de la littérature.

1. Valider les fondements expérientiels du prototype, en évaluant plusieurs dimensions clés : la présence, l'agentivité et la cinétose (condition de contrôle de la RVI).
2. Observer les effets de l'immersion/contrôle sur l'empathie et les aspects liés à la compréhension et mémorisation de connaissances factuelles et à la stéréotypisation (amorçage du changement attitudes).

Pour le premier objectif, il était nécessaire de mesurer les dimensions clés de la RVI. Comme le souligne Gorisse (2019), la présence physique et l'incarnation sont des dimensions essentielles en environnement immersif. Parmi les composantes de l'incarnation décrites par Kiltani et al. (2012), l'agentivité a été identifiée comme un axe prioritaire pour cette première itération, sans pour autant négliger le sentiment de possession du corps virtuel ou la localisation de soi.

Le second objectif central concernait tout d'abord la génération d'empathie envers l'avatar, considérée comme une dimension essentielle dans les dispositifs de RVI visant à modifier les attitudes sociales (cf. travaux sur l'attitude en RVI). Cette approche s'inscrit dans la continuité d'expériences comme « Dans la peau d'une personne autiste » de Lanovaz (2023) ou « *The Displaced* », qui ont montré la capacité de la RVI à susciter de l'empathie. L'autre partie de cet objectif portait sur l'évaluation de la compréhension (l'acquisition de connaissances factuelles) et la stéréotypisation (amorçage du changement attitudes). La littérature souligne en effet que la RVI peut être moins performante que des supports traditionnels, tels que les présentations type diaporama, en matière de transmission de connaissances factuelles,

Parong et Mayer (2018). Bien que cette variable n'ait pas été manipulée de manière directe, un contrôle de la rétention des connaissances a été mis en place pour estimer l'efficacité du dispositif proposé. Pour la stéréotypisation, l'idée était qu'une meilleure compréhension du TSA pouvait initier une évolution des représentations (changement des stéréotypes).

Le protocole incluait également un test du système de mesure développé pour l'étude inspiré de la littérature existante mais substantiellement adapté, afin d'en évaluer la pertinence.

Conditions expérimentales : Les quatre modalités expérimentales se distinguent avant tout par le niveau de présence et d'agentivité qu'elles offrent à l'utilisateur. La modalité RVI propose une immersion technologique forte : l'utilisateur dispose de mains virtuelles et d'interactions en six degrés de liberté, ce qui favorise une véritable incarnation et un engagement actif dans l'environnement. À l'inverse, les modalités Diaporama (DP) et Diaporama Audio (DP-A) ne permettent aucun contrôle direct ni navigation, plaçant l'utilisateur dans une posture strictement passive face au contenu. La seule différence réside dans l'ajout d'un support audio pour DP-A. Entre ces deux extrêmes, la modalité écran plat (FS) propose un compromis : l'utilisateur agit dans un environnement projeté, avec contrôle sur son avatar.

La manipulation de l'audio constitue un levier expérimental intéressant. L'audio contribue à l'immersion sensorielle et soutient la transmission d'indices narratifs, tandis que son absence (DP) permet d'examiner l'effet d'un environnement visuel dépourvu de toute profondeur sonore. Ces contrastes mettent ainsi en évidence l'éventail des expériences offertes, allant d'une immersion riche et interactive (RVI) à une exposition strictement passive et silencieuse (DP).

Tableau 3
Modalités et conditions de l'expérience 1

Mod.	Support	Présence / Agentivité	Audio
<i>RVI</i>	Meta Quest 3 (HMD)	Haut niveau (mains virtuelles, interactions 6- <i>DoF</i>)	Oui
<i>FS</i>	Écran 15" + casque audio	Bas niveau (navigation souris, mains virtuelles non contrôlables)	Oui
<i>DP-A</i>	Écran 15" + casque audio	Très bas niveau (pas de navigation clavier-souris, aucune main virtuelle contrôlable)	Oui
<i>DP</i>	Écran 15"	Très bas niveau (pas de navigation clavier-souris, aucune main virtuelle contrôlable)	Non

2. Hypothèses

À partir de la revue de littérature et du protocole mis en place, quatre hypothèses principales ont été formulées :

H1 : Le score de cinétose devrait être supérieur en condition RVI par rapport aux autres modalités (condition de contrôle de la RVI, pas de lien direct avec l'amorce de changement d'attitude et de compréhension du TSA).

H2 : Les scores de présence (H2.1), d'agentivité (H2.2) et d'empathie (H2.3) devraient être supérieurs en condition RVI par rapport aux autres modalités. Par ailleurs, le score de la modalité *Flatscreen* (RV non immersive) sera supérieur à la modalité DP-A (diaporama audio) qui sera elle-même supérieur à la modalité DP (diaporama).

H3 : Les scores de compréhension et de mémorisation de connaissances factuelles sur le TSA devraient être supérieurs en RVI par rapport aux autres conditions, et resteront plus élevés dans le temps (H3.1, mesure différée à deux semaines). Par ailleurs, une progression significative est attendue entre la mesure pré-test et post-test en condition RVI (H3.2). Par ailleurs, le score de la modalité *Flatscreen* (RV non immersive) devrait être supérieur à la modalité DP-A (diaporama audio) qui sera elle-même supérieur à la modalité DP (diaporama).

H4 : Les scores de stéréotypisation (amorce du changement d'attitude, mesure du comportement des enseignants) devraient être supérieurs en condition RVI par rapport aux autres conditions et se maintiendront dans le temps (H4.1). Une progression est également attendue entre le pré-test et le post-test en RVI (H4.2). Par ailleurs, le score de la modalité *Flatscreen* (RV non immersive) devrait être supérieur à la modalité DP-A (diaporama audio) qui sera elle-même supérieur à la modalité DP (diaporama).

3. Matériel et méthode

Cette étude a été validée par le Comité d'éthique de la recherche de Toulouse. Numéro de dossier : 2024_823.

3.1 Participants

L'échantillon était composé de 112 participants ($N = 112$), un effectif permettant d'assurer une puissance statistique satisfaisante (puissance $> 0,80$ pour un effet de taille moyen, $\eta^2 = 0,10$. Évalué avec *G power*). Les participants ont été répartis aléatoirement en quatre groupes expérimentaux. Les critères d'inclusion retenus étaient les suivants : être âgé de 18 à 65 ans (majorité et période d'activité professionnelle maximum) et travaillant dans l'enseignement et l'accompagnement, ne présenter aucun trouble vestibulaire sévère et être novice en RV (moins de trois séances antérieures d'exposition).

Sur 118 participants initialement recrutés, 112 participants ont été retenus (6 exclus dont 2 exclus du traitement des données après passation. Le premier à cause d'un bug majeur et le second à cause d'un abandon en cours de procédure. Les quatre autres étaient des désistements). Les participants étaient âgés de 22 à 62 ans ($M = 36.6$; $ET = 10.0$). L'échantillon étant composé de 53 femmes et 59 hommes. Le niveau scolaire était situé entre le bac et le HDR ($M = 4.3$; $ET = 1.9$). L'ensemble des participants a eu une expérience avec l'enseignement ou la transmission et l'accompagnement de personnes en situation d'apprentissage (notamment 59 enseignants - écoles, lycées et universités). Le reste des participants est composé de métiers proches de l'enseignement, comme des formateurs, des éducateurs...). Comme abordé précédemment, la formation des enseignants au TSA est essentielle car elle améliore la connaissance des particularités cognitives et comportementales des personnes TSA, permettant d'ajuster les pratiques pédagogiques de manière appropriée. Ensuite, elle renforce le sentiment d'auto-efficacité des enseignants, en les aidant à se sentir compétents face aux défis de la scolarisation inclusive. Enfin, les actions de formation cohérentes sont déterminantes pour tenir compte des besoins spécifiques liés au TSA, favoriser une scolarisation inclusive de qualité et soutenir les enseignants dans leur posture professionnelle. 29 participants ont intégré la modalité diaporama (DP), 31 la modalité écran plat ou *flatscreen* (FS), 27 la modalité réalité virtuelle incarnée (RV) et 25 pour la modalité diaporama audio (DP-A). La répartition des sujets a été faite de manière aléatoire. Tous les participants ont répondu volontairement à l'appel de participation contre remise d'un certificat de participation à une sensibilisation au TSA.

Ensuite, 87 participants sur 112 ont répondu au post-test différé (partie 2) avec une répartition de 26 participants pour la modalité diaporama (DP), 24 pour la modalité écran plat ou *flatscreen* (FS), 20 pour la modalité réalité virtuelle incarnée (RV) et 19 pour la modalité diaporama audio (DP-A).

3.2 Application

L'application développée au laboratoire est composée de deux types d'éléments distincts : les éléments inanimés (mobilier) et les éléments animés (agents et avatar). Les éléments inanimés ne sont pas nécessairement inertes car certains d'entre eux seront actifs durant certaines tâches : 6 activités scénarisées sont implémentées et s'enchaînent selon un ordre défini. Ces activités vont permettre d'explorer différentes particularités du TSA. Des pauses pédagogiques ont été intégrées au dispositif, selon un protocole équilibré et conçu pour réduire l'impact des ruptures de l'illusion de plausibilité, Slater (2009) afin de palier au déficit d'acquisition de connaissances factuelles comme le suggèrent Parong et Mayer (2018).

3.3 Équipement

Le matériel utilisé pour cette phase est un casque occlusif stand alone (ou HMD) Meta Quest 3 pour la RVI et un ordinateur portable MSI 15'' pour le *flatscreen* et le DP-A et DP. Pour rappel, le casque occlusif Meta Quest 3 fonctionne sans ordinateur pilote, ce qui renforce son utilisabilité au niveau expérimental. Le dispositif est léger en termes ergonomique et sa mise en place ne requiert que le casque et ses contrôleurs. L'ordinateur portable MSI a été sélectionné pour ses performances sur des activités de type *Gaming* afin d'obtenir une expérience fluide avec un minimum de 60fps (garantie d'une fluidité optimale). Les deux appareils sélectionnés ont des performances et une autonomie qui assurent le bon déroulement des passations sans crainte de coupure ou de rupture de la simulation. Voici les dispositifs :

- Meta Quest 3.
- MSI 15'' (Ecran 15'' et Résolution full HD 1920 x1080px, Processeur : Intel I5 9eme génération. GPU : Nvidia 3060 6Go. RAM : 16 Go. Mémoire interne : 512 Go).

3.4 Procédures

Les passations se sont déroulées sur le schéma suivant : après l'accueil des participants et une brève présentation de rappel, le formulaire de consentement leur a été présenté. Une fois informé de leurs droits et de la possibilité d'arrêter à tout moment et sans justification

l'expérience, les participants étaient invités à prendre en main leurs dispositifs respectifs. Une fois le dispositif connu des participants, la passation pouvait commencer. L'expérience s'est déroulée sur le lieu de travail des participants.

Le temps total de passation était en moyenne de 30 minutes avec un temps maximum en RVI de 10 minutes soit 50% de moins que la recommandation d'exposition maximale (actuellement recommandation à 20 minutes maximum, Anses, 2021).

Les consignes données aux participants étaient les suivantes :

« Avant de commencer, nous allons vous expliquer le déroulement de l'expérience. Vous allez être équipé d'un casque de réalité virtuelle. Si vous n'avez jamais utilisé ce type de matériel, ne vous inquiétez pas : tout est prévu pour que vous soyez accompagné étape par étape. L'objectif est de vous immerger dans un environnement virtuel afin d'étudier certains aspects dans ce contexte. Vous serez assis confortablement pendant toute l'expérience, et vous pourrez interrompre la séance à tout moment si vous ressentez un inconfort, par exemple lié au port du casque ou à une fatigue visuelle. Pendant la tâche, nous vous demanderons simplement de suivre les instructions qui apparaîtront à l'écran ou qui vous seront données par une voix, et d'interagir avec l'environnement virtuel de la manière la plus naturelle possible. Il n'y a pas de bonne ou de mauvaise réponse, l'important est de participer normalement. À la fin de l'expérience, nous vous poserons quelques questions simples pour recueillir vos impressions et vos ressentis. »

- Accueil des participants et présentation du dispositif - 3 minutes.
 - 4 modalités : groupe test DP - groupe DP-A - casque occlusif RVI - écran plat FS.
- Pré-questionnaire sur l'attitude initiale des participants (16 questions - TSA) - 5 minutes.
- Tutoriel - 4 minutes.
- Expérimentation - 6 minutes (voir tableau récapitulatif).
- Questionnaire sur l'attitude post expérience des participants (40 questions - Cinétose - Présence - agentivité - TSA) - 10 minutes.
- Clôture et remerciement - prise de contact pour troisième questionnaire Post-immersion, mémorisation moyen terme - 2 minutes.

3.5 Mesures

Afin d'assurer la qualité des mesures, une fiabilité interne minimale de $\alpha \geq 0,70$ a été ciblée pour chaque bloc de questionnaires, qui n'a malheureusement pas été atteinte pour tous les blocs (éléments abordés dans la partie « procédures »).

Les questionnaires mesuraient 5 variables¹⁷. Le pré-questionnaire mesurait uniquement les connaissances factuelles et la stéréotypisation (amorce du changement d'attitudes) avec 16 questions. Ensuite le questionnaire immédiat mesurait tout d'abord la cinétose, à savoir si l'utilisateur avait été affecté par l'un ou plusieurs des symptômes de la cinétose. Il mesurait ensuite le degré de présence, à savoir si l'utilisateur s'était senti immergé dans la simulation. Ensuite venait la mesure d'agentivité, à savoir si l'utilisateur s'était senti maître de ses mouvements et avait senti avoir un impact sur l'environnement grâce à ses actions. L'empathie envers les agents et l'avatar des sujets était ensuite mesurée. Enfin, les connaissances factuelles et la stéréotypisation (amorce du changement d'attitudes) étaient de nouveau mesurées pour voir si la simulation avait eu un impact sur la compréhension et la perception de l'utilisateur concernant le TSA.

Le post-questionnaire mesurait lui aussi uniquement les connaissances factuelles et la stéréotypisation (amorce du changement attitudes) mais après un délai minimum de deux semaines après la simulation pour mesurer la différence de rétention selon les modalités.

3.5.1 Mesures de la cinétose : *Simulator Sickness Questionnaire*

Une adaptation du *Simulator Sickness Questionnaire* (SSQ) de Kennedy et al. (1993) et traduit par Bouchard et al. (2007) a été utilisé pour mesurer la cinétose avec suppression de certains items (le choix de recueillir le degré de cinétose avec moins d'items a été fait car l'étude ne porte pas sur la cinétose directement mais vérifie néanmoins sa présence). Les mesures sont faites sur une échelle de Likert en quatre point : Pas du tout ; Un peu ; Modérément ; Sévèrement. Nous avons conservé sept items.

Les items suivants ont été supprimés du questionnaire pour des raisons de durée. En effet, l'ensemble des items de tous les questionnaires réunis ne permettaient pas de rentrer dans l'emploi du temps accordé par les participants, souvent sur leur lieu de travail et en activité.

¹⁷ Voir annexe D.1 : [Questionnaires de référence issus de la littérature](#) p. 354.

Voici les items supprimés (après concertation lors des ateliers itératifs et avec les experts du projets) : 5. Difficulté à faire le focus. 6. Augmentation de la salivation. 7. Transpiration. 10. Impression de lourdeur dans la tête. 12. Étourdissement les yeux ouverts. 13. Étourdissement les yeux fermés. 14. Vertige. 15. Conscience de l'estomac. 16. Rots.

Voici un exemple d'item conservé : Consignes : Encerclez à quel point chaque symptôme ci-dessous vous affecte en ce moment : Items 2 « fatigue » - item 5 « nausée ».

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.877.

3.5.2 *Mesure de la présence et Sense of Embodiment - Incarnation*

Fribourg (2020) propose, pour mesurer la présence, une adaptation de plusieurs échelles comme celle de Botvinick et Cohen (1998). Une traduction en français adaptée au projet a été effectué. Pour l'incarnation, les travaux de Falcone (2022) proposent une approche simplifiée pour obtenir une bonne mesure du SoE avec des mesures explicites qui correspondent aux entretiens et questionnaires de Peck et Gonzalez-Franco (2021). Une adaptation de certains items du questionnaire de Peck et Gonzalez-Franco (2021) est utilisée, ainsi que des items issus des travaux de Fribourg afin de mixer présence et SoE. Une adaptation et traduction non validée a été effectuée pour pouvoir utiliser ces items en français. Les participants répondaient sur une échelle de Likert en 7 point pour chaque item.

Voici deux exemples d'items : item 1. Dans quelle mesure y a-t-il eu des moments, au cours de l'expérience, où la classe virtuelle était la réalité pour vous ? (Est ce que, par moment, la classe virtuelle vous a paru réelle, comme si c'était la réalité ?) - item 5. Pendant l'expérience, est ce que vous avez pensé être réellement dans la salle de classe (même ponctuellement) ?

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.627.

3.5.3 *Mesure de l'agentivité*

Pour les mesures d'agentivités, Fribourg (2020) propose des adaptations de plusieurs sous-échelles dans ses travaux qui permettent d'obtenir une version courte. Une division entre l'agentivité du corps et des bras est effectuée. Ces items permettent d'observer la présence d'agentivité ou non. Le but n'étant pas de mesurer le niveau exact d'agentivité mais plutôt de vérifier son existence. Il sera éventuellement envisageable de proposer une version plus complète pour mesurer l'agentivité. A cause de l'absence de corps de l'avatar, l'agentivité des mains virtuelles sera mesurée uniquement. Une traduction non validée a été effectuée pour

pouvoir utiliser ces items en français. L'échelle de Likert en 7 point est appliquée à toutes les questions.

Voici deux exemples d'items : 2. J'ai eu l'impression de contrôler le mouvement des mains virtuelles - 5. J'ai senti que les mains virtuelles n'étaient pas les miennes.

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.791.

3.5.4 *Mesure de l'empathie*

Pour l'empathie, une adaptation de certains items de la *Basic Empathie Scale A*, BES-A a été mise en place. Elle est issue d'une adaptation française par D'Ambrosio et al. (2009) de la BES de Jolliffe et Farrington (2006) sur une échelle de Likert en 7 points (Q3 - questions 6 à 12). La BES-A est un questionnaire auto-rapporté de 20 items évaluant deux dimensions de l'empathie : cognitive (capacité à comprendre les émotions d'autrui) et affective (tendance à partager ou ressentir ces émotions). Dans ce travail, la structure d'origine a été adaptée tout comme les items afin de correspondre à une application en environnement virtuel. Les items traitant d'autres types d'environnements ne concernant pas l'étude ont été supprimés tout comme les items traitant de personnes. Une adaptation non validée a ainsi été effectuée pour pouvoir utiliser ces items. L'échelle de Likert en 7 point est appliquée à toutes les questions. Les questions s'enchaînent directement après celles sur l'agentivité.

Voici deux exemples d'items : 7. A la fin de l'expérience, j'ai été affecté par la réaction de mon avatar face aux autres élèves - 9. J'ai ressenti les émotions de mon avatar comme si je vivais vraiment cette expérience.

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.935.

3.5.5 *Les mesures du TSA : Connaissances factuelles et stéréotypisation (amorce du changement d'attitudes)*

L'analyse de fiabilité révèle une cohérence interne insuffisante, avec un α de Cronbach de 0,550 pour les items QCM et de 0,489 pour les questionnaires d'attitudes sous échelles de Likert. Ces valeurs, nettement en dessous du seuil acceptable de 0,70, limitent la robustesse interprétative des résultats. La batterie de questionnaires complète est accessible¹⁸.

¹⁸ Voir annexe D.2 : [Batterie de mesures - expérimentation 1](#) D.2p. 359.

Dans le cadre du TSA, il est difficile de mesurer un changement d'attitude. Les mesures utilisées ici n'ont pas pour vocation de mesurer précisément les changements mais d'observer s'il y a une amorce de ces changements avec une évolution de la perception des stéréotypes (pour permettre une amorce du changement d'attitudes). Le choix le plus pertinent est une passation avec peu d'items, qui permettront de montrer l'amorce d'un changement et ainsi d'ouvrir la voie à des recherches plus approfondies sur le niveau et l'implication de ces changements. Le but est de montrer que le protocole permet une meilleure compréhension du TSA, ce qui est la première étape avant d'amorcer un changement d'attitude. La compréhension du TSA passera par une évaluation des connaissances factuelles puis le niveau de stéréotypisation sera mesuré (pour voir s'il y a une amorce du changement d'attitudes).

Pour mesurer les connaissances factuelles et le niveau de stéréotypisation (amorce de changement d'attitudes) des participants vis-à-vis du TSA, une adaptation de certains items de l'AAAS (*Autism Attitude Acceptance Scale*) de Kim (2020) a été effectuée. Elle était composée de 21 items divisés en 3 catégories (cognitive, comportementale et affective) et proposant deux axes de résultats : connaissances et stéréotypisation. Les modalités de réponses étaient un QCM (3 réponses possibles) et une échelle d'attitude sous échelle de Likert (en 7 point). L'AAAS de Kim (2020) est un instrument de mesure de l'acceptation du TSA par les adultes neurotypiques sur la base du cadre de neurodiversité. Il permet d'examiner les associations entre l'acceptation du TSA et les variables présumées. Une adaptation a été effectuée afin de pouvoir utiliser le test en français. Cette traduction n'est pas validée et il conviendra de proposer une version standardisée ultérieurement.

Le reste des items a été conçu en lien avec le contenu des connaissances. Ils sont issus des travaux effectués dans le cadre de cette thèse en ateliers de design thinking du GEM la bulle de Rodez et des rencontres avec des experts TSA comme Mesdames Willay et Bouvet. Quatre versions ont ainsi été proposées avant d'obtenir la version finale. Chacune de ses versions a été testée sur le terrain avec des utilisateurs volontaires représentatifs de la population cible (personnels enseignants).

Ces items supplémentaires ont été conçus et ajoutés pour représenter le plus fidèlement la perception et les connaissances du TSA dans le cadre d'une pratique d'enseignement. Après analyse des items existants, il en a été conclu que certains aspects spécifiques n'étaient pas pris en compte et qu'il fallait en ajouter certains. Voici un exemple d'item ajouté (en souligné, la proposition juste) : Lorsqu'un enseignant voit qu'une personne avec TSA SDI est concentrée

sur une tâche, il est conseillé de : a : L'interpeller, même avec bienveillance. b : L'encourager pour qu'elle avance plus vite. c : La laisser faire et attendre qu'elle ait finie.

Les items ont été divisés en deux thèmes : les connaissances factuelles du TSA (incluant les connaissances des comportements à adopter avec une personne avec TSA en tant qu'enseignant) et le niveau de stéréotypisation (incluant les représentations des capacités des personnes étudiantes).

La démarche d'élaboration a été la suivante :

- Recueil d'information auprès des personnes avec TSA.
- Proposition de question.
- Tests des questions.
- Amélioration des questions.
- Tests des version améliorées.
- Répétition du processus (4 itérations).

Le test est composé d'une série de questions à choix multiples et d'un questionnaire d'attitude sous échelles de Likert sur les contenus abordés.

Les différents types de connaissances sont classées en deux modalités représentant chacune un ensemble de concepts :

- C = Connaissances factuelles du TSA.
- S = Niveau de stéréotypisation.

4. Résultats attendus et contribution

Plan d'analyse statistique :

Afin de comparer les quatre conditions indépendantes (RVI, FS, DP-A et DP), des analyses de la variance (ANOVA) à mesures indépendantes ont été conduites sur les scores de cinétose, de présence, d'agentivité et d'empathie. En ce qui concerne l'évolution des connaissances factuelles et niveau de stéréotypisation, des ANOVA à mesures répétées ont été conduites avec les quatre conditions d'apprentissage comme mesure indépendante et le temps de passation comme mesure dépendante (pré-test, test immédiat, test différé). Les tailles d'effet sont rapportées avec l'éta au carré (η^2_p) de Cohen (2013). Les seuils d'interprétation sont les suivants : un η^2 entre 0,01 et 0,059 indique un effet de petite taille, entre 0,06 et 0,139 un effet

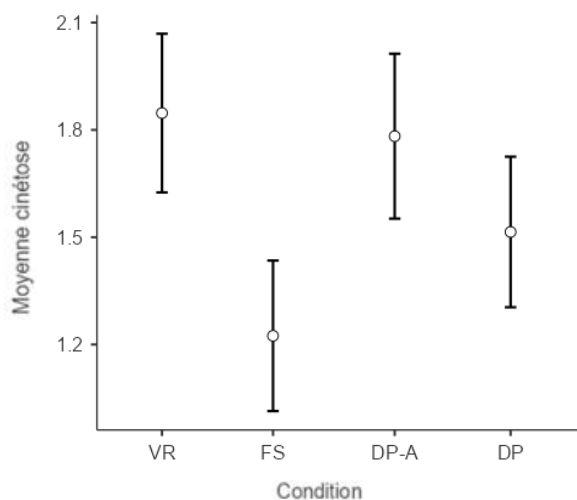
moyen, et à partir de 0,14 un effet de grande taille. Les analyses statistiques issues de Jamovi sont accessibles¹⁹.

L'analyse des données issue de cette première expérimentation a été conduite à l'aide d'une analyse de variance (ANOVA). L'attention a été portée, dans un premier temps, sur les dimensions de la cinétose, de la présence, de l'agentivité et de l'empathie, afin de vérifier si la simulation permettait effectivement de les induire, conformément aux hypothèses H1 et H2. Par la suite, les scores de connaissances factuelles et de niveau de stéréotypisation du TSA ont été mesurés et comparés à l'aide des réponses obtenues au questionnaire à choix multiples (QCM) ainsi qu'au questionnaire d'attitude sous échelle de Likert, portant sur les connaissances factuelles et le niveau de stéréotypisation sous les hypothèses H3 à H4.

4.1 Cinétose

Figure 28

Score de cinétose



Les scores de cinétose sont tous inférieurs à 2 (score inférieur au segment « un peu », considéré comme très convenable pour une expérience de RVI (fig. 28). On peut observer une différence significative soit une Anova avec $F(2, 108) = 7.04, p < .001, \eta^2p = 0.164$, ce qui indique un effet important, Cohen (2013). Les comparaisons post-hoc montrent une différence entre le groupe VR et les groupes FS (différence significative entre VR ($M = 1,85$) et FS ($M =$

¹⁹ Voir annexe E.1 : [Résultats expérimentation 1](#) p. 374.

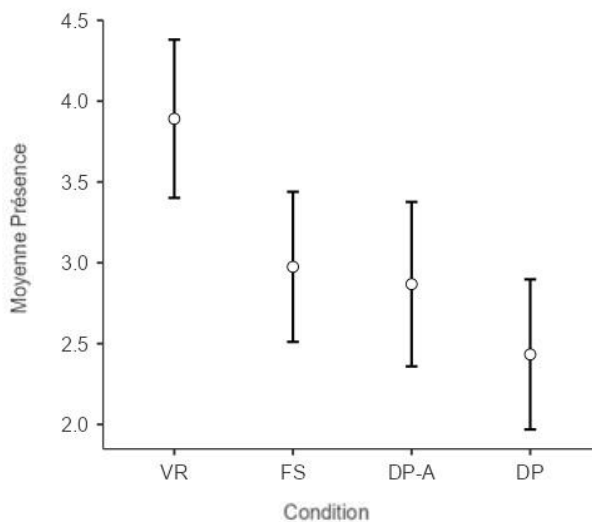
1,22), $p < .001$), ainsi qu'entre le groupe FS et le groupe DP-A (différence significative entre FS ($M = 1,2$) et DP-A ($M = 1,78$), $p = 0.02$).

Cette différence était attendue car bien que le score de RV soit inférieur à 2, l'inconfort provoqué par le port du casque a été relevé par les sujets. Cet inconfort était minime mais néanmoins présent. Ces résultats confirment H1 mais permettent de montrer que la simulation a un effet minime sur la cinétose due essentiellement à de l'inconfort matériel. La différence avec le groupe DP-A était inattendue tout comme l'absence de différences entre DP et VR ; et entre DP-A et VR.

4.2 Présence

Figure 29

Score présence



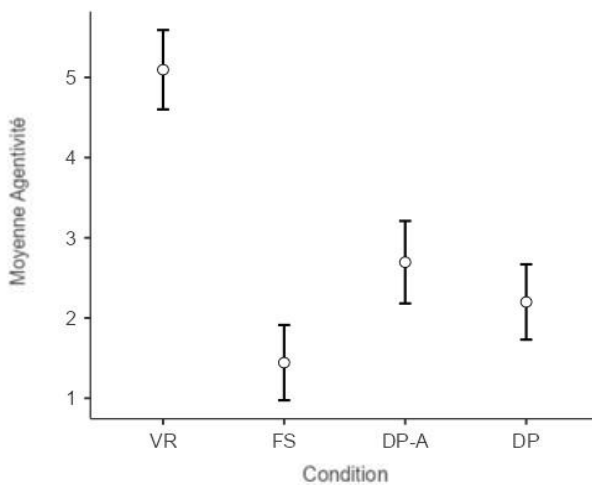
Une ANOVA a révélé une différence significative entre les groupes soit, $F(2, 108) = 6,34$ et $p < .001$). La taille de l'effet, η^2p , était de 0.15, ce qui indique un effet important. Les comparaisons post-hoc montrent une différence entre le groupe VR et tous les autres groupes (différence significative entre VR ($M = 3,89$) et FS ($M = 2,96$) et DP-A ($M = 2,87$) et DP ($M = 2,43$), $p < .05$). Le constat d'un score de présence nettement plus élevé en condition de RVI (fig. 29) conforte H2.1, ce qui correspond à la littérature. D'après les résultats, l'induction de présence est supérieure en condition de RVI par rapport aux autres conditions. Ces résultats confirment H2.1, la présence générée par un simulateur de RVI semble bien être supérieur par rapport à une présentation classique.

Ces résultats semblent valider les choix de conception sur l'implémentation de présence au sein du dispositif, comme le recommande la littérature avec les travaux de Markranski et Petersen (2021) avec une vue en première personne et les travaux de Gorisse (2019) pour une réactivité fluide afin que l'utilisateur se sente investi dans son avatar.

4.3 Agentivité

Figure 30

Score d'agentivité



Une ANOVA a révélé une différence significative entre les groupes soit, $F(2, 108) = 41,1$ et $p < .001$). La taille de l'effet, éta au carré partiel (η^2p), était de 0.534, ce qui indique un effet important. Les comparaisons post-hoc montrent une différence entre le groupe VR et tous les autres groupes (différence significative entre VR ($M = 5,10$) et FS ($M = 1,47$) et DP-A ($M = 2,2$) et DP ($M = 2,70$), $p < .001$).

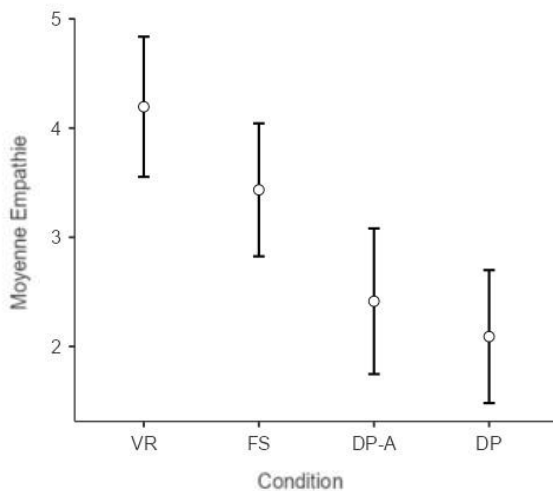
Le constat d'un score d'agentivité nettement plus élevé en condition de RVI (fig. 30) conforte H2.2, ce qui correspond à la littérature. D'après ces résultats, l'induction d'agentivité est supérieure en condition de RVI par rapport aux autres conditions. Ces résultats confirment donc une partie de H2.2, l'agentivité générée par un simulateur de RVI semble bien être supérieur par rapport à une présentation classique. Le score de la question 3 (« J'ai eu l'impression que les mains virtuelles avaient une volonté propre ») a soulevé une interrogation. Avec une légère tendance supérieure (proche de 5, démontrant une volonté de réponse proche du « je ne sais pas » par la majorité des participants), cette question a été mal comprise (plusieurs participants ont posé des questions lors des passations, optant pour une réponse

neutre dans la plupart des cas, ce qui impacte l'analyse de fiabilité de cet item). Elle sera revue et adaptée pour la prochaine passation.

4.4 Empathie

Figure 31

Score d'empathie



Une ANOVA a révélé une différence significative entre les groupes soit $F(2, 108) = 9,52$ et $p < .001$). La taille de l'effet, éta au carré partiel (η^2p), était de 0.202, ce qui indique un effet important. Les comparaisons post-hoc montrent une différence entre le groupe VR et les groupes DP-A et DP (différence significative entre VR ($M = 4,19$) et DP-A ($M = 2,41$) et DP ($M = 2,03$), $p < .001$), ainsi qu'entre le groupe FS et le groupe DP (différence significative entre FS ($M = 3,44$) et DP ($M = 2,03$), $p = 0.013$).

Le constat d'un score d'empathie plus élevé en condition de RVI (fig. 31) semble conforter H2.3. Néanmoins, bien que ces résultats confirment que l'induction d'empathie envers l'avatar est supérieure en condition de RVI par rapport aux conditions non immersives sur l'ensemble des items, il est important de remarquer que la modalité FS montre également une différence significative par rapport à la modalité DP.

4.5 Compréhension et stéréotypisation du TSA

Au vu des α de Cronbach de l'analyse de fiabilité, les résultats de la partie compréhension (connaissances factuelles) et stéréotypisation (amorces du changement d'attitudes) du TSA doivent être considérés avec précaution. Pour rappel, l'analyse de fiabilité

révèle une cohérence interne insuffisante, avec un α de Cronbach de 0,550 pour les items QCM et de 0,489 pour les questionnaires d'attitudes sous échelles de Likert. Ces valeurs, nettement en dessous du seuil acceptable de 0,70 limitent la robustesse interprétative des résultats.

Les items ont été répartis selon 2 modalités :

- Une partie questionnaire à choix multiple (QCM).
- Une partie questionnaire d'attitudes sous échelle de Likert.

Pour rappel, le test a été administré en trois temps :

- Pré questionnaire : les questions initiales de positionnement sur le TSA.
- Questionnaire immédiat : les mêmes questions après simulation, pour mesurer l'évolution des connaissances.
- Questionnaire différé : les mêmes questions, après un minimum de deux semaines d'intervalle, pour mesurer le maintien de ces connaissances dans le temps.

4.5.1 Questionnaire de connaissances (QCM)

Tableau 4

Moyennes et écarts-types des scores de connaissances par groupe et moment de mesure

	Groupe Pré-questionnaire		Questionnaire immédiat		Questionnaire différé	
	M	ET	M	ET	M	ET
VR	2,08	0,47	2,41	0,36	2,42	0,38
FS	1,66	0,67	2,48	0,32	2,38	0,42
DP-A	2,34	0,43	2,39	0,38	1,82	0,49
DP	1,79	0,72	2,49	0,34	2,37	0,45

Note. VR = Réalité virtuelle ; FS = Flatscreen ; DP-A = Diaporama audio ; DP = Diaporama.

L'ANOVA à mesures répétées révèle une interaction significative entre la condition expérimentale et le moment de mesure, $F(6, 182) = 7,82, p < .001, \eta^2p = 0,205$, indiquant un effet de taille importante. Cette interaction suggère que l'évolution des performances entre les différents moments de mesure varie selon la condition expérimentale (tableau 4).

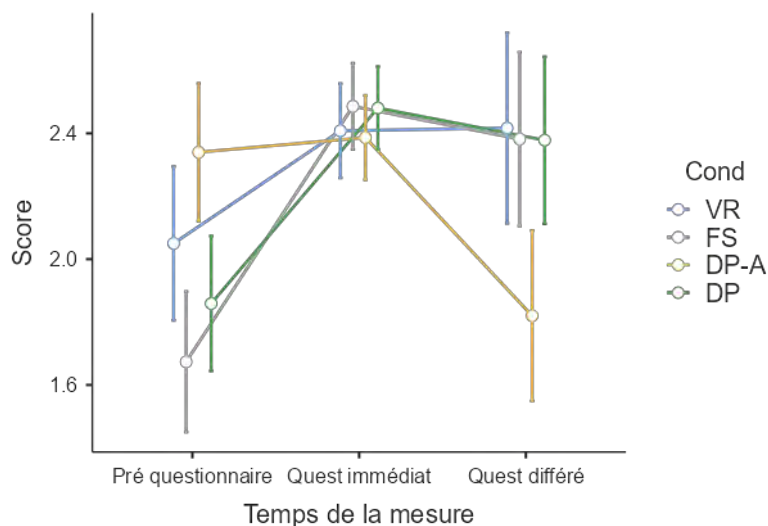
Au niveau des gains de connaissances en mesure immédiate (pré-test vs post-test immédiat), les comparaisons post-hoc révèlent des patterns d'apprentissage différentiels selon les conditions. On peut observer une amélioration significative pour le groupe FS ($p < .001$) et le groupe DP ($p < .001$). En revanche, aucune augmentation significative n'a été observée pour les groupes VR et DP-A. De plus il n'y a pas de différence en test immédiat pour chaque condition, indiquant que la condition VR n'as pas eu d'effets particuliers (fig. 32).

Au niveau de l'analyse de la rétention à long terme (pré-test vs post-test différé), on peut observer une amélioration qui se maintient à long terme pour certaine condition. Ainsi, pour le groupe FS, on observe une amélioration significative maintenue entre le pré-questionnaire et le questionnaire différé ($M = 2,38$), $p = .002$. Pour le groupe DP on observe une amélioration significative maintenue entre le pré-questionnaire ($M = 1,79$) et le questionnaire différé ($M = 2,37$), $p = .006$. Enfin, pour le groupe VR, on observe une amélioration significative entre le pré-questionnaire ($M = 2,08$) et le questionnaire différé ($M = 2,42$), $p = .009$. Le groupe DP-A présente une tendance à la baisse entre le pré-questionnaire ($M = 2,34$) et le questionnaire différé ($M = 1,82$), sans toutefois atteindre le seuil de significativité statistique ($p = .064$).

En observant ensuite les comparaisons entre conditions au post-test différé, on peut voir une différence significative qui émerge au questionnaire différé entre le groupe DP-A et les autres groupes (avec des différences DP-A vs FS : $p = .002$; DP-A vs VR : $p = .014$ et enfin DP-A vs DP : $p = .002$). Le groupe DP-A présente des performances inférieures aux autres groupes au post-test différé, ce qui suggère un pattern de rétention différent.

Figure 32

Score du QCM



Ces résultats indiquent que les conditions FS et DP facilitent un apprentissage immédiat significatif qui se maintient dans le temps. La condition VR ne montre pas d'amélioration immédiate significative (peut-être due à un niveau de base plus élevé) mais une amélioration sur le long terme. Le groupe DP-A présente un pattern particulier avec des performances initiales élevées mais sans amélioration significative immédiate, suggérant un effet plafond potentiel dû à son niveau initial déjà élevé, puis qui diminuent dans le temps. Comme on peut le voir sur le graphique (fig. 32), le groupe DP-A possède un pattern différent des autres groupes. On peut voir également l'amélioration des groupes FS et DP en immédiat et à long terme, ainsi que l'amélioration à long terme du groupes VR. Ces résultats ne permettent pas d'accepter H3 : la RV n'a pas d'effet accrue sur les connaissances par rapport aux autres modalités (excepté DP-A mais qui présente un comportement atypique).

4.5.2 Questionnaire d'attitude sous échelle de Likert

Tableau 5

Moyennes et écarts-types des scores de connaissances par groupe et moment de mesure

Groupe	Pré-questionnaire	Questionnaire immédiat		Questionnaire différé		
	M	ET	M	ET	M	ET
VR	4,85	0,57	5,01	0,36	5,01	0,48
FS	4,65	0,67	5,03	0,80	4,36	0,57
DP-A	4,80	0,68	4,85	0,60	4,85	0,56
DP	4,76	0,57	4,85	0,66	4,82	0,56

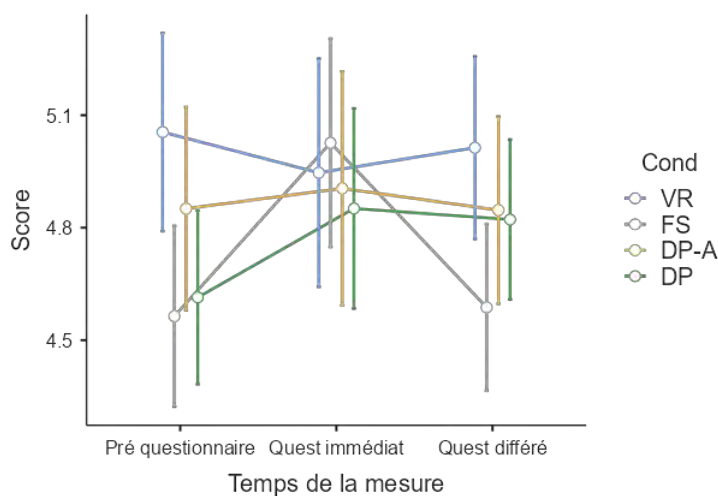
Note. VR = Réalité virtuelle ; FS = Flatscreen ; DP-A = Diaporama audio ; DP = Diaporama.

L'ANOVA à mesures répétées révèle l'absence d'effets significatifs : ni effet principal du temps ($F(2, 170) = 2,75, p = .067, \eta^2p = 0,031$), ni effet principal de la condition ($F(3, 85) = 1,49, p = .224, \eta^2p = 0,050$), ni interaction temps \times condition ($F(6, 170) = 2,10, p = .056, \eta^2p = 0,069$). Bien que l'interaction approche le seuil de significativité ($p = .056$), elle demeure statistiquement non significative. L'analyse des comparaisons post-hoc montre que, malgré l'absence de significativité globale, il existe une différence significative isolée. En effet, le

groupe FS montre une différence significative entre le questionnaire immédiat et le questionnaire différé ($p = .018$), indiquant une diminution des performances au post-test différé. Malgré cet élément isolé, toutes les autres comparaisons sont non significatives ($p > .05$). Cela inclut toutes les comparaisons pré-test vs post-test immédiat pour chaque condition, toutes les comparaisons pré-test vs post-test différé pour chaque condition, toutes les comparaisons entre conditions à chaque moment de mesure ainsi que toutes les autres comparaisons (questionnaire immédiat vs questionnaire différé). Le tableau 5 montre les moyennes et écarts types.

Figure 33

Score questionnaire d'attitude



Ces résultats indiquent que la RVI ne présente ni avantage ni désavantage significatif par rapport aux modalités traditionnelles (H4 doit être rejetée). Cette équivalence d'efficacité pourrait être interprétée positivement, suggérant que la RVI constitue une alternative viable aux approches pédagogiques conventionnelles sans pour autant améliorer l'apprentissage dans le domaine testé.

Comme on peut le voir sur le graphique (fig. 33), la condition FS est la seule condition présentant une tendance à la hausse (non significative) en questionnaire immédiat. Les autres conditions ne présentent pas de différence significative. De plus, les résultats des groupes VR, DP-A et DP reviennent à leur état initial lors du questionnaire différé tandis que le groupe FS présente une diminution significative par rapport au questionnaire immédiat et se situe en dessous de son niveau initial.

5. Discussion

Il est important de préciser que le bilan des résultats relatifs aux gains de connaissances et au changement d'attitudes aucune conclusion robuste en raison de l'analyse de fiabilité insuffisante pour les items de connaissance et d'attitude ($\alpha < .70$). La portée interprétative de ces résultats est donc à considérer avec une certaine prudence.

5.1 Synthèse de la validation des hypothèses

H1 - Cinétose supérieure en condition RVI - État : validée. Conformément aux attentes, les scores de cinétose se révèlent supérieurs en condition RVI par rapport aux autres modalités, bien que celle-ci ait été globalement maîtrisée. Des inconforts liés au port du casque ont néanmoins été relevés, confirmant les préoccupations ergonomiques classiques associées à la RVI.

H2 - Supériorité de la RVI sur présence, agentivité et empathie - État : validée. Les scores de présence (H2.1), d'agentivité (H2.2) et d'empathie (H2.3) sont significativement supérieurs en condition RVI, avec des effets cohérents avec la littérature. La hiérarchie attendue entre modalités (RVI > FS > DP-A > DP) se confirme partiellement.

H3 - Connaissances factuelles supérieures en RVI - État : non validée. L'analyse révèle une absence de différences significatives entre les conditions pour l'acquisition et le maintien des connaissances factuelles sur le TSA. Aucune progression significative n'est observée entre pré-questionnaire et questionnaire immédiat en condition RVI (H3.2) et aucune supériorité n'émerge au questionnaire différé (H3.1). Néanmoins, les conditions ont amélioré le gain au regard du test différé mais pas pour le DP-A (certainement car les résultats initiaux sont trop élevés au pré-test, réduisant donc les chances de progression).

H4 - Réduction de la stéréotypisation en RVI - État : non validée. Malgré une réduction partielle de la stéréotypisation observée (H4.1), avec une amélioration des connaissances pratiques concernant les enseignants (H4.2), on ne peut pas valider H4 en raison d'un effet d'interaction entre condition d'apprentissage et temps de mesure marginalement significatif.

5.2 Interprétation des résultats

Les résultats sur la présence, l'agentivité et l'empathie corroborent les méta-analyses montrant la capacité supérieure de la RVI à générer ces états psychologiques par rapport aux modalités non immersives comme le suggère Makransky et Petersen (2021). Ces dimensions constituent les effets les plus robustes de cette expérimentation.

La présence est sujette à une gradation conforme aux attentes. La RVI génère un sentiment de présence significativement supérieur à toutes les autres modalités, confirmant pleinement H2.1. La gradation observée ($RVI > FS > DP-A > DP$) démontre que le degré d'immersion technologique influence directement l'intensité de cette expérience subjective. Fait notable, même les modalités non immersives génèrent des niveaux mesurables de présence, suggérant que des facteurs narratifs ou le design pédagogique peuvent également contribuer à l'émergence de ce sentiment, même en l'absence d'immersion technologique complète. Ces résultats valident les choix de conception du dispositif, notamment l'utilisation d'une vue en première personne et d'une réactivité fluide (Makransky & Petersen, 2021 ; Gorisse, 2019).

L'agentivité possède l'effet le plus spectaculaire. L'agentivité présente l'effet le plus marqué de l'expérimentation ($\eta^2p = 0,534$), avec une supériorité massive de la RVI par rapport à toutes les autres modalités. Cette différence s'explique par la congruence sensori-motrice unique offerte par la RVI : l'utilisation de mains virtuelles contrôlées naturellement crée une boucle perception-action impossible à reproduire dans les autres modalités. Les trois autres conditions présentent des scores faibles et homogènes, suggérant qu'en l'absence d'interaction corporelle directe, le sentiment d'agentivité reste limité quelle que soit la richesse visuelle. L'agentivité apparaît donc comme une dimension spécifiquement dépendante de l'interactivité motrice. Ces résultats confirment que l'utilisateur doit être à la source de ses actions pour développer un SoE comme le suggèrent Haggard (2017) et Kilteni et al. (2012).

L'empathie présente un pattern intermédiaire révélateur et plus nuancé. Si la RVI génère significativement plus d'empathie que les modalités traditionnelles (DP-A et DP), la modalité FS occupe une position intermédiaire intéressante, produisant également un niveau d'empathie significativement supérieur à la modalité la moins immersive. Cette gradation suggère que l'empathie ne dépend pas exclusivement du niveau d'immersion maximal. Ces résultats rejoignent les travaux de Riva et al. (2007) sur la capacité de la RV à engendrer des émotions.

Pour synthétiser, ces résultats démontrent que la RVI excelle dans la génération d'expériences psychologiques immersives, avec une hiérarchie claire pour la présence et l'agentivité, et un continuum plus nuancé pour l'empathie. Les choix de conception (simulation sans vection, contrôle naturel des interactions, tutoriel épuré) ont permis de générer ces effets tout en maintenant un niveau de cinétose acceptable, conformément aux recommandations d'Arnaldi et al. (2018).

Au niveau de l'apprentissage et de la rétention, l'étude révèle une équivalence qui interroge. En effet, on observe une absence d'effets différentiels entre modalités. L'analyse révèle une absence d'effets principaux significatifs sur l'acquisition et la rétention des connaissances factuelles. Cette absence doit être interprétée en tenant compte des différences initiales entre groupes au pré-test, avec notamment le groupe DP-A présentant le niveau le plus élevé et le groupe FS le plus faible. Trois interprétations principales émergent. Le premier est une équivalence réelle des modalités pédagogiques, s'inscrivant dans la lignée des travaux questionnant l'avantage des technologies immersives (Clark, 1983, revisité). Le second suggère une puissance statistique limitée, la taille d'effet modérée pour l'interaction ($\eta^2_p = 0,069$) suggérant que des effets pourraient exister sans être détectables avec l'échantillon actuel. Enfin, le dernier montre une variabilité inter-individuelle importante masquant potentiellement des effets de groupe.

Ces résultats demeurent encourageants malgré tout. L'absence d'effet sur les connaissances factuelles rejoint les préoccupations de Parong et Mayer (2018) concernant les déficits en RVI. Cependant, contrairement à leurs conclusions, aucun effet délétère n'est observé, suggérant une implémentation appropriée du modèle CAMIL. Alors que la littérature souligne les difficultés de transmission des connaissances en RV, le dispositif conçu dans ce travail parvient à maintenir une efficacité équivalente aux modalités traditionnelles. L'absence d'effet négatif contraste avec les inquiétudes de Poupard et al. (2024) concernant la surcharge cognitive, indiquant une gestion appropriée de la charge cognitive, probablement facilitée par les pauses pédagogiques intégrées. Il convient néanmoins de relever le cas particulier de la modalité *Flatscreen*. Le groupe FS présente un pattern unique avec une diminution significative entre le post-test immédiat et différé, suggérant une rétention moins stable. Cette observation interroge la position intermédiaire de cette modalité.

Enfin, on peut observer une remise en question du continuum d'immersion. L'hypothèse d'une hiérarchie claire RVI > FS > DP-A > DP n'est que partiellement vérifiée. Cette

observation suggère que le degré d'immersion ne constitue pas un prédicteur linéaire de l'efficacité pédagogique, et que d'autres facteurs pourraient jouer un rôle déterminant.

Au niveau de l'amorce du changement d'attitudes, les résultats sont mitigés. Tout d'abord il a été observé un effet plafond déterminant. Les résultats révèlent une absence d'effets significatifs sur le changement d'attitudes envers le TSA malgré les niveaux élevés d'empathie générés par la RVI. L'analyse des valeurs initiales révèle des scores déjà très élevés (entre 4,65 et 4,85 sur l'échelle), indiquant un effet plafond majeur. Les participants présentaient déjà des attitudes très favorables avant l'intervention, laissant une marge de progression extrêmement limitée. Cette configuration constitue une limite méthodologique importante expliquant largement l'absence d'amélioration observable.

Au-delà de l'effet plafond et des problèmes de fiabilité des mesures ($\alpha < .70$), plusieurs facteurs peuvent expliquer cette absence d'effet. En premier lieu, existe une temporalité inadaptée. Les changements d'attitudes profonds, particulièrement concernant des représentations sociales ancrées, nécessitent potentiellement des expositions répétées. Une session unique pourrait être insuffisante pour déclencher une transformation durable. Ensuite, la dissociation empathie-attitude pourraient apporter certains éléments de réponses. L'empathie situationnelle générée durant l'expérience ne se traduit pas automatiquement en changement d'attitude stable. Il existe un écart entre l'émotion ressentie in situ et l'intégration cognitive de nouvelles représentations. Un autre paramètre important est la complexité des stéréotypes. Les stéréotypes concernant le TSA sont multidimensionnels et profondément enracinés (Willay, 2020). Une intervention unique, même innovante, pourrait nécessiter un accompagnement complémentaire pour produire des effets mesurables. Enfin, une inadéquation des mesures pourrait être une cause essentielle. Les items pourraient ne pas capturer les nuances du changement attitudinal ou mesurer des dimensions trop générales pour détecter des évolutions fines.

Au niveau des implications théoriques, une validation partielle du modèle CAMIL est envisageable. En effet, le modèle CAMIL semble avoir permis d'éviter les écueils de distraction couramment observés en RVI, mais ne génère pas l'avantage d'apprentissage escompté. Cette observation interroge les conditions d'efficacité du modèle au-delà de la simple prévention des effets négatifs. Le CAMIL constitue un cadre pertinent pour concevoir des expériences qui ne nuisent pas à l'acquisition de connaissances mais son potentiel d'amélioration significative des performances reste à démontrer.

Néanmoins, la médiation par l'empathie est un mécanisme non confirmé. L'hypothèse selon laquelle l'empathie générée par la RVI constituerait un levier de changement d'attitude n'est pas confirmée empiriquement. Cette absence de médiation suggère que le modèle théorique postulant un lien direct doit être complexifié pour intégrer des variables médiatrices ou modératrices supplémentaires (engagement cognitif, réflexivité post-expérience, accompagnement pédagogique, expositions répétées).

5.3 Limites

Plusieurs éléments d'interprétation sont envisageables. D'une part, des limites dans la conception et le développement du contenu pédagogique, en particulier pour la version RVI. Il est possible que le contenu DP et DP-A (bien qu'ils n'aient pas été spécifiquement pensés pour ce support) aient été mieux structurés, tandis que la version RVI, conçue sur le même socle de contenu, n'aurait pas tiré parti des spécificités du média immersif. Cette observation rejoint les préconisations pour une conception adaptée, comme le suggère Fowler (2015) sur le modèle de Dalgarno et Lee ; un modèle alternatif plus ancien au CAMIL d'apprentissage en environnement virtuel en trois dimensions, cités par Fowler (2015).

Par ailleurs, certaines questions ont été mal comprises (notamment la question 3 des échelles de présence et d'agentivité). Cela a probablement introduit des biais dans les réponses. Ce point a été amélioré pour l'étude expérimentale suivante. Enfin, sur le plan technique, bien que la cinétose ait été globalement maîtrisée, des inconforts liés au port du casque de réalité virtuelle ont été relevés. Ce facteur, même s'il reste secondaire, pourrait affecter indirectement d'autres dimensions de l'expérience, telles que l'attention ou l'engagement. Il conviendra de surveiller ce point dans les itérations futures du dispositif.

5.4 Perspectives d'amélioration et de recherche

Ces résultats, bien que mitigés, ouvrent plusieurs perspectives concrètes d'amélioration. Sur le plan méthodologique, une refonte des outils de mesure relatifs à la compréhension du TSA est nécessaire. Il conviendrait, notamment, de mieux distinguer les dimensions évaluées de les valider à travers des études pilotes et des analyses exploratoires. De plus, l'absence d'effets immédiats n'exclut pas des bénéfices à plus long terme, particulièrement pour des changements attitudeux complexes nécessitant une maturation cognitive.

Du point de vue de la conception pédagogique, il apparaît indispensable d'adapter plus finement le contenu aux spécificités de la RVI. Le modèle CAMIL fournit un cadre pertinent

qui nécessite une implémentation rigoureuse. Les éléments narratifs, interactifs et sensoriels, bien que pensés nativement pour la RVI, souffrent peut-être de certains défauts de conception dus aux contraintes initiales. Par ailleurs, la mise à jour logiciel de certaines tâches pourrait améliorer également certains résultats.

Enfin, si les questionnaires n'ont pas permis de saisir pleinement l'impact de l'exposition en RVI, les retours informels des utilisateurs après les séances offrent une base de réflexion intéressante pour la préparation d'un nouveau protocole expérimental. Une refonte des questions connaissances et attitudes (stéréotype) pourrait ainsi permettre de mieux évaluer l'évolution des attitudes dans le temps. Pour finir, Il pourrait être intéressant d'envisager une approche d'évaluation intégrées et plus complexes (*eyetracking*, EEG, etc.) pour saisir la complexité des effets de la RVI sur l'apprentissage et les attitudes.

5.5 Conclusion

Cette première étape de recherche expérimentale montre que, malgré certaines limites de conception et de mesure, l'utilisation de la RVI peut être envisagée comme un levier pertinent dans les dispositifs de formation à la compréhension du TSA. Le résultat le plus concluant de cette première expérimentation est la génération d'empathie accrue en condition RVI qui ouvre des perspectives encourageantes. À terme, l'objectif est d'exploiter le potentiel de l'empathie générée par la RVI pour initier une amorce de changement d'attitudes, ce qui nécessitera une amélioration des outils et une meilleure contextualisation des contenus.

Chapitre 3 - Expérimentation 2 : test du MVP de sensibilisation au TSA - les effets de l'immersion sur la compréhension du TSA - V2

1. Présentation / Introduction

Cette seconde itération expérimentale vise à tester, dans le cadre d'un dispositif de sensibilisation au TSA, la dernière version du logiciel, le MVP. Ce protocole permet ainsi de dissocier l'influence du format immersif (et de ses effets attendus) du contenu informationnel lui-même. Voici les deux objectifs principaux de ce nouveau protocole :

1. Reproduire les résultats de la première expérience sur le prototype dans cette version MVP, en évaluant plusieurs dimensions clés : la présence, l'agentivité, et la cinétose (condition de contrôle de la RVI).
2. Évaluer les effets de l'immersion/contrôle sur l'empathie et les aspects liés à la compréhension (acquisition de connaissances factuelles) et à la stéréotypisation (amorce du changement attitudes) avec le nouveau protocole de questionnaire revisité (problématique de fiabilité de la première expérience).

Sur le plan théorique, cette seconde expérimentation poursuit l'hypothèse de la valeur ajoutée de la réalité virtuelle pour les dispositifs de sensibilisation. Cette valeur ne tient pas uniquement à ses propriétés immersives isolées mais à la conjonction de composantes qui agissent en synergie pour favoriser un engagement subjectif profond sans compromettre la compréhension des contenus. Comme dans la première expérimentation, cette perspective s'inscrit dans la continuité des travaux de Parong et Mayer (2018) sur l'apport des pauses pédagogiques ainsi que ceux de Slater (2009), Gorisse (2019) ou encore Taupiac (2020) qui soulignent que l'efficacité pédagogique de la VR repose autant sur ses dimensions affectives et expérientielles que cognitives.

Conditions expérimentales :

Afin d'assurer une stricte comparabilité entre les conditions expérimentales, le contenu pédagogique (comprenant à la fois les sous-titres explicatifs et les pauses pédagogiques) a été rigoureusement conservé à l'identique entre les deux formats testés : un support vidéo traditionnel de type *flatscreen* sonore, et une version en RVI sur HMD (tableau 6). Les deux modalités expérimentales se distinguent avant tout par le niveau de présence et d'agentivité qu'elles offrent à l'utilisateur, illustrant la place donnée à l'immersion. La condition RVI propose une immersion forte : l'utilisateur dispose de mains virtuelles et d'interactions en six degrés de liberté, ce qui favorise une véritable incarnation et un engagement actif dans l'environnement. À l'inverse, la nouvelle modalité FS ne permet aucun contrôle direct ni navigation, plaçant l'utilisateur dans une posture strictement passive face au contenu. Les différences avec l'ancienne modalité FS réside dans la suppression du contrôle (agentivité) du participant sur l'expérience ainsi que la réduction de son niveau de présence. Le nouveau support FS est une vidéo (par capture d'écran) d'une séance de RVI qui reproduit strictement le contenu de la séance de RVI. Le but est de reproduire strictement le contenu en supprimant tout effet de présence et agentivité et en plaçant le participant dans une posture totalement passive. Ces modifications mettent ainsi en évidence le contraste entre les expériences offertes, allant d'une immersion riche et interactive (RVI) à une exposition strictement passive (FS).

Tableau 6

Modalités et conditions de l'expérience 2

Mod.	Support	Présence / Agentivité	Audio & Sous-titres	Pauses pédagogiques
<i>RVI</i>	Meta Quest 3 (HMD)	Haut niveau (mains virtuelles, interactions 6-DoF)	Oui	Oui
<i>FS</i>	Écran 15" + casque audio	Bas niveau (pas de navigation clavier-souris, aucune main virtuelle contrôlable)	Oui	Oui

2. Hypothèses

H1.1 : Le score de cinétose sera supérieur en condition RVI par rapport à la condition FS (condition de contrôle de la RVI - pas de lien direct avec l'amorce de changement d'attitude et de compréhension du TSA).

H1.2 : Le score de charge cognitive sera supérieur en condition RVI par rapport à la condition FS. La réalité virtuelle est connue pour augmenter la charge cognitive de ses utilisateurs par rapport à un support non immersif, Frederiksen et al. (2020).

H2 : Les scores de présence (H2.1), d'agentivité (H2.2) et d'empathie (H2.3) seront supérieurs en condition RVI par rapport à la condition FS. En effet, La RVI permet d'augmenter la présence, l'agentivité et l'empathie par rapport à des supports traditionnels, Han et al. (2022).

H3 : Les scores d'apprentissage de connaissances factuelles sur le TSA seront supérieurs en RVI par rapport à la condition FS (H3.1) et resteront plus élevés dans le temps (H3.2, mesure différée à deux semaines).

H4 : Les scores d'attitude envers le TSA (amorce du changement d'attitude) seront supérieurs en condition RVI par rapport à la condition FS (H4.1) et se maintiendront dans le temps (H4.2).

3. Matériel et Méthodes

Cette étude a été validée par le Comité d'éthique de la recherche de Toulouse. Numéro de dossier : Avenant du 9 avril 2025 - 2024_823.

3.1 Participants

L'échantillon était composé de 67 participants ($N = 67$), un effectif permettant d'assurer une puissance statistique satisfaisante (puissance $> 0,80$ pour un effet de taille moyen, $\eta^2 = 0,10$). Les participants ont été répartis aléatoirement en deux groupes expérimentaux : RVI ($n = 33 - 31$ après exclusion) et FS ($n = 34 - 33$ après exclusion), Les critères d'inclusion retenus étaient les suivants : être âgé de 18 à 65 ans, ne présenter aucun trouble vestibulaire sévère et être novice en réalité virtuelle (moins de trois séances antérieures d'exposition).

Les participants étaient âgés de 22 à 60 ans ($M = 32,09$; $ET = 11,09$). L'échantillon était composé de 44 femmes et 23 hommes (avant exclusion - 42 femmes et 22 hommes après). Le niveau scolaire était situé entre le bac+2 et le bac +8 ($M = 4,81$; $ET = 1,6$). Comme lors de la première expérimentation, l'ensemble des participants a eu une expérience avec l'enseignement ou la transmission et l'accompagnement de personnes en situation d'apprentissage : 51 enseignants et 16 professionnels de l'encadrement (formation professionnelle et les établissements de santé), allant du secondaire au supérieur pour les enseignants. Très peu de participants ont eu un contact préalable avec la réalité virtuelle novices en VR (< 3 séances). Moins de 20% des participants jouent à des jeux vidéo de manière régulière (pratique d'une activité vidéo ludique plus d'une fois par semaine). La répartition des sujets a été faite de manière aléatoire. Tous les participants ont répondu volontairement à l'appel de participation contre remise d'un certificat de participation à une sensibilisation au TSA. Ensuite, il a été possible de recontacter 60 participants pour la post passation (partie 2) avec une répartition de 31 pour la modalité FS et 29 pour la modalité RVI (7 participants n'ont pas répondu au rappel et n'ont donc pas répondu au dernier questionnaire différé).

Pour le traitement statistique, 33 participants ont été retenue pour la modalité FS (après exclusion) et 31 participants ont été retenue la modalité RVI (après exclusion). Les participants VR 02 et 06 ont été exclue du dispositif à cause d'un score > 3 au niveau de la cinétose et un score > 3 en nausée ou vision embrouillé. Ces participants ont terminé l'expérience, n'ont pas fait de remarque particulière à l'expérimentateur concernant leur état et on remplit le questionnaire. Ce n'est que lors du traitement que le constat d'un score de cinétose accru a été

constaté. Un score de cinétose trop élevé ne permet pas de considérer les résultats de ces participants comme valide (il empêche le participant d'être réceptif à la RV, Arnaldi et al., 2018). Ils ont donc été exclus du dispositif après examen de leurs résultats (scores généraux négatifs, pas d'effet d'attitude et de comportement, présence, agentivité et empathie très basse, ce qui a engendré des scores très bas en attitudes et connaissances). Le sujet FS 10 a été exclu avec un score de cinétose >3.00 et avec un score de nausée >3 (il n'a pas répondu non plus au rappel).

3.2 Application

L'application est une mise à jour MVP, exportée en *build flatscreen* et en *build VR*. Le correctif détaillé est visible dans le chapitre 3, partie 6 les modifications de l'outils en cours de thèse.

3.3 Équipement

Comme pour la première expérimentation, le matériel utilisé est un casque occlusif Meta Quest 3 pour la RVI et un ordinateur portable MSI 15'' pour le *flatscreen*.

3.4 Procédure (≈ 30 min)

Tout comme la première expérimentation, les passations se sont déroulées sur le schéma suivant : après un accueil personnalisé et un bref rappel du cadre de l'étude, les participants se voyaient présenter le formulaire de consentement éclairé. L'ensemble des droits afférents à leur participation, y compris la possibilité de se retirer à tout moment sans avoir à se justifier, leur était alors clairement exposé. Une fois le consentement obtenu, les participants étaient invités à se familiariser avec leur dispositif respectif (*flatscreen* ou RVI). Dès que cette phase d'appropriation était jugée suffisante, la passation de l'expérimentation pouvait débuter.

La durée totale de la session s'élevait en moyenne à 30 minutes, avec un temps d'immersion en RVI plafonné à 10 minutes, soit la moitié de la limite actuellement préconisée par l'Anses (2021), fixée à 20 minutes.

Les consignes données aux participants étaient les même que l'expérimentation 1 (voir partie procédure de la partie matérielle et méthode de la première expérimentation).

- Accueil des participants et présentation du dispositif - 3 minutes.
 - 2 modalités : RVI casque occlusif - FS écran plat.

- Pré-questionnaire (21 questions - TSA) - 5 minutes.
- Tutoriel - 4 minutes.
- Expérimentation - 6 minutes (voir tableau récapitulatif).
- Questionnaire (47 questions - 7 Cinétose - 5 Présence - 5 Agentivité - 7 Empathie - 2 Charge cognitive - 21 TSA) - 10 minutes.
- Clôture et remerciement - prise de contact pour troisième questionnaire Post-immersion mémorisation moyen terme - 2 minutes.

3.5 Mesures

Afin d'assurer la qualité des mesures, une fiabilité interne minimale de $\alpha \geq 0,70$ a été ciblée pour chaque bloc de questionnaires, ce qui a été vérifié au préalable lors d'un pré-test mené auprès de 15 participants. Cette fois ci, la validité interne a été atteinte (voir partie procédure) Les tests complets sont accessibles²⁰.

3.5.1 Les mesures d'immersion technologique

Une partie des instruments de mesure employés dans cette seconde expérimentation sont rigoureusement identiques à ceux utilisés lors de la première sur le plan des dimensions cinétose, présence, agentivité et empathie envers l'avatar. Cette continuité méthodologique permet d'assurer la comparabilité des résultats entre les deux études. Les quatre dimensions évaluées ont ainsi été mesurées à l'aide des mêmes échelles validées, dans leur version française. Une mesure de la charge cognitive a été ajoutée afin de mesurer succinctement sa présence dans l'expérience immersive. L'ensemble des items a été mesuré sur une échelle d'attitude de Likert en 7 point. Voici un bref rappel de ces mesures avec ajout de la mesure de charge cognitive :

Cinétose : La cinétose a été évaluée à l'aide de la version abrégée du *Simulator Sickness Questionnaire* (SSQ), composée de 7 items. Cette version permet une évaluation synthétique mais sensible des symptômes typiques de la cinétose en RV.

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.854.

Charge cognitive : La charge cognitive a été évaluée via deux questions indépendantes conçues pour cette expérimentation. L'objectif était de recueillir une brève mesure auto-

²⁰ Voir annexe D.3 : [Batterie de mesures expérimentation 2](#) p. 365.

rapportée de la charge cognitive ressentie par les participants lors de l'immersion sans alourdir le questionnaire. Voici les questions ajoutées :

1. J'ai trouvé qu'il était difficile de comprendre dans cette activité comment un apprenant avec Trouble du Spectre de l'Autisme Sans Déficience Intellectuelle (TSA SDI) pouvait traiter l'information et apprendre.
2. J'ai dû investir un effort mental important dans cette activité pour apprendre comment procède un apprenant TSA SDI pour acquérir des connaissances.

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.678.

Présence : Le sentiment de présence a été mesuré à l'aide d'un questionnaire composé de 5 items, combinant ceux issus de Fribourg (2020) et de Peck et Gonzalez-Franco (2021), adaptés en français.

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.442 (ou 0.712 sans la question 4 qui a posé problème). Cette fois-ci, l'analyse de fiabilité plus basse a contraint de retirer l'item des analyses.

Agentivité : L'agentivité perçue vis-à-vis des mains virtuelles était mesurée à l'aide de la version courte du questionnaire Fribourg (2020), comprenant 5 items. Cette mesure vise à capturer le sentiment de contrôle et d'appropriation corporelle dans l'environnement virtuel.

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.881.

Empathie envers l'avatar / PNJ : Enfin, l'empathie ressentie à l'égard de l'avatar ou des PNJ était évaluée à l'aide d'une version adaptée du *Basic Empathy Scale - A* (BES-A), composée de 7 items.

L' α de Cronbach pour l'évaluation de la cohérence interne de la mesure est de 0.902.

3.5.2 Les mesures d'apprentissage et d'attitudes vis-à-vis du TSA

L'analyse de fiabilité révèle une cohérence interne suffisante, avec un α de Cronbach de 0,772 pour les items de connaissances factuelles et de 0,838 pour les items sur les stéréotypes (amorces des changements d'attitudes). Ces valeurs, nettement plus acceptables que celles de la première expérimentation montrent une meilleure robustesse interprétative des résultats.

Pour les deux parties, la démarche d'élaboration a été la même que lors de la première expérimentation. Différentes versions ont été élaborées lors des ateliers de design thinking

(GEM de Rodez) et des rencontres avec des experts TSA. Trois versions ont ainsi été proposées avant d'obtenir la version finale. Chacune de ses versions a été testée sur le terrain avec des utilisateurs volontaires représentatifs de la population cible (personnels enseignants, huit volontaires). Voici le process :

- Recueil d'informations auprès des personnes avec TSA (utilisateurs initiateurs des concepts centraux et d'étalonnage du dispositif afin de vérifier en ateliers itératifs sous process CCU la cohérence des items).
- Proposition de question.
- Test des questions.
- Amélioration des questions.
- Test des version améliorées.
- Répétition du processus (4 itérations).

3.5.3 Les mesures de connaissances du TSA (connaissances factuelles - bloc A) :

Toujours sous le postulat qu'une acquisition de connaissances solides permet une meilleure compréhension du TSA, ces mesures examinent d'abord l'appropriation de connaissances a ce sujet. Comme dans la première expérimentation, un protocole comprenant un nombre réduit d'items a été choisi. Le nombre d'items reste suffisant pour détecter un gain éventuel en compréhension. L'objectif principal est donc de démontrer que le dispositif favorise une meilleure compréhension du TSA via l'acquisition de connaissances, condition indispensable à tout changement ultérieur d'attitude, et, potentiellement, de comportement envers les personnes concernées.

Le QCM initial de la première expérimentation a été supprimé mais quelques items fiables ont été conservés en les adaptant à un format de mesure d'attitude sous échelle de Likert (les items problématiques ont été totalement exclus). Les 13 items réadaptés traitent désormais uniquement de l'aspect connaissances factuelles.

Voici des exemples des items sélectionnés : 2a. Les apprenants avec TSA SDI peuvent ne pas remarquer que quelqu'un leur parle. 4a. Pour résoudre un problème, les apprenants avec TSA SDI préfèrent souvent le diviser en plusieurs petites parties.

3.5.4 Les mesures de l'attitudes envers le TSA (amorces des changements d'attitude - bloc B)

Pour l'attitude, le questionnaire standardisé de Lanowaz (2024) a été choisi. Seuls les items de la seconde partie du questionnaire (soit 8 items) ont été conservés. Ces items mesurent l'attitude générale du participant envers des personnes avec TSA en contexte scolaire. Ils se présentent sous la forme d'une échelle de Likert sur 7 points. L'ensemble des items de la première passation a été supprimé et remplacé par ceux de Lanowaz (2024). Les items du questionnaire traitant du rapport au TSA hors contexte d'apprentissage (première partie) ont été exclus. En effet, ces items se rapportaient à des situations sans rapport avec le contexte de l'étude (rapport individuel familiaux en contexte privé notamment).

Voici un exemple des items sélectionnés : 3a. Les personnes autistes devraient être institutionnalisées pour leur sécurité et celle des autres. 8a. Les élèves autistes qui sont intégrés dans des classes ordinaires constituent une distraction pour les élèves non autistes de cette classe.

Un seul item issu des entretiens avec le GEM a été ajouté afin de recueillir un aspect supplémentaire (la notion de réussite universitaire). Il a été ajouté à la demande des adhérents car cette dimension semblait sous-évaluée selon eux : 1a. Les personnes autistes peuvent tout à fait réussir à l'université.

4. Résultats attendus et contribution

Plan d'analyse statistique :

- Modèle mixte 2 (Mod.) \times 3 (Temps) ; corrections Greenhouse-Geisser ; post-hoc (Tukey).
- Effets de taille : η^2p , d de Cohen.

Les premières analyses visent l'évaluation des effets des conditions cinétose, présence, agentivité et empathie (hypotheses H1 et H2). Enfin, les scores d'attitudes envers le TSA (amorçage du changement d'attitude) puis de compréhension du TSA (connaissances factuelles) ont été mesurés et comparés avec une ANOVA à mesures répétées sous les hypothèses H3 et H4. Les analyses statistiques issues de Jamovi sont accessibles²¹.

4.1 Cinétose

La mesure de cinétose représente une mesure à part. Pour cette mesure, les scores des participants exclus des autres mesures ont été conservés. L'objectif était d'inclure les scores des participants ayant plus de 3 au score de cinétose (participants 02, 06, et 10) afin de proposer une mesure réelle de l'état de cinétose induit par le dispositif.

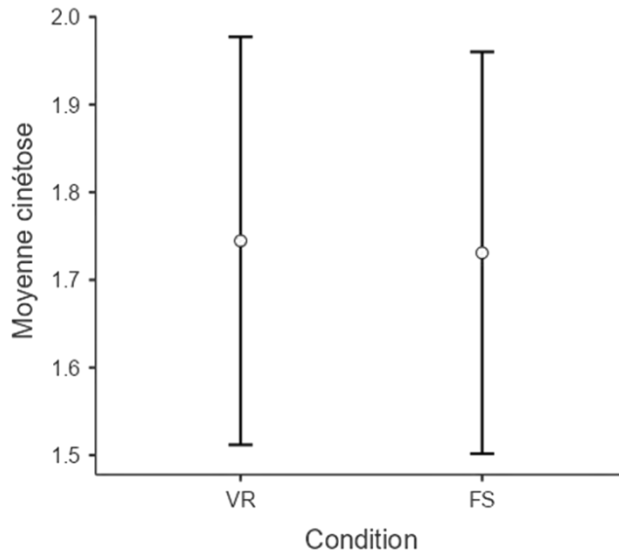
Bien que les participants avec score élevé aient été inclus, les scores de cinétose demeurent inférieurs à 2 (score inférieur au segment « un peu », considéré comme très convenable pour une expérience de RVI). L'ANOVA n'a révélé aucune différence significative soit $F(1, 65) = 0.007$, $p = 0.93$, $\eta^2p = 0.00$, ce qui indique un effet nul. Les comparaisons post-hoc ne montrent aucune différence entre le groupe VR et le groupe FS.

²¹ Voir annexe E.2 : [Résultats expérimentation 2](#) p. 380.

Cette absence de différence n'était pas attendue. Ces résultats rejettent H1.1 et permettent de montrer que la simulation n'a pas d'effet sur la cinétose (fig. 34).

Figure 34

Score de cinétose



Comme pour la première expérience, le protocole de conception en RVI a permis de réduire significativement l'incidence de la cinétose chez les participants en s'appuyant sur les recommandations établies pour les environnements immersifs proposé par Fuchs et al. (2006), avec une suppression de toute vection notamment. Cette démarche méthodologique rejoint les travaux de Makransky et Petersen (2021) qui soulignent l'importance d'une conception adaptée pour limiter les effets indésirables de la RV.

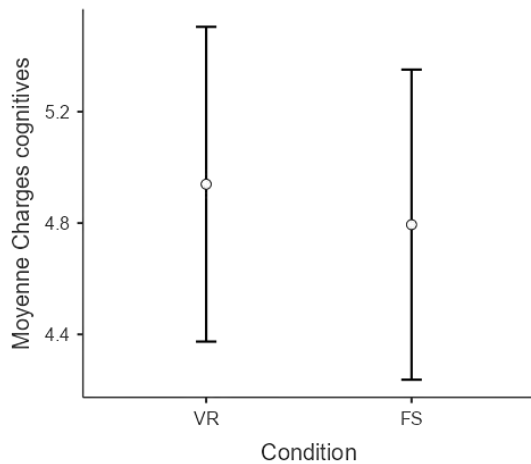
4.2 Charge cognitive

L'ANOVA n'a révélé aucune différence significative. Cette absence de différence n'était pas attendue car la RV est connue pour augmenter la charge cognitive (Makransky & Lilleholt, 2018) bien que certains travaux, comme ceux de Wenk et al. (2023), suggèrent qu'il puisse y avoir des cas où il n'y a pas de différence entre la RVI et la 2D. Dans le cas présent, il n'y a pas de différence et le score moyen de CL est modéré dans l'ensemble (score de 1 à 7 sur échelle de Likert). La charge cognitive provoquée par la simulation, bien que légère, reste présente (moins de 1 points au-dessus du score neutre qui est de 4 sur l'échelle de 1 à 7).

Ces résultats rejettent H1.2 (fig. 35). La simulation en modalité de RVI va provoquer une charge cognitive supérieure à la modalité FS et permettent de montrer que la simulation a un effet minime sur la charge cognitive.

Figure 35

Score de charge cognitive

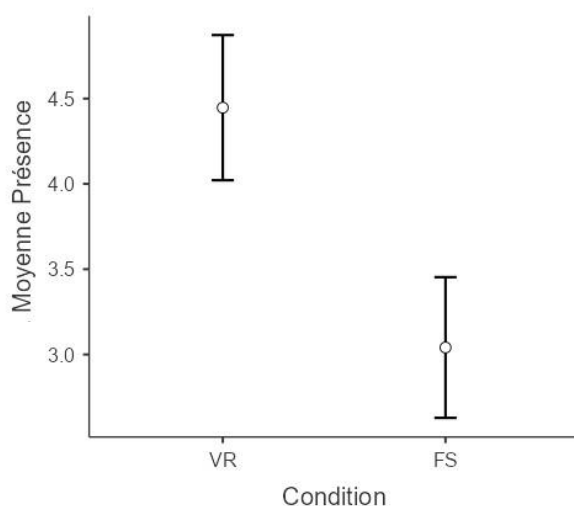


4.3 Présence

Une ANOVA a révélé une différence significative entre les groupes soit, $F(1, 62) = 22,5$ et $p < .001$). La taille de l'effet, η^2p , était de 0.266, ce qui indique un effet important. D'après ces résultats (fig. 36) l'induction de présence est supérieure en condition de RVI par rapport à la condition FS. Ces résultats confirment donc une partie de H.2.1, la présence générée par un simulateur de RVI semble bien être supérieure par rapport à une présentation classique. Les résultats de la première expérimentation sont ici répliqués.

Figure 36

Score de présence



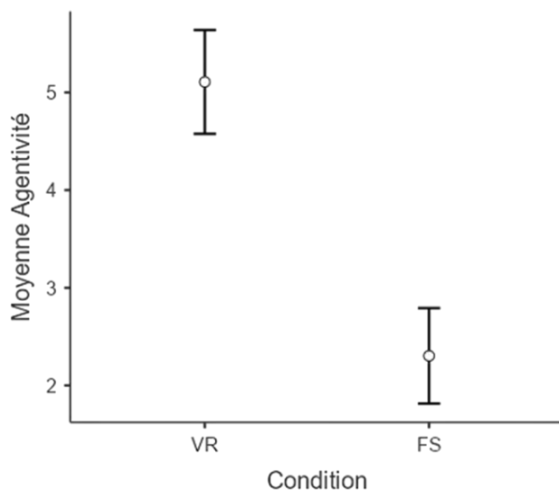
Le protocole de conception mis en place a favorisé l'émergence d'un sentiment de présence élevé avec une moyenne de 4,59 (score de 1 à 7 sur échelle de Likert) chez les participants, dimension cruciale pour la validité écologique des expériences en RV.

4.4 Agentivité

Une ANOVA a révélé une différence significative entre les groupes soit, $F(1, 62) = 48.9$ et $p < .001$). La taille de l'effet, η^2 , était de 0.441, ce qui indique un effet important. Le constat d'un score d'agentivité nettement plus élevé en condition de RVI correspond à la littérature. D'après ces résultats, l'induction d'agentivité est supérieure en condition de RVI par rapport à la condition FS. Ces résultats (fig. 37) confirment donc H2.2, l'agentivité générée par un simulateur de RVI semble bien être supérieur par rapport à une présentation classique. Les résultats de la première expérimentation sont ici répliqués.

Figure 37

Score agentivité



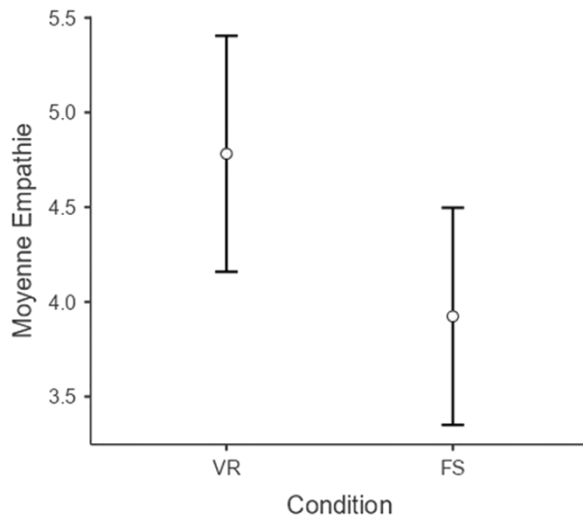
Le protocole de conception développé a particulièrement privilégié le renforcement du sentiment d'agentivité des participants avec une moyenne de 5 (score de 1 à 7 sur échelle de Likert).

4.5 Empathie

Une ANOVA a révélé une différence significative entre les groupes soit, $F(1, 62) = 6,72$ et $p = 0.012$). La taille de l'effet, η^2p , était de 0.098, ce qui indique un effet moyen.

Figure 38

Score empathie



Le constat d'un score d'empathie nettement plus élevé en condition de RVI semble confirmer une H2.3. La figure 38 montre que l'induction d'empathie est supérieure en condition de RVI avec une moyenne de 4,80 par rapport à la condition FS (score de 1 à 7 sur échelle de Likert). Les résultats de la première expérimentation ont été répliqués ici aussi.

4.6 Performances d'apprentissage

Tableau 7

Moyennes et écarts-types des groupes VR et FS aux questionnaires de compréhension envers le TSA selon le moment de mesure

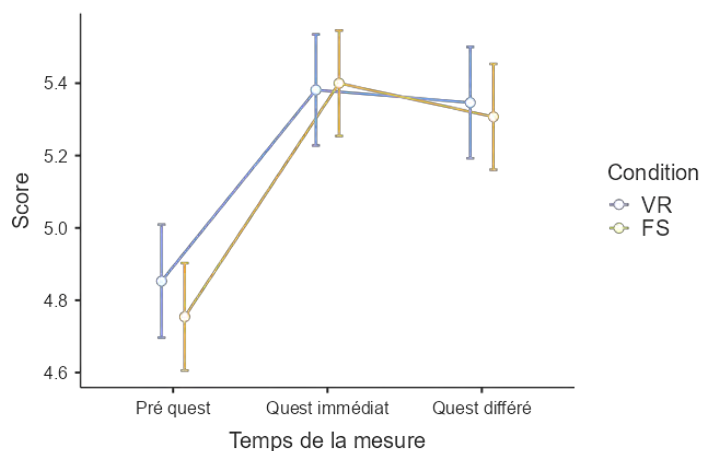
	Groupe pré-questionnaire		Questionnaire immédiat		Questionnaire différé	
	M	ET	M	ET	M	ET
VR	5.05	0.56	5.81	0.54	5.90	0.58
FS	4.87	0.43	5.36	0.43	5.58	0.60

Note. VR = réalité virtuelle ; FS = film de sensibilisation ; *M* = moyenne ; *ET* = écart-type. Les scores sont mesurés sur une échelle de Likert.

L'analyse ne révèle pas d'interaction significative entre la condition expérimentale et le moment de mesure. L'effet principal du temps est significatif, $F(2, 114) = 76,6, p < 0,001, \eta^2p = 0,564$, pour une taille d'effet importante, indiquant un gain dans les deux conditions. Le tableau 7 montre les moyennes et écarts types.

Figure 39

Score connaissances



L'effet principal de la condition n'est pas significatif (fig. 39). H3 est rejetée, la RV ne semble pas avoir d'effet particulier par rapport à un support non immersif.

4.7 Attitude

Tableau 8

Moyennes et écarts-types des groupes VR et FS aux questionnaires d'attitudes envers le TSA selon le moment de mesure

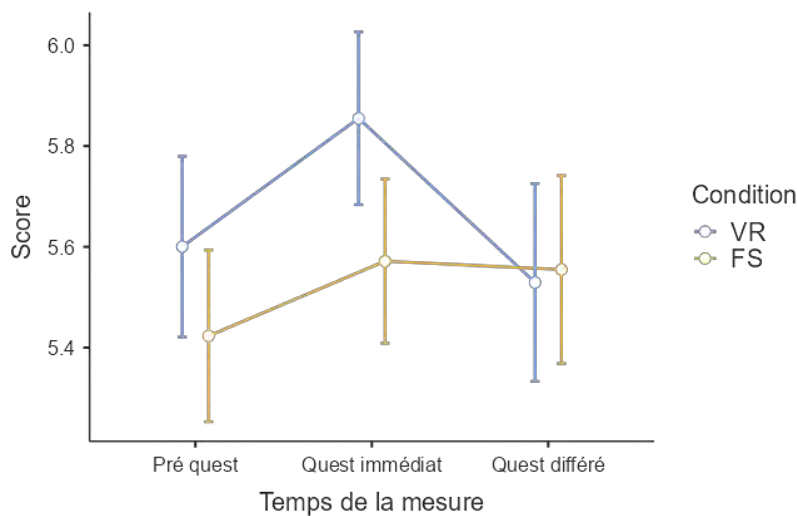
	Groupe pré-questionnaire		Questionnaire immédiat		Questionnaire différé	
	M	ET	M	ET	M	ET
VR	5.52	0.43	5.77	0.38	5.56	0.54
FS	5.41	0.50	5.57	0.51	5.55	0.50

Note. VR = réalité virtuelle ; FS = film de sensibilisation ; *M* = moyenne ; *ET* = écart-type. Les scores sont mesurés sur une échelle de Likert.

L'effet principal du temps est significatif, $F(2, 114) = 6,34, p = 0,002, \eta^2 = 0,1$ pour une taille d'effet moyenne, semblant indiquer une évolution globale des attitudes. L'effet principal de la condition n'est pas significatif, suggérant une absence de différence globale entre les modalités. Cependant, l'analyse révèle une interaction significative entre la condition expérimentale et le moment de mesure sur l'attitude envers le TSA, $F(2, 114) = 3,31, p = 0,04, \eta^2 = 0,055$ pour une taille d'effet petite. Le tableau 8 montre les moyennes et écarts types.

Figure 40

Score attitudes



Les analyses post-hoc révèlent des patterns d'évolution des attitudes distincts selon les modalités. Dans la condition FS, les participants ne montrent pas d'évolution significative de leur attitude entre le pré-test et le post-test immédiat (différence moyenne = -0,14, $p = 0,31$), ni entre le pré-test et le post-test différé (différence moyenne = -0,13, $p = 0,75$). Dans la condition RV, on observe une amélioration significative de l'attitude du pré-questionnaire au questionnaire immédiat (différence moyenne = -0,25, $p = 0,015$, semblant confirmer H4.1), qui ne se maintient pas au post-test différé (différence pré-test/post-différé : $p = 0,98$, retour proche du niveau initial et rejet de H4.2).

Ces résultats suggèrent que la réalité virtuelle produit un effet positif temporaire sur l'attitude envers le TSA contrairement à la modalité traditionnelle qui ne génère aucun changement d'attitude (la figure 40 semble confirmer H4.1, les scores d'attitude sur le TSA sont significativement supérieurs en RVI par rapport à la condition FS lors du questionnaire immédiat). Cependant, cet effet bénéfique de la RV s'estompe dans le temps. H4 est rejetée : globalement, ces résultats montrent que les scores d'attitudes sur le TSA ne sont pas significativement supérieurs en RVI par rapport à la condition FS, les attitudes revenant à leur

niveau initial lors du post-test différé. Néanmoins, cette capacité de la RV à modifier temporairement les attitudes pourrait être liée à son pouvoir immersif et à sa capacité à générer de l'empathie, bien que des stratégies additionnelles semblent obligatoirement nécessaires pour maintenir ces changements à long terme.

5. Discussion

Les améliorations apportées suite à la première expérimentation (reformulation des items, nouvelle échelle de Lanowaz, 2024 et fiabilité des échelles utilisées) ont permis d'obtenir des résultats plus concluants et exploitables. Ces ajustements ont ainsi permis d'obtenir un score d'analyse de fiabilité globalement satisfaisants ($\geq 0,70$), permettant d'envisager que les outils de mesure sont plus pertinents. Toutefois, certaines limites persistent.

Synthèse de la validation des hypothèses :

- **H1 - Cinétose supérieure en condition RVI - État : non validée.** Contrairement aux attentes, les scores de cinétose ne se révèlent pas supérieurs en condition RVI par rapport à la condition FS. Comme lors de la première expérimentation, des inconforts liés au port du casque ont néanmoins été relevés, confirmant les préoccupations ergonomiques classiques associées à la RVI.
- **H2 - Supériorité de la RVI sur présence, agentivité et empathie - État : validée.** Les scores de présence (H2.1), d'agentivité (H2.2) et d'empathie (H2.3) sont significativement supérieurs en condition RVI, avec des effets cohérents avec la littérature. Les résultats de la première expérimentation sont ici répliqués.
- **H3 - Connaissances factuelles supérieures en RVI - État : non validée.** L'analyse révèle une absence d'effets significatifs entre les conditions pour l'acquisition et le maintien des connaissances factuelles sur le TSA (malgré des tendances observables et des différences significatives partielles sur certains points). Une progression significative est observée entre pré-questionnaire et questionnaire immédiat en condition de RVI (H3.1), mais aucune supériorité n'émerge au questionnaire différé (H3.2 rejetée).
- **H4 - Niveau d'attitudes envers le TSA (amorce du changement d'attitude) en RVI - État : non validée.** L'analyse révèle une absence d'effets significatifs entre les conditions pour un changement de niveau d'attitude envers le TSA sur le sujet et ce, malgré une différence significative observée entre le pré-questionnaire et le

questionnaire immédiat (entre RVI et FS - H4.1). Aucune différence n'émerge au questionnaire différé (H4.2 rejetée). Le bref changement observé ne reste pas dans le temps.

Contrairement aux attentes théoriques, cette expérimentation ne révèle pas de cinétose supérieure en condition RVI par rapport à la modalité traditionnelle. Ce résultat diffère de nombreuses études qui identifient la cinétose comme l'un des principaux obstacles à l'adoption de la RVI, Stanney et al. (2020). Cette absence de différence significative pourrait s'expliquer par l'utilisation d'un tutoriel adapté à l'introduction des participants en RV et l'application des recommandations d'utilisabilité en RV d'Arnaldi et al. (2018). Néanmoins, les inconforts ergonomiques liés au port du casque demeurent présents, confirmant les préoccupations classiques soulevées dans la littérature concernant l'acceptabilité physique des dispositifs immersifs, Jerald (2016).

Au niveau de la présence, l'agentivité et l'empathie (H2), les résultats confirment pleinement la supériorité de la RVI validant H2 dans son ensemble. Ces éléments s'alignent parfaitement avec le corpus théorique établi. La supériorité en termes de présence (H2.1) corrobore les travaux fondamentaux de Slater et Wilbur (1997) et les méta-analyses récentes démontrant que l'immersion technologique favorise le sentiment de présence, Cummings et Bailenson (2016). En optimisant la conception, cette méthodologie a permis d'atteindre un niveau d'immersion technologique favorisant l'illusion de « être là » dans l'environnement virtuel (Slater, 2009). Cette recherche de présence s'inscrit dans la lignée des travaux de Fuchs et Moreau (2003) qui démontrent l'impact direct de la présence sur les processus cognitifs étudiés. Les résultats sur l'agentivité (H2.2) confirment les théories de l'*embodiment* virtuel et du *sense of agency* de Kilteni et al. (2012) particulièrement pertinentes dans les contextes d'interaction avec des agents virtuels. En implémentant des modalités d'interactions naturelles et intuitives, le dispositif a permis aux utilisateurs de développer un fort sentiment de contrôle sur leurs actions et leurs conséquences dans l'espace virtuel. Cette approche rejoint les préoccupations de Gorisse et al. (2018) concernant l'importance de préserver l'intentionnalité motrice pour maintenir un niveau d'agentivité comparable à celui observé en situation réelle. L'utilisateur doit être, au cours de son expérience de RV, à la source de ses actions, Haggard (2017). Enfin, les scores d'empathie supérieurs (H2.3) convergent avec les travaux pionniers d'Herrera et al. (2018) sur le potentiel de la RVI pour induire des réponses empathiques, notamment dans les contextes de sensibilisation sociale. Le protocole de conception développé a favorisé l'émergence d'empathie chez les participants, dimension fondamentale pour l'étude

des processus sociocognitifs en environnement virtuel. Le dispositif a permis de susciter des réactions émotionnelles, Riva et al. (2007). Cette approche s'aligne sur les travaux de Fuchs et Moreau (2006) qui soulignent l'importance de la fidélité représentationnelle pour déclencher des mécanismes de partage émotionnel en RV. Les mesures d'empathie collectées démontrent que le protocole de conception mis en place a effectivement permis aux participants d'éprouver une connexion empathique significative avec leur avatar, validant ainsi l'utilisation de la RVI pour l'investigation des processus empathiques.

Au niveau de l'acquisition et du maintien des connaissances factuelles, les résultats présentent un pattern nuancé reflétant la complexité de la littérature. L'absence de supériorité globale de la RVI s'inscrit dans la continuité des préoccupations soulevées par plusieurs méta-analyses concernant l'efficacité pédagogique pour l'apprentissage de contenus factuels (Radianti et al., 2020). L'équivalence observée entre les modalités contraste cependant avec les déficits d'apprentissage factuel habituellement rapportés en RVI (Parong & Mayer, 2018). Cette performance équivalente peut être attribuée à l'intégration de pauses pédagogiques spécifiquement conçues pour pallier les limitations cognitives de la RVI sans compromettre l'illusion de plausibilité de Slater (2009), représentant ainsi une innovation méthodologique prometteuse. Ces résultats démontrent que la transmission et la rétention de connaissances en RVI peuvent être équivalentes à celles d'un support traditionnel, équivalence rarement observée dans la littérature.

Pour le changement d'attitude envers le TSA, les résultats révèlent un pattern caractéristique des interventions en RVI : un effet positif immédiat suivi d'une décroissance temporelle. La différence significative observée au post-test immédiat confirme la capacité de la RVI à induire des changements attitudeux à court terme, particulièrement dans les domaines de l'empathie et de la sensibilisation sociale (Herrera et al., 2018). Cette amélioration s'explique par les mécanismes d'immersion et de présence qui favorisent l'identification et l'engagement émotionnel comme le suggèrent Slater et Wilbur (1997). Cependant, l'absence de maintien de cet effet au post-test différé confirme les préoccupations concernant la durabilité des changements attitudeux post-exposition en RVI (Maister et al., 2015). Ce pattern suggère que si la RVI est fonctionnelle pour initier une amorce de changement, des stratégies complémentaires de consolidation sont nécessaires pour assurer leur pérennité, soulignant l'importance de développer des approches hybrides combinant l'impact immédiat de la RVI avec des interventions de suivi.

5.1 Limites

Plusieurs éléments d'interprétation sont envisageables. D'une part, des limites dans la conception et le développement du contenu pédagogique, en particulier pour la version RVI. Il est possible que malgré la mise à jour, le contenu soit encore perfectible pour utiliser le plein potentiel de la RVI. La manipulation expérimentale des pauses pédagogiques, pourtant centrale dans la transmission des connaissances, n'a pas été isolée entre les conditions. Bien que cette homogénéité permette une comparaison directe des modalités, elle empêche d'attribuer clairement les effets de rétention aux pauses elles-mêmes. Il conviendra de proposer une variante avec manipulation des pauses pédagogiques.

Ensuite, la taille de l'échantillon reste limitée, ce qui peut réduire la puissance statistique pour détecter des effets de faible amplitude, notamment sur les attitudes et les connaissances.

Enfin, sur le plan technique, bien que la cinétose ait été globalement maîtrisée, des inconforts liés au port du casque de réalité virtuelle ont été relevés comme dans la première expérimentation. Ce facteur, même s'il reste secondaire, pourrait affecter indirectement d'autres dimensions de l'expérience, telles que l'attention ou l'engagement. Il conviendra de surveiller ce point et de tester le contenu avec les casques occlusifs à venir (réduction du poids, optimisation du confort, etc.).

5.2 Perspectives d'améliorations et de recherches

Ces résultats encouragent à approfondir les travaux sur les mécanismes cognitifs et affectifs de l'immersion, en particulier pour mieux dissocier les effets du format (RV vs FS) de ceux liés au scénario pédagogique (pauses, interactivité, narration). La manipulation future des pauses comme variable indépendante serait essentielle pour clarifier leur rôle dans l'acquisition de connaissances durables. De plus, si l'empathie semble modérément sensible à l'immersion technologique, les attitudes générales paraissent plus résistantes au changement. Cela ouvre des pistes vers des formats hybrides (RV + débriefing collectif ou accompagnement pédagogique) susceptibles d'ancrer plus durablement ces effets. Enfin, l'introduction d'un suivi longitudinal à plus long terme, ainsi que des mesures qualitatives (entretiens post-expérience), pourrait enrichir l'interprétation des évolutions attitudeles.

5.3 Conclusion

Cette seconde expérimentation valide les améliorations méthodologiques apportées à l'outil, tout en confirmant les bénéfices de la RVI sur plusieurs composantes clés : présence, agentivité, et, dans une moindre mesure, empathie. Bien que les effets sur les connaissances soient comparables dans les deux modalités, leur maintien dans le temps témoigne d'un effet pédagogique robuste, probablement soutenu par les pauses pédagogiques.

En revanche, les attitudes apparaissent plus fluctuantes. Cela suggère qu'un changement de perspective sur des thématiques sociales sensibles nécessite plus qu'une immersion ponctuelle. L'outil développé présente néanmoins un certain potentiel en contexte éducatif à condition de l'intégrer dans une démarche pédagogique plus globale (formation complémentaire, etc.).

Discussion générale

1. Rappel des objectifs et de la démarche scientifique

Cette recherche doctorale avait pour objectif principal d'évaluer l'impact d'un outil de réalité virtuelle incarnée (RVI) conçu et développé en interdisciplinarité sur la sensibilisation aux troubles du spectre de l'autisme (TSA) en comparant plusieurs modalités de présentation d'un même contenu pédagogique. L'enjeu consistait à déterminer dans quelle mesure l'immersion technologique et l'incarnation d'avatar peuvent constituer des leviers efficaces pour modifier les représentations et attitudes envers une population stigmatisée, question centrale dans le contexte actuel des politiques d'inclusion scolaire et sociale.

La démarche adoptée s'est structurée autour d'un travail de conception et de deux expérimentations complémentaires, conçues selon une logique méthodologique progressive. Cette approche visait à isoler graduellement les facteurs technologiques responsables des effets observés, conformément aux recommandations méthodologiques pour les technologies éducatives immersives (Taupiac, 2020).

Dans cette recherche, le travail de conception occupe une place fondamentale et constitue une contribution majeure en soi. L'approche interdisciplinaire adoptée, articulant sciences cognitives, informatique et design centré utilisateur, a permis de concevoir et de développer un outil de RVI spécifiquement adapté aux besoins empiriques de la sensibilisation au TSA. Ce dispositif résulte d'un processus rigoureux intégrant plusieurs dimensions complémentaires : l'ancrage théorique dans la littérature scientifique sur le TSA et la RVI, l'implémentation des recommandations techniques pour optimiser l'expérience immersive, le développement informatique avancé et surtout une démarche de conception centrée utilisateur impliquant des ateliers itératifs avec des personnes avec TSA.

La première expérimentation adoptait une perspective exploratoire large, visant à évaluer l'effet d'une expérience immersive en RV par rapport à différents formats moins immersifs : une version *flatscreen* semi-immersive et deux conditions non immersives (diaporama et diaporama audio).

La seconde expérimentation, s'appuyant sur les résultats de la première, avait pour but de tester les modifications apportées au logiciel ainsi qu'aux modalités de recueil de données

via une comparaison stricte entre les deux supports (version RVI immersive vs version vidéo *flatscreen* totalement passive).

1.1 Rappel des hypothèses générales

Fondements théoriques de la recherche - le modèle CAMIL comme cadre de référence :

L'architecture théorique de cette recherche s'articule principalement autour du modèle CAMIL (*Cognitive-Affective Model of Immersive Learning*) développé par Makransky et Petersen (2021). Ce modèle intégrateur constitue le cadre de référence central de ce travail doctoral, postulant une chaîne causale reliant les caractéristiques technologiques de la réalité virtuelle aux *outcomes* d'apprentissages via des médiateurs cognitifs et affectifs spécifiques.

Le modèle CAMIL distingue deux voies principales d'influence : une voie cognitive, où l'immersion technologique favorise la présence et l'engagement cognitif, facilitant ainsi l'acquisition de connaissances ; et une voie affective, où l'immersion technologique génère des émotions positives et de la motivation intrinsèque, soutenant l'apprentissage par des mécanismes motivationnels. Cette architecture théorique intègre harmonieusement les apports des théories établies de la présence (Witmer & Singer, 1998), de l'incarnation d'avatar (Kilteni et al., 2012) et du changement d'attitude par contact virtuel (Yee & Bailenson, 2007), en les structurant dans une perspective cohérente appliquée aux environnements d'apprentissages immersifs.

Un objectif de cette thèse consistait à valider empiriquement le modèle CAMIL dans le contexte spécifique de la sensibilisation au TSA, un domaine d'application original encore inexploré sous le CAMIL. Cette validation impliquait de tester les relations causales postulées entre immersion technologique, médiateurs psychologiques et *outcomes* d'apprentissages socio-affectifs (compréhension, attitudes, réduction des stéréotypes).

L'approche théorique adoptée s'inscrit également dans le prolongement des modèles de Slater (2009) sur l'illusion de plausibilité en RV, qui distinguent les processus perceptuels (présence spatiale) des processus cognitifs (crédibilité de l'expérience). Cette distinction conceptuelle, intégrée dans le modèle CAMIL, a guidé la formulation des hypothèses sur les mécanismes par lesquels l'immersion technologique peut influencer les attitudes sociales.

Par ailleurs, cette recherche s'appuie sur les travaux récents traitant de l'apprentissage immersif de Parong & Mayer (2018) qui soulignent l'importance de certains aspects de

conception (les pauses pédagogiques) dans l'efficacité pédagogique des environnements virtuels, éléments centraux du modèle CAMIL.

Hypothèses générales et mécanismes postulés : Les hypothèses générales de cette recherche posaient que l'exposition à une expérience immersive favorise, plus qu'un support non immersif, trois dimensions interconnectées de la sensibilisation aux TSA : la compréhension des connaissances factuelles, la réduction des stéréotypes et l'amélioration des attitudes envers les personnes avec TSA.

Cette prédiction principale s'appuyait sur l'hypothèse que l'immersion technologique, par la richesse sensorielle qu'elle procure et l'incarnation qu'elle permet, génère une forme d'apprentissage expérientiel plus engageante et mémorable que les modalités traditionnelles de présentation d'informations. Conformément aux théories du traitement de l'information de Mayer (2014), l'activation simultanée de multiples canaux sensoriels était supposée faciliter l'encodage et la rétention des contenus pédagogiques.

L'originalité théorique de cette recherche résidait dans l'hypothèse de médiation par des variables spécifiques à l'expérience en RV. Il était postulé que les effets observés sur la compréhension et les attitudes seraient médiés par trois processus psychologiques distincts mais complémentaires :

- La présence, définie comme le sentiment subjectif d'être présent dans l'environnement virtuel, était conceptualisée comme élément clé. Conformément aux modèles de Witmer et Singer (1998), la présence était supposée résulter de l'interaction entre les caractéristiques immersives du système et les facteurs individuels. Cette variable était hypothétiquement liée à une attention soutenue et à un engagement cognitif renforcé dans l'apprentissage des contenus sur le TSA.
- L'agentivité, comprise comme le sentiment de contrôle dans l'environnement virtuel, était postulée également comme élément clé. S'appuyant sur la théorie de l'autodétermination de Deci et Ryan (2000), cette variable était supposée favoriser une motivation intrinsèque pour l'apprentissage et une persévérance accrue dans l'exploration des contenus pédagogiques.
- L'empathie, opérationnalisée comme la capacité à adopter la perspective et à comprendre les états émotionnels d'autrui, constituait l'élément central du modèle théorique. L'incarnation d'un avatar représentant une personne avec TSA était supposée activer les mécanismes de prise de perspective (*perspective-taking*), conformément aux

modèles neurocognitifs de l'empathie comme celui de Decety et Jackson (2004). Cette activation empathique était hypothétiquement à l'origine des changements d'attitude observés, par une réduction de la distance psychologique et une humanisation accrue des personnes avec TSA.

Articulation des hypothèses avec les objectifs expérimentaux : L'architecture expérimentale a été conçue pour tester progressivement ces hypothèses générales à travers des comparaisons contrôlées. La première expérimentation visait à établir la supériorité générale de la modalité immersive sur les formats traditionnels tandis que la seconde se concentrait sur la validation des modifications apportées au premier dispositif (expérience immersive et recueil de données).

Cette approche permettait de répondre aux questions de l'efficacité (la RVI est-elle plus efficace ?) et des mécanismes (pourquoi et comment la RVI est-elle efficace ?), contribution essentielle pour le développement théorique du domaine et les applications pratiques en contexte éducatif.

Les hypothèses spécifiques de médiation, bien que non testées dans le cadre de cette recherche en raison de contraintes d'échantillonnage, constituent la perspective théorique centrale qui guide l'interprétation des résultats obtenus et oriente les développements futurs de cette ligne de recherche.

Ce positionnement et la présentation des résultats empiriques dans leur contexte théorique et méthodologique prépare l'analyse critique des données recueillies et leur contribution à la validation du modèle CAMIL. Les résultats de cette recherche doctorale valident partiellement le modèle de Makransky et Petersen (2021), confirmant certaines relations théoriques postulées tout en révélant certaines limites. Cette étude démontre notamment que la RVI, bien qu'efficace sur certaines dimensions, ne constitue pas la solution la plus performante pour le changement d'attitude lorsqu'elle est utilisée de manière isolée et dans de courts formats d'intervention. Cette nuance critique contribue à une compréhension plus fine des conditions d'efficacité des technologies immersives dans l'apprentissage socio-affectif, enrichissant ainsi le modèle théorique initial et orientant les développements futurs de cette ligne de recherche vers des approches plus intégrées et longitudinales.

2. Apports empiriques et théoriques : une validation partielle du modèle CAMIL

2.1 Confirmation des mécanismes de présence, d'agentivité et d'empathie

Les données collectées au fil des deux études fournissent des résultats nuancés qui permettent une validation différentielle du modèle CAMIL. D'un côté, les variables immédiatement liées à l'expérience immersive (en particulier la présence, l'agentivité et l'empathie) sont significativement plus élevées dans les conditions RVI. Ces résultats confirment empiriquement les prédictions théoriques du modèle CAMIL concernant la voie affective de l'apprentissage immersif et valident l'opérationnalisation réussie des facteurs immersifs ciblés par le protocole expérimental.

Concernant la présence et l'agentivité, les résultats s'inscrivent en cohérence avec la littérature établie sur les environnements virtuels éducatifs. L'augmentation significative des scores de présence en condition RVI confirme les prédictions de Witmer et Singer (1998) à propos de l'influence des caractéristiques technologiques immersives sur le sentiment « d'être-là ». De manière similaire, les niveaux d'agentivité plus élevés observés dans les conditions immersives, avec un effet particulièrement robuste, corroborent les travaux de Slater (2009) sur l'importance du contrôle sensori-moteur dans l'appropriation de l'expérience virtuelle. Ces résultats suggèrent que l'environnement immersif développé est bien perçu comme tel par les participants et qu'il favorise une implication sensorimotrice plus forte, conformément aux mécanismes postulés par la voie cognitive du modèle CAMIL.

L'empathie constitue un résultat particulièrement significatif qui valide les prédictions théoriques concernant les mécanismes d'identification affective en RVI. L'augmentation observée de l'empathie en condition RVI confirme les travaux de Yee et Bailenson (2007) sur l'effet Proteus et ceux de Herrera et al. (2018) sur la construction de l'empathie par les environnements virtuels. Cette validation empirique de la voie affective du modèle CAMIL dans le contexte spécifique des TSA constitue un apport original de cette recherche, démontrant que l'incarnation d'un avatar représentant une personne avec TSA peut effectivement activer les mécanismes de prise de perspective et de contagion émotionnelle postulés par les modèles neurocognitifs de l'empathie.

2.2 Résultats contrastés sur les résultats d'apprentissages : limites du modèle CAMIL

En revanche, les résultats concernant les effets sur les connaissances et les attitudes sont plus contrastés et ne confirment pas totalement les hypothèses dérivées du modèle CAMIL. Cette divergence révèle des limites importantes dans l'application du modèle au domaine de l'apprentissage socio-affectif, particulièrement dans le contexte de sensibilisation aux différences neurodéveloppementales.

Les connaissances factuelles - un paradoxe de l'immersion : Sur le plan des connaissances factuelles sur le TSA, les deux expérimentations convergent sur un résultat inattendu : aucune supériorité de la RVI n'est observée comparativement aux autres formats. Ce résultat contraste avec les prédictions de la voie cognitive du modèle CAMIL, mais s'avère néanmoins particulièrement encourageant au regard de la littérature. En effet, les travaux de Parong et Mayer (2018) documentent des déficits d'apprentissage factuel en RVI dus aux distractions et au déficit de focalisation, tandis que Poupard et al. (2024) rapportent que 38% des études montrent des effets négatifs contre seulement 15% d'effets positifs. Les résultats démontrent non pas une infériorité, mais une équivalence entre modalités, attribuable à l'implémentation rigoureuse de pauses pédagogiques conformes au modèle CAMIL. Cette observation constitue un apport empirique original : contrairement aux craintes initiales concernant la surcharge cognitive (Sweller et al., 2011), il est désormais possible d'égaliser les supports traditionnels sur leur terrain de prédilection. La transmission de connaissances factuelles en environnement virtuel immersif, longtemps considérée comme une limite de la technologie, devient viable grâce à une conception appropriée. Si l'on considère les autres atouts de la RVI (présence, agentivité et empathie), cet apport empirique constitue à lui seul un résultat majeur de ce travail.

L'amorce de changement d'attitude - une étape encore inaccessible : Concernant les attitudes et stéréotypes envers les personnes avec TSA, les résultats révèlent un pattern différentiel entre les deux expérimentations. L'expérimentation 1 ne montre aucun effet significatif, probablement en raison d'un effet plafond (scores initiaux entre 4,65 et 4,85). L'expérimentation 2, en revanche, démontre que la RVI produit une amélioration significative immédiate qui ne se maintient pas au post-test différé. Ce pattern confirme les travaux de Herrera et al. (2018) sur les changements attitudeux à court terme en RVI et valide les préoccupations de Maister et al. (2015) concernant l'absence de persistance. Cette limitation constitue un défi majeur pour la validation du modèle CAMIL dans le domaine du changement

d'attitude, car elle suggère que la RVI, en l'état, ne suffit pas à modifier en profondeur les représentations sociales.

Cette résistance au changement peut s'expliquer par plusieurs mécanismes théoriques. Une seule session immersive semble insuffisante pour provoquer une transformation durable des attitudes, comme le suggèrent des travaux en psychologie sociale sur la résistance des stéréotypes implicites (Devine et al., 2012). La méta-analyse de Martingano et al. (2021) éclaire ce phénomène en distinguant empathie émotionnelle (augmentée par la RV) et empathie cognitive (non affectée), expliquant le paradoxe observé entre forte empathie générée et absence de transformation attitudinale durable. Par ailleurs, l'effet plafond chez les enseignants, population déjà sensibilisée aux enjeux d'inclusion, peut limiter la détection de changements significatifs et questionner la généralisation des résultats à d'autres populations. La spécificité du domaine TSA, où les représentations sociales sont particulièrement ancrées (Willay, 2020), constitue également un facteur de résistance notable.

Contributions théoriques - vers une version enrichie du modèle CAMIL : Sur le plan théorique, ces résultats conduisent à nuancer significativement les modèles qui postulent un effet direct et systématique de la RV sur la sensibilisation sociale. La présente recherche démontre que la RV n'est pas une solution universelle ni automatiquement efficace, mais qu'elle agit de manière conditionnelle, via l'engagement sensorimoteur et affectif, et probablement en interaction complexe avec des variables individuelles (sensibilité à l'immersion, empathie dispositionnelle, connaissances initiales sur les TSA).

Ces observations enrichissent le modèle CAMIL en révélant l'importance de variables modératrices non prises en compte dans sa formulation initiale. L'efficacité des voies cognitive et affective du modèle dépend probablement de facteurs contextuels (durée d'exposition, design expérientiel, population cible) et individuels (traits empathiques, attitudes initiales, familiarité technologique) qui mériteraient d'être intégrés dans une version enrichie du modèle. Cette recherche met également en lumière la nécessité d'un design expérientiel sophistiqué, intégrant la narration immersive, la durée d'exposition optimale (ou expositions répétées), la personnalisation de l'expérience, ainsi que des dispositifs de consolidation post-expérience (réflexivité guidée, interventions complémentaires). Ces éléments, bien qu'absents du modèle CAMIL original, s'avèrent probablement cruciaux pour l'efficacité des interventions de sensibilisation par la réalité virtuelle incarnée, particulièrement pour obtenir des changements attitudinaux durables.

3. Limites et perspectives de recherche

3.1 Limites méthodologiques liées à la conception expérimentale

Tâches limitées et problématiques de conception : Cette recherche présente plusieurs limites inhérentes à la conception de l'environnement virtuel et aux tâches proposées aux participants. Les scénarios développés, bien qu'inspirés de situations réelles vécues par les personnes avec TSA, demeurent nécessairement simplifiés et ne peuvent reproduire l'ensemble de la complexité des interactions sociales authentiques. Cette simplification, bien que nécessaire pour le contrôle expérimental, peut limiter la généralisation des résultats aux contextes naturels d'interaction auprès des personnes avec TSA. De plus, la durée limitée de l'exposition (une session unique) ne permet pas d'évaluer la persistance des effets observés, contrairement aux interventions de sensibilisation traditionnelles qui s'étalent généralement sur plusieurs séances.

Les choix de conception technologique, notamment en termes d'interactions disponibles, constituent également des compromis entre réalisme et faisabilité technique. Ces limitations peuvent affecter l'immersion technologique et l'incarnation de l'avatar, variables centrales dans le modèle théorique utilisé, et, par conséquent, influencer l'efficacité globale de l'intervention.

La charge cognitive - une limite intrinsèque de la réalité virtuelle : Selon la théorie de la charge cognitive de Sweller (1988), l'utilisation de la RV génère une charge cognitive extrinsèque liée à l'appropriation de l'interface et à la navigation dans l'environnement virtuel. Cette surcharge peut mobiliser une partie des ressources cognitives nécessaires aux processus d'empathie et d'apprentissage conceptuel. Bien que les résultats semblent présenter une charge cognitive faible, cette dernière pourrait constituer un facteur modérateur des effets de l'intervention. Cette limite soulève des questions importantes sur l'optimisation de l'ergonomie cognitive des environnements virtuels éducatifs et la nécessité de phases d'adaptation préalables pour minimiser la charge extrinsèque.

3.2 Limites liées aux participants et au contexte

Un échantillon spécifique - les enseignants et leur profil atypique : Le recrutement d'enseignants comme participants constitue à la fois une force et une limite de cette recherche. Si ce choix se justifie par la pertinence écologique de cette population cible, il introduit également des biais spécifiques. Les enseignants présentent des niveaux de connaissances

initiales sur les TSA supérieures à la population générale, ce qui se traduit par des scores de départ relativement élevés sur les mesures d'attitude et de compréhension. Cette situation « effet plafond » limite la marge de progression observable et peut expliquer certains effets modérés constatés dans les analyses. Par ailleurs, les enseignants constituent une population professionnellement sensibilisée aux enjeux d'inclusion et potentiellement soumise à des pressions de désirabilité sociale particulièrement importantes sur les questions liées au handicap. Cette spécificité peut influencer leurs réponses aux questionnaires auto-rapportés et limiter la détection de changements d'attitude authentiques.

La question de l'évolution des points de vue sur le TSA : L'évolution récente des représentations sociales du TSA, notamment à travers les médias et les campagnes de sensibilisation, constitue un facteur contextuel important qui complexifie l'interprétation des résultats. Les scores élevés observés en pré-test témoignent d'une sensibilisation déjà avancée de l'échantillon, ce qui réduit la fenêtre d'amélioration possible et peut expliquer certains effets de seuil dans les analyses. Cette évolution sociétale positive constitue paradoxalement un défi méthodologique pour les recherches sur la réduction des préjugés envers le TSA.

Limites des conditions expérimentales - Les conditions écologiques : Bien que ce travail ait cherché à reproduire des situations authentiques d'interaction auprès de personnes avec TSA, les conditions expérimentales demeurent artificielles. La durée limitée de l'exposition et le caractère explicite de l'objectif pédagogique diffèrent significativement des contextes naturels d'apprentissage social. Cette limite questionne la validité écologique de ces résultats et leur transférabilité aux situations réelles d'inclusion scolaire.

3.3 Limites méthodologiques et instrumentales

Le protocole de recherche du présent travail s'appuie principalement sur des mesures auto-rapportées par questionnaires, approche qui présente des limitations importantes. Ces mesures explicites sont susceptibles de biais de désirabilité sociale, particulièrement prégnants sur des sujets sensibles comme les attitudes envers le handicap. De plus, elles ne permettent pas d'accéder aux processus automatiques et implicites qui sous-tendent les attitudes et comportements réels.

Nécessité d'utiliser des mesures implicites : L'utilisation de mesures implicites, telles que les tests d'association implicite (IAT) ou les mesures de temps de réaction, permettrait de compléter les données auto-rapportées en accédant aux associations automatiques et aux biais cognitifs

non conscients. Ces mesures sont particulièrement pertinentes pour évaluer les changements d'attitude authentiques et moins soumis aux effets de désirabilité sociale.

Améliorer la mesure de la charge cognitive : Bien que le modèle théorique utilisé intègre la charge cognitive comme variable modératrice potentielle, il n'a pas été possible de déployer des outils spécifiques pour sa mesure objective. L'utilisation de techniques psychophysiologiques (EEG, mesure de la dilatation pupillaire, analyse de la variabilité cardiaque) ou de mesures comportementales (tâches duales, temps de réaction) permettrait une évaluation plus précise de cette variable cruciale pour l'optimisation des environnements virtuels éducatifs.

Ajouter des mesures de la motivation plus élaborées : La motivation, identifiée comme variable médiatrice dans le modèle théorique principal, n'a pas été évaluée. Des instruments sophistiqués, inspirés de la théorie de l'autodétermination de Deci et Ryan (2000), permettraient de mesurer et distinguer les différents types de motivation (intrinsèque et extrinsèque) et leurs effets différentiels sur l'apprentissage empathique.

3.4 Limites des analyses

Les hypothèses de médiation non abordées : Une limite importante de cette recherche réside dans l'impossibilité de tester rigoureusement les hypothèses de médiation formulées dans le modèle théorique. La taille insuffisante de l'échantillon ($N < 200$) ne permet pas de conduire les analyses de médiation complexes nécessaires pour valider les mécanismes psychologiques postulés. Cette limitation empêche de comprendre finement comment et pourquoi l'intervention produit ces effets, informations cruciales pour l'optimisation future des dispositifs. Les analyses de médiation séquentielle et parallèle, particulièrement pertinentes pour tester les chaînes causales incarnation → empathie → compréhension → changement d'attitude, nécessitent des échantillons plus importants et des designs expérimentaux spécifiquement conçus à cet effet.

Variables de contrôle insuffisamment prises en compte : Ce protocole n'a pas systématiquement contrôlé plusieurs variables susceptibles d'influencer les effets observés :

- **Attitudes initiales envers le TSA :** La variabilité des attitudes pré-existantes peut modérer l'efficacité de l'intervention, les participants avec des attitudes initialement négatives présentant potentiellement une marge de progression plus importante.
- **Expérience préalable avec la RV :** La familiarité technologique influence probablement la charge cognitive et l'engagement dans l'expérience virtuelle.

- **Anxiété technologique** : Cette variable pourrait constituer un frein à l'incarnation.
- **Traits d'empathie dispositionnelle** : Les différences individuelles en matière de capacités empathiques constituent probablement des variables modératrices importantes des effets observés.

3.5 Perspectives de recherche

Améliorations méthodologiques - Découvrir la RV via la puissance d'un tutoriel : Les différents travaux de la littérature suggèrent l'importance de développer des phases de familiarisation optimisées. Des recherches futures pourraient explorer l'efficacité de tutoriels immersifs spécifiquement conçus pour favoriser l'appropriation rapide des fondamentaux de la RVI. L'investigation de différentes modalités pédagogiques (guidage progressif, apprentissage par exploration, feedback adaptatif) dans ces phases préparatoires constitue un axe de recherche prometteur. L'étude des courbes d'apprentissage technologique et de leur impact sur l'efficacité des interventions éducatives en RV pourrait engendrer le développement de protocoles d'adaptation personnalisés, tenant compte des profils et expériences individuelles des utilisateurs.

Approches intégratives - L'inclusion dans un programme plus complet : Les limites conceptuelles identifiées soulignent la nécessité de concevoir la RVI comme composante de programmes pédagogiques intégrés plutôt que comme intervention isolée. Des recherches longitudinales pourraient explorer l'efficacité de dispositifs hybrides combinant expériences virtuelles, rencontres de personnes avec TSA, formations théoriques et accompagnement réflexif. L'investigation de synergies entre différentes modalités pédagogiques (RV, simulation, jeu de rôle, témoignages) constitue une perspective particulièrement prometteuse pour optimiser l'efficacité et la durabilité des changements d'attitude. Ces approches multimodales permettraient de bénéficier des avantages spécifiques de chaque méthode tout en compensant leurs limitations respectives.

Développements technologiques et méthodologiques : Les perspectives technologiques incluent l'exploration de nouvelles modalités d'interaction (haptique), l'intégration d'intelligence artificielle pour personnaliser les expériences et le développement d'environnements virtuels collaboratifs permettant des interactions authentiques entre utilisateurs neurotypiques et personnes avec TSA. Sur le plan méthodologique, l'intégration de mesures psychophysiologicals en temps réel (EEG, GSR, *eye-tracking*) ouvre des perspectives d'évaluations objectives des processus cognitifs et émotionnels. Ces approches permettraient de

développer des systèmes adaptatifs ajustant automatiquement les paramètres de l'expérience virtuelle en fonction de l'état de l'utilisateur.

Élargissement des populations et contextes : Des recherches futures pourraient étendre l'investigation à d'autres populations (étudiants, grand public) et contextes (formation dans le monde professionnel, sensibilisation communautaire). L'exploration des facteurs culturels et sociodémographiques modulant l'efficacité des interventions constitue également un axe important pour le développement d'outils inclusifs et universellement accessibles.

L'investigation longitudinale des effets, incluant des mesures de rétention à moyen et long terme, permettrait d'évaluer la durabilité des changements induits et d'identifier les facteurs prédictifs de leur maintien dans le temps. Ces perspectives de recherche, en abordant les limitations identifiées de manière systématique, contribueraient à l'émergence d'une nouvelle génération d'outils techno-pédagogiques pour l'inclusion, plus efficaces, adaptatifs et intégrés dans des démarches éducatives globales.

4. Conclusion générale

Cette recherche doctorale avait pour ambition d'évaluer l'efficacité de la RVI pour la sensibilisation au TSA, en s'appuyant sur une validation empirique du modèle CAMIL de Makransky et Petersen (2021). Les résultats obtenus invitent à une compréhension nuancée des potentialités des technologies immersives en contexte éducatif. Si cette thèse confirme l'efficacité de la RVI pour générer les mécanismes psychologiques proximaux (présence, agentivité et empathie), elle révèle également que ces processus ne se traduisent pas automatiquement par une supériorité pédagogique sur les modalités traditionnelles de sensibilisation.

La RVI n'est pas la solution miracle espérée pour transformer les attitudes envers les personnes avec TSA. Elle constitue plutôt un outil complémentaire dont l'efficacité dépend de conditions spécifiques : design expérientiel sophistiqué, intégration dans des programmes pédagogiques plus larges et adaptation aux caractéristiques des publics cibles.

Cette recherche contribue ainsi à tempérer l'optimisme technologique parfois excessif entourant les applications éducatives de la RV, tout en identifiant des pistes prometteuses pour son utilisation raisonnée. L'avenir de la sensibilisation aux différences neurodéveloppementales réside probablement dans l'articulation réfléchie entre innovations technologiques et approches pédagogiques éprouvées, au service d'une société plus inclusive.

Au terme de ce parcours doctoral, cette thèse espère avoir contribué, avec la rigueur et la nuance nécessaires, à une meilleure compréhension des conditions dans lesquelles la technologie peut servir l'inclusion, sans jamais perdre de vue que l'humain et ses diversités demeurent au cœur de toute démarche éducative authentique.

Vivre la vie d'autrui, voir à travers ses yeux. A travers la réalité virtuelle, c'est une fiction devenue réalité, mais surtout un outil fantastique pour tente de mieux comprendre l'autre.

Propos recueillis au GEM la bulle de Rodez, octobre 2022.

Références

Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1992). *Dispositif CAVE*.

<https://www.evl.uic.edu/documents/cacm92-cave-cruz-neira.pdf>

Deweerd, S. (2018, 31 août). Les révisions proposées sur l'autisme pour la CIM-11 soulèvent des préoccupations. *Mediapart*. <https://blogs.mediapart.fr/jean-vincot/blog/310818/les-revisions-proposees-sur-lautisme-pour-la-cim-11-soulevent-des-preoccupations>

Drifter Entertainment. (s.d.). *Lies Beneath* [Tutoriel vidéo]. YouTube.

<https://www.youtube.com/watch?v=6KOxD9eBUYE>

Electronic Arts. (s.d.). *Mass Effect Andromeda*. <https://www.ea.com/fr-fr/games/mass-effect/mass-effect-andromeda>

Haute Autorité de Santé. (2018). *TSA : Des signes d'alerte à la consultation dédiée en soins primaires (1^{re} ligne) - Synthèse*. https://www.has-sante.fr/upload/docs/application/pdf/2018-02/tsa_-_des_signes_dalerte_a_la_consultation_dediee_en_soins_primaires_1er_ligne_-_synthese.pdf

Haute Autorité de Santé. (2023). *Note de cadrage : Trouble du spectre de l'autisme (TSA) : Interventions et parcours de vie de l'enfant et de l'adolescent*.

<https://www.cnape.fr/documents/haute-autorite-de-sante-note-de-cadrage-sur-trouble-du-spectre-de-lautisme-tsa-interventions-et-parcours-de-vie-de-lenfant-et-de-ladolescent/>

History of Information. (s.d.). *Le Sensorama*.

<https://www.historyofinformation.com/detail.php?id=2785>

Ifop. (s.d.). *Le regard des Français sur l'autisme*. <https://www.ifop.com/publication/le-regard-des-francais-sur-lautisme/>

In the cell. (s.d.). Oculus. https://www.oculus.com/experiences/gear-vr/1108839405883293/?locale=fr_FR

KAT VR. (s.d.). *KAT Walk C 2 Core*. <https://www.kat-vr.com/products/kat-walk-c-2-core>

Le Figaro Santé. (s.d.). *Suivez en direct la première opération chirurgicale assistée par réalité virtuelle*. <http://sante.lefigaro.fr/article/suivez-en-direct-la-premiere-operation-chirurgicale-assistee-par-realite-virtuelle/>

Mathieu, M. (2019, 26 avril). Réalité virtuelle avec les yeux d'un enfant autiste. *La Presse*. <https://www.lapresse.ca/societe/famille/2019-04-26/realite-virtuelle-avec-les-yeux-d-un-enfant-autiste>

Meta. (s.d.). *Présentation tutoriel Meta* [Vidéo]. YouTube. <https://www.youtube.com/watch?v=VZydMRa4RZE>

Ministère de l'Éducation nationale. (2005). Circulaire interministérielle n° 2005-124 du 8 mars 2005 relative à la prise en charge des personnes atteintes d'autisme et de troubles envahissants du développement. *Bulletin Officiel de l'Éducation Nationale*, 15.

NVIDIA. (s.d.). *Omniverse Audio2Face*. <https://www.nvidia.com/en-us/omniverse/apps/audio2face/>

Réalité Virtuelle. (s.d.). *Vtopia : Surgical VR chirurgie*. <https://www.realite-virtuelle.com/vtopia-surgical-vr-chirurgie-3110>

Réseau national d'expertise en troubles du spectre de l'autisme. (s.d.). *Dans la peau d'une personne autiste*. <https://www.rnetsa.ca/fr/centre-de-documentation/182/dans-la-peau-d-une-personne-autiste>

The Body VR – Journey inside a cell. (s.d.). Steam. https://store.steampowered.com/app/451980/The_Body_VR_Journey_Inside_a_Cell/?l=french

The displaced. (s.d.). MIT Docubase. <https://docubase.mit.edu/project/the-displaced/>

Université de Montréal. (2023, 12 avril). *Une nouvelle application pour se mettre dans la peau d'une personne autiste*. <https://nouvelles.umontreal.ca/article/2023/04/12/une-nouvelle-application-pour-se-mettre-dans-la-peau-d-une-personne-autiste/>

SAIS92. (s.d.). *Guide « Scolariser les élèves autistes ou présentant des troubles envahissants du développement »*. Consulté le 31 octobre 2022, à l'adresse <https://www.sais92.fr/scolarisation/22-guides/72-guide-scolariser-les-eleves-autistes-ou-presentant-des-troubles-envahissants-du-developpement>

VR Expert. (s.d.). *Les différences entre l'Oculus Quest 1 et 2*. <https://vr.vr-expert.com/fr/les-differences-entre-oculus-quest-1-et-2/>

Bibliographie

- Ackerman, E. (2016). Before virtual reality was cool. *IEEE Spectrum*, 53(12), 68.
<https://doi.org/10.1109/MSPEC.2016.7761893>
- Ai-Lim Lee, E., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, 55(4), 1424–1442. <https://doi.org/10.1016/j.compedu.2010.06.006>
- Ajzen, I., & Fishbein, M. (2000). Attitudes and the attitude-behavior relation: Reasoned and automatic processes. *European Review of Social Psychology*, 11(1), 1–33.
<https://doi.org/10.1080/14792779943000116>
- Albus, P., Vogt, A., & Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education*, 166.
<https://doi.org/10.1016/j.compedu.2021.104154>
- Allcoat, D., & von Mühlennen, A. (2018). Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology*, 26.
<https://doi.org/10.25304/rlt.v26.2140>
- Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798–844). Clark University Press.
- Altinpulluk, H. (2019). Determining the trends of using augmented reality in education between 2006-2016. *Education and Information Technologies*, 24(2), 1089–1114.
<https://doi.org/10.1007/s10639-018-9806-3>
- American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.; DSM-5-TR). American Psychiatric Association Publishing. <https://doi.org/10.1176/appi.books.9780890425787>
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369–406.
<https://doi.org/10.1037/0033-295X.89.4.369>
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

- Anses. (2021). *Expositions aux technologies de réalité virtuelle et/ou augmentée : Avis de l'Anses : Rapport d'expertise collective* (Saisine n° 2017-SA-0076).
<https://www.anses.fr/fr/system/files/AP2017SA0076Ra.pdf>
- Arnaldi, B., Guitton, P., & Moreau, G. (2018). *Réalité virtuelle et réalité augmentée : Mythes et réalités*. ISTE. <https://doi.org/10.51926/ISTE.9781784054670>
- Arzy, S., Thut, G., Mohr, C., Michel, C. M., & Blanke, O. (2006). Neural basis of embodiment: Distinct contributions of temporoparietal junction and extrastriate body area. *Journal of Neuroscience*, 26(31), 8074–8081.
<https://doi.org/10.1523/JNEUROSCI.0745-06.2006>
- Attwood, T. (2022). *Le syndrome d'Asperger et l'autisme de haut niveau : Approche comportementaliste*. Dunod.
- Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. W. W. Norton & Company.
- Banakou, D., Kishore, S., & Slater, M. (2018). Virtually being Einstein results in an improvement in cognitive task performance and a decrease in age bias. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.00917>
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1), 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bandura, A. (2019). *Auto-efficacité : Comment le sentiment d'efficacité personnelle influence notre qualité de vie*. De Boeck Supérieur.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21(1), 37–46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8)
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, 6(6), 248–254. [https://doi.org/10.1016/s1364-6613\(02\)01904-6](https://doi.org/10.1016/s1364-6613(02)01904-6)
- Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T., & Chakrabarti, B. (2009). Talent in autism: Hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1522), 1377–1383. <https://doi.org/10.1098/rstb.2008.0337>

- Baron-Cohen, S. (2017). Editorial perspective: Neurodiversity – a revolutionary concept for autism and psychiatry. *Journal of Child Psychology and Psychiatry*, 58(6), 744–747. <https://doi.org/10.1111/jcpp.12703>
- Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., Klein, T. R., & Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, 72(1), 105–118. <https://doi.org/10.1037/0022-3514.72.1.105>
- Belmonte, M. K., Allen, G., Beckel-Mitchener, A., Boulanger, L. M., Carper, R. A., & Webb, S. J. (2004). Autism and abnormal development of brain connectivity. *Journal of Neuroscience*, 24(42), 9228–9231. <https://doi.org/10.1523/JNEUROSCI.3340-04.2004>
- Bennett, M., & Goodall, E. (2016). A meta-analysis of DSM-5 autism diagnoses in relation to DSM-IV and DSM-IV-TR. *Review Journal of Autism and Developmental Disorders*, 3(2), 119–124. <https://doi.org/10.1007/s40489-016-0070-4>
- Bickman, L. (1972). Environmental attitudes and actions. *The Journal of Social Psychology*, 87(2), 323–324. <https://doi.org/10.1080/00224545.1972.9922533>
- Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments [1]. *Journal of Computer-Mediated Communication*, 3(2), JCMC324. <https://doi.org/10.1111/j.1083-6101.1997.tb00070.x>
- Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, 6(6), 237–242. [https://doi.org/10.1016/S1364-6613\(02\)01907-1](https://doi.org/10.1016/S1364-6613(02)01907-1)
- Bled, C. (2022). *Imagerie mentale et style cognitif dans le spectre de l'autisme* [Thèse de doctorat, Université Toulouse 2 Jean Jaurès]. <https://www.theses.fr/s219133>
- Boberg, M., Piippo, P., & Ollila, E. (2008). Designing avatars. *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts*, 232–239. <https://doi.org/10.1145/1413634.1413679>
- Bochner, S., & Insko, C. A. (1966). Communicator discrepancy, source credibility, and opinion change. *Journal of Personality and Social Psychology*, 4(6), 614–621. <https://doi.org/10.1037/h0021192>

- Boldsen, S. (2018). Toward a phenomenological account of embodied subjectivity in autism. *Culture, Medicine, and Psychiatry*, 42(4), 893–913. <https://doi.org/10.1007/s11013-018-9590-y>
- Bolt, H. M., & Hengstler, J. G. (2018). Contemporary trends in toxicological research on arsenic. *Archives of Toxicology*, 92(11), 3251–3253. <https://doi.org/10.1007/s00204-018-2311-z>
- Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, 391(6669), 756. <https://doi.org/10.1038/35784>
- Bouchard, S., Bernier, F., Boivin, E., Dumoulin, S., Laforest, M., Guitard, T., Robillard, G., Monthuy-Blanc, J., & Renaud, P. (2013). Empathy toward virtual humans depicting a known or unknown person expressing pain. *Cyberpsychology, Behavior and Social Networking*, 16(1), 61–71. <https://doi.org/10.1089/cyber.2012.1571>
- Bouffard, T., & Vezeau, C. (2010). Intention d'apprendre, motivation et apprentissage autorégulé : Le rôle de la perception de compétence et des émotions. Dans M. Crahay & M. Dutrévis (Dir.), *Psychologie des apprentissages scolaires* (pp. 66–84). De Boeck.
- Bourgeron, T. (2016). Current knowledge on the genetics of autism and propositions for future research. *Comptes Rendus Biologies*, 339(7), 300–307. <https://doi.org/10.1016/j.crv.2016.05.004>
- Boyd, J., Barnett, W. S., Bodrova, E., Leong, D. J., & Gomby, D. (2005). *Promoting children's social and emotional development through preschool education* (NIEER Preschool Policy Brief, Issue 7). National Institute for Early Education Research, Rutgers University. <https://nieer.org/research-library/promoting-childrens-social-emotional-development-through-preschool-education>
- Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.
- Bruckert, L., & Sanguin-Bruckert, C. (2013). *Le rôle de la voix de l'enseignant dans les situations d'apprentissage* [Communication]. <https://hal-univ-tlse2.archives-ouvertes.fr/hal-01018635>

- Brugada-Ramentol, V., Clemens, I., & de Polavieja, G. G. (2019). Active control as evidence in favor of sense of ownership in the moving virtual hand illusion. *Consciousness and Cognition, 71*, 123–135. <https://doi.org/10.1016/j.concog.2019.04.003>
- Brun, C., Demazeux, S., Vittorio, P. D., Gonon, F., Gorry, P., Konsman, J. P., Lung, F., Lung, Y., Minard, M., Montalban, M., Rumeau, N., & Smith, A. (2015). La construction des catégories diagnostiques de maladie mentale. *Revue de la Régulation. Capitalisme, Institutions, Pouvoirs, 17*. <https://doi.org/10.4000/regulation.11299>
- Buttussi, F., & Chittaro, L. (2023). Acquisition and retention of spatial knowledge through virtual reality experiences: Effects of VR setup and locomotion technique. *International Journal of Human-Computer Studies, 177*. <https://doi.org/10.1016/j.ijhcs.2023.103067>
- Cadet, L. B., & Chainay, H. (2021). How preadolescents and adults remember and experience virtual reality: The role of avatar incarnation, emotion, and sense of presence. *International Journal of Child-Computer Interaction, 29*. <https://doi.org/10.1016/j.ijcci.2021.100299>
- Cao, Z., Jerald, J., & Kopper, R. (2018). Visually-induced motion sickness reduction via static and dynamic rest frames. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 105–112. <https://doi.org/10.1109/VR.2018.8446210>
- Caron, M.-J., Mottron, L., Berthiaume, C., & Dawson, M. (2006). Cognitive mechanisms, specificity and neural underpinnings of visuospatial peaks in autism. *Brain, 129*(7), 1789–1802. <https://doi.org/10.1093/brain/awl072>
- Carruth, D. W. (2017). Virtual reality for education and workforce training. *15th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, 1–6. <https://doi.org/10.1109/ICETA.2017.8102472>
- Caspar, E. A., Cleeremans, A., & Haggard, P. (2015). The relationship between human agency and embodiment. *Consciousness and Cognition, 33*, 226–236. <https://doi.org/10.1016/j.concog.2015.01.007>
- Champney, R. K., Stanney, K. M., Hash, P. A. K., Malone, L. C., Kennedy, R. S., & Compton, D. E. (2007). Recovery from virtual environment exposure: Expected time course of symptoms and potential readaptation strategies. *Human Factors, 49*(3), 491–506. <https://doi.org/10.1518/001872007X200120>

- Chang, C.-Y., Chen, D. C.-L., & Chang, W.-K. (2019). Research on immersion for learning using virtual reality, augmented reality and mixed reality. *Enfance*, 3(3), 413–426. <https://doi.org/10.3917/enf2.193.0413>
- Chen, A., Hanna, J. J., Manohar, A., & Tobia, A. (2018). Teaching empathy: The implementation of a video game into a psychiatry clerkship curriculum. *Academic Psychiatry*, 42(3), 362–365. <https://doi.org/10.1007/s40596-017-0862-6>
- Chen, J., Peñagarikano, O., Belgard, T., Swarup, V., & Geschwind, D. (2015). The emerging picture of autism spectrum disorder: Genetics and pathology. *Annual Review of Pathology*, 10, 111–144. <https://doi.org/10.1146/annurev-pathol-012414-040405>
- Chen, Y., Smith, T. J., York, C. S., & Mayall, H. J. (2020). Google Earth virtual reality and expository writing for young English learners from a funds of knowledge perspective. *Computer Assisted Language Learning*, 33(1–2), 1–25. <https://doi.org/10.1080/09588221.2018.1544151>
- Cheng, L. K., Chieng, M. H., & Chieng, W. H. (2014). Measuring virtual experience in a three-dimensional virtual reality interactive simulator environment: A structural equation modeling approach. *Virtual Reality*, 18(3), 173–188. <https://doi.org/10.1007/s10055-014-0244-2>
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42(2), 21–29. <https://doi.org/10.1007/BF02299088>
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Collins, J., Regenbrecht, H., Langlotz, T., Said Can, Y., Ersoy, C., & Butson, R. (2019). Measuring cognitive load and insight: A methodology exemplified in a virtual reality learning context. *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 351–352. <https://doi.org/10.1109/ISMAR.2019.00033>
- Cook, M., Hall, N., Lischer-Katz, Z., Hardesty, J., Johnson, J., McDonald, R. H., & Carlisle, T. (2019). Challenges and strategies for educational virtual reality: Results of an expert-led forum on 3D/VR technologies across academic institutions. *Information Technology and Libraries*, 38(4), 25–48. <https://doi.org/10.6017/ital.v38i4.11075>

- Cooper, G., Park, H., Nasr, Z., Thong, L. P., & Johnson, R. (2019). Using virtual reality in the classroom: Preservice teachers' perceptions of its use as a teaching and learning tool. *Educational Media International*, 56(1), 1–13.
<https://doi.org/10.1080/09523987.2019.1583461>
- Crisp, R. J., & Turner, R. N. (2009). Can imagined interactions produce positive perceptions? Reducing prejudice through simulated social contact. *American Psychologist*, 64(4), 231–240. <https://doi.org/10.1037/a0014718>
- Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19(2), 272–309. <https://doi.org/10.1080/15213269.2015.1015740>
- D'Ambrosio, F., Olivier, M., Didon, D., & Besche, C. (2009). The basic empathy scale: A French validation of a measure of empathy in youth. *Personality and Individual Differences*, 46(2), 160–165. <https://doi.org/10.1016/j.paid.2008.09.020>
- Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, 41(1), 10–32.
<https://doi.org/10.1111/j.1467-8535.2009.01038.x>
- Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, 9(1), 28–30.
<https://doi.org/10.1038/nn1611>
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85–104.
- Day, B., Ebrahimi, E., Hartman, L. S., Pagano, C. C., Robb, A. C., & Babu, S. V. (2019). Examining the effects of altered avatars on perception-action in virtual reality. *Journal of Experimental Psychology: Applied*, 25(1), 1–24.
<https://doi.org/10.1037/xap0000192>
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3(2), 71–100.
<https://doi.org/10.1177/1534582304267187>

- Deci, E. L., & Ryan, R. M. (2015). Self-determination theory. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., pp. 486–491). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.26036-4>
- Demitriadou, E., Stavroulia, K.-E., & Lanitis, A. (2020). Comparative evaluation of virtual and augmented reality for teaching mathematics in primary education. *Education and Information Technologies*, 25(1), 381–401. <https://doi.org/10.1007/s10639-019-09973-5>
- Denis, M. (1989). *Image et cognition*. Presses Universitaires de France.
- Descroix, E. (2021). *Spontaneous blinking in communication and conscious processing of information* [Thèse de doctorat, Université Grenoble Alpes]. <https://theses.hal.science/tel-03594601>
- Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, 48(6), 1267–1278. <https://doi.org/10.1016/j.jesp.2012.06.003>
- Dewez, D., Fribourg, R., Argelaguet, F., Hoyet, L., Mestre, D., Slater, M., & Lécuyer, A. (2019). Influence of personality traits and body awareness on the sense of embodiment in virtual reality. *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 123–134. <https://doi.org/10.1109/ISMAR.2019.00-12>
- Dicheva, D., & Dichev, C. (2015). Gamification in education: Where are we in 2015? In *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 1445–1454). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/primary/p/152186/>
- Domínguez, L. G., Velázquez, J. L. P., & Galán, R. F. (2013). A model of functional brain connectivity and background noise as a biomarker for cognitive phenotypes: Application to autism. *PLOS ONE*, 8(4). <https://doi.org/10.1371/journal.pone.0061493>
- Dyer, E., Swartzlander, B. J., & Gugliucci, M. R. (2018). Using virtual reality in medical education to teach empathy. *Journal of the Medical Library Association*, 106(4), 498–500. <https://doi.org/10.5195/jmla.2018.518>
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>

- Ehrsson, H. H., Wiech, K., Weiskopf, N., Dolan, R. J., & Passingham, R. E. (2007). Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Proceedings of the National Academy of Sciences*, *104*(23), 9828–9833. <https://doi.org/10.1073/pnas.0610011104>
- Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., & Grafton, S. T. (2008). The angular gyrus computes action awareness representations. *Cerebral Cortex*, *18*(2), 254–261. <https://doi.org/10.1093/cercor/bhm050>
- Fagnant, A., Demonty, I., Dierendonck, C., Dupont, V., & Marcoux, G. (2014). Chapitre 11. Résolution de tâches complexes, évaluation « en phases » et compétence en mathématiques. In C. Dierendonck (Ed.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel* (pp. 179–189). De Boeck Supérieur. <https://doi.org/10.3917/dbu.dinck.2014.01.0179>
- Falcone, S., Brouwer, A.-M., Heylen, D., Van Erp, J., Zhang, L., Pradhan, S. S., Stuldreher, I. V., Cocu, I., Heuvel, M., de Vries, P. S., Gijbertse, K., & Englebienne, G. (2022). Pupil diameter as implicit measure to estimate sense of embodiment. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*(44). <https://escholarship.org/uc/item/8rt963b9>
- Favrod, J., Mcquillan, A., Pomini, V., & Ferrero, F. P. (1998). Training interpersonal problem-solving skills in French-speaking Switzerland. *International Review of Psychiatry*, *10*(1), 30–34. <https://doi.org/10.1080/09540269875050>
- Fernandes, A. S., & Feiner, S. K. (2016). Combating VR sickness through subtle dynamic field-of-view modification. *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, 201–210. <https://doi.org/10.1109/3DUI.2016.7460053>
- Festinger, L., & Maccoby, N. (1964). On resistance to persuasive communications. *The Journal of Abnormal and Social Psychology*, *68*(4), 359–366. <https://doi.org/10.1037/h0049073>
- Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741. <https://doi.org/10.1007/s10648-015-9348-9>
- Fleming, R., Mohler, B. J., Romero, J., Black, M. J., & Breidt, M. (2017). Appealing avatars from 3D body scans: Perceptual effects of stylization. In J. Braz, N. Magnenat-Thalmann, P. Richard, L. Linsen, A. Telea, S. Battiato, & F. Imai (Eds.), *Computer*

- vision, imaging and computer graphics theory and applications* (pp. 175–196). Springer International Publishing. https://doi.org/10.1007/978-3-319-64870-5_9
- Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann Publishers Inc. <https://doi.org/10.1016/C2009-0-21512-1>
- Fointiat, V., & Barbier, L. (2015). Persuasion et influence : Changer les attitudes, changer les comportements. Regards de la psychologie sociale. *Journal d'Interaction Personne-Système*, 4(1). <https://doi.org/10.46298/jips.1301>
- Fowler, C. (2015). Virtual reality and learning: Where is the pedagogy? *British Journal of Educational Technology*, 46(2), 412–422. <https://doi.org/10.1111/bjet.12135>
- Frederiksen, J. G., Sørensen, S. M. D., Konge, L., Svendsen, M. B. S., Nobel-Jørgensen, M., Bjerrum, F., & Andersen, S. A. W. (2020). Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: A randomized trial. *Surgical Endoscopy*, 34(3), 1244–1252. <https://doi.org/10.1007/s00464-019-06887-8>
- Fribourg, R. (2020). *Contribution à l'étude des facteurs influençant le sentiment d'incarnation envers un avatar en réalité virtuelle* [Thèse de doctorat, Université Rennes 1]. <https://tel.archives-ouvertes.fr/tel-03191307>
- Frith, U. (1989). *Autism: Explaining the enigma*. Blackwell Publishing.
- Fuchs, P. (1996). *Les interfaces de la réalité virtuelle*. AJIIMD. <https://minesparis-psl.hal.science/hal-00785588>
- Fuchs, P. (Dir.). (2006). *Le traité de la réalité virtuelle : Vol. 1. L'Homme et l'environnement virtuel* (G. Moreau, Coord.). Presses des Mines.
- Fuchs, P. (Dir.). (2006). *Le traité de la réalité virtuelle : Vol. 2. L'interfaçage, l'immersion et l'interaction en environnement virtuel* (J.-M. Burkhardt & S. Coquillart, Coord.). Presses des Mines.
- Gall, D., Roth, D., Stauffert, J.-P., Zarges, J., & Latoschik, M. E. (2021). Embodiment in virtual reality intensifies emotional responses to virtual stimuli. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.674179>

- Gauthier, G., Charette, E., Castell, P., Ringuet, S., Fontaine, J., Munn, Y., & Leroux, J. L. (2016). *Définition d'une famille de tâches*. <https://eduq.info/xmlui/handle/11515/35756>
- Gao, T., Zhang, T., Zhu, L., Gao, Y., & Qiu, L. (2019). Exploring psychophysiological restoration and individual preference in the different environments based on virtual reality. *International Journal of Environmental Research and Public Health*, 16(17). <https://doi.org/10.3390/ijerph16173102>
- Gegenfurtner, A., Quesada-Pallarès, C., & Knogler, M. (2014). Digital simulation-based training: A meta-analysis. *British Journal of Educational Technology*, 45(6), 1097–1114. <https://doi.org/10.1111/bjet.12188>
- Girandola, F., & Roussiau, N. (2003). L'engagement comme source de modifications à long terme. *Cahiers Internationaux de Psychologie Sociale*, 57, 83–101.
- Gohier, C., Anadón, M., Bouchard, Y., Charbonneau, B., & Chevrier, J. (2001). La construction identitaire de l'enseignant sur le plan professionnel : Un processus dynamique et interactif. *Revue des Sciences de l'Éducation*, 27(1), 3–32. <https://doi.org/10.7202/000304ar>
- Goldin-Meadow, S. (2011). Learning through gesture. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(6), 595–607. <https://doi.org/10.1002/wcs.132>
- Gorisse, G. (2019). *Impact du point de vue et de la fidélité des avatars sur les sentiments de présence et d'incarnation en environnement virtuel immersif* [Thèse de doctorat, Paris, ENSAM]. <https://www.theses.fr/2019ENAM0004>
- Gorisse, G., Christmann, O., Amato, E. A., & Richir, S. (2017). First- and third-person perspectives in immersive virtual environments: Presence and performance analysis of embodied users. *Frontiers in Robotics and AI*, 4. <https://doi.org/10.3389/frobt.2017.00033>
- Gorisse, G., Christmann, O., & Richir, S. (2018). De la présence à l'incarnation : Proposition d'un méta-modèle pour la réalité virtuelle. *Interfaces Numériques*, 7(1). <https://doi.org/10.25965/interfaces-numeriques.3295>
- Grandin, T. (2009). How does visual thinking work in the mind of a person with autism? A personal account. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1522), 1437–1442. <https://doi.org/10.1098/rstb.2008.0297>

- Gutiérrez-Maldonado, J., Alsina-Jurnet, I., & Rangel-Gómez, M.-V. (2011). The role of presence in the level of anxiety experienced in clinical virtual environments. *Computers in Human Behavior*, 27(1), 504–512.
<https://doi.org/10.1016/j.chb.2010.09.018>
- Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4), 196–207. <https://doi.org/10.1038/nrn.2017.14>
- Han, I., Shin, H. S., Ko, Y., & Shin, W. S. (2022). Immersive virtual reality for increasing presence and empathy. *Journal of Computer Assisted Learning*, 38(4), 1115–1126.
<https://doi.org/10.1111/jcal.12669>
- Happé, F., & Vital, P. (2009). What aspects of autism predispose to talent? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1522), 1369–1375.
<https://doi.org/10.1098/rstb.2008.0332>
- Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy Insights from the Behavioral and Brain Sciences*, 3(2), 220–227. <https://doi.org/10.1177/2372732216655542>
- Harm, D. L., Taylor, L. C., Reschke, M. F., Somers, J. T., & Bloomberg, J. J. (2008). Sensorimotor coordination aftereffects of exposure to a virtual environment. *The Visual Computer*, 24(11), 995–999. <https://doi.org/10.1007/s00371-008-0277-1>
- Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human–Computer Interaction*, 13(4), 481–499.
https://doi.org/10.1207/S15327590IJHC1304_07
- Hedberg, J. G., Harper, B., & Brown, C. (1993). Reducing cognitive load in multimedia navigation. *Australasian Journal of Educational Technology*, 9(2).
<https://doi.org/10.14742/ajet.2106>
- Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE*, 13(10). <https://doi.org/10.1371/journal.pone.0204494>
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127.
https://doi.org/10.1207/s15326985ep4102_4

- Higgins, C., & Walker, R. (2012). Ethos, logos, pathos: Strategies of persuasion in social/environmental reports. *Accounting Forum*, 36(3), 194–208.
<https://doi.org/10.1016/j.accfor.2012.02.003>
- Hite, R. (2022). Virtual reality—Flight of fancy or feasible? Ways to use virtual reality technologies to enhance students' science learning. *The American Biology Teacher*, 84(2), 106–108. <https://doi.org/10.1525/abt.2022.84.2.106>
- Hocine, N., Gouaich, A., Di Loreto, I., & Abrouk, L. (2011). Techniques d'adaptation dans les jeux ludiques et sérieux. *Revue d'Intelligence Artificielle*, 25(2), 253–280.
<https://doi.org/10.3166/ria.25.253-280>
- Hojat, M. (2007). The Jefferson Scale of Physician Empathy. In M. Hojat (Ed.), *Empathy in patient care: Antecedents, development, measurement, and outcomes* (pp. 87–115). Springer. https://doi.org/10.1007/0-387-33608-7_7
- Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. (1949). *Experiments on mass communication* (Vol. 3). Princeton University Press.
- Iarocci, G., & McDonald, J. (2006). Sensory integration and the perceptual experience of persons with autism. *Journal of Autism and Developmental Disorders*, 36(1), 77–90.
<https://doi.org/10.1007/s10803-005-0044-3>
- Ichbiah, D., & Lefranc, J. -M. (2014). *Les rebelles numériques*. First Éditions.
- Jaarsma, P., & Welin, S. (2012). Autism as a natural human variation: Reflections on the claims of the neurodiversity movement. *Health Care Analysis*, 20(1), 20–30.
<https://doi.org/10.1007/s10728-011-0169-9>
- Jarjour, I. (2014). Neurodevelopmental outcome after extreme prematurity: A review of the literature. *Pediatric Neurology*, 52(2), 143–152.
<https://doi.org/10.1016/j.pediatrneurol.2014.10.027>
- Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, 23(4), 1515–1529. <https://doi.org/10.1007/s10639-017-9676-0>
- Jerald, J. (2016). *The VR book: Human-centered design for virtual reality*. Association for Computing Machinery and Morgan & Claypool. <https://doi.org/10.1145/2792790>

- Jeunet, C., Albert, L., Argelaguet, F., & Lécuyer, A. (2018). "Do you feel in control?": Towards novel approaches to characterise, manipulate and measure the sense of agency in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 24(4), 1486–1495. <https://doi.org/10.1109/TVCG.2018.2794598>
- Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? *Journal of Child Psychology and Psychiatry*, 38(5), 527–534. <https://doi.org/10.1111/j.1469-7610.1997.tb01539.x>
- Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, 29(4), 589–611. <https://doi.org/10.1016/j.adolescence.2005.08.010>
- Joseph, R. M., Keehn, B., Connolly, C., Wolfe, J. M., & Horowitz, T. S. (2009). Why is visual search superior in autism spectrum disorder? *Developmental Science*, 12(6), 1083–1096. <https://doi.org/10.1111/j.1467-7687.2009.00855.x>
- Joule, R.-V., Girandola, F., & Bernard, F. (2007). How can people be induced to willingly change their behavior? The path from persuasive communication to binding communication. *Social and Personality Psychology Compass*, 1(1), 493–505. <https://doi.org/10.1111/j.1751-9004.2007.00018.x>
- Jourabloo, A., De la Torre, F., Saragih, J., Wei, S.-E., Lombardi, S., Wang, T.-L., Belko, D., Trimble, A., & Badino, H. (2022). Robust egocentric photo-realistic facial expression transfer for virtual reality. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (p. 20323–20332). IEEE. <https://doi.org/10.1109/CVPR52688.2022.01968>
- Kätsyri, J., Förger, K., Mäkräinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00390>
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology*, 3(3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3

- Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4), 373–387. https://doi.org/10.1162/PRES_a_00124
- Kim, C. S., Jung, M., Kim, S. Y., & Kim, K. (2020). Controlling the sense of embodiment for virtual avatar applications: Methods and empirical study. *JMIR Serious Games*, 8(3). <https://doi.org/10.2196/21879>
- Kim, S., Nussbaum, M. A., & Gabbard, J. L. (2019). Influences of augmented reality head-worn display type and user interface design on performance and usability in simulated warehouse order picking. *Applied Ergonomics*, 74, 186–193. <https://doi.org/10.1016/j.apergo.2018.08.026>
- Kim, S. Y. (2020). The development and pilot-testing of the Autism Attitude Acceptance Scale: An instrument measuring autism acceptance. *Autism in Adulthood*, 2(3), 204–215. <https://doi.org/10.1089/aut.2019.0066>
- Klingenberg, S., Jørgensen, M., Dandanell, G., Skriver, K., Mottelson, A., & Makransky, G. (2020). Investigating the effect of teaching as a general learning strategy when learning through desktop and immersive VR: A media and methods experiment. *British Journal of Educational Technology*, 51(6), 2219–2236. <https://doi.org/10.1111/bjet.13029>
- Knogler, M., Harackiewicz, J. M., Gegenfurtner, A., & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology*, 43, 39–50. <https://doi.org/10.1016/j.cedpsych.2015.08.004>
- Kokkinara, E., & Slater, M. (2014). Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception*, 43(1), 43–58. <https://doi.org/10.1068/p7545>
- Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education*, 14(1), 23–40. <https://doi.org/10.1007/BF03173109>
- Lanier, J., & Biocca, F. (1992). An insider's view of the future of virtual reality. *Journal of Communication*, 42(4), 150–172. <https://doi.org/10.1111/j.1460-2466.1992.tb00816.x>

- Lee, K. M. (2004). Presence, explicated. *Communication Theory*, 14(1), 27–50.
<https://doi.org/10.1111/j.1468-2885.2004.tb00302.x>
- Lee, C. K., & Shea, M. (2020). Exploring the use of virtual reality by pre-service elementary teachers for teaching science in the elementary classroom. *Journal of Research on Technology in Education*, 52(2), 163–177.
<https://doi.org/10.1080/15391523.2020.1726234>
- Lehmann-Grube, S. K., Tobisch, A., & Dresel, M. (2024). Changing preservice teacher students' stereotypes and attitudes and reducing judgment biases concerning students of different family backgrounds: Effects of a short intervention. *Social Psychology of Education*, 27(4), 1621–1658. <https://doi.org/10.1007/s11218-023-09862-3>
- Leveau, P. H., & Camus, S. (2022). *Le rôle de l'incarnation dans les émotions suscitées lors des visites touristiques en réalité virtuelle* [Communication]. 9ème conférence de l'AFMAT (Association Française de Management du Tourisme). <https://hal.univ-angers.fr/hal-03688697>
- Lewin, K. (1951). *Field theory in social science: Selected theoretical papers* (1st ed.). Harper.
- Lewis, F., Plante, P., & Lemire, D. (2021). Pertinence, efficacité et principes pédagogiques de la réalité virtuelle et augmentée en contexte scolaire. *Médiations et Médiatisations*, 5, 43–64. <https://doi.org/10.52358/mm.vi5.161>
- Leyens, J. P., Yzerbyt, V., & Schadron, G. (1996). *Stéréotypes et cognition sociale*. Mardaga.
- Li, P., Legault, J., Klippel, A., & Zhao, J. (2020). Virtual reality for student learning: Understanding individual differences. *Human Behaviour and Brain*, 1(1), 28–36.
<https://doi.org/10.37716/HBAB.2020010105>
- Lin, J. J. W., & Parker, D. E. (2007). User experience modeling and enhancement for virtual environments that employ wide-field displays. In V. G. Duffy (Ed.), *Digital human modeling* (pp. 423–433). Springer. https://doi.org/10.1007/978-3-540-73321-8_49
- Lindner, M., Leutritz, T., Backhaus, J., König, S., & Mühlhling, T. (2025). Knowledge gain and the impact of stress in a fully immersive virtual reality–based medical emergencies training with automated feedback: Randomized controlled trial. *Journal of Medical Internet Research*, 27(1). <https://doi.org/10.2196/67412>

- Lombard, M., Ditton, T. B., & Weinstein, L. (2009). *Measuring presence: The Temple Presence Inventory*. Proceedings of the 12th Annual International Workshop on Presence (pp. 1–15). International Society for Presence Research.
http://matthewlombard.com/ISPR/Proceedings/2009/Lombard_et_al.pdf
- Lombard, M., Weinstein, L., & Ditton, T. B. (2011). *Measuring telepresence: The validity of the Temple Presence Inventory (TPI) in a gaming context*. Proceedings of the International Society for Presence Research Annual Conference (ISPR 2011). International Society for Presence Research.
- Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, 31(4), 557–564. <https://doi.org/10.3758/BF03200735>
- Lopez, C., Halje, P., & Blanke, O. (2008). Body ownership and embodiment: Vestibular and multisensory mechanisms. *Neurophysiologie Clinique/Clinical Neurophysiology*, 38(3), 149–161. <https://doi.org/10.1016/j.neucli.2007.12.006>
- Mac Cárthaigh, S., & López, B. (2020). Factually based autism awareness campaigns may not always be effective in changing attitudes towards autism: Evidence from British and South Korean nursing students. *Autism*, 24(5), 1177–1190.
<https://doi.org/10.1177/1362361319898362>
- Madden, T. J., Hewett, K., & Roth, M. S. (2000). Managing images in different cultures: A cross-national study of color meanings and preferences. *Journal of International Marketing*, 8(4), 90–107. <https://doi.org/10.1509/jimk.8.4.90.19795>
- Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2015). Changing bodies changes minds: Owning another body affects social cognition. *Trends in Cognitive Sciences*, 19(1), 6–12. <https://doi.org/10.1016/j.tics.2014.11.001>
- Makowski, D., Sperduti, M., Nicolas, S., & Piolino, P. (2017). "Being there" and remembering it: Presence improves memory encoding. *Consciousness and Cognition*, 53, 194–202.
<https://doi.org/10.1016/j.concog.2017.06.015>
- Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*, 113(4), 719–735. <https://doi.org/10.1037/edu0000473>

- Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, 66(5), 1141–1164. <https://doi.org/10.1007/s11423-018-9581-2>
- Makransky, G., Mayer, R. E., Nørremølle, A., Lopez Cordoba, A., Wandall, J., & Bonde, M. T. (2020). Investigating the feasibility of using assessment and explanatory feedback in desktop virtual reality simulations. *Educational Technology Research and Development*, 68(1), 293–317. <https://doi.org/10.1007/s11423-019-09690-3>
- Makransky, G., & Petersen, G. B. (2019). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. *Computers & Education*, 134, 15–30. <https://doi.org/10.1016/j.compedu.2019.02.002>
- Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 33(3), 937–958. <https://doi.org/10.1007/s10648-020-09586-2>
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225–236. <https://doi.org/10.1016/j.learninstruc.2017.12.007>
- Markram, K., & Markram, H. (2010). The Intense World Theory – A unifying theory of the neurobiology of autism. *Frontiers in Human Neuroscience*, 4. <https://doi.org/10.3389/fnhum.2010.00224>
- Martingano, A. J., Hererra, F., & Konrath, S. (2021). *Virtual reality improves emotional but not cognitive empathy: A meta-analysis*. <https://hdl.handle.net/1805/46938>
- Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369>
- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.015>

- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 316–344). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.016>
- Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, 70, 29–40. <https://doi.org/10.1016/j.compedu.2013.07.033>
- Messinger, P., Stroulia, E., & Lyons, K. (2008). A typology of virtual worlds: Historical overview and future directions. *Journal of Virtual Worlds Research*, 1(1). <https://www.learntechlib.org/p/178155/>
- Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, 140. <https://doi.org/10.1016/j.compedu.2019.103603>
- Mildner, P., Stamer, N., & Effelsberg, W. (2015). From game characteristics to effective learning games. In S. Göbel, M. Ma, J. Baalsrud Hauge, M. F. Oliveira, J. Wiemeyer, & V. Wendel (Eds.), *Serious games* (pp. 51–62). Springer International Publishing. https://doi.org/10.1007/978-3-319-19126-3_5
- Miltcheva, N. (2021). *Projet EDI : Empathie narrative et avatar tangible pour dispositif de réalité virtuelle* [Mémoire de maîtrise, Université du Québec à Chicoutimi]. <https://constellation.uqac.ca/id/eprint/8111/>
- Mishina, Y., Block, E. S., & Mannor, M. J. (2012). The path dependence of organizational reputation: How social judgment influences assessments of capability and character. *Strategic Management Journal*, 33(5), 459–477. <https://doi.org/10.1002/smj.958>
- Moesgaard, T. G., Witt, M., Fiss, J., Warming, C., Klubien, J., & Schönau-Fog, H. (2015). Implicit and explicit information mediation in a virtual reality museum installation and its effects on retention and learning outcomes. Dans R. Munkvold & L. Kolås (Dir.), *Proceedings of the 9th European Conference on Games-Based Learning (ECGBL 2015)* (p. 387–394). Academic Conferences International.

- Moreno, R. (2006). Learning in high-tech and multimedia environments. *Current Directions in Psychological Science*, 15(2), 63–67. <https://doi.org/10.1111/j.0963-7214.2006.00408.x>
- Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, 94(3), 598–610. <https://doi.org/10.1037/0022-0663.94.3.598>
- Mori, M. (2012). La vallée de l'étrange (I. Yaya, Trad.). *Gradhiva. Revue d'Anthropologie et d'Histoire des Arts*, 15, 26–33. <https://doi.org/10.4000/gradhiva.2311>
- Mottron, L., Belleville, S., & Ménard, E. (1999). Local bias in autistic subjects as evidenced by graphic tasks: Perceptual hierarchization or working memory deficit? *Journal of Child Psychology and Psychiatry*, 40(5), 743–755. <https://doi.org/10.1111/1469-7610.00490>
- Mottron, L., & Burack, J. A. (2001). Enhanced perceptual functioning in the development of autism. Dans J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Dir.), *The development of autism: Perspectives from theory and research* (p. 131–148). Routledge. <https://doi.org/10.4324/9781410600196>
- Mottron, L., & Bzdok, D. (2020). Autism spectrum heterogeneity: Fact or artifact? *Molecular Psychiatry*, 25(12), 3178–3185. <https://doi.org/10.1038/s41380-020-0748-y>
- Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36(1), 27–43. <https://doi.org/10.1007/s10803-005-0040-7>
- Muletier, C., Bertholet, G., & Lang, T. (2014). *La gamification : Ou l'art d'utiliser les mécaniques du jeu dans votre business*. Éditions Eyrolles.
- Müller, R. A., Shih, P., Keehn, B., Deyoe, J. R., Leyden, K. M., & Shukla, D. K. (2011). Underconnected, but how? A survey of functional connectivity MRI studies in autism spectrum disorders. *Cerebral Cortex*, 21(10), 2233–2243. <https://doi.org/10.1093/cercor/bhq296>

- Newport, R., Pearce, R., & Preston, C. (2010). Fake hands in action: Embodiment and control of supernumerary limbs. *Experimental Brain Research*, 204(3), 385–395. <https://doi.org/10.1007/s00221-009-2104-y>
- Nielsen, J. (1994). *Usability engineering* (lastest ed.). Morgan Kaufmann.
- Organisation mondiale de la santé. (1993). *Classification statistique internationale des maladies et des problèmes de santé connexes, dixième révision (CIM-10)*. <https://www.who.int/standards/classifications/classification-of-diseases>
- Organisation mondiale de la santé. (2022). *Classification internationale des maladies, onzième révision (CIM-11)*. <https://icd.who.int/browse11/l-m/fr>
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry*, 32(7), 1081–1105. <https://doi.org/10.1111/j.1469-7610.1991.tb00351.x>
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2016). Cognitive load measurement as a means to advance cognitive load theory. Dans F. Paas, A. Renkl, & J. Sweller (Dir.), *Cognitive load theory: A special issue of Educational Psychologist* (p. 63–71). Routledge. <https://doi.org/10.4324/9780203764770-8>
- Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review*, 22(2), 115–121. <https://doi.org/10.1007/s10648-010-9133-8>
- Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science*, 29(4), 394–398. <https://doi.org/10.1177/0963721420922183>
- Park, H. O. (2023). Autism spectrum disorder and savant syndrome: A systematic literature review. *Journal of the Korean Academy of Child and Adolescent Psychiatry*, 34(2), 76–92. <https://doi.org/10.5765/jkacap.230003>
- Parong, J., & Mayer, R. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6), 785–797. <https://doi.org/10.1037/edu0000241>

- Peck, T. C., & Gonzalez-Franco, M. (2021). Avatar embodiment: A standardized questionnaire. *Frontiers in Virtual Reality, 1*.
<https://doi.org/10.3389/frvir.2020.575943>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences, 16*(10), 504–510.
<https://doi.org/10.1016/j.tics.2012.08.009>
- Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology, 38*(6), 922–934. <https://doi.org/10.1002/ejsp.504>
- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- Petty, R. E., & Cacioppo, J. T. (1986). Methodological factors in the ELM. In R. E. Petty & J. T. Cacioppo (Eds.), *Communication and persuasion: Central and peripheral routes to attitude change* (pp. 25–59). Springer. https://doi.org/10.1007/978-1-4612-4964-1_2
- Poirier, N., & Des Rivières-Pigeon, C. (2013). *Le trouble du spectre de l'autisme : État des connaissances*. Presses de l'Université du Québec.
<https://doi.org/10.1515/9782760538580>
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. Dans H. Bouma & D. G. Bouwhuis (Dir.), *Attention and performance X: Control of language processes* (p. 531–556). Lawrence Erlbaum Associates.
- Poupard, M., Larrue, F., Sauzéon, H., & Tricot, A. (2024). A systematic review of immersive technologies for education: Learning performance, cognitive load, and intrinsic motivation. *British Journal of Educational Technology, 55*(6), 2415–2443.
<https://doi.org/10.1111/bjet.13503>
- Qian, N., & Lipkin, R. M. (2011). A learning-style theory for understanding autistic behaviors. *Frontiers in Human Neuroscience, 5*. <https://doi.org/10.3389/fnhum.2011.00077>

- Radiani, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*.
<https://doi.org/10.1016/j.compedu.2019.103778>
- Ramaseri Chandra, A. N., El Jamiy, F., & Reza, H. (2022). A systematic survey on cybersickness in virtual environments. *Computers, 11*(4).
<https://doi.org/10.3390/computers11040051>
- Ramirez, E. J., & LaBarge, S. (2018). Real moral problems in the use of virtual reality. *Ethics and Information Technology, 20*(4), 249–263. <https://doi.org/10.1007/s10676-018-9473-5>
- Reason, J. T., & Brand, J. J. (1975). *Motion sickness*. Academic Press.
- Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays. *Virtual Reality, 20*(2), 101–125. <https://doi.org/10.1007/s10055-016-0285-9>
- Rettinger, M., Schmaderer, C., & Rigoll, G. (2022). Do you notice me? How bystanders affect the cognitive load in virtual reality. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 77–82. <https://doi.org/10.1109/VR51125.2022.00025>
- Riva, G., Davide, F., & IJsselsteijn, W. A. (Eds.). (2003). *Being there: Concepts, effects and measurements of user presence in synthetic environments*. IOS Press.
- Riva, G., Mantovani, F., Capideville, C., Preziosa, A., Morganti, F., Villani, D., Gaggioli, A., Botella, C., & Alcañiz Raya, M. (2007). Affective interactions using virtual reality: The link between presence and emotions. *Cyberpsychology & Behavior, 10*(1), 45–56.
<https://doi.org/10.1089/cpb.2006.9993>
- Riva, G., Waterworth, J. A., Waterworth, E. L., & Mantovani, F. (2011). From intention to action: The role of presence. *New Ideas in Psychology, 29*(1), 24–37.
<https://doi.org/10.1016/j.newideapsych.2009.11.002>
- Robertson, C., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience, 18*(11), 671–684. <https://doi.org/10.1038/nrn.2017.112>
- Robertson, L. C., & Lamb, M. R. (1991). Neuropsychological contributions to theories of part/whole organization. *Cognitive Psychology, 23*(2), 299–330.
[https://doi.org/10.1016/0010-0285\(91\)90012-D](https://doi.org/10.1016/0010-0285(91)90012-D)

- Rogers, S. J., Hepburn, S., & Wehner, E. (2003). Parent reports of sensory symptoms in toddlers with autism and those with other developmental disorders. *Journal of Autism and Developmental Disorders*, 33(6), 631–642.
<https://doi.org/10.1023/B:JADD.0000006000.38991.a7>
- Rothbaum, B. O., & Hodges, L. F. (1999). The use of virtual reality exposure in the treatment of anxiety disorders. *Behavior Modification*, 23(4), 507–525.
<https://doi.org/10.1177/0145445599234001>
- Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, 2(5), 255–267.
<https://doi.org/10.1034/j.1601-183X.2003.00037.x>
- Runswick-Cole, K. (2016). Understanding this thing called autism. In K. Runswick-Cole, R. Mallett, & S. Timimi (Eds.), *Re-thinking autism: Diagnosis, identity and equality* (pp. 19–29). Jessica Kingsley Publishers.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67.
<https://doi.org/10.1006/ceps.1999.1020>
- Schubert, T. W. (2003). The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Zeitschrift für Medienpsychologie*, 15(2), 69–71. <https://doi.org/10.1026//1617-6383.15.2.69>
- Schraw, G., Flowerday, T., & Lehman, S. (2001). Increasing situational interest in the classroom. *Educational Psychology Review*, 13(3), 211–224.
<https://doi.org/10.1023/A:1016619705184>
- Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory in education. Dans K. R. Wentzel & D. B. Miele (Dir.), *Handbook of motivation at school* (2e éd., p. 34–54). Routledge. <https://doi.org/10.4324/9781315773384>
- Schutte, N. S., & Stilinović, E. J. (2017). Facilitating empathy through virtual reality. *Motivation and Emotion*, 41(6), 708–712. <https://doi.org/10.1007/s11031-017-9641-7>
- Schwind, V., Wolf, K., & Henze, N. (2018). Avoiding the uncanny valley in virtual character design. *Interactions*, 25(5), 45–49. <https://doi.org/10.1145/3236673>

- Seymour, M., Riemer, K., & Kay, J. (2017). Interactive realistic digital avatars—Revisiting the uncanny valley. *Proceedings of the 50th Hawaii International Conference on System Sciences*. <https://aisel.aisnet.org/hicss-50/cl/hci/4>
- Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments*, 1(1), 120–125. <https://doi.org/10.1162/pres.1992.1.1.120>
- Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. Yale University Press.
- Shin, D. (2017). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, 78, 64–73. <https://doi.org/10.1016/j.chb.2017.09.012>
- Shin, D., Biocca, F., & Choo, H. (2013). Exploring the user experience of three-dimensional virtual learning environments. *Behaviour & Information Technology*, 32(2), 203–216. <https://doi.org/10.1080/0144929X.2011.606334>
- Sinatra, G. M., & Pintrich, P. R. (Eds.). (2003). *Intentional conceptual change*. (1st ed.). Routledge. <https://doi.org/10.4324/9781410606716>
- Slater, M. (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 8(5), 560–565. <https://doi.org/10.1162/105474699566477>
- Slater, M. (2009). 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. <https://doi.org/10.1098/rstb.2009.0138>
- Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, 3. <https://doi.org/10.3389/frobt.2016.00074>
- Slater, M., & Wilbur, S. (1997). 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. <https://doi.org/10.1162/pres.1997.6.6.603>
- Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 3(2), 130–144. <https://doi.org/10.1162/pres.1994.3.2.130>

- Stanney, K., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., Weech, S., & Fulvio, J. M. (2020). Identifying causes of and solutions for cybersickness in immersive technology: Reformulation of a research and development agenda. *International Journal of Human-Computer Interaction*, 36(19), 1783–1803. <https://doi.org/10.1080/10447318.2020.1828535>
- Steed, A., Pan, Y., Zisch, F., & Steptoe, W. (2016). The impact of a self-avatar on cognitive load in immersive virtual reality. *2016 IEEE Virtual Reality (VR)*, 67–68. <https://doi.org/10.1109/VR.2016.7504689>
- Stevenson, J. L., & Nonack, M. B. (2018). Gender differences in mental rotation strategy depend on degree of autistic traits. *Autism Research*, 11(7), 1024–1037. <https://doi.org/10.1002/aur.1958>
- Steptoe, W., Steed, A., & Slater, M. (2013). Human tails: Ownership and control of extended humanoid avatars. *IEEE Transactions on Visualization and Computer Graphics*, 19(4), 583–590. <https://doi.org/10.1109/TVCG.2013.32>
- Stone, R. (2001). Virtual reality for interactive training: An industrial practitioner's viewpoint. *International Journal of Human-Computer Studies*, 55(4), 699–711. <https://doi.org/10.1006/ijhc.2001.0497>
- Sutherland, I. E. (1968). *A head-mounted three dimensional display*. Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I (AFIPS '68 (Fall, part I)) (p. 757–764). Association for Computing Machinery. <https://doi.org/10.1145/1476589.1476686>
- Sweller, J. (2023). The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. *Educational Psychology Review*, 35(4). <https://doi.org/10.1007/s10648-023-09817-2>
- Tardif, C., & Gepner, B. (2010). *L'autisme* (3e éd.). Armand Colin.
- Tardif, J. (1992). *Pour un enseignement stratégique : L'apport de la psychologie cognitive*. Éditions Logiques.
- Taupiac, J.-D. (2020). *Étude et conception d'environnements de formation pertinents en réalité virtuelle* [Thèse de doctorat, Université de Montpellier]. <https://www.theses.fr/2020MONT040>

- Tcha-Tokey, K., Loup-Escande, E., Christmann, O., Canac, G., Farin, F., & Richir, S. (2015). Vers un modèle de l'expérience utilisateur en environnement virtuel immersif : Une analyse de la littérature. *IHM'15*, 27–30. <https://doi.org/10.1145/2820619.2825006>
- Terrade, F., Pasquier, H., Reerinck-Boulanger, J., Guingouain, G., & Somat, A. (2009). L'acceptabilité sociale : La prise en compte des déterminants sociaux dans l'analyse de l'acceptabilité des systèmes technologiques. *Le Travail Humain*, 72(4), 383–395. <https://doi.org/10.3917/th.724.0383>
- Tordo, F., & Binkley, C. (2016). L'auto-empathie, ou le devenir de l'autrui-en-soi : Définition et clinique du virtuel. *L'Évolution Psychiatrique*, 81(2), 293–308. <https://doi.org/10.1016/j.evopsy.2014.02.002>
- Van Etten, H. M., Kaur, M., Srinivasan, S. M., Cohen, S. J., Bhat, A., & Dobkins, K. R. (2017). Increased prevalence of unusual sensory behaviors in infants at risk for, and teens with, autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 47(11), 3431–3445. <https://doi.org/10.1007/s10803-017-3227-9>
- Ventura, S., Badenes-Ribera, L., Herrero, R., Cebolla, A., Galiana, L., & Baños, R. (2020). Virtual reality as a medium to elicit empathy: A meta-analysis. *Cyberpsychology, Behavior and Social Networking*, 23(10), 667–676. <https://doi.org/10.1089/cyber.2019.0681>
- Viaud-Delmon, I. (2007). Corps, action et cognition : La réalité virtuelle au défi des sciences cognitives. *Intellectica*, 45(1), 37–58. <https://doi.org/10.3406/intel.2007.1266>
- Viaud-Delmon, I., Ivanenko, Y. P., Grasso, R., & Israël, I. (1999). Non-specific directional adaptation to asymmetrical visual-vestibular stimulation. *Cognitive Brain Research*, 7(4), 507–510. [https://doi.org/10.1016/S0926-6410\(98\)00052-4](https://doi.org/10.1016/S0926-6410(98)00052-4)
- Voinescu, A., Matu, S., Sava, F., & David, D. (2016). Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: A meta-analytic approach. *Computers in Human Behavior*, 54, 414–424. <https://doi.org/10.1016/j.chb.2015.08.029>
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4(1), 45–69. [https://doi.org/10.1016/0959-4752\(94\)90018-3](https://doi.org/10.1016/0959-4752(94)90018-3)

- Weiner, B. (1986). *An attributional theory of motivation and emotion*. Springer-Verlag.
<https://doi.org/10.1007/978-1-4612-4948-1>
- Wenk, N., Penalver-Andres, J., Buetler, K. A., Nef, T., Müri, R. M., & Marchal-Crespo, L. (2023). Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment. *Virtual Reality*, 27(1), 307–331.
<https://doi.org/10.1007/s10055-021-00565-8>
- Wiemann, J. M. (1977). Explication and test of a model of communicative competence. *Human Communication Research*, 3(3), 195–213. <https://doi.org/10.1111/j.1468-2958.1977.tb00518.x>
- Willay, E. (2020). *Concevoir des formations utilisant la vidéo pour faire évoluer les croyances des professionnels sur les capacités des personnes autistes* [Thèse de doctorat, Université Toulouse 2 Jean Jaurès]. <https://dante.univ-tlse2.fr/s/fr/item/11166>
- Williams, J. M., & Tolmie, A. (2000). Conceptual change in biology: Group interaction and the understanding of inheritance. *British Journal of Developmental Psychology*, 18(4), 625–649. <https://doi.org/10.1348/026151000165896>
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), 225–240. <https://doi.org/10.1162/105474698565686>
- Zell, E., Aliaga, C., Jarabo, A., Zibrek, K., Gutierrez, D., McDonnell, R., & Botsch, M. (2015). To stylize or not to stylize? The effect of shape and material stylization on the perception of computer-generated faces. *ACM Transactions on Graphics*, 34(6).
<https://doi.org/10.1145/2816795.2818126>

Annexes

ANNEXE A - CADRE THÉORIQUE : TROUBLE DU SPECTRE DE L'AUTISME

Cette section présente plus en détail les fondements théoriques et cliniques du trouble du spectre de l'autisme (TSA). Elle regroupe les critères diagnostiques internationaux (DSM-5 et CIM-11), les principaux modèles explicatifs issus de la recherche cognitive et perceptive, ainsi qu'une présentation plus détaillée du modèle *Enhanced Perceptual Functioning* (EPF) de Mottron et al. (2006), qui constitue l'un des cadres théoriques centraux de cette thèse.

A.1 Définitions et classifications diagnostiques

A.1.1 Critères diagnostiques DSM-5

Selon le DSM-5 (American Psychiatric Association, 2013), le TSA (code 299.00 / F84.0) est défini par les critères diagnostiques suivants.

Critère A : Déficiences persistantes de la communication et des interactions sociales

Les déficiences persistantes de la communication et des interactions sociales sont observées dans des contextes variés. Ces déficiences se manifestent actuellement ou dans les antécédents par l'ensemble des éléments suivants :

1. Déficiences de la réciprocité socio-émotionnelle, allant par exemple d'une approche sociale anormale et d'un échec de la conversation bidirectionnelle normale, à une réduction du partage des intérêts, des émotions ou de l'affect, jusqu'à l'échec d'initiation ou de réponse aux interactions sociales.

2. Déficiences des comportements de communication non verbaux utilisés pour l'interaction sociale, allant par exemple d'une communication verbale et non verbale mal intégrée, à des anomalies du contact visuel et du langage corporel, ou à des déficiences dans la compréhension et l'utilisation des gestes, jusqu'à une absence totale d'expressions faciales et de communication non verbale.

3. Déficiences du développement, du maintien et de la compréhension des relations, allant par exemple de difficultés à ajuster le comportement à divers contextes sociaux, à des difficultés à partager des jeux imaginatifs ou à se faire des amis, jusqu'à l'absence d'intérêt pour les pairs.

Critère B : Caractère restreint et répétitif des comportements, des intérêts ou des activités

Le caractère restreint et répétitif des comportements, des intérêts ou des activités se manifeste par au moins deux des éléments suivants, actuels ou dans les antécédents :

1. Mouvements moteurs, utilisation des objets, ou parole stéréotypés ou répétitifs (par exemple, stéréotypies motrices simples, alignement des jouets ou rotation des objets, écholalie, phrases idiosyncrasiques).

2. Insistance sur l'adhésion inflexible à des routines ou à des modes comportementaux verbaux ou non verbaux ritualisés (par exemple, détresse extrême provoquée par des changements mineurs, difficultés avec les transitions, modes de pensée rigides, ritualisation des formules de salutation, nécessité de prendre le même chemin ou de manger les mêmes aliments chaque jour).

3. Intérêts extrêmement restreints et circonscrits dont l'intensité ou le sujet sont anormaux (par exemple, attachement intense ou préoccupation concernant des objets inhabituels, intérêts excessivement circonscrits ou persévérants).

4. Hyper ou hyporéactivité aux stimulations sensorielles ou intérêt inhabituel pour les aspects sensoriels de l'environnement (par exemple, indifférence apparente à la douleur/à la température, réaction négative à des sons ou à des textures spécifiques, action de sentir ou de toucher excessivement des objets, fascination visuelle pour des lumières ou des mouvements).

Critères complémentaires

Critère C : Les symptômes doivent être présents dès les étapes précoces du développement (mais ils peuvent ne devenir pleinement manifestes que lorsque les demandes sociales dépassent les capacités limitées, ou peuvent être masqués par des stratégies apprises plus tard dans la vie).

Critère D : Les symptômes occasionnent un retentissement cliniquement significatif en termes de fonctionnement actuel social, scolaire/professionnel ou dans d'autres domaines importants.

Critère E : Ces perturbations ne sont pas mieux expliquées par une déficience intellectuelle (trouble du développement intellectuel) ou un retard global de développement. La déficience intellectuelle et le TSA sont fréquemment associés ; pour poser un diagnostic de

comorbidité de TSA et de déficience intellectuelle, le niveau de communication sociale doit être inférieur à celui attendu pour le niveau de développement général.

Spécification de la sévérité actuelle

Le niveau de sévérité se fonde sur l'importance des déficits de la communication sociale et des modes comportementaux restreints et répétitifs :

Niveau 3 : Nécessitant un soutien très important. Les déficits sévères des compétences de communication sociale, verbale et non verbale, entraînent des altérations sévères du fonctionnement, une initiation très limitée des interactions sociales et une réponse minimale aux ouvertures sociales d'autrui. Par exemple, une personne ayant peu de mots intelligibles qui initie rarement les interactions et qui, lorsqu'elle le fait, a des approches inhabituelles visant uniquement à répondre à des besoins et réagit seulement à des approches sociales très directes.

Niveau 2 : Nécessitant un soutien important. Déficiences marquées des compétences de communication sociale, verbale et non verbale ; altérations sociales apparentes même en présence d'un soutien ; initiation limitée des interactions sociales ; et réponse réduite ou anormale aux ouvertures sociales d'autrui. Par exemple, une personne qui ne prononce que des phrases simples, dont les interactions sont limitées à des intérêts restreints particuliers, et qui a une communication non verbale très étrange.

Niveau 1 : Nécessitant un soutien. En l'absence de soutien, des déficits de la communication sociale entraînent des altérations notables. Difficulté d'initiation des interactions sociales et exemples clairs de réponses atypiques ou inefficaces aux ouvertures sociales d'autrui. La personne peut sembler présenter un intérêt réduit pour les interactions sociales. Par exemple, une personne qui est capable de parler en faisant des phrases complètes et qui engage la communication mais dont la conversation bidirectionnelle avec d'autres personnes échoue, et dont les tentatives pour se faire des amis sont étranges et généralement sans succès.

A.1.2 Critères diagnostiques CIM-11

La Classification Internationale des Maladies, 11e révision (CIM-11 ; Organisation Mondiale de la Santé, 2018), catégorise le TSA sous le code 6A02. Cette classification propose une organisation dimensionnelle qui prend en compte à la fois le niveau de déficience intellectuelle et le degré d'altération du langage fonctionnel.

Sous-catégories du code 6A02

6A02.0 : TSA sans déficience intellectuelle et avec altération légère ou absente du langage fonctionnel. Cette catégorie correspond aux personnes avec TSA présentant des capacités intellectuelles dans la norme ou supérieures à la norme, et dont le langage expressif et réceptif est fonctionnel, bien que pouvant présenter des particularités pragmatiques.

6A02.1 : TSA avec déficience intellectuelle et avec altération légère ou absente du langage fonctionnel. Cette catégorie inclut les personnes avec TSA présentant une déficience intellectuelle concomitante ($QI < 70$), mais dont les capacités langagières de base restent préservées.

6A02.2 : TSA sans déficience intellectuelle et avec altération du langage fonctionnel. Cette catégorie correspond aux personnes avec TSA dont les capacités intellectuelles sont préservées mais qui présentent des difficultés significatives dans l'utilisation fonctionnelle du langage pour la communication sociale.

6A02.3 : TSA avec déficience intellectuelle et avec altération du langage fonctionnel. Cette catégorie regroupe les personnes présentant à la fois une déficience intellectuelle et des altérations significatives du langage fonctionnel.

6A02.4 : TSA sans déficience intellectuelle et avec absence de langage fonctionnel. Cette catégorie inclut les personnes avec TSA non verbales ou minimalement verbales, malgré des capacités intellectuelles préservées.

6A02.5 : TSA avec déficience intellectuelle et avec absence de langage fonctionnel. Cette catégorie correspond aux personnes présentant à la fois une déficience intellectuelle et une absence de langage fonctionnel.

6A02.Y : Autre TSA spécifié. Cette catégorie est utilisée pour les présentations atypiques ou les cas nécessitant une spécification particulière.

6A02.Z : TSA, sans précision. Cette catégorie est utilisée lorsque l'information disponible est insuffisante pour permettre une classification plus spécifique.

Caractéristiques selon la HAS (2023)

Selon la Haute Autorité de Santé (HAS, 2023), le TSA se caractérise par un ensemble de particularités neurodéveloppementales qui affectent quatre domaines principaux du fonctionnement :

1. Particularités dans les interactions sociales : Difficultés dans la compréhension et l'utilisation des codes sociaux implicites, dans l'ajustement du comportement au contexte social, et dans l'établissement et le maintien de relations réciproques.

2. Modes de communication verbale et non-verbale particuliers : Particularités dans l'utilisation pragmatique du langage (tour de parole, adaptation au contexte, compréhension des sous-entendus), dans l'interprétation et l'utilisation de la communication non verbale (expressions faciales, gestes, prosodie), et dans la compréhension des intentions communicatives d'autrui.

3. Centres d'intérêt restreints et comportements répétitifs : Intérêts circonscrits d'intensité ou de focalisation inhabituelle, adhésion rigide à des routines spécifiques, comportements moteurs répétitifs (stéréotypés), et besoin de maintien de l'identité de l'environnement (résistance au changement).

4. Particularités sensorielles : Hyper ou hyposensibilité à certaines stimulations sensorielles (visuelles, auditives, tactiles, olfactives, gustatives, proprioceptives, vestibulaires), conduisant à des réponses sensorielles atypiques et à des difficultés de modulation sensorielle.

La HAS souligne que ces caractéristiques s'inscrivent dans un continuum de sévérité et que leur expression varie considérablement d'une personne à l'autre, justifiant ainsi la conceptualisation en termes de « spectre ».

A.2 Modèles théoriques explicatifs du TSA

La recherche scientifique sur le TSA a développé de nombreux modèles théoriques visant à expliquer les particularités cognitives, perceptives et comportementales observées chez les personnes avec TSA. Ces modèles peuvent être organisés en quatre grandes catégories en fonction de leur niveau d'analyse et de leur focus explicatif : les théories perceptives, les théories cognitives, les théories généralistes et computationnelles, et les théories bayésiennes. Chacune de ces approches offre un éclairage complémentaire sur le fonctionnement du TSA.

A.2.1 Théories perceptives

Les théories perceptives placent les particularités du traitement sensoriel et perceptif au cœur de l'explication du fonctionnement du TSA. Ces théories proposent que de nombreuses caractéristiques du TSA, y compris certaines difficultés sociales et communicationnelles,

découlent de différences fondamentales dans la manière dont l'information sensorielle est traitée.

Weak Central Coherence (cohérence centrale faible)

La théorie de la cohérence centrale faible, proposée par Frith (1989), postule que les personnes avec TSA présentent un style cognitif caractérisé par un traitement préférentiel de l'information locale plutôt que globale. Selon cette théorie, alors que les personnes neurotypiques ont tendance à traiter spontanément l'information de manière holistique, en privilégiant le sens général et le contexte, les personnes avec TSA privilégient le traitement des détails au détriment de l'intégration globale.

Cette particularité se manifeste dans diverses tâches expérimentales. Par exemple, dans le test des figures enchâssées (*Embedded Figures Test*), les personnes avec TSA excellent à identifier rapidement des formes simples dissimulées dans des configurations complexes, précisément parce qu'elles sont moins influencées par le contexte global. De même, dans les tâches de mémoire, elles tendent à mieux se souvenir des détails spécifiques que du sens général d'une histoire.

Frith suggère que cette particularité perceptive pourrait expliquer certaines difficultés sociales : comprendre une situation sociale requiert souvent d'intégrer de multiples indices contextuels et de saisir le sens général plutôt que de se focaliser sur des détails spécifiques. Toutefois, cette théorie a été critiquée pour ne pas expliquer tous les aspects du TSA et pour conceptualiser le traitement local comme un déficit plutôt que comme une particularité pouvant être avantageuse dans certains contextes.

Enhanced Perceptual Functioning (fonctionnement perceptif amélioré)

Le modèle du fonctionnement perceptif amélioré (*Enhanced Perceptual Functioning*), développé par Mottron et Burack (2001), propose une reconceptualisation du traitement perceptif dans le TSA. Contrairement à la théorie de la cohérence centrale faible qui présente le traitement local comme une conséquence d'un déficit d'intégration globale, le modèle EPF postule que les personnes avec TSA présentent des capacités perceptives supérieures qui résultent d'un investissement neural accru dans les régions perceptives.

Selon ce modèle, le traitement perceptif dans le TSA se caractérise par une autonomie accrue vis-à-vis des influences descendantes (*top-down*). Autrement dit, la perception est moins biaisée par les connaissances préalables, les attentes ou le contexte, ce qui permet un traitement

plus fidèle et précis de l'information sensorielle brute. Cette particularité conduit à des performances supérieures dans diverses tâches perceptives : discrimination de hauteur de son (*pitch*), détection de patterns visuels complexes, discrimination tactile fine.

Le modèle EPF suggère également que les régions cérébrales typiquement dédiées au traitement perceptif sont recrutées de manière plus extensive dans le TSA, y compris pour des tâches cognitives de haut niveau. Cette allocation de ressources explique à la fois les capacités perceptives exceptionnelles et potentiellement certaines particularités dans les domaines cognitifs qui, dans le développement neurotypique, s'appuient moins sur les processus perceptifs.

Veridical Mapping (cartographie véridique)

L'approche de la cartographie véridique, proposée par Bouvet et al. (2014), suggère que les personnes avec TSA établissent une correspondance plus directe et moins médiatisée entre les stimuli sensoriels et leurs représentations mentales. Dans le traitement neurotypique, les percepts sont automatiquement et inconsciemment filtrés, catégorisés et interprétés selon des schémas conceptuels préexistants, ce qui peut introduire des biais et des distorsions mais facilite le traitement rapide et l'intégration contextuelle.

Dans le TSA, selon cette théorie, le *mapping* entre sensation et représentation serait plus véridique (plus fidèle à la réalité sensorielle brute), avec moins de filtrage conceptuel automatique. Cette particularité expliquerait pourquoi les personnes avec TSA sont souvent sensibles à des détails sensoriels que les personnes neurotypiques ne remarquent pas, et pourquoi elles peuvent être perturbées par des stimuli que d'autres filtrent automatiquement comme non pertinents.

Cette approche rejoint le modèle EPF dans son emphase sur la fidélité perceptive, mais met davantage l'accent sur les mécanismes de correspondance entre niveaux de représentation (sensoriel, perceptif, conceptuel) plutôt que sur la performance perceptive en tant que telle.

A.2.2 Théories cognitives

Les théories cognitives expliquent le TSA par des particularités dans les fonctions cognitives de haut niveau telles que la planification, la flexibilité mentale, l'inférence des états mentaux d'autrui, ou l'imitation. Ces théories se focalisent sur les mécanismes computationnels et les représentations mentales plutôt que sur le traitement sensoriel.

Théorie du dysfonctionnement exécutif

La théorie du dysfonctionnement exécutif, proposée par Ozonoff et al. (1991), suggère que de nombreuses caractéristiques du TSA peuvent être expliquées par des difficultés dans les fonctions exécutives. Les fonctions exécutives désignent un ensemble de processus cognitifs de haut niveau qui permettent le contrôle et la régulation du comportement orienté vers un but : planification, flexibilité cognitive, inhibition des réponses automatiques, mise à jour de la mémoire de travail, et initiation de nouvelles actions.

Les recherches ont effectivement montré que les personnes avec TSA présentent fréquemment des difficultés dans certaines tâches évaluant les fonctions exécutives, notamment les tâches de flexibilité cognitive (*Wisconsin Card Sorting Test*) et de planification (Tour de Londres). Ces difficultés se manifestent dans la vie quotidienne par une rigidité comportementale, des difficultés à s'adapter à des situations nouvelles ou changeantes, et une tendance à la persévération.

Toutefois, cette théorie présente certaines limites. D'abord, les difficultés exécutives ne sont ni universelles ni spécifiques au TSA : certaines personnes avec TSA n'en présentent pas, et des difficultés similaires s'observent dans d'autres troubles neurodéveloppementaux. Ensuite, cette théorie peine à expliquer les particularités perceptives et les intérêts circonscrits qui caractérisent le TSA. Enfin, il n'est pas toujours clair si les difficultés exécutives sont primaires (causales) ou secondaires (conséquences d'autres particularités).

Théorie du cerveau masculin extrême

Baron-Cohen (2002) a proposé la théorie du cerveau masculin extrême (*Extreme Male Brain*), qui suggère que le TSA représente une forme extrême de la cognition typiquement masculine. Cette théorie repose sur la distinction entre deux dimensions psychologiques :

L'empathie (*empathizing*) : la capacité à identifier les états mentaux et émotionnels d'autrui et à y répondre de manière appropriée. Les études psychométriques montrent qu'en moyenne, les femmes obtiennent des scores plus élevés sur cette dimension.

La systématisation (*systemizing*) : la capacité et la motivation à analyser, construire et comprendre des systèmes gouvernés par des règles (systèmes techniques, abstraits, naturels, etc.). Les études montrent qu'en moyenne, les hommes obtiennent des scores plus élevés sur cette dimension.

Selon Baron-Cohen, les personnes avec TSA se caractériseraient par une hypo-empathie (scores faibles sur les mesures d'empathie) et une hyper-systématisation (scores très élevés sur les mesures de systématisation), représentant ainsi un profil cognitif type « masculin extrême ».

Cette théorie reste controversée pour plusieurs raisons. D'abord, elle s'appuie sur des différences de groupe (hommes vs femmes) qui, bien que statistiquement significatives, présentent un chevauchement considérable, ce qui limite leur pertinence pour expliquer un phénomène catégoriel comme le TSA. Ensuite, de nombreuses personnes avec TSA, y compris des femmes avec TSA, ne correspondent pas à ce profil. Enfin, le concept d'empathie est multidimensionnel (empathie cognitive vs affective), et les difficultés dans le TSA concernent principalement la dimension cognitive (comprendre les états mentaux) plutôt que la dimension affective (résonance émotionnelle).

Théorie de la cécité mentale (théorie de l'esprit altérée)

La théorie de la cécité mentale ou *mindblindness*, proposée par Baron-Cohen et al. (1985), suggère que les personnes avec TSA présentent une altération spécifique de la théorie de l'esprit (*Theory of Mind*, ToM), c'est-à-dire de la capacité à attribuer des états mentaux (croyances, désirs, intentions, connaissances) à soi-même et à autrui, et à comprendre que ces états mentaux guident les comportements.

Les expériences pionnières de Baron-Cohen ont utilisé le paradigme de la fausse croyance (*false belief task*), notamment le test de Sally-Anne. Dans ce test, l'enfant observe une scène où Sally place un objet dans un panier puis quitte la pièce ; Anne déplace alors l'objet dans une boîte ; Sally revient. L'enfant doit prédire où Sally va chercher l'objet. Les enfants neurotypiques de plus de 4 ans répondent correctement (dans le panier, là où Sally croit que l'objet se trouve), alors que les enfants avec TSA, même plus âgés et avec un QI normal, ont tendance à répondre incorrectement (dans la boîte, là où l'objet se trouve réellement), suggérant une difficulté à se représenter la fausse croyance de Sally.

Cette théorie a été très influente et a généré une vaste littérature de recherche. Elle offre une explication élégante de nombreuses difficultés sociales dans le TSA : si l'on ne peut pas inférer de manière automatique et intuitive ce que l'autre pense, croit ou ressent, les interactions sociales deviennent extrêmement complexes et imprévisibles.

Cependant, plusieurs limites ont été identifiées. D'abord, certaines personnes avec TSA réussissent les tâches de fausse croyance, surtout avec l'âge, suggérant que la difficulté n'est peut-être pas absolue mais plutôt un retard développemental ou l'utilisation de stratégies compensatoires explicites. Ensuite, des difficultés de théorie de l'esprit s'observent dans d'autres troubles, ce qui limite la spécificité explicative. Enfin, cette théorie n'explique pas les particularités perceptives, les comportements répétitifs, et les intérêts restreints qui sont pourtant des caractéristiques cardinales du TSA.

Théorie du miroir brisé

La théorie du miroir brisé (*Broken Mirror Theory*), proposée par Dapretto et al. (2006), s'appuie sur la découverte des neurones miroirs dans le cortex prémoteur et le cortex pariétal. Ces neurones s'activent à la fois lorsqu'un individu réalise une action et lorsqu'il observe cette même action réalisée par autrui. Ce système de résonance motrice serait impliqué dans la compréhension des actions d'autrui, l'imitation, et potentiellement l'empathie.

Des études en neuroimagerie ont montré que, chez les personnes avec TSA, le système des neurones miroirs présente une activation réduite lors de l'observation et de l'imitation d'actions, ainsi que lors de la perception d'émotions sur des visages. Selon cette théorie, un dysfonctionnement du système miroir expliquerait plusieurs caractéristiques du TSA : difficultés d'imitation, difficultés à comprendre les intentions d'autrui, et altération de l'empathie.

Cette théorie reste débattue. D'une part, tous les résultats ne sont pas répliqués de manière cohérente, et certaines études ne trouvent pas de différences dans l'activation du système miroir. D'autre part, même en admettant un dysfonctionnement du système miroir, il n'est pas clair si celui-ci est primaire (causal) ou secondaire (conséquence d'autres particularités développementales, comme une attention réduite aux stimuli sociaux). Enfin, le lien entre système miroir et cognition sociale de haut niveau reste lui-même débattu dans la littérature neuroscientifique générale.

A.2.3 Théories généralistes et computationnelles

Les théories généralistes proposent des explications plus transversales du fonctionnement autistique, cherchant à identifier des mécanismes fondamentaux qui pourraient sous-tendre l'ensemble ou une large part des particularités observées.

Théorie du style d'apprentissage

La théorie du style d'apprentissage, proposée par Qian et Lipkin (2011), suggère que les personnes avec TSA présentent un style d'apprentissage particulier qui diffère qualitativement du style neurotypique. Cette théorie s'inscrit dans une perspective neurodiversitaire, considérant les particularités autistiques non comme des déficits mais comme une organisation cognitive alternative.

Selon cette approche, les personnes avec TSA privilégient l'apprentissage par règles explicites plutôt que par induction implicite. Dans l'apprentissage neurotypique, de nombreuses compétences (langage pragmatique, codes sociaux, habiletés motrices) s'acquièrent de manière largement implicite par exposition répétée et extraction automatique de régularités statistiques. Les personnes avec TSA, en revanche, bénéficieraient davantage d'un enseignement explicite des règles et des principes sous-jacents.

De plus, cette théorie suggère que les personnes avec TSA privilégient l'information concrète, spécifique et contextualisée plutôt que l'abstraction et la généralisation. Elles excellent dans l'apprentissage lorsque les exemples sont nombreux, variés et clairement reliés aux principes qu'ils illustrent, mais peuvent rencontrer des difficultés lorsqu'on attend d'elles qu'elles généralisent spontanément des apprentissages d'un contexte à un autre.

Ce style d'apprentissage expliquerait plusieurs observations empiriques : l'efficacité des interventions structurées et explicites (comme les approches comportementales structurées), les difficultés de généralisation fréquemment observées, et le fait que de nombreuses personnes avec TSA deviennent expertes dans des domaines circonscrits où elles peuvent accumuler des connaissances détaillées et explicites.

Cette approche a des implications pédagogiques importantes et souligne la nécessité d'adapter les méthodes d'enseignement aux particularités cognitives de l'apprenant plutôt que de considérer un style d'apprentissage comme supérieur ou normatif.

A.2.4 Théories bayésiennes

Les théories bayésiennes du TSA représentent une approche computationnelle récente qui conceptualise le fonctionnement cérébral en termes d'inférence probabiliste. Selon le cadre théorique du codage prédictif (*predictive coding*), le cerveau fonctionne comme une machine à prédire qui génère constamment des hypothèses sur les causes des stimuli sensoriels (priors ou prédictions a priori) et met à jour ces hypothèses en fonction des erreurs de prédiction (écart entre prédiction et input sensoriel réel).

Théorie du codage prédictif dans le TSA

Le cadre général du codage prédictif propose que la perception et la cognition émergent d'une hiérarchie de prédictions : les niveaux supérieurs du système génèrent des prédictions sur l'activité des niveaux inférieurs, et seules les erreurs de prédiction (la différence entre prédiction et réalité) sont transmises vers le haut. Ce mécanisme permet un traitement efficace de l'information en ne transmettant que ce qui est nouveau ou inattendu.

Dans ce cadre, le TSA peut être conceptualisé comme résultant d'un déséquilibre dans la pondération relative des priors (prédictions descendantes) et des données sensorielles (information ascendante). Deux hypothèses principales, apparemment opposées mais potentiellement complémentaires, ont été proposées.

Hypothèse des hypopriors

L'hypothèse des hypopriors (ou priors faibles), proposée par Pellicano et Burr (2012), suggère que les personnes avec TSA ont des prédictions a priori moins influentes, ce qui les rend plus dépendantes des données sensorielles brutes et moins biaisées par les attentes contextuelles.

Cette particularité expliquerait plusieurs observations empiriques. Sur le plan perceptif, elle rendrait compte de la moindre influence des illusions visuelles chez les personnes avec TSA : si les priors perceptifs sont faibles, les illusions (qui dépendent de prédictions implicites sur la structure du monde visuel) ont moins d'impact. De même, la supériorité dans certaines tâches de discrimination perceptive s'expliquerait par une moindre contamination du percept par les attentes.

Sur le plan social, cette théorie expliquerait pourquoi les personnes avec TSA sont moins influencées par les informations contextuelles lors de l'interprétation de situations sociales : elles se fient davantage aux indices explicites actuels qu'à leurs connaissances préalables sur ce qui est typiquement attendu dans ce type de situation.

L'hypothèse des hypopriors est cohérente avec le modèle EPF car celle postule une moindre influence des processus descendants sur la perception. Elle offre également un cadre unificateur potentiel pour expliquer à la fois les particularités perceptives et certaines difficultés sociales.

A.3 Modèle Enhanced Perceptual Functioning (EPF)

Le modèle du Fonctionnement Perceptif Amélioré (EPF), développé par Mottron et al. (2006), constitue l'un des cadres théoriques les plus influents pour comprendre le traitement de l'information dans le TSA. Contrairement aux approches déficitaires qui conceptualisent le TSA en termes de dysfonctionnements ou d'incapacités, le modèle EPF adopte une perspective neurodiversitaire en postulant que les personnes avec TSA présentent une organisation cognitive alternative caractérisée par un investissement accru dans les processus perceptifs.

Ce modèle s'appuie sur une vaste littérature expérimentale documentant des performances supérieures des personnes avec TSA dans diverses tâches perceptives (discrimination auditive, détection de patterns visuels, mémoire perceptive), ainsi que sur des données de neuroimagerie montrant une activation accrue et plus étendue des régions perceptives cérébrales, y compris lors de tâches cognitives de haut niveau.

A.3.1 *Les Huit principes du modèle EPF*

Le modèle EPF repose sur huit principes fondamentaux qui décrivent les particularités du fonctionnement perceptif et cognitif des personnes avec TSA. Ces principes sont interconnectés et forment un cadre théorique cohérent.

Principe 1 : Supériorité perceptive de bas niveau

Les personnes avec TSA présentent des performances supérieures dans les tâches de discrimination sensorielle élémentaire. Cette supériorité a été documentée dans plusieurs modalités sensorielles. Dans le domaine auditif, les personnes avec TSA montrent une meilleure discrimination de hauteur de son (pitch), une meilleure détection de changements dans des séquences auditives complexes, et des seuils de détection plus bas pour certaines fréquences. Dans le domaine visuel, elles excellent dans la détection de patterns subtils, la discrimination d'orientations, et l'identification rapide d'éléments cibles dans des configurations complexes.

Cette supériorité reflète un traitement perceptif optimisé au niveau des processus de bas niveau, c'est-à-dire au niveau du traitement sensoriel précoce avant l'intervention de processus cognitifs de haut niveau. Des études psychophysiques ont montré que ces avantages ne se limitent pas aux stimuli simples mais s'étendent à des stimuli complexes et écologiquement valides.

Principe 2 : Orientation vers les stimuli perceptifs

Les personnes avec TSA manifestent une préférence spontanée pour le traitement des aspects perceptifs des stimuli plutôt que pour leur signification sociale, symbolique ou catégorielle. Cette orientation se manifeste dans les tâches d'attention sélective : lorsqu'un stimulus possède à la fois des propriétés perceptives saillantes et une signification sociale ou symbolique, les personnes avec TSA ont tendance à se focaliser spontanément sur les propriétés perceptives.

Par exemple, lors de la présentation de visages, les personnes avec TSA peuvent porter davantage attention aux textures, contours et patterns visuels qu'à l'expression émotionnelle ou à l'identité de la personne. De même, dans les tâches de mémoire, elles ont tendance à mieux encoder les détails perceptifs spécifiques que la signification conceptuelle ou catégorielle des items.

Cette orientation n'est pas inflexible : lorsqu'on attire explicitement l'attention sur les aspects sociaux ou symboliques, les personnes avec TSA peuvent y accéder. Cependant, l'orientation spontanée, automatique, privilégie le niveau perceptif.

Principe 3 : Autonomie accrue des opérations perceptives

Le traitement perceptif opère de manière plus indépendante des influences descendantes (top-down) telles que les connaissances préalables, le contexte ou les attentes. Dans le traitement neurotypique, la perception est fortement modulée par les processus cognitifs de haut niveau : ce que nous percevons est influencé par ce que nous savons, par nos attentes, par le contexte dans lequel le stimulus apparaît.

Dans le TSA, cette influence descendante est réduite, ce qui permet un traitement plus fidèle de l'information sensorielle brute, moins biaisé par les prédictions et les schémas conceptuels préexistants. Cette autonomie explique plusieurs observations empiriques : moindre sensibilité à certaines illusions visuelles qui dépendent d'hypothèses implicites sur le monde physique, meilleure performance dans les tâches nécessitant d'ignorer le contexte, et perception moins catégorielle (moins influencée par les catégories conceptuelles).

Cette particularité est cohérente avec les théories bayésiennes postulant des priors faibles au niveau perceptif : si les prédictions descendantes sont moins influentes, le système perceptif s'appuie davantage sur les données sensorielles brutes (*likelihood*).

Principe 4 : Rôle accru des régions perceptives dans les tâches cognitives

Les études de neuroimagerie fonctionnelle montrent que les personnes avec TSA recrutent davantage les régions perceptives (cortex visuel primaire et associatif, cortex auditif) lors de tâches cognitives de haut niveau, y compris pour des tâches abstraites ou sociales qui, chez les personnes neurotypiques, activent principalement des régions associatives ou préfrontales.

Par exemple, lors de tâches de raisonnement sur des matrices (Raven's Progressive Matrices), les personnes avec TSA montrent une activation accrue du cortex visuel, suggérant qu'elles résolvent ces problèmes en s'appuyant davantage sur des stratégies visuo-perceptives. De même, lors de tâches de jugement social à partir de visages ou de scènes sociales, on observe une activation plus importante des régions visuelles et moins importante des régions impliquées dans la cognition sociale de haut niveau.

Ce recrutement différentiel suggère que les personnes avec TSA utilisent des stratégies cognitives alternatives qui capitalisent sur leurs forces perceptives. Cette observation a des implications pour la compréhension des profils neuropsychologiques : un résultat identique à une tâche peut être atteint par des voies neurales et des stratégies différentes.

Principe 5 : Traitement perceptif étendu

Le traitement perceptif approfondi s'étend à une gamme plus large de stimuli et se poursuit plus longtemps dans le temps. Les personnes avec TSA montrent une tendance à traiter de manière détaillée non seulement les stimuli focaux mais aussi les stimuli périphériques ou contextuels. Cette extension du traitement perceptif peut conduire à une richesse d'information perceptive mais aussi, dans certains contextes, à une saturation des ressources attentionnelles.

Les études de suivi oculaire montrent que les personnes avec TSA explorent les scènes visuelles de manière plus exhaustive, avec des patterns de fixation plus distribués et moins guidés par les conventions ou les zones typiquement informatives (comme les yeux dans un visage). Cette exploration étendue reflète un traitement moins sélectif, où tous les éléments de la scène reçoivent une analyse perceptive approfondie.

De même, des études électrophysiologiques montrent que les réponses perceptives cérébrales (potentiels évoqués précoces) persistent plus longtemps et sont moins rapidement modulées par les processus attentionnels, suggérant un traitement perceptif plus prolongé.

Cette extension du traitement perceptif peut être avantageuse dans des contextes où une analyse exhaustive est bénéfique (détection d'anomalies, analyse de patterns complexes), mais peut aussi conduire à une surcharge sensorielle dans des environnements riches en stimulations.

Principe 6 : Pics de compétence dans les domaines perceptifs

Les talents savants et les compétences exceptionnelles observées chez certaines personnes avec TSA émergent du surinvestissement et de l'optimisation des mécanismes perceptifs. Les domaines dans lesquels ces talents se manifestent typiquement sont ceux qui sollicitent fortement les capacités perceptives : musique (perception absolue de la hauteur, mémoire auditive exceptionnelle), arts visuels (dessin précis, perception des détails et des proportions), et compétences visuo-spatiales (architecture, design).

Le modèle EPF suggère que ces capacités exceptionnelles ne sont pas des îlots isolés de compétence mais représentent l'expression extrême du fonctionnement perceptif amélioré. Elles émergent de la combinaison d'une supériorité perceptive de bas niveau, d'une orientation préférentielle vers les aspects perceptifs, et d'un investissement intensif (pratique extensive) dans des domaines qui valorisent ces compétences perceptives.

Cette perspective suggère que le potentiel pour développer des expertises dans des domaines perceptuels est largement répandu dans le TSA, bien que sa réalisation dépende de facteurs environnementaux (opportunités, encouragements, ressources) et motivationnels (intérêts spécifiques).

Principe 7 : Correspondance accrue entre performances perceptives et cognitives

Dans la population neurotypique, les performances dans les tâches perceptives de bas niveau et les performances dans les tâches cognitives de haut niveau (raisonnement, résolution de problèmes) sont relativement indépendantes : un individu peut exceller dans l'une sans exceller dans l'autre. Dans le TSA, en revanche, ces domaines sont plus fortement corrélés : les capacités cognitives de haut niveau sont davantage prédites par les capacités perceptives de bas niveau.

Cette correspondance accrue reflète le fait que, dans le TSA, les processus cognitifs de haut niveau s'appuient davantage sur les mécanismes perceptifs. Par exemple, le QI non verbal (mesuré par des tâches comme les Matrices de Raven) corrèle plus fortement avec les capacités de discrimination perceptive chez les personnes avec TSA que chez les personnes neurotypiques.

Cette observation a des implications pour l'évaluation cognitive : les tests d'intelligence traditionnels, qui minimisent les aspects perceptifs au profit des aspects conceptuels et verbaux, peuvent sous-estimer les capacités réelles des personnes avec TSA. Des mesures tenant compte du potentiel perceptif pourraient fournir une évaluation plus fidèle.

Principe 8 : Développement neural atypique des systèmes perceptifs

Le modèle EPF s'ancre dans une architecture neuronale modifiée, caractérisée par une connectivité locale accrue dans les régions perceptives et une connectivité à longue distance réduite entre régions perceptives et associatives. Cette organisation neuronale sous-tend les particularités fonctionnelles décrites dans les principes précédents.

Les études de connectivité anatomique (imagerie en tenseur de diffusion) et fonctionnelle (corrélations entre activités de différentes régions au repos) convergent pour montrer ce pattern de connectivité modifiée. L'augmentation de la connectivité locale dans les régions perceptives permettrait un traitement plus riche et plus détaillé de l'information au sein de ces régions. La réduction de la connectivité à longue distance expliquerait l'autonomie accrue des processus perceptifs vis-à-vis des influences descendantes depuis les régions associatives.

Ce principe neurobiologique fournit une base mécanistique au modèle EPF et suggère que les particularités fonctionnelles observées émergent d'une organisation cérébrale alternative plutôt que de dysfonctionnements localisés.

Implications pour la Conception d'Environnements de Réalité Virtuelle Immersive

Le modèle EPF a des implications directes et importantes pour la conception d'environnements de réalité virtuelle destinés à représenter l'expérience perceptive autistique et à sensibiliser des personnes neurotypiques à cette expérience.

Premièrement, les environnements virtuels devraient représenter une richesse accrue des détails perceptifs dans toutes les modalités sensorielles (visuelles, auditives, et dans la mesure du possible tactiles et proprioceptives). Chaque élément de l'environnement devrait être rendu avec un niveau de détail élevé, reflétant le fait que, dans l'expérience autistique, ces détails sont perçus et traités de manière approfondie.

Deuxièmement, les environnements devraient réduire ou éliminer les filtres attentionnels qui, dans l'expérience neurotypique, atténuent automatiquement les stimuli jugés non pertinents. Par exemple, dans une salle de classe, tous les sons (conversations

périphériques, bruit de ventilation, bruits extérieurs) devraient être présents et saillants, plutôt que d'être automatiquement filtrés pour ne laisser passer que la voix de l'enseignant.

Troisièmement, la simulation devrait modéliser l'absence de hiérarchisation automatique entre information centrale et périphérique. Visuellement, cela pourrait se traduire par un focus attentionnel moins restreint (pas de flou périphérique), de sorte que l'ensemble du champ visuel soit traité avec une résolution élevée. Auditivement, tous les flux sonores pourraient avoir un niveau de saillance comparable, sans hiérarchisation automatique en fonction de leur pertinence sociale.

Quatrièmement, la simulation pourrait modéliser une persistance temporelle prolongée des traces perceptives. Les stimuli visuels pourraient laisser des traces rémanentes plus longues, et les stimuli auditifs pourraient se superposer de manière plus marquée, reflétant un traitement perceptif étendu dans le temps.

Enfin, l'environnement virtuel devrait intégrer des mécanismes permettant d'illustrer la saturation sensorielle : lorsque la quantité d'information perceptive dépasse les capacités de traitement, cela pourrait se traduire par une intensification des stimuli, une distorsion, ou d'autres marqueurs phénoménologiques de la surcharge sensorielle.

Ces principes de conception, inspirés du modèle EPF, visent à permettre aux utilisateurs neurotypiques d'expérimenter, dans un cadre contrôlé et éthique, certains aspects de l'expérience perceptive TSA, favorisant ainsi une compréhension empathique et incarnée des particularités sensorielles dans le TSA.

ANNEXE B - CADRE THÉORIQUE : RÉALITÉ VIRTUELLE IMMERSIVE

Cette section expose les concepts théoriques et techniques essentiels à la conception d'environnements de réalité virtuelle immersive (RVI). Elle aborde les mécanismes psychologiques de l'*embodiment* et du sentiment de présence et agentivité, les facteurs de cinétose et leurs solutions de conception, ainsi que les principes ergonomiques d'interfaçage en environnement virtuel tridimensionnel. Ces éléments constituent le socle méthodologique sur lequel repose la conception du dispositif expérimental utilisé dans cette recherche. La compréhension de ces principes est essentielle pour concevoir des environnements virtuels qui maximisent l'immersion et le sentiment de présence tout en minimisant l'inconfort et en respectant les particularités sensorielles des populations cibles.

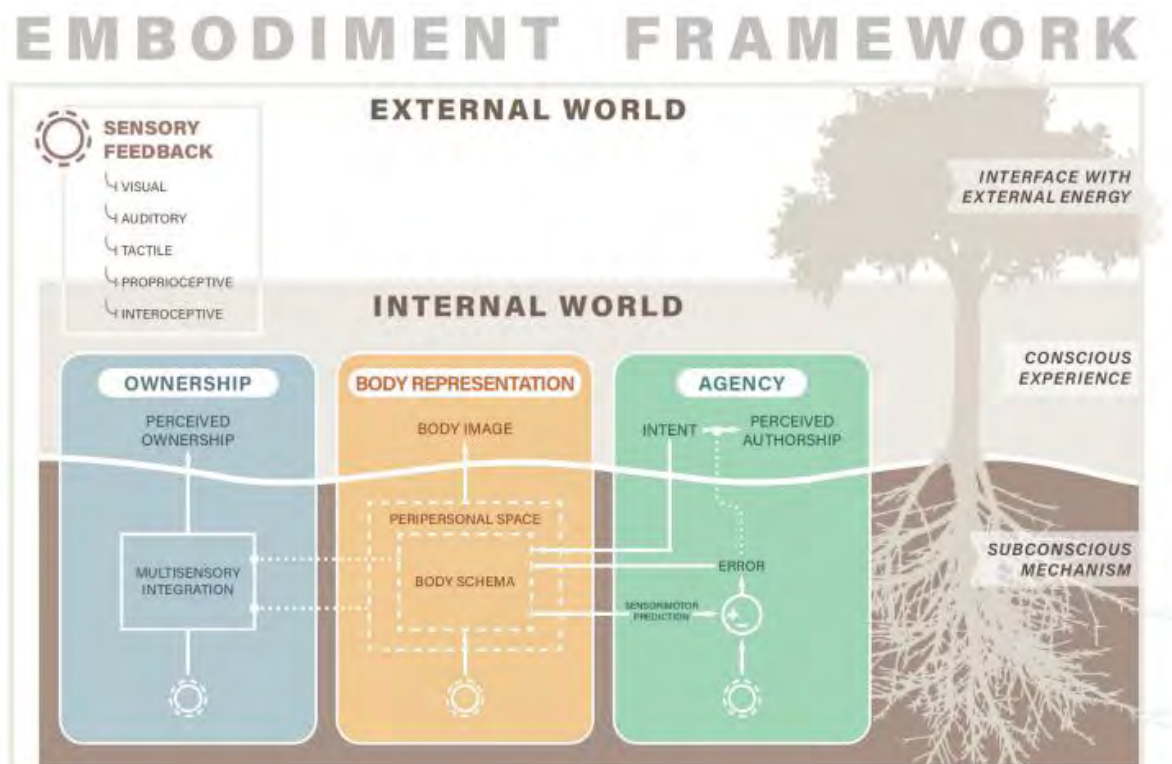
B.1 Embodiment et sentiment d'incarnation

L'embodiment, ou sentiment d'incarnation, constitue l'un des mécanismes psychologiques centraux de l'expérience en réalité virtuelle. Il désigne le sentiment subjectif d'habiter un corps virtuel et de l'expérimenter comme le sien propre. Ce phénomène repose sur l'intégration multisensorielle et la cohérence entre les signaux visuels, proprioceptifs, vestibulaires et moteurs. *L'embodiment* n'est pas un phénomène unitaire mais comprend plusieurs composantes interdépendantes qui contribuent conjointement au sentiment global d'incarnation dans un corps virtuel.

B.1.1 Embodiment framework de Segil et al. (2022)

Figure 41

Embodiment Framework de Segil et al. (2022)



Le cadre conceptuel de l'*embodiment* proposé par Segil et al. (2022) offre une conceptualisation intégrative et multidimensionnelle du sentiment d'incarnation en réalité virtuelle. Ce modèle distingue trois domaines principaux qui, bien qu'interconnectés, peuvent être dissociés empiriquement et correspondent à des mécanismes neurocognitifs partiellement distincts.

Body Ownership (sentiment de propriété corporelle)

Le sentiment de propriété corporelle désigne l'expérience subjective selon laquelle le corps virtuel observé m'appartient, qu'il fait partie intégrante de mon schéma corporel, qu'il est « mon » corps. Ce sentiment émerge principalement de la congruence multisensorielle entre différentes sources d'information sensorielle, notamment :

La correspondance visuo-proprioceptive : Lorsque je vois un corps virtuel occuper l'espace que mon corps physique occupe (par exemple, lorsque je baisse les yeux et vois des jambes virtuelles à l'endroit où je sens mes jambes physiques), cette correspondance spatiale favorise le sentiment de propriété.

La synchronie visuo-tactile : Lorsque je vois le corps virtuel être touché au moment exact où je ressens moi-même un toucher (comme dans l'illusion de la main en caoutchouc classique), cette synchronie temporelle renforce puissamment le sentiment que ce corps est le mien. En réalité virtuelle, cette synchronie peut être créée via des dispositifs haptiques, mais elle peut aussi émerger simplement de la correspondance visuo-proprioceptive lors des mouvements.

La congruence de forme et d'apparence : Un corps virtuel qui ressemble au corps réel (en termes de proportions, de couleur de peau, de morphologie) facilite le sentiment de propriété, bien que cette correspondance ne soit pas strictement nécessaire. Les recherches montrent qu'il est possible de ressentir un sentiment de propriété envers des corps virtuels très différents du corps réel, y compris des corps d'apparence non humaine, à condition que la correspondance visuo-proprioceptive et motrice soit maintenue.

Le sentiment de propriété corporelle est typiquement mesuré via des questionnaires subjectifs comprenant des items tels que « J'avais l'impression que le corps virtuel était mon corps » ou « Il me semblait que je regardais mon propre corps ». Des mesures objectives incluent la réaction électrodermale lors de menaces dirigées vers le corps virtuel (body threat illusion) et les biais proprioceptifs (dérive du jugement de la position des membres réels vers la position des membres virtuels).

Body Representation (représentation corporelle)

La représentation corporelle désigne les représentations mentales, explicites et implicites, du corps dans l'espace. Segil et al. (2022) distinguent deux types de représentations corporelles :

Le schéma corporel (*body schema*) : Une représentation sensori-motrice dynamique et largement implicite du corps, qui sous-tend le contrôle moteur et l'action. Le schéma corporel se met à jour automatiquement en fonction du retour sensoriel et permet la planification et l'exécution fluides des mouvements. En réalité virtuelle, le schéma corporel peut être rapidement recalibré pour intégrer le corps virtuel : les utilisateurs ajustent spontanément leurs mouvements en fonction des dimensions du corps virtuel, démontrant une plasticité remarquable du schéma corporel.

L'image du corps (*body image*) : Une représentation cognitive, perceptive et affective plus explicite du corps, incluant la connaissance conceptuelle du corps, les perceptions corporelles, et les attitudes envers le corps. L'image du corps est moins directement impliquée dans le contrôle moteur mais joue un rôle important dans la conscience de soi corporelle et l'identité.

En réalité virtuelle, des modifications du corps virtuel (par exemple, allongement des bras, modification des proportions) peuvent induire des recalibrations du schéma corporel qui persistent brièvement après le retrait du casque, démontrant la plasticité de ces représentations corporelles.

Agency (sentiment d'agentivité)

Le sentiment d'agentivité désigne l'expérience subjective d'être l'agent causal des mouvements et actions du corps virtuel, le sentiment que c'est moi qui contrôle ce corps, que ses mouvements sont initiés par mes intentions motrices et correspondent à mes volitions. L'agentivité repose sur plusieurs facteurs :

La correspondance temporelle entre intention motrice et mouvement observé : Un délai minimal (idéalement < 20ms, acceptable jusqu'à environ 75-100ms) entre l'initiation d'un mouvement et sa manifestation visuelle dans le corps virtuel est crucial. Des latences plus importantes brisent le sentiment d'agentivité et créent une expérience d'asynchronie dérangeante.

La correspondance spatiale entre commande motrice et mouvement résultant : Le mouvement du corps virtuel doit correspondre spatialement et directionnellement au mouvement effectué. Par exemple, si je lève mon bras droit, le bras droit du corps virtuel doit se lever, avec une amplitude et une trajectoire correspondante.

La prédictibilité : Le système moteur génère constamment des prédictions (*efference copy*) sur les conséquences sensorielles attendues de nos actions. L'agentivité émerge lorsque les conséquences observées correspondent aux prédictions. En réalité virtuelle, tout comportement imprévisible du corps virtuel (mouvements non intentionnels, réponses erratiques) diminue l'agentivité.

Le sentiment d'agentivité est mesuré via des questionnaires incluant des items tels que « J'avais l'impression de contrôler les mouvements du corps virtuel » ou « Les mouvements du corps virtuel correspondaient à mes intentions ». Des mesures comportementales incluent le paradigme de binding temporel (les actions volontaires et leurs conséquences sont perçues comme temporellement plus proches qu'elles ne le sont objectivement) et la sensibilité aux perturbations sensorimotrices.

Intégration des trois domaines

Le modèle de Segil et al. (2022) souligne que ces trois domaines, bien que conceptuellement distincts, interagissent de manière complexe pour générer le sentiment global d'*embodiment*. Le sentiment de propriété corporelle est facilité par l'agentivité : lorsque je contrôle les mouvements du corps virtuel, je suis plus susceptible de le ressentir comme mien. Réciproquement, le sentiment de propriété renforce l'agentivité : lorsque je ressens le corps virtuel comme étant mon corps, j'ai plus facilement le sentiment de contrôler ses actions.

Les représentations corporelles (schéma corporel et image du corps) sont modifiées par l'expérience d'*embodiment* : un *embodiment* prolongé dans un corps virtuel aux proportions différentes peut induire des recalibrations persistantes du schéma corporel. Ces recalibrations, à leur tour, influencent les sentiments de propriété et d'agentivité dans les expositions subséquentes.

Le modèle distingue également entre l'expérience consciente de l'*embodiment* (les sentiments subjectifs rapportés dans les questionnaires) et les mécanismes conscients (les ajustements automatiques du schéma corporel, les recalibrations sensorimotrices) qui sous-tendent cette expérience. Cette distinction est importante car elle suggère que l'*embodiment* peut influencer le comportement et la cognition même en l'absence d'une prise de conscience explicite.

Implications pour la conception de dispositifs de RVI

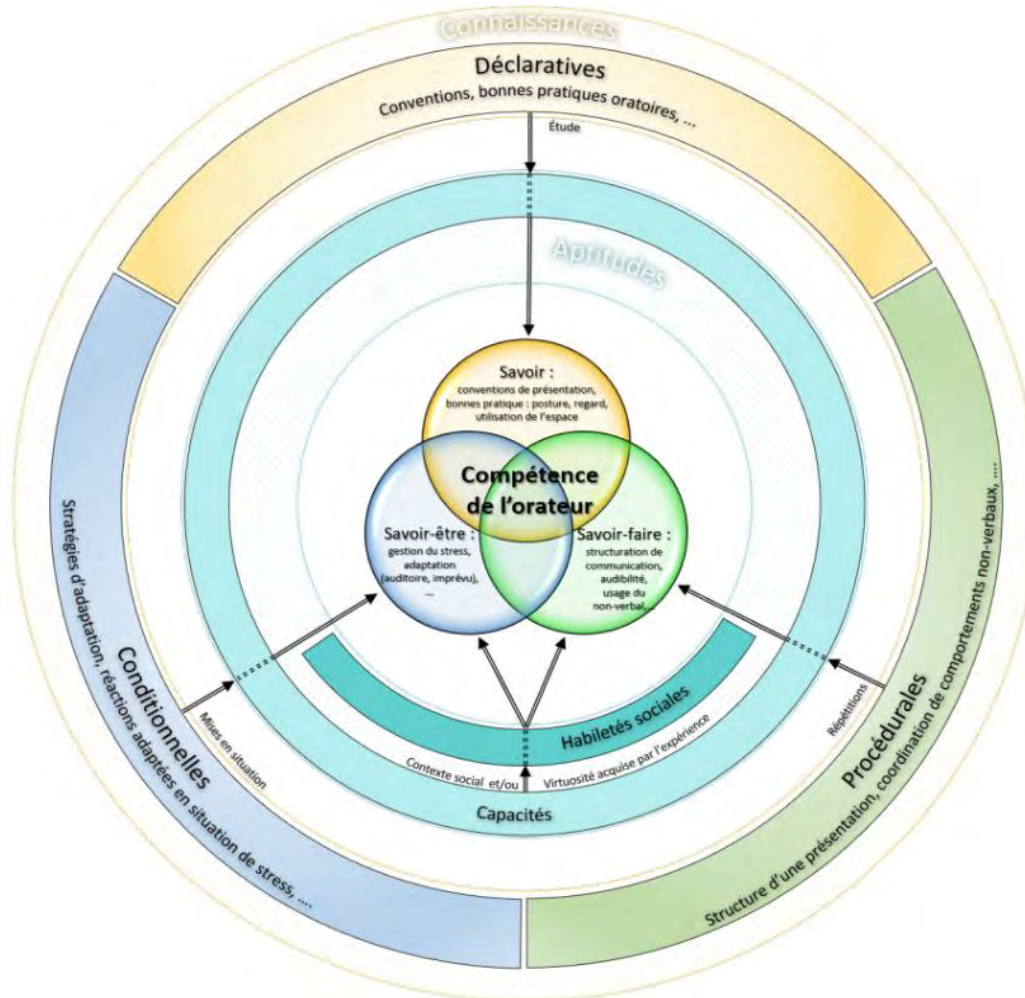
Le cadre de Segil et al. (2022) fournit des principes directeurs pour maximiser l'*embodiment* en réalité virtuelle :

1. Assurer une latence minimale (< 20ms idéalement, < 75-100ms au maximum) entre les mouvements réels et leur représentation virtuelle pour préserver l'agentivité.
2. Maintenir une correspondance visuo-proprioceptive précise, en alignant la position du corps virtuel avec la position ressentie du corps physique.
3. Fournir un retour haptique ou vibrotactile synchronisé avec les événements visuels lorsque c'est possible, pour renforcer la correspondance multisensorielle.
4. Concevoir des corps virtuels dont les proportions et les possibilités d'action sont compatibles avec les capacités motrices réelles de l'utilisateur, pour faciliter la recalibration du schéma corporel.
5. Éviter les comportements imprévisibles ou non intentionnels du corps virtuel, qui brisent l'agentivité et peuvent générer une expérience perturbante (*uncanny valley* corporelle).

B.1.2 Schéma des compétences sociales de l'orateur (Taupiac, 2020)

Figure 42

Schéma des compétences sociales de l'orateur de Taupiac (2020)



Ce schéma de Taupiac (2020) dans le contexte de la formation à la prise de parole en public offre un cadre conceptuel pertinent pour comprendre les multiples dimensions de compétence impliquées dans les situations de communication sociale, particulièrement dans les contextes éducatifs formels. Bien que ce modèle ait été initialement développé pour la population générale, il présente un intérêt particulier pour comprendre les défis spécifiques auxquels font face les personnes avec TSA dans les situations de communication sociale.

Compétences techniques

Les compétences techniques désignent la maîtrise des outils et supports de communication. Dans le contexte éducatif, cela inclut :

- La maîtrise des supports de présentation (diaporamas, tableaux, outils numériques)
- La gestion de l'espace physique (déplacement dans la salle, utilisation du tableau)
- La manipulation des objets et des outils pédagogiques
- La gestion du temps et du rythme de présentation

Pour les personnes avec TSA, ces compétences techniques peuvent représenter à la fois une source de difficulté (en raison de particularités motrices ou de rigidité dans l'utilisation des outils) et une force (capacité à maîtriser en profondeur les aspects techniques, préférence pour les aspects structurés et prévisibles de la présentation).

Compétences communicationnelles

Les compétences communicationnelles englobent les aspects verbaux et non verbaux de la communication :

Composante verbale : Clarté de l'articulation, débit de parole approprié, volume et projection vocale, prosodie (variation de l'intonation, accentuation), utilisation d'un vocabulaire adapté à l'audience, structuration logique du discours.

Composante non verbale : Expressions faciales, contact visuel, gestes illustratifs, posture corporelle, proxémique (utilisation de l'espace par rapport à l'audience), synchronisation entre verbal et non-verbal.

Les recherches documentent que les personnes avec TSA présentent fréquemment des particularités dans ces domaines communicationnels. La prosodie peut être atypique (monotone ou avec des patterns d'intonation inhabituels), le contact visuel peut être réduit ou vécu comme inconfortable, et la synchronisation entre communication verbale et non verbale peut être moins automatique. Cependant, ces particularités ne constituent pas nécessairement des déficits mais plutôt des différences de style communicationnel qui peuvent être fonctionnelles dans certains contextes.

Compétences émotionnelles

Les compétences émotionnelles concernent la reconnaissance, la compréhension, et la régulation des émotions, tant les siennes propres que celles d'autrui :

- Reconnaissance et gestion du stress et de l'anxiété de performance
- Régulation des émotions intenses (excitation, frustration, découragement)
- Conscience émotionnelle (capacité à identifier et nommer ses propres états émotionnels)
- Récupération émotionnelle suite à des difficultés ou des erreurs

Les situations de prise de parole en public sont typiquement génératrices d'anxiété, même chez les personnes neurotypiques. Pour les personnes avec TSA, cette anxiété peut être amplifiée par l'imprévisibilité des réactions de l'audience, la nécessité de traiter simultanément de multiples flux d'information (contenu verbal, langage corporel de l'audience, gestion de ses propres états internes), et la surcharge sensorielle potentielle (bruits, éclairage, présence de nombreuses personnes).

De plus, les particularités dans la reconnaissance des émotions d'autrui (alexithymie, difficultés de lecture des expressions faciales subtiles) peuvent compliquer l'ajustement en temps réel de la communication en fonction des réactions de l'audience.

Compétences sociales

Les compétences sociales désignent la capacité à adapter son comportement au contexte social, à lire et interpréter les signaux sociaux, et à ajuster sa communication en fonction de l'audience :

- Adaptation du registre de langue et du contenu au niveau de l'audience
- Lecture des signaux d'intérêt, de compréhension ou d'incompréhension de l'audience
- Ajustement du rythme et du niveau de détail en fonction des réactions observées
- Gestion des interruptions et des questions
- Compréhension des normes sociales implicites de la situation (formalité, durée attendue, conventions de prise de parole)

Ces compétences sociales sont précisément celles qui posent souvent le plus de défis aux personnes avec TSA. Les normes sociales implicites peuvent ne pas être intuitives et nécessiter un apprentissage explicite. L'ajustement en temps réel basé sur la lecture des signaux sociaux subtils de l'audience peut être cognitivement coûteux. Les interruptions et les changements de plan peuvent être perturbants en raison d'une préférence pour la prévisibilité et les routines.

Interrelations entre les domaines de compétence

Le modèle de Taupiac (2020) souligne que ces quatre domaines de compétence sont interdépendants. Une difficulté dans un domaine peut être compensée par des forces dans d'autres domaines. Par exemple, une personne qui maîtrise très bien les aspects techniques et la structuration logique du contenu peut compenser partiellement des difficultés dans la lecture des signaux sociaux en s'appuyant sur une présentation très claire et structurée.

Inversement, des difficultés cumulées dans plusieurs domaines peuvent interagir de manière multiplicative pour créer une expérience de communication particulièrement difficile. Par exemple, l'anxiété (compétence émotionnelle) peut interférer avec la fluidité verbale (compétence communicationnelle), ce qui peut à son tour augmenter l'anxiété dans un cycle d'amplification.

Pertinence pour la conception du dispositif de RVI

Ce modèle est particulièrement pertinent pour la conception du dispositif de réalité virtuelle immersive visant à sensibiliser à l'expérience autistique dans le contexte éducatif. L'environnement virtuel simule précisément une situation de prise de parole en classe, contexte dans lequel les quatre domaines de compétence identifiés par Taupiac (2020) sont sollicités simultanément.

En plaçant l'utilisateur neurotypique dans la position d'un étudiant avec TSA confronté à ces exigences multidimensionnelles, tout en simulant les particularités perceptives et sensorielles caractéristiques du TSA (surcharge sensorielle, difficulté à filtrer les stimuli non pertinents, anxiété), le dispositif permet une prise de conscience incarnée de la complexité et de la charge cognitive associées à ces situations pour les personnes avec TSA.

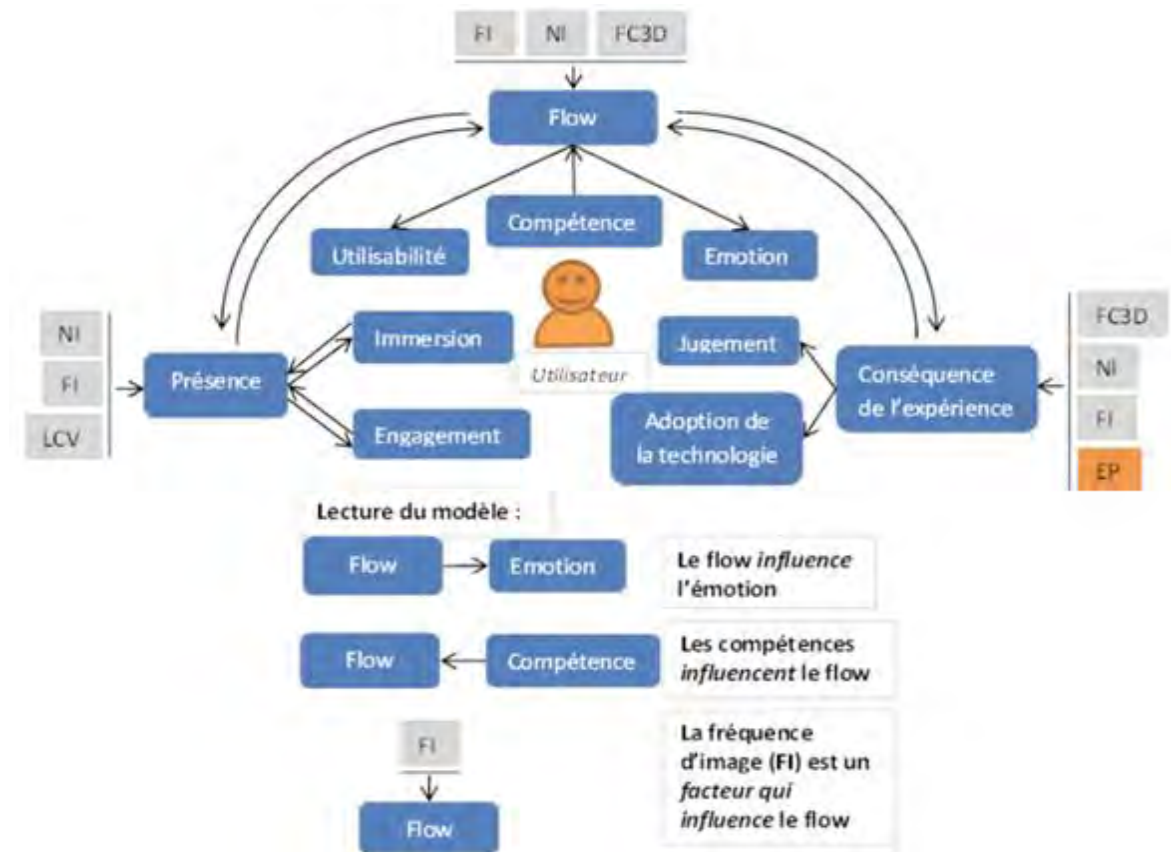
Le modèle suggère également des pistes pour structurer l'expérience de RVI en différentes phases qui ciblent successivement différents domaines de compétence, permettant à

l'utilisateur de prendre conscience progressivement de la manière dont ces multiples exigences s'accroissent et interagissent.

B.1.3 Modèle heuristique de Tcha-Tokey et al. (2015)

Figure 43

Modèle Heuristique de Tcha-Tokey et al. (2015)



Le modèle heuristique proposé par Tcha-Tokey et al. (2015) offre un cadre intégratif pour comprendre les multiples facteurs qui influencent l'expérience utilisateur en réalité virtuelle et leurs conséquences sur les variables d'intérêt telles que l'apprentissage, le changement d'attitude, et l'adoption de la technologie. Ce modèle est particulièrement précieux pour guider la conception de dispositifs de RVI à visée pédagogique ou thérapeutique, car il explicite les chaînes causales reliant les caractéristiques techniques du système aux *outcomes* psychologiques et comportementaux.

Variables d'expérience subjective

Le modèle identifie plusieurs variables psychologiques qui caractérisent l'expérience subjective de l'utilisateur en RV :

Flow (état d'engagement optimal) : Le flow désigne un état psychologique caractérisé par une absorption totale dans l'activité, une perte de la conscience de soi, une distorsion de la perception du temps, et un sentiment de contrôle et de maîtrise. Cet état émerge lorsque le niveau de défi de la tâche est optimalement ajusté aux compétences de l'individu : ni trop facile (ce qui conduirait à l'ennui) ni trop difficile (ce qui conduirait à l'anxiété).

En réalité virtuelle, le flow est favorisé par des environnements qui fournissent des objectifs clairs, un retour immédiat sur les actions, et un niveau de difficulté progressif. Le flow est associé à une expérience intrinsèquement gratifiante, à un apprentissage profond, et à une motivation persistante.

Immersion : Dans le modèle de Tcha-Tokey et al. (2015), l'immersion désigne l'immersion système, c'est-à-dire les caractéristiques objectives du système technologique qui sont susceptibles de créer un sentiment d'être enveloppé par l'environnement virtuel. Ces caractéristiques incluent la qualité graphique, la richesse sensorielle (modalités sensorielles engagées), le champ de vision, et le niveau d'interactivité.

L'immersion système est une propriété de la technologie et peut être mesurée objectivement. Elle se distingue de la présence (voir ci-dessous), qui est la réponse psychologique subjective à l'immersion système.

Présence : La présence, également appelée sentiment d'être là (*sense of being there*), désigne l'expérience subjective de se sentir localisé et présent dans l'environnement virtuel plutôt que dans l'environnement physique où le corps se trouve réellement. Witmer et Singer (1998) définissent la présence comme « *l'expérience subjective d'être dans un lieu ou un environnement, même lorsque l'on est physiquement situé dans un autre* ».

La présence est influencée par l'immersion système (meilleure immersion technique favorise la présence) mais aussi par des facteurs individuels (capacité d'attention, tendance à l'absorption imaginative) et contextuels (distracteurs dans l'environnement physique, familiarité avec la technologie).

La présence est généralement considérée comme un médiateur important des effets de la réalité virtuelle : c'est en partie parce que les utilisateurs se sentent réellement présents dans

l'environnement virtuel que cet environnement peut induire des réactions émotionnelles, cognitives et comportementales similaires à celles qu'induirait un environnement réel correspondant.

Engagement : L'engagement désigne le degré d'implication cognitive, émotionnelle et comportementale de l'utilisateur dans l'expérience virtuelle. Un utilisateur engagé est attentif, investi, et activement impliqué dans les interactions avec l'environnement virtuel.

L'engagement est influencé par la pertinence et l'intérêt du contenu, par la qualité de l'interactivité (réactivité du système, richesse des possibilités d'action), et par le sentiment de contrôle et d'efficacité. L'engagement est crucial pour l'apprentissage et le transfert : un utilisateur désengagé traitera l'expérience de manière superficielle et en retirera peu de bénéfices durables.

Émotion : Les émotions ressenties pendant l'expérience de réalité virtuelle constituent une variable centrale, particulièrement dans les applications visant la sensibilisation ou le changement d'attitude. La réalité virtuelle a la capacité d'induire des réponses émotionnelles intenses et écologiquement valides, précisément parce que les utilisateurs se sentent présents dans l'environnement virtuel.

Les émotions peuvent être variées : curiosité et intérêt (favorisant l'exploration et l'apprentissage), anxiété ou peur (dans les expositions thérapeutiques ou les simulations de situations stressantes), empathie et compassion (dans les expériences de prise de perspective), satisfaction et joie.

La valence et l'intensité des émotions doivent être soigneusement considérées dans la conception : des émotions trop intenses peuvent être contre-productives (surcharge émotionnelle, réactions défensives), tandis que des émotions trop faibles peuvent indiquer un manque d'engagement ou de présence.

Jugement : Le jugement désigne l'évaluation cognitive que l'utilisateur fait de l'expérience virtuelle : réalisme perçu, crédibilité, utilité perçue, satisfaction globale. Ces évaluations influencent l'attitude envers la technologie et la volonté de l'utiliser à nouveau ou de la recommander.

Le jugement est influencé par les attentes préalables, par la comparaison avec d'autres expériences (médias traditionnels, réalité), et par l'adéquation perçue entre les objectifs de l'expérience et sa réalisation effective.

Facteurs technologiques modulateurs

Le modèle de Tcha-Tokey et al. (2015) identifie plusieurs paramètres technologiques qui modulent l'expérience subjective :

FI : Fréquence d'images (Frame Rate) : La fréquence d'images désigne le nombre d'images affichées par seconde. Une fréquence élevée (90-120 Hz dans les HMD modernes) est essentielle pour la fluidité du mouvement, la réduction de la latence perçue, et la minimisation de la cinétose. Une fréquence insuffisante crée des saccades, augmente la latence motion-to-photon, et dégrade l'expérience de présence et d'immersion.

NI : Niveau d'interactivité : Le niveau d'interactivité désigne la richesse et la réactivité des interactions possibles avec l'environnement virtuel. Une interactivité élevée signifie que l'utilisateur peut agir sur de nombreux éléments de l'environnement, que ces actions produisent des conséquences visibles et cohérentes, et que le système répond rapidement et de manière prévisible.

L'interactivité est cruciale pour l'agentivité et le sentiment de contrôle. Elle favorise l'engagement actif (par opposition à l'observation passive) et supporte l'apprentissage par l'action et l'exploration.

FC3D : Feedback du contenu 3D : Ce paramètre désigne la qualité graphique, le réalisme visuel, et la richesse des détails de l'environnement virtuel. Cela inclut la résolution des textures, la complexité géométrique des modèles 3D, l'éclairage dynamique, les effets de post-traitement, et la cohérence physique (physique des objets, animations réalistes).

Un feedback 3D de haute qualité favorise l'immersion et la présence en créant un environnement visuellement convaincant. Cependant, la relation n'est pas strictement linéaire : un hyper-réalisme peut parfois créer un effet de Vallée de l'étrange (*uncanny valley*) de Mori (1970), et un style visuel stylisé mais cohérent peut être préférable dans certains contextes (particulièrement pour des applications pédagogiques).

LCV : Largeur du champ de vision : Le champ de vision désigne l'étendue angulaire de l'environnement virtuel visible à travers le HMD. Un champ de vision large (110° ou plus) favorise l'immersion en englobant la vision périphérique et en réduisant la conscience du cadre limitant du dispositif.

Cependant, un champ de vision très large peut aussi augmenter la cinétose lors des mouvements virtuels, créant un compromis entre immersion et confort. Les techniques de

réduction dynamique du champ de vision lors des déplacements (*vignetting*) permettent de gérer ce compromis.

Facteurs individuels

Le modèle reconnaît également l'importance des différences individuelles :

EP : Expérience précédente avec la RV : Les utilisateurs novices et expérimentés ont des expériences qualitativement différentes. Les novices peuvent être plus susceptibles à la cinétose, peuvent nécessiter plus de temps pour maîtriser les interfaces et les contrôles, et peuvent être plus facilement impressionnés par l'immersion. Les utilisateurs expérimentés, inversement, sont plus à l'aise techniquement, moins susceptibles à la cinétose, mais peuvent aussi être plus critiques et moins facilement immergés.

Cette variable suggère l'importance d'adapter les dispositifs de RVI au niveau d'expérience de la population cible, notamment via des tutoriels appropriés et des ajustements du niveau de complexité.

Compétence perçue : Le sentiment de compétence et d'efficacité dans l'utilisation de la technologie influence l'anxiété, l'engagement, et la volonté de persister face aux difficultés. Les utilisateurs qui se perçoivent comme compétents sont plus susceptibles d'explorer activement, de prendre des risques, et de s'engager profondément dans l'expérience.

Conséquences de l'expérience : Le modèle identifie trois types principaux de conséquences résultant de l'expérience de réalité virtuelle :

Adoption de la Technologie : L'expérience utilisateur influence l'attitude envers la technologie de RVI elle-même et la volonté de l'utiliser à nouveau. Une expérience positive (forte présence, engagement, émotions appropriées, jugement favorable) favorise l'adoption et l'utilisation continue.

Apprentissage et transfert : Pour les applications pédagogiques ou de formation, la qualité de l'expérience influence l'apprentissage (acquisition de connaissances, développement de compétences) et le transfert (application des apprentissages dans des contextes réels hors de la RVI).

Le modèle suggère que les variables d'expérience (flow, engagement, présence) médient la relation entre les caractéristiques techniques et les *outcomes* d'apprentissage : une technologie

immersive favorise l'apprentissage dans la mesure où elle induit un état de flow, d'engagement profond, et de présence.

Changement d'attitude : Dans les applications de sensibilisation, l'objectif principal est souvent le changement d'attitude : modification des croyances, des stéréotypes, et des intentions comportementales envers le groupe cible (ici, les personnes avec TSA).

Le modèle suggère que les émotions ressenties pendant l'expérience (notamment l'empathie) et la présence jouent un rôle médiateur crucial : c'est parce que les utilisateurs se sentent réellement présents dans la situation simulée et ressentent des émotions authentiques qu'ils sont susceptibles de modifier durablement leurs attitudes.

Implications pour la conception de dispositifs de sensibilisation au TSA

Le modèle de Tcha-Tokey et al. (2015) fournit un cadre pour optimiser la conception de dispositif de sensibilisation :

1. Maximiser l'immersion système (qualité graphique, interactivité, richesse sensorielle) pour favoriser la présence.
2. Concevoir des scénarios et des mécaniques d'interaction qui favorisent le flow et l'engagement actif plutôt que l'observation passive.
3. Calibrer soigneusement l'intensité émotionnelle : suffisamment intense pour créer une expérience marquante et empathique, mais pas au point de générer une défense psychologique ou une surcharge.
4. Fournir des tutoriels et un scaffolding appropriés pour maximiser la compétence perçue et minimiser l'anxiété technologique, particulièrement pour les utilisateurs novices.
5. Mesurer systématiquement les variables d'expérience (présence, agentivité, empathie, cinétose) pour vérifier que le dispositif induit bien l'expérience subjective visée.
6. Évaluer les conséquences à court et long terme sur les attitudes, les connaissances, et les intentions comportementales pour valider l'efficacité du dispositif.

B.2 Cinétose et confort en RV

La cinétose, également appelée *cybersickness* dans le contexte spécifique de la réalité virtuelle, constitue l'un des défis majeurs de l'utilisation des environnements virtuels immersifs. Elle désigne l'ensemble des symptômes d'inconfort physiologique et psychologique ressentis lors d'une exposition à un environnement virtuel, incluant nausée, désorientation, fatigue oculaire, vertiges, et malaise général. La compréhension des mécanismes de la cinétose et des stratégies pour la minimiser est essentielle pour concevoir des dispositifs de RVI utilisables, confortables, et éthiques, particulièrement lorsque les utilisateurs cibles incluent des populations potentiellement plus vulnérables comme les personnes avec TSA qui peuvent présenter des particularités sensorielles.

B.2.1 Glossaire de la cinétose (Arnaldi et al., 2018)

Tableau 7

Types de cinétose : stimuli, technologies et symptômes caractéristiques

Type	Stimulus déclencheur	Technologies concernées	Symptômes principaux
<i>Simulator Sickness</i>	Discordance entre stimulation visuelle (mouvement perçu) et stimulation vestibulaire/kinesthésique (absence de mouvement physique correspondant) dans les simulateurs	Simulateurs de vol, simulateurs de conduite automobile, simulateurs d'engins lourds	Nausée, désorientation, sudation, pâleur, salivation excessive, sensation de lourdeur de la tête
<i>Motion Sickness (Mal des transports)</i>	Mouvement physique réel répété ou prolongé (oscillations, accélérations, décélérations) affectant le système vestibulaire	Transports maritimes (bateau), transports routiers (voiture, bus), transports aériens, manèges	Nausée pouvant aller jusqu'au vomissement, pâleur, sueurs froides, hypersalivation, apathie, somnolence

Type	Stimulus déclencheur	Technologies concernées	Symptômes principaux
<i>Cybersickness</i>	Conflit visuo-vestibulaire induit par l'environnement virtuel : le système visuel détecte un mouvement tandis que le système vestibulaire ne détecte aucun mouvement physique correspondant	Casques de réalité virtuelle (HMD), CAVE (environnements immersifs multi-projecteurs), écrans larges, jeux vidéo en première personne	Fatigue oculaire (<i>eyestrain</i>), désorientation spatiale, nausée, vertiges, difficulté de concentration, flou visuel, maux de tête
<i>VIMS (Visually Induced Motion Sickness)</i>	Stimulation visuelle de mouvement intense en l'absence de mouvement physique, particulièrement via le flux optique périphérique	Grands écrans de cinéma (IMAX), contenus vidéo en point de vue subjectif, écrans courbés, jeux vidéo immersifs sur grand écran	Désorientation, sensation de vertige, inconfort général, nausée légère à modérée, instabilité posturale perçue

Le glossaire proposé par Arnaldi et al. (2018) distingue quatre types de cinétose en fonction des stimuli déclencheurs, des technologies impliquées, et des profils symptomatiques. Cette taxonomie permet de mieux comprendre les mécanismes spécifiques à chaque contexte et d'adapter les stratégies de prévention en conséquence (voir tableau B.2.1).

Distinctions conceptuelles importantes

Bien que ces quatre types de cinétose partagent certains symptômes communs (notamment la nausée et la désorientation), ils se distinguent par leurs mécanismes déclencheurs et leurs profils symptomatiques spécifiques.

Le *simulator sickness* et le *cybersickness* partagent un mécanisme commun (conflit visuo-vestibulaire) mais se distinguent par le contexte technologique : les simulateurs impliquent typiquement une plateforme physique qui peut elle-même générer des mouvements

(créant ainsi des patterns de conflit sensoriel complexes), tandis que le *cybersickness* pur concerne des dispositifs où l'utilisateur reste physiquement immobile.

Le *motion sickness* classique (mal des transports) implique un mouvement physique réel mais peut néanmoins générer des conflits sensoriels, notamment lorsque les informations visuelles ne correspondent pas aux mouvements vestibulaires (par exemple, lire dans une voiture en mouvement : le système vestibulaire détecte les accélérations tandis que le système visuel, focalisé sur le texte stable, ne les détecte pas).

Le VIMS représente un cas où la stimulation visuelle seule, sans aucune stimulation vestibulaire ni mouvement physique, suffit à déclencher des symptômes. Ce phénomène démontre la puissance de la vision dans la genèse de la cinétose et souligne l'importance du flux optique périphérique dans le déclenchement des symptômes.

Mécanismes explicatifs de la cinétose

Plusieurs théories ont été proposées pour expliquer les mécanismes sous-jacents de la cinétose. Ces théories ne sont pas mutuellement exclusives et peuvent être partiellement complémentaires.

Théorie du conflit sensoriel (Sensory Conflict Theory)

La théorie du conflit sensoriel, proposée par Reason et Brand (1975) et développée par Oman (1990), constitue l'explication dominante de la cinétose. Selon cette théorie, le système nerveux central maintient constamment un modèle interne de la configuration spatiale attendue du corps et de ses relations avec l'environnement. Ce modèle génère des prédictions sur les patterns de stimulation sensorielle attendus.

La cinétose émergerait lorsqu'il existe une discordance entre les patterns de stimulation sensorielle attendus (prédictions du modèle interne) et les patterns effectivement reçus. Plus spécifiquement, trois types de conflits ont été identifiés :

1. Conflit inter-sensoriel : Discordance entre les informations fournies par différents systèmes sensoriels. Par exemple, en RV, le système visuel signale un mouvement tandis que le système vestibulaire signale une immobilité. Ce type de conflit est le plus fréquent dans le *cybersickness*.

2. Conflit intra-sensoriel : Discordance au sein d'un même système sensoriel. Par exemple, dans le système vestibulaire, les canaux semi-circulaires (qui détectent les rotations)

et les organes otolithiques (qui détectent les accélérations linéaires et la gravité) peuvent fournir des informations contradictoires dans certaines situations de mouvement complexe.

3. Conflit avec le modèle interne : Même en l'absence de conflit entre les systèmes sensoriels actuels, une discordance peut exister entre les patterns sensoriels reçus et ceux attendus sur la base de l'expérience passée. Par exemple, lors d'une première exposition à un environnement de microgravité, les patterns vestibulaires sont radicalement différents de tous ceux précédemment expérimentés, générant un conflit avec le modèle interne construit dans un environnement gravitationnel.

La théorie du conflit sensoriel explique bien pourquoi la cinétose diminue typiquement avec l'exposition répétée (habituation) : le modèle interne se recalibre progressivement pour accommoder les nouveaux patterns sensoriels, réduisant ainsi le conflit. Elle explique également pourquoi certaines personnes sont plus sensibles que d'autres : les différences dans la flexibilité du modèle interne et dans la pondération relative accordée aux différentes sources d'information sensorielle créent une variabilité inter-individuelle.

Théorie de l'instabilité posturale (Postural Instability Theory)

La théorie de l'instabilité posturale, proposée par Riccio et Stoffregen (1991), offre une perspective alternative. Selon cette théorie, la cinétose émerge non pas directement du conflit sensoriel mais des difficultés à maintenir un contrôle postural stable face à des informations sensorielles inhabituelles ou conflictuelles.

Le système de contrôle postural s'appuie sur l'intégration multisensorielle (vision, système vestibulaire, proprioception) pour maintenir l'équilibre et la stabilité. Lorsque les informations sensorielles sont conflictuelles ou inhabituelles, le système de contrôle postural doit déployer des stratégies compensatoires coûteuses sur le plan énergétique et attentionnel. Selon Riccio et Stoffregen (1991), c'est cet effort prolongé de stabilisation posturale, plutôt que le conflit sensoriel per se, qui génère les symptômes de cinétose.

Cette théorie prédit que les mesures objectives d'instabilité posturale (oscillations du centre de pression, variabilité de la posture) devraient précéder et prédire l'apparition des symptômes subjectifs de cinétose. Plusieurs études ont effectivement montré que les individus présentant une plus grande instabilité posturale en condition de conflit sensoriel sont plus susceptibles de développer des symptômes de cinétose.

La théorie de l'instabilité posturale offre également des prédictions sur les facteurs de protection : les individus ayant un meilleur contrôle postural (par exemple, les athlètes, les danseurs) devraient être moins susceptibles à la cinétose, prédiction qui est partiellement supportée par les données empiriques.

Théorie évolutionniste (Toxin Theory ou Poison Hypothesis)

La théorie évolutionniste, proposée par Treisman (1977), offre une explication phylogénétique de la cinétose. Selon cette théorie, la nausée et le vomissement induits par les conflits sensoriels seraient un vestige évolutif d'un mécanisme adaptatif de détection d'intoxication.

De nombreuses toxines et poisons affectent le système nerveux central et peuvent perturber le fonctionnement du système vestibulaire et de l'intégration sensorielle, créant des sensations de vertige et des conflits sensoriels. Au cours de l'évolution, les organismes ayant développé une réponse réflexe de nausée et de vomissement face à ces signaux de conflit sensoriel auraient eu un avantage sélectif : l'expulsion du contenu gastrique permettait d'éliminer rapidement la toxine ingérée.

Selon cette théorie, la cinétose moderne (liée aux transports, aux simulateurs, à la réalité virtuelle) résulterait d'une activation inadaptée de ce mécanisme ancestral : le système nerveux interprète incorrectement le conflit sensoriel comme un signal d'intoxication et déclenche la réponse émétique, bien qu'aucune toxine ne soit effectivement présente.

Cette théorie est difficile à tester empiriquement mais offre une explication plausible de pourquoi un conflit sensoriel, qui ne représente aucune menace directe, déclenche néanmoins une réponse physiologique aussi intense et désagréable. Elle explique également pourquoi les symptômes incluent des marqueurs du système nerveux autonome (sudation, pâleur, salivation) typiques des réponses à l'intoxication.

Intégration des théories

Ces trois théories ne sont pas mutuellement exclusives et peuvent être intégrées dans un cadre explicatif plus complet. Le conflit sensoriel (théorie 1) génère une instabilité posturale (théorie 2), et cette instabilité, persistante et non résolue, active un mécanisme de détection d'intoxication ancestral (théorie 3), conduisant aux symptômes autonomes et à la nausée.

Cette intégration permet également de comprendre la progression temporelle typique des symptômes : les premiers symptômes sont typiquement la désorientation et l'inconfort (reflet du conflit sensoriel et de l'instabilité posturale), suivis, si l'exposition persiste, par des symptômes autonomes plus intenses (sudation, pâleur) et éventuellement la nausée (activation du mécanisme de détection d'intoxication).

B.2.2 Facteurs technologiques influençant la cinétose

Plusieurs paramètres technologiques des systèmes de réalité virtuelle modulent significativement l'intensité de la cinétose. La compréhension de ces facteurs permet d'optimiser la conception des dispositifs pour minimiser l'inconfort tout en préservant l'immersion et la qualité de l'expérience.

Latence Motion-to-Photon

La latence motion-to-photon désigne le délai entre le mouvement de la tête de l'utilisateur et la mise à jour correspondante de l'affichage visuel. Cette latence résulte de la chaîne complète de traitement : détection du mouvement par les capteurs, transmission des données, calcul de la nouvelle perspective, rendu graphique, et affichage.

La recherche a établi que la latence est un facteur critique de la cinétose. Une méta-analyse de Weech et al. (2019) conclut qu'une latence supérieure à 20 millisecondes augmente significativement les symptômes de *cybersickness*, avec un effet dose-dépendant : plus la latence est élevée, plus les symptômes sont intenses. À 50-75 millisecondes, les symptômes deviennent marqués pour une majorité d'utilisateurs, et au-delà de 100 millisecondes, l'expérience devient inconfortable voire intolérable pour la plupart.

Le mécanisme par lequel la latence induit la cinétose implique un conflit entre le retour proprioceptif (mouvement de la tête détecté immédiatement par le système proprioceptif et vestibulaire) et le retour visuel (changement de perspective visuelle retardé). Ce décalage temporel crée une incohérence multisensorielle qui génère un conflit sensoriel.

Les systèmes de RVI modernes (Meta Quest 3) visent une latence inférieure à 20 millisecondes, et les systèmes haut de gamme atteignent des latences de 10-15 millisecondes. Les techniques d'optimisation incluent le rendu asynchrone (*asynchronous timewarp*), qui re-projette la dernière *frame* en fonction du mouvement le plus récent de la tête pendant que la prochaine *frame* est calculée, réduisant ainsi la latence perçue.

Fréquence d'images (Frame Rate / Refresh Rate)

La fréquence d'images désigne le nombre d'images affichées par seconde (mesurée en Hertz, Hz). Une fréquence élevée est essentielle pour la fluidité du mouvement, la réduction de la latence perçue, et la minimisation de la cinétose.

Les premières générations de HMD fonctionnaient typiquement à 60-75 Hz, ce qui s'est avéré insuffisant pour de nombreux utilisateurs. Les systèmes actuels fonctionnent à 90 Hz (standard minimum), 120 Hz (confortable), voire 144 Hz ou plus pour les systèmes haut de gamme. La recherche suggère un effet de seuil : en dessous de 90 Hz, la cinétose augmente significativement, tandis qu'au-delà de 90 Hz, les bénéfices additionnels deviennent plus marginaux, bien que toujours présents.

Une fréquence insuffisante crée plusieurs problèmes. D'abord, elle augmente la latence effective : si le système affiche 60 images par seconde, il y a en moyenne 16,7 millisecondes entre chaque *frame*, ajoutant de la latence. Ensuite, elle crée un effet stroboscopique ou de saccade (*judder*) lors des mouvements, particulièrement visible dans la vision périphérique. Enfin, elle peut créer une persistance rétinienne désagréable, où l'image précédente est encore visible lorsque la nouvelle image apparaît.

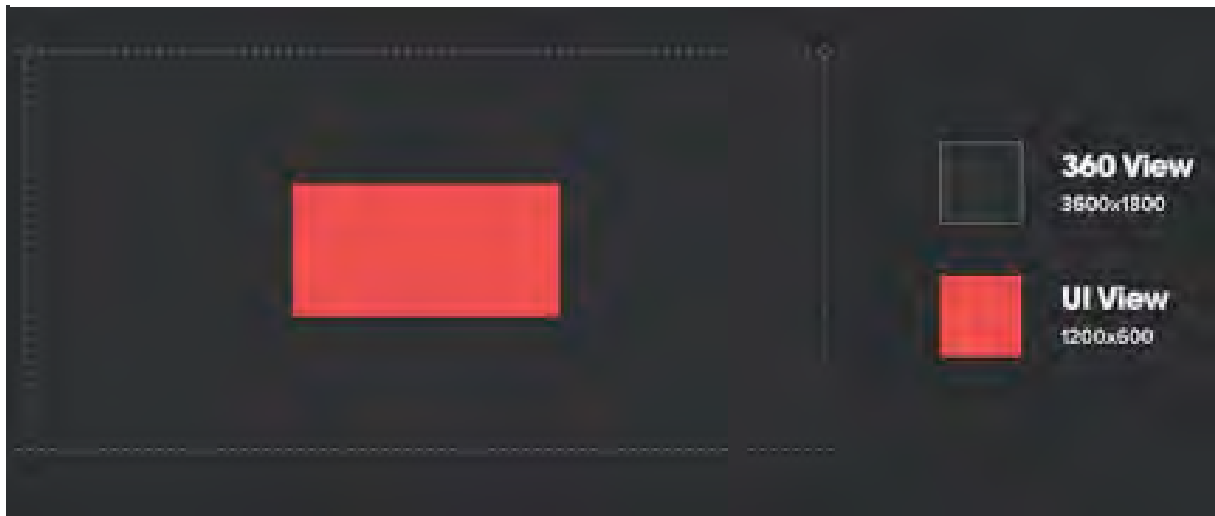
Il est crucial que la fréquence d'images soit maintenue constante. Des chutes de frame rate (frame drops), où le système ne parvient pas à générer les images assez rapidement et où la fréquence descend ponctuellement, sont particulièrement nauséogènes. C'est pourquoi l'optimisation graphique pour maintenir un frame rate stable est prioritaire sur l'augmentation de la complexité visuelle.

Champ de vision (Field of View, FOV)

Le champ de vision désigne l'étendue angulaire de l'environnement virtuel visible à travers le HMD, typiquement mesurée en degrés horizontaux. Le champ de vision humain naturel est d'environ 200-220 degrés horizontalement (incluant la vision périphérique extrême) et environ 135 degrés verticalement.

Figure 44

Field of view



Les HMD actuels offrent typiquement un champ de vision de 90-110 degrés (systèmes grand public) à 120-130 degrés (systèmes haut de gamme). Un champ de vision plus large favorise l'immersion en englobant la vision périphérique et en réduisant la conscience du cadre limitant du dispositif (effet "binoculaire").

Cependant, la relation entre champ de vision et cinétose est complexe et non linéaire. Un champ de vision large augmente le sentiment de présence mais augmente aussi la cinétose lors des mouvements virtuels, particulièrement lors de mouvements de translation ou de rotation. Cet effet s'explique par le rôle de la vision périphérique dans la détection du mouvement propre : le flux optique périphérique est un signal puissant de mouvement, et lorsque ce flux est présent visuellement mais non accompagné de stimulation vestibulaire correspondante, il crée un conflit sensoriel intense.

Cette relation complexe crée un compromis entre immersion et confort, qui peut être géré de plusieurs manières. Les techniques de réduction dynamique du champ de vision (*vignetting* ou "effet tunnel") réduisent temporairement le champ de vision périphérique lors

des mouvements virtuels, préservant ainsi le confort tout en maintenant un champ large lors de la vision statique. Des études (Fernandes & Feiner, 2016) ont montré que cette technique réduit significativement la cinétose sans affecter substantiellement le sentiment de présence.

Qualité optique et résolution

La qualité optique des lentilles et la résolution de l'affichage influencent également la cinétose, bien que de manière moins directe que la latence ou le frame rate.

Les aberrations chromatiques (où les différentes longueurs d'onde de la lumière sont réfractées différemment par les lentilles, créant des franges colorées) et la distorsion optique (courbure des lignes droites) créent une fatigue oculaire qui peut contribuer à l'inconfort général et éventuellement à la nausée. Les systèmes modernes utilisent des algorithmes de correction logicielle pour compenser ces aberrations optiques, appliquant une distorsion inverse à l'image rendue pour que l'image finale, après passage par les lentilles, soit correcte.

La résolution insuffisante crée un effet de "porte moustiquaire" (*screen door effect*) où les espaces entre pixels sont visibles, et rend difficile la lecture de textes ou la perception de détails fins. Ceci crée une fatigue oculaire par l'effort constant d'accommodation et peut contribuer à des maux de tête et à l'inconfort général. Les HMD modernes haut de gamme atteignent des résolutions de 2000x2000 pixels par œil ou plus, réduisant significativement ces problèmes.

Inadéquation Interpupillaire (IPD Mismatch)

La distance interpupillaire (IPD) est la distance entre les centres des pupilles des deux yeux, qui varie d'un individu à l'autre (typiquement 54-74 mm, avec une moyenne autour de 63-64 mm). Les HMD présentent deux lentilles séparées, et la distance entre ces lentilles doit idéalement correspondre à l'IPD de l'utilisateur pour que les yeux regardent à travers le centre optique des lentilles.

Une inadéquation IPD (où la distance entre les lentilles ne correspond pas à l'IPD de l'utilisateur) crée plusieurs problèmes. Elle peut causer une vision binoculaire inconfortable, avec un effort de vergence (convergence ou divergence des yeux) inapproprié. Elle peut créer une perception de l'échelle incorrecte dans l'environnement virtuel. Et elle peut causer une fatigue oculaire significative.

Les systèmes haut de gamme offrent des ajustements IPD manuels (ajustement physique de la distance entre les lentilles) ou logiciels. L'ajustement physique est préférable car il assure

que les yeux regardent à travers le centre optique des lentilles, minimisant les aberrations optiques.

Persistance de l'affichage

La persistance de l'affichage désigne la durée pendant laquelle chaque pixel reste allumé après son activation. Les écrans à cristaux liquides (LCD) traditionnels ont une persistance relativement élevée, ce qui peut créer un effet de flou de mouvement (*motion blur*) lors des mouvements rapides de la tête.

Les écrans OLED, utilisés dans de nombreux HMD modernes, ont une persistance très faible (les pixels s'allument et s'éteignent très rapidement), réduisant le *motion blur*. Certains systèmes utilisent également un mode d'affichage "*low persistence*" où chaque pixel n'est allumé que brièvement lors de chaque frame plutôt que de rester allumé en continu, réduisant encore le flou de mouvement.

La réduction de la persistance améliore la clarté des images pendant les mouvements de tête, ce qui réduit la charge cognitive de traitement visuel et peut contribuer à réduire la fatigue oculaire et la cinétose.

B.2.3 Solutions de conception pour atténuer la cinétose

Les recommandations de Fuchs et al. (2011) et les recherches subséquentes ont établi un ensemble de principes de conception visant à minimiser la cinétose tout en préservant l'immersion et la qualité de l'expérience. Ces principes peuvent être organisés en stratégies techniques et en stratégies de design d'interaction :

Optimisation de la latence et du frame rate

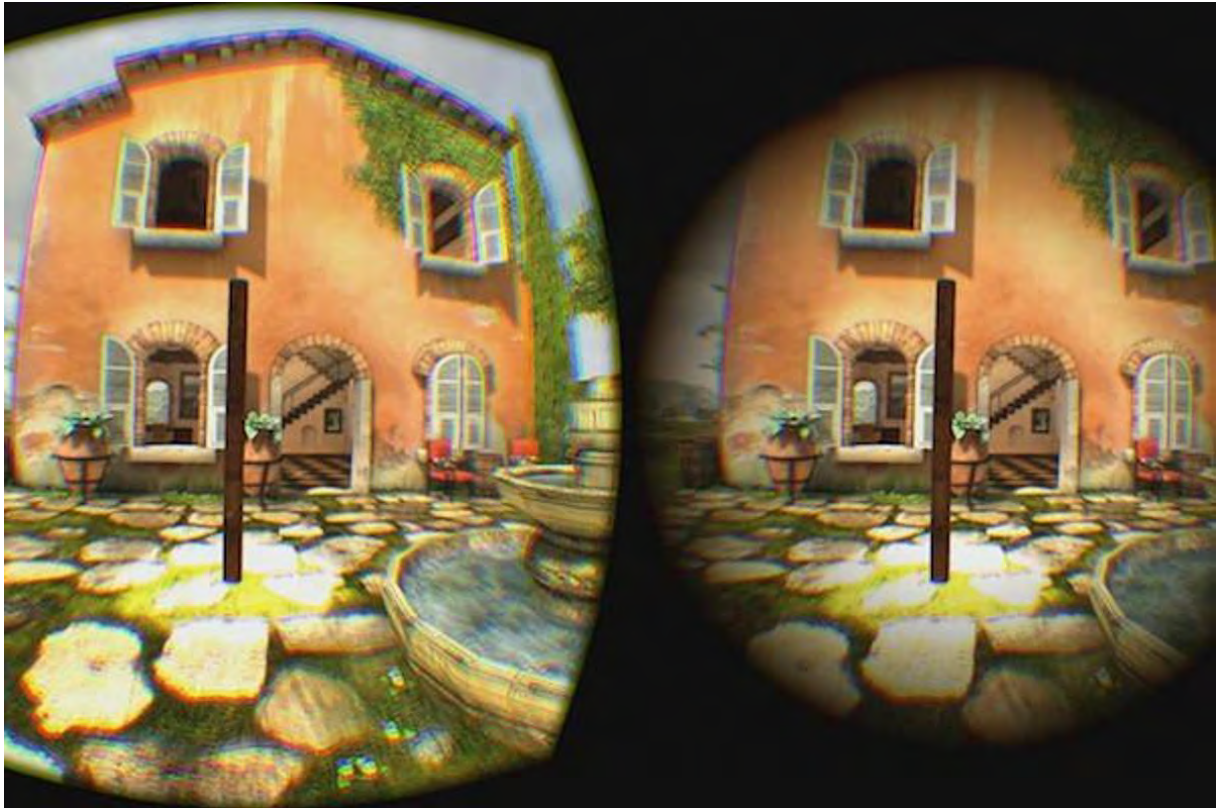
Comme discuté précédemment, maintenir une latence motion-to-photon inférieure à 20 millisecondes et un frame rate stable d'au moins 90 Hz constitue la base technique indispensable. Ceci nécessite une optimisation graphique rigoureuse, incluant :

- Utilisation de techniques de rendu optimisées (*level of detail* dynamique, *frustum culling*, *occlusion culling*)
- Limitation de la complexité graphique pour assurer un *frame rate* stable plutôt que de maximiser la fidélité visuelle au risque de frame drops
- Utilisation de techniques de rendu asynchrone (*timewarp*, *spacewarp*) pour compenser les variations de frame rate

- Priorisation du rendu de la zone fovéale centrale (*foveated rendering*) lorsque la technologie de *eye-tracking* est disponible

Figure 45

Réduction dynamique du champ de vision (Vignetting)



La technique du *vignetting* (fig. 45), également appelée "effet tunnel" ou "réduction dynamique du FOV", consiste à réduire progressivement le champ de vision périphérique lors des mouvements virtuels (translations ou rotations), en assombrissant ou en floutant graduellement les bords de l'image.

Fernandes et Feiner (2016) ont démontré l'efficacité de cette technique : la réduction du champ de vision de 110° à 60-90° lors des déplacements réduit significativement les symptômes de cinétose (réduction de 25-40% selon les mesures) sans affecter substantiellement le sentiment de présence ou la performance dans les tâches de navigation.

Le mécanisme explicatif repose sur le rôle du flux optique périphérique dans l'induction de la cinétose : en réduisant le champ de vision périphérique, on élimine le flux optique

périphérique (qui est le signal le plus puissant de mouvement propre) tout en préservant la vision centrale fonctionnelle pour la navigation et l'interaction.

L'implémentation optimale de cette technique nécessite plusieurs ajustements fins :

- La réduction doit être graduelle et progressive (sur 200-500 millisecondes) plutôt qu'abrupte, pour éviter un effet de distraction
- L'ampleur de la réduction doit être proportionnelle à la vitesse du mouvement virtuel : mouvements lents = réduction légère, mouvements rapides = réduction marquée
- La zone centrale préservée doit être suffisamment large (60-70° minimum) pour maintenir la fonctionnalité
- La transition visuelle (le bord entre zone claire et zone assombrie) doit être douce (gradient progressif) plutôt que nette

HMD partiellement occlusif

Figure 46

Exemple de HMD partiellement occlusif



réelle.

Le principe repose sur la résolution du conflit sensoriel : si la périphérie visuelle montre l'environnement physique réel (qui est stationnaire), le conflit entre vision centrale (environnement virtuel en mouvement) et système vestibulaire (immobilité physique) est partiellement résolu par la vision périphérique (environnement réel stationnaire).

Cette approche présente toutefois des compromis significatifs en termes d'immersion : la conscience persistante de l'environnement physique réduit le sentiment de présence dans l'environnement virtuel. Elle est donc plus appropriée pour des applications où le confort prime sur l'immersion maximale.

Point de référence stationnaire (*rest frame*)

Une approche alternative consiste à concevoir des HMD qui laissent intentionnellement entrevoir le monde réel en périphérie, plutôt que d'occlure complètement le champ visuel. Cette approche, explorée notamment dans les dispositifs de réalité augmentée mais applicable aussi à certaines configurations de RVI, fournit des indices visuels persistants de l'immobilité physique

L'inclusion d'un cadre de référence visuel stationnaire dans l'environnement virtuel constitue une stratégie efficace pour réduire la cinétose lors des mouvements virtuels. Ce cadre de référence peut prendre diverses formes :

- Cockpit d'un véhicule (dans les simulateurs de vol ou de conduite) : les éléments du cockpit restent visuellement stationnaires par rapport à l'utilisateur pendant que l'environnement extérieur défile
- Cadre de fenêtre ou portail : l'utilisateur voit l'environnement virtuel à travers un cadre fixe
- Éléments d'interface utilisateur persistants : un HUD (*head-up display*) ou des éléments de menu qui restent fixes dans le champ de vision

Le mécanisme explicatif repose sur la fourniture d'un ancrage visuel stable : même si l'environnement virtuel extérieur est en mouvement, les éléments du *rest frame* restent visuellement stationnaires, signalant au système visuel que l'observateur lui-même ne bouge pas. Ceci réduit l'intensité du signal visuel de mouvement propre et donc le conflit avec le système vestibulaire.

Cao et al. (2018) ont montré que la présence d'un *rest frame* réduit la cinétose lors de mouvements passifs (où l'utilisateur est déplacé dans l'environnement virtuel sans initier activement le mouvement). L'efficacité est maximale lorsque le *rest frame* occupe environ 30-40% du champ visuel central.

Sollicitation minimale du système vestibulaire

Fuchs et al. (2011) recommandent de concevoir les mouvements virtuels de manière à minimiser la sollicitation du système vestibulaire, c'est-à-dire à minimiser les signaux visuels qui, dans le monde réel, seraient typiquement associés à une stimulation vestibulaire intense.

- Éviter les accélérations et décélérations brusques : Les mouvements à vitesse constante génèrent moins de cinétose que les mouvements accélérés. Le système vestibulaire est particulièrement sensible aux accélérations (c'est un accéléromètre biologique), donc les mouvements virtuels à vitesse constante créent moins de conflit sensoriel.
- Éviter les rotations non contrôlées : Les rotations de la vue (particulièrement autour de l'axe de roulis - *roll*) sont extrêmement nauséogènes car elles correspondent, dans le monde réel, à des mouvements vestibulaires intenses. Les rotations de la caméra

virtuelle devraient être exclusivement contrôlées par les mouvements réels de la tête de l'utilisateur, jamais imposées par le système.

- Privilégier les mouvements horizontaux : Les mouvements verticaux (ascenseurs, chutes) sont plus nauséogènes que les mouvements horizontaux car ils stimuleraient fortement les organes otolithiques si le mouvement était réel. Lorsque des mouvements verticaux sont nécessaires, ils devraient être lents et accompagnés d'un rest frame.

Cohérence multisensorielle

Assurer la cohérence entre les différentes modalités sensorielles réduit les conflits sensoriels et la cinétose :

- Retour audio spatialisé : Le son doit être spatialisé en 3D de manière cohérente avec l'environnement visuel. Lorsque l'utilisateur tourne la tête, les sources sonores doivent rester localisées correctement dans l'espace virtuel.
- Retour haptique synchronisé : Lorsque des interactions impliquent un contact (saisir un objet, toucher une surface), un retour haptique synchronisé temporellement avec l'événement visuel renforce la cohérence multisensorielle.
- *Wind Effect* : Dans les situations de mouvement rapide (vol, conduite rapide), un retour de vent (ventilateur dirigé vers le visage) synchronisé avec la vitesse visuelle peut réduire la cinétose en fournissant une stimulation somatosensorielle cohérente avec le mouvement visuel.

Stratégies de design d'interaction

Au-delà des aspects techniques, les choix de design d'interaction influencent significativement la cinétose comme par exemple le contrôle utilisateur vs mouvement passif.

Un principe fondamental est que les mouvements initiés et contrôlés par l'utilisateur génèrent moins de cinétose que les mouvements passifs (où l'utilisateur est déplacé sans contrôle). Ce phénomène s'explique par le rôle des prédictions motrices : lorsque nous initions un mouvement, notre système moteur génère une copie de la commande motrice (*efference copy*) qui prédit les conséquences sensorielles attendues. Cette prédiction permet d'anticiper et de "précompenser" partiellement le conflit sensoriel.

Les implications pratiques sont :

- Privilégier les modes de déplacement actifs (marcher physiquement dans le *roomscale*, téléportation initiée par l'utilisateur) plutôt que les déplacements automatiques ou scriptés
- Lorsque des déplacements passifs sont nécessaires (cinématiques, transports virtuels), les rendre aussi courts et prévisibles que possible
- Fournir un sentiment de contrôle même lors de déplacements partiellement automatisés (par exemple, permettre à l'utilisateur d'ajuster la vitesse ou d'interrompre le déplacement)

Acclimatation et exposition graduelle

L'exposition répétée à la RV produit typiquement une habitude, avec une réduction progressive des symptômes de cinétose sur plusieurs sessions. Cette adaptation résulte probablement d'une recalibration du modèle interne pour accommoder les patterns sensoriels spécifiques de la RV.

Pour faciliter cette acclimatation, il est recommandé de :

- Commencer par des sessions courtes (10-15 minutes) lors des premières expositions, et augmenter progressivement la durée
- Débuter par des expériences à faible potentiel nauséogène (mouvements lents, contrôle utilisateur complet) avant de progresser vers des expériences plus intenses
- Encourager les utilisateurs à interrompre immédiatement l'expérience dès l'apparition des premiers symptômes de cinétose, plutôt que de "persévérer" (la persévération peut aggraver les symptômes et créer un conditionnement aversif)
- Prévoir des périodes de récupération adéquates entre les sessions (au moins 15-30 minutes pour les utilisateurs sensibles)

Recommandations de l'ANSES (2021)

L'Agence Nationale de Sécurité Sanitaire de l'alimentation, de l'environnement et du travail (ANSES, 2021) a publié des recommandations spécifiques concernant l'utilisation de la RV incluant des précautions relatives à la cinétose :

Mise en garde des utilisateurs : Les fabricants et développeurs doivent informer clairement les utilisateurs des risques potentiels de cinétose, de ses symptômes (pour permettre

une reconnaissance précoce), et des comportements recommandés (interruption immédiate dès l'apparition des symptômes).

Durée limitée des sessions : Pour les utilisateurs novices ou sensibles, limiter les sessions initiales à 20-30 minutes maximum, avec une surveillance de l'apparition de symptômes.

Périodes de pause et réadaptation : Après le retrait du casque, prévoir une période de transition (5-10 minutes) avant de reprendre des activités nécessitant concentration ou coordination (conduite automobile, opération de machinerie), car des effets résiduels (perturbation de l'équilibre, persistance de l'adaptation oculomotrice) peuvent persister brièvement.

Contre-indications temporaires : Éviter l'utilisation de la RV en cas de fatigue importante, de maladie (particulièrement affections vestibulaires ou troubles de l'équilibre), de consommation d'alcool ou de substances psychoactives, ou de prise de médicaments affectant l'équilibre ou la vision.

Populations vulnérables : Précaution particulière pour les enfants (dont le système vestibulaire est encore en développement), les personnes âgées (dont le système d'équilibre peut être plus fragile), et les personnes ayant des antécédents de troubles vestibulaires, d'épilepsie photosensible, ou de migraines.

Différences de genre et facteurs individuels

Munafo et al. (2017) ont conduit une méta-analyse de 28 études examinant les différences de genre dans la susceptibilité à la cinétose. Leurs résultats montrent que les femmes présentent statistiquement des scores de cinétose significativement plus élevés que les hommes, avec une taille d'effet petite à moyenne (d de Cohen $\approx 0,3-0,5$ selon les mesures).

Plusieurs facteurs explicatifs ont été proposés :

Facteurs hormonaux : Les fluctuations hormonales du cycle menstruel influencent la susceptibilité à la cinétose, avec une sensibilité accrue pendant la phase lutéale et autour des menstruations.

Facteurs anatomiques : Les différences dans le centre de gravité et dans la stratégie de contrôle postural entre hommes et femmes pourraient contribuer à une susceptibilité différentielle selon la théorie de l'instabilité posturale.

Facteurs perceptifs : Certaines études suggèrent des différences dans la pondération relative accordée aux différents canaux sensoriels dans le contrôle postural et l'orientation spatiale, les femmes s'appuyant relativement plus sur les informations visuelles.

Au-delà du genre, d'autres facteurs individuels modulent la susceptibilité :

- Historique de mal des transports (prédicteur fort de susceptibilité au *cybersickness*)
- Âge (les enfants et adolescents tendent à être plus susceptibles que les adultes jeunes ; les personnes âgées peuvent également présenter une susceptibilité accrue)
- Ethnicité (certaines études suggèrent des variations, potentiellement liées à des différences dans les distributions d'IPD ou dans l'expérience culturelle du mouvement)
- Traits de personnalité (tendance à l'anxiété associée à une susceptibilité accrue)

Ces différences individuelles soulignent l'importance de concevoir des expériences de RVI avec des options d'ajustement permettant à chaque utilisateur d'optimiser son confort (ajustements de vitesse de mouvement, intensité du *vignetting*, etc.).

B.2.4 Modes de déplacement en environnement virtuel

Le choix du mode de déplacement constitue l'un des facteurs les plus critiques pour le confort de l'utilisateur en réalité virtuelle, car c'est lors des déplacements que le conflit visuo-vestibulaire est le plus intense. Différentes approches ont été développées, chacune présentant des compromis spécifiques entre confort, sentiment de présence, et fonctionnalité.

Roomscale restreint (physical locomotion)

Le mode *roomscale* restreint permet à l'utilisateur de se déplacer physiquement dans un espace limité (typiquement 2m × 2m à 4m × 4m), ses mouvements étant traqués en temps réel par des capteurs externes (stations de base) ou par les caméras intégrées du HMD (*inside-out tracking*).

Avantages : Ce mode offre la congruence sensorielle maximale puisque les mouvements virtuels correspondent exactement aux mouvements physiques réels : le système vestibulaire détecte le même mouvement que le système visuel. En conséquence, ce mode génère la cinétose minimale (quasi nulle pour les déplacements purement physiques) et maximise le sentiment de présence et d'incarnation.

Limites : L'espace physique disponible est typiquement très limité (quelques mètres carrés), ce qui restreint sévèrement les possibilités de navigation dans des environnements

virtuels de grande échelle. Les utilisateurs peuvent également se heurter aux limites physiques de l'espace ou aux meubles, nécessitant des systèmes de gardiennage (*guardian ou chaperone systems*) qui affichent des murs virtuels lorsque l'utilisateur approche des limites physiques.

Applications optimales : Expériences conçues pour se dérouler entièrement dans un espace restreint (salle unique, poste de travail), ou en combinaison avec d'autres modes de déplacement pour les mouvements de grande échelle.

Téléportation

La téléportation permet à l'utilisateur de pointer vers une destination et d'être instantanément transféré à cet emplacement, sans traverser visuellement l'espace intermédiaire. L'utilisateur vise typiquement avec un rayon (*ray*) émis depuis le contrôleur, et un marqueur de destination apparaît à l'endroit visé.

Avantages : La téléportation élimine complètement le flux optique continu qui génère la cinétose. En l'absence de mouvement visuel, il n'y a pas de conflit visuo-vestibulaire. Cette technique est donc extrêmement confortable même pour les utilisateurs très sensibles.

Limites : La téléportation peut réduire le sentiment de présence spatiale et la continuité de l'expérience. Les utilisateurs peuvent avoir des difficultés à maintenir une représentation spatiale cohérente de l'environnement lorsqu'ils "sautent" d'un endroit à l'autre. De plus, la téléportation peut permettre des stratégies d'évitement irréalistes dans les jeux (téléportation à travers des obstacles ou derrière des ennemis).

Applications optimales : Applications d'exploration où la cinétose est une préoccupation majeure, expériences destinées à des utilisateurs novices ou sensibles, environnements où le réalisme de la locomotion n'est pas critique.

Téléportation avec Eyelid (Descroix, 2021)

Figure 47

Eyelid



La téléportation avec *eyelid* (fig. 47), documentée par Descroix (2021), constitue une variante améliorée de la téléportation classique. Lors du déplacement, un fondu au noir rapide (typiquement 150-300 millisecondes) simule un clignement d'yeux, masquant le changement de position instantané.

Avantages : Cette technique préserve les avantages de confort de la téléportation tout en améliorant la continuité temporelle perçue. Le clignement virtuel imite un processus naturel et fournit une explication perceptible pour la discontinuité spatiale, rendant l'expérience moins *jarring*. Le mécanisme imite également la suppression saccadique qui se produit naturellement lors des mouvements oculaires rapides ou des clignements réels, période durant laquelle le cerveau supprime partiellement le traitement visuel.

Implémentation : La durée du fondu doit être soigneusement calibrée : trop court (<100ms) et l'effet est à peine perceptible ; trop long (>500ms) et l'expérience devient frustrante. Descroix recommande une durée de 150-250ms, avec un fondu symétrique (*fade out* et *fade in* de durée égale).

Applications Optimales : Toute application utilisant la téléportation peut bénéficier de cette amélioration. Particulièrement utile dans les expériences narratives où la continuité temporelle est importante.

Locomotion artificielle continue (artificial continuous locomotion)

La locomotion artificielle continue, également appelée "*smooth locomotion*" ou "*joystick locomotion*", permet à l'utilisateur de se déplacer de manière continue dans l'environnement virtuel en utilisant un joystick ou un touchpad sur les contrôleurs. La direction du déplacement est typiquement déterminée par la direction du regard ou du contrôleur, et la vitesse par l'amplitude de l'inclinaison du joystick.

Avantages : Ce mode offre une liberté de mouvement maximale et une continuité spatiale complète. Il permet une navigation fluide et intuitive dans des environnements de grande échelle. Pour les utilisateurs habitués aux jeux vidéo traditionnels, le schéma de contrôle est familier.

Limites : C'est le mode le plus générateur de cinétose car il crée un conflit visuo-vestibulaire maximal : le système visuel détecte un mouvement continu et souvent rapide (flux optique) tandis que le système vestibulaire ne détecte aucun mouvement physique correspondant. De nombreux utilisateurs, particulièrement les novices, trouvent ce mode inconfortable voire intolérable.

Optimisations : Plusieurs techniques peuvent atténuer la cinétose en locomotion artificielle continue :

- Réduction dynamique du champ de vision (*vignetting*) proportionnelle à la vitesse de déplacement
- Limitation de la vitesse maximale
- Accélération et décélération progressives plutôt qu'instantanées
- Direction du mouvement alignée avec le regard (plutôt qu'indépendante) pour que le flux optique soit centré devant l'utilisateur plutôt que périphérique

Applications optimales : Jeux d'action ou d'exploration destinés à des utilisateurs expérimentés, applications où la fluidité et la rapidité de navigation sont prioritaires.

Walk-in-Place

Le *walk-in-place* est une technique où l'utilisateur marche sur place (sans se déplacer réellement dans l'espace physique), et le système détecte ce mouvement (via les contrôleurs tenus en main qui oscillent, ou via l'accéléromètre du HMD qui détecte le mouvement de la tête) pour produire un déplacement virtuel proportionnel.

Avantages : Ce mode fournit un retour proprioceptif et kinesthésique (le sentiment de marcher, les oscillations du corps) qui correspond partiellement au mouvement visuel, réduisant ainsi le conflit sensoriel par rapport à la locomotion purement artificielle. Il permet également un engagement physique qui peut être bénéfique pour l'immersion.

Limites : Marcher sur place pendant des périodes prolongées peut être physiquement fatigant. De plus, la correspondance entre le mouvement physique et le mouvement virtuel n'est qu'approximative (marcher sur place produit des sensations vestibulaires différentes de la marche réelle avec translation), donc le conflit sensoriel n'est que partiellement résolu.

Applications optimales : Expériences d'exploration à rythme modéré, applications d'exercice physique, situations où l'engagement corporel est souhaitable.

Redirected Walking

Le *redirected walking* est une technique sophistiquée qui manipule subtilement la perception de l'utilisateur pour lui faire parcourir physiquement un chemin différent de celui qu'il perçoit virtuellement, permettant ainsi de naviguer dans un environnement virtuel de grande échelle tout en restant dans un espace physique restreint.

La technique exploite les seuils de détection : de petites rotations ou courbures du chemin virtuel (en dessous du seuil de détection consciente) sont introduites, de sorte que l'utilisateur marche réellement en cercle ou en pattern complexe dans l'espace physique tout en ayant l'impression de marcher en ligne droite dans l'environnement virtuel.

Avantages : Lorsqu'il fonctionne bien, le *redirected walking* offre les avantages de la locomotion physique réelle (congruence sensorielle maximale, cinétose minimale) tout en permettant l'exploration d'espaces virtuels arbitrairement grands.

Limites : La technique nécessite un espace physique d'au moins 6m × 6m pour être efficace. Elle fonctionne mieux lorsque l'environnement virtuel contient des distracteurs visuels (murs, obstacles) qui justifient naturellement les changements de direction. Les gains de

redirection trop agressifs peuvent être détectés consciemment et briser l'immersion. Enfin, la technique est computationnellement complexe et nécessite des algorithmes sophistiqués de planification de trajectoire.

Applications optimales : Installations permanentes dans des espaces dédiés (arcades VR, laboratoires de recherche), expériences d'exploration architecturale ou environnementale.

Modes hybrides

De nombreuses applications combinent plusieurs modes de déplacement pour optimiser le compromis confort-fonctionnalité :

- *Roomscale* pour les mouvements locaux + téléportation pour les déplacements de grande échelle
- Téléportation pour les déplacements longue distance + locomotion continue pour les ajustements fins de position
- *Walk-in-place* pour la navigation exploratoire + téléportation pour les déplacements rapides intentionnels

B.3 Principes d'interfaçage en RV

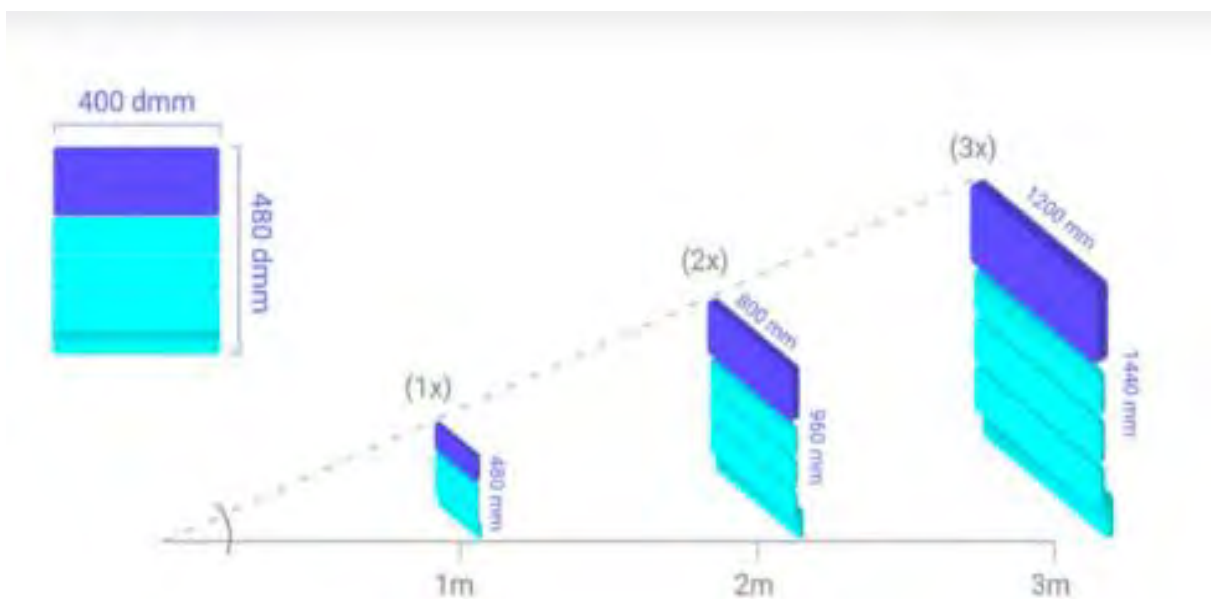
La conception d'interfaces utilisateur en RV nécessite des principes ergonomiques spécifiques, radicalement différents de ceux des interfaces bidimensionnelles traditionnelles (écrans d'ordinateur, smartphones). En environnement virtuel tridimensionnel, la distance, l'angle de vision, la profondeur, et la taille physique (plutôt que la taille en pixels) des éléments interactifs jouent un rôle critique pour le confort visuel, la lisibilité, et l'utilisabilité. De plus, les modalités d'interaction (pointage spatial, saisie gestuelle) diffèrent fondamentalement des interactions 2D (souris, tactile), nécessitant des ajustements dans le dimensionnement et le positionnement des zones interactives.

B.3.1 Distance Matters Measurement (DMM) : Échelle et distances de confort

Le système *Distance Matters Measurement* (DMM), développé dans le contexte du design d'interfaces pour la RV, fournit un cadre standardisé pour dimensionner les éléments d'interface en fonction de leur distance de l'utilisateur. Ce système repose sur le principe que la taille apparente (angle visuel) d'un objet est déterminée par la combinaison de sa taille physique et de sa distance, selon la géométrie de la vision.

Figure 48

DMM



Définition de l'unité DMM

L'unité de base du système DMM est définie comme suit :

1 DMM = 1 millimètre à 1 mètre de distance

Plus précisément, un objet de 1mm de hauteur, situé à 1m de l'observateur sous-tend un angle visuel d'environ 0.057 degrés, soit approximativement 1 milliradian. Cette relation découle de la géométrie de base : pour de petits angles, angle (en radians) \approx taille / distance.

Implications pratiques : Un élément dimensionné à 40 DMM aura les propriétés suivantes selon sa distance :

- À 1 mètre : 40 millimètres (4 centimètres) de hauteur physique

- À 2 mètres : 80 millimètres (8 centimètres) de hauteur physique pour le même angle visuel
- À 5 mètres : 200 millimètres (20 centimètres) de hauteur physique pour le même angle visuel

Le système DMM permet donc de spécifier des tailles d'éléments d'interface en termes d'angle visuel constant, indépendamment de la distance à laquelle ils seront placés, simplifiant ainsi considérablement la conception d'interfaces spatiales.

Zones de confort visuel

La recherche en ergonomie visuelle et en optique physiologique a établi des zones de distance optimales pour le positionnement des éléments d'interface en réalité virtuelle, basées sur les contraintes du système visuel humain.

Zone interdite : 0m - 2m (distance minimale)

Les éléments interactifs ne devraient pas être positionnés à moins de 2 mètres de distance virtuelle de l'utilisateur.

Cette restriction repose sur plusieurs facteurs physiologiques :

Accommodation excessive : L'accommodation désigne le processus par lequel le cristallin de l'œil change de forme pour focaliser sur des objets à différentes distances. Pour focaliser sur un objet proche, le cristallin doit devenir plus convexe, ce qui nécessite une contraction du muscle ciliaire. Une accommodation soutenue sur des objets proches (< 2m) est fatigante et peut conduire à une fatigue visuelle, des maux de tête, et à long terme potentiellement à une myopie induite.

Vergence excessive : La vergence désigne le mouvement coordonné des deux yeux pour converger vers un point focal commun. Pour des objets très proches, les yeux doivent converger fortement (tourner vers l'intérieur), ce qui est également fatigant pour les muscles oculomoteurs.

Conflit accommodation-vergence : En RV, les écrans des HMD sont physiquement situés à environ 5-7 centimètres des yeux mais, grâce aux lentilles optiques, l'image est focalisée à l'infini optique (typiquement 1.5-2 mètres de distance équivalente). Ceci crée une situation inhabituelle où l'accommodation (qui reste fixée à la distance focale des lentilles) et la vergence (qui suit la distance virtuelle des objets) sont dissociées. Dans la vision naturelle, accommodation et vergence sont couplées (*accommodation-vergence coupling*). Ce découplage

peut générer un inconfort, particulièrement lorsque des objets virtuels sont placés très près ou très loin de la distance focale des lentilles.

Exception : Pour des éléments d'interface critiques nécessitant une interaction immédiate (par exemple, un menu contextuel attaché au contrôleur tenu en main), un positionnement à 0.5-1 mètre peut être acceptable pour de courtes durées, mais ne devrait pas être maintenu.

Zone acceptable : 2m - 5m (distance utilisable)

Les éléments peuvent être placés dans cette zone, particulièrement pour des interactions secondaires ou des informations de contexte. À ces distances, l'accommodation et la vergence restent dans des plages confortables, bien que non optimales.

Considérations : Les éléments placés dans cette zone doivent être légèrement agrandis par rapport à ceux placés dans la zone optimale pour maintenir la même lisibilité apparente. Par exemple, un texte destiné à être lu à 2 mètres devrait être dimensionné à environ 48-50 DMM pour offrir une lisibilité équivalente à du texte de 24 DMM à 1 mètre (en tenant compte des non-linéarités de la perception de taille).

Zone optimale : 5m - 10m (distance idéale)

Cette zone représente le *sweet spot* pour le positionnement des éléments d'interface principaux, des menus, et des contenus textuels. À ces distances :

Accommodation minimale : L'œil est proche de son état de repos (accommodation pour l'infini), minimisant l'effort du muscle ciliaire. La plupart des HMD ont leur plan focal optique dans cette zone (typiquement équivalent à 1.5-3 mètres, certains systèmes plus avancés à 5+ mètres).

Vergence confortable : La vergence requise est minimale, les axes visuels des deux yeux étant presque parallèles, ce qui est l'état de repos des muscles oculomoteurs.

Résolution apparente : À ces distances, la résolution angulaire des HMD actuels (typiquement 10-20 pixels par degré) est suffisante pour afficher du texte lisible et des détails d'interface fins sans artefacts de pixellisation trop visibles.

Conflit accommodation-vergence minimal : Le conflit entre accommodation (fixée par les lentilles du HMD) et vergence (suivant la distance virtuelle) est minimal dans cette zone de distance.

Zone éloignée : >10m (distance périphérique)

Les éléments placés à plus de 10 mètres peuvent être utilisés pour des informations non critiques, des éléments de décor, ou des indicateurs directionnels. À ces distances :

Taille requise : Les éléments doivent être significativement agrandis pour rester visibles et lisibles. Par exemple, un texte à 20 mètres nécessiterait une taille de l'ordre de 200-300 DMM pour être confortablement lisible.

Résolution limitée : La résolution angulaire limitée des HMD actuels fait que les détails fins deviennent difficiles à percevoir au-delà de 10-15 mètres. Cette limitation est analogue au niveau de détail (LOD) en graphisme 3D et suggère que les éléments d'interface distants devraient éviter les détails fins et privilégier des formes simples et des contrastes élevés.

Usage approprié : Cette zone est appropriée pour des indicateurs de navigation (*waypoints*, flèches directionnelles), des éléments de HUD (*head-up display*) informationnels, ou des panneaux de grande taille (équivalents à des panneaux d'affichage dans le monde réel).

Considérations sur la profondeur et la parallaxe Bbinoculaire

En vision stéréoscopique, la perception de la profondeur repose en partie sur la disparité binoculaire (les légères différences entre les images perçues par l'œil gauche et l'œil droit). Les HMD de réalité virtuelle exploitent ce mécanisme en affichant des images légèrement différentes à chaque œil, correspondant aux perspectives que ces yeux auraient dans le monde réel.

Disparité binoculaire excessive : Pour des objets très proches (< 1m), la disparité binoculaire devient très importante, ce qui peut être difficile à fusionner perceptivement pour certains utilisateurs, conduisant à une diplopie (vision double) ou à un effort de fusion inconfortable.

Limite de fusion stéréoscopique : Au-delà d'environ 20-30 mètres, la disparité binoculaire devient si faible qu'elle ne contribue plus significativement à la perception de la profondeur. À ces distances, d'autres indices de profondeur (perspective, occultation, taille relative, parallaxe de mouvement) deviennent dominants.

Implications pour le design : Les éléments d'interface placés très près ou très loin devraient éviter de s'appuyer exclusivement sur la disparité binoculaire pour communiquer des

informations spatiales importantes. Des indices de profondeur redondants (ombres, perspective, taille relative) devraient être fournis.

B.3.2 Dimensionnement des éléments et polices

Le dimensionnement approprié des éléments textuels et graphiques est crucial pour la lisibilité et l'utilisabilité des interfaces en réalité virtuelle (fig. 49). Contrairement aux interfaces 2D où la taille est spécifiée en pixels ou en points, les interfaces VR nécessitent une spécification en unités physiques ou angulaires (comme les DMM) qui tiennent compte de la distance de visualisation.

Figure 49

Dimensionnement des éléments et polices



Hiérarchie typographique recommandée

Une hiérarchie typographique claire facilite la navigation visuelle et la compréhension de la structure informationnelle (voir tableau B.3.2). Les tailles recommandées suivantes sont spécifiées en DMM pour des éléments placés dans la zone optimale (5-10 mètres), et peuvent être mises à l'échelle proportionnellement pour d'autres distances.

Tableau 8
Hierarchie typographique pour interfaces RV (Tailles en DMM)

Niveau typographique	Taille (DMM)	Usage recommandé	Remarques
<i>Headline Regular</i>	40 dmm	Titres principaux, en-têtes de section, information critique nécessitant attention immédiate	Devrait être utilisé avec parcimonie pour maintenir la hiérarchie visuelle
<i>Title Medium</i>	32 dmm	Sous-titres, titres de menus principaux, catégories majeures	Poids de police medium (semi-bold) pour distinction sans excès d'emphase
<i>Subheading Regular</i>	28 dmm	Sections secondaires, divisions de contenu, labels de groupes d'éléments	Poids régulier, peut utiliser une couleur secondaire pour différenciation
<i>Body 2 Medium</i>	24 dmm	Texte de corps nécessitant une emphase légère, items importants dans des listes	<i>Medium weight</i> se distingue légèrement du body text régulier
<i>Body 1 Regular</i>	24 dmm	Texte de corps standard, contenu principal de paragraphes, descriptions	Taille de base pour lisibilité confortable du texte continu
<i>Caption Regular</i>	20 dmm	Légendes d'images, annotations, notes de bas de page, métadonnées	Limite inférieure de lisibilité confortable ; éviter pour texte long
<i>Button Medium</i>	24 dmm	Libellés de boutons et éléments interactifs	<i>Medium weight</i> améliore la lisibilité sur surfaces colorées

Principes d'application de la hiérarchie typographique

Ratio de taille minimal : Pour assurer une distinction perceptive claire entre les niveaux hiérarchiques, maintenir un ratio minimum de 1.25 entre les niveaux adjacents. Par exemple, si le *Body text* est à 24 DMM, le niveau immédiatement supérieur (*Subheading*) devrait être au

moins à 30 DMM. Cette règle est généralement respectée dans le tableau ci-dessus, avec des ratios variant de 1.14 à 1.33 entre niveaux adjacents.

Limitation des niveaux : Éviter d'utiliser plus de 3-4 niveaux hiérarchiques dans une même vue ou un même écran. Une hiérarchie trop complexe crée une confusion visuelle et dilue l'efficacité de la différenciation des niveaux. Si plus de niveaux sont nécessaires, envisager d'utiliser d'autres dimensions de variation (couleur, espacement, placement spatial) en complément de la taille.

Consistance : Maintenir une utilisation consistante de chaque niveau à travers l'interface. Par exemple, tous les titres principaux devraient utiliser *Headline Regular*, créant ainsi des patterns visuels prévisibles qui facilitent la navigation et la compréhension.

Considérations de lisibilité

Au-delà de la taille, plusieurs facteurs influencent la lisibilité du texte en RV :

Contraste : Le ratio de contraste entre le texte et l'arrière-plan doit respecter les standards d'accessibilité. Pour du texte de corps (Body 1 Regular, 24 DMM), un ratio de contraste minimum de 4.5:1 est recommandé (standard WCAG AA). Pour du texte large (>28 DMM ou bold >24 DMM), un ratio de 3:1 est acceptable. En environnement VR, où la résolution effective est limitée par la densité de pixels du HMD, il est prudent de viser des ratios plus élevés (7:1 pour texte de corps, 4.5:1 pour texte large) lorsque c'est possible.

Choix de police : Les polices sans-serif (sans empattement) sont généralement préférables en VR en raison de leur lisibilité à basse résolution. Les empattements (serifs) peuvent créer des artefacts de pixellisation dérangeants. Polices recommandées : Roboto, Open Sans, Lato, Source Sans Pro. Éviter les polices condensées ou ultra-light qui perdent en lisibilité. Pour les interfaces exigeant une lisibilité maximale (applications médicales, professionnelles critiques), des polices spécifiquement optimisées pour l'affichage à basse résolution peuvent être utilisées.

Interligne (*Line Spacing*) : Prévoir un interligne généreux pour le texte continu, typiquement 1.5 à 1.8 fois la hauteur de police. Par exemple, pour du Body 1 Regular à 24 DMM, l'interligne devrait être de 36-43 DMM. Un interligne insuffisant crée une impression de densité difficile à lire, tandis qu'un interligne excessif brise la cohérence visuelle des paragraphes.

Longueur de ligne : Limiter la longueur des lignes de texte à 50-75 caractères (10-15 mots) pour le confort de lecture. Des lignes trop longues nécessitent de longs mouvements oculaires horizontaux et rendent difficile le repérage du début de la ligne suivante. Des lignes trop courtes créent une lecture hachée avec des retours à la ligne trop fréquents. En VR, où le texte est lu à distance et où l'espace horizontal n'est pas aussi contraint que sur un écran de smartphone, il est tentant de créer de très longues lignes, mais ceci doit être évité.

Anti-aliasing : Le texte doit impérativement utiliser de l'anti-aliasing (lissage) pour réduire les effets d'escalier (*aliasing*) sur les contours des lettres. Les techniques de *signed distance field* (SDF) rendering sont particulièrement efficaces pour le texte en VR car elles maintiennent des contours nets même lors de changements de taille ou de distance.

Couleur et lisibilité : Les combinaisons de couleurs doivent être choisies avec soin. Le noir sur blanc (ou blanc sur noir) offre le contraste maximum mais peut créer un éblouissement (*glare*) désagréable. Des combinaisons légèrement atténuées (par exemple, texte gris foncé #333333 sur fond blanc #FFFFFF, ou texte blanc sur fond gris sombre #222222) réduisent l'éblouissement tout en maintenant un excellent contraste. Éviter les combinaisons de couleurs pures saturées (rouge sur bleu, par exemple) qui créent un effet de "vibration chromatique" désagréable.

B.3.3 Zones d'interaction et hit zones

Les zones cliquables (*hit zones ou touch targets* – fig. 50) désignent les aires des éléments interactifs sur lesquelles l'utilisateur peut pointer et déclencher une action.

Figure 50

Text and Hit size



En RV, où l'interaction se fait via un pointage spatial avec les contrôleurs (ray casting) plutôt que via une souris précise ou un doigt sur un écran tactile, le dimensionnement approprié des hit zones est encore plus critique qu'en interface 2D.

Dimensionnement des hit zones

Le dimensionnement des zones interactives (tableau B.3.3) doit prendre en compte la précision limitée du pointage spatial en RV. Contrairement à une souris d'ordinateur (précision sub-millimétrique sur l'écran) ou à un doigt sur un écran tactile (précision de quelques millimètres), le pointage avec un contrôleur RV présente une variabilité plus importante due à :

- Microtremblements naturels de la main (tremor physiologique)
- Fatigue du bras lors du maintien prolongé d'une pose de pointage
- Précision limitée du *tracking* des contrôleurs (typiquement 1-2mm de précision de position, mais cette précision se traduit en imprécision angulaire amplifiée par la distance)
- Absence de *feedback* tactile au moment du contact (contrairement au tactile) rendant difficile la détection fine du moment où le pointeur entre dans la zone cible

Tableau 9

Dimensionnement des hit zones en RV

Catégorie	Dimensions minimales	Padding	Usage recommandé	Notes additionnelles
<i>Minimum</i>	64 × 64 mm	16 dmm (sur tous les côtés)	Éléments secondaires, interfaces denses où l'espace est limité, utilisateurs expérimentés	À la limite inférieure d'utilisabilité ; les erreurs de sélection sont fréquentes (~15-20% miss rate)
<i>Confortable</i>	96 × 96 mm	16 dmm	Éléments principaux, interactions fréquentes, interface standard	Taille optimale pour équilibre entre utilisabilité et efficacité

Catégorie	Dimensions minimales	Padding	Usage recommandé	Notes additionnelles
<i>Large</i>	128 × 128 mm	24 dmm	Cibles critiques, boutons de confirmation/annulation, interfaces pour utilisateurs novices ou à mobilité réduite	<p>spatiale ; miss rate <5%</p> <p>Maximise la probabilité de sélection réussie ; miss rate <2% ; peut être perçu comme excessivement grand dans certains contextes</p>
<i>Extra-Large</i>	160 × 160 mm ou plus	32 dmm	Cibles primaires dans des interfaces simplifiées, accessibilité maximale	<p>Réservé aux cas où la facilité d'interaction prime absolument sur l'efficacité spatiale</p>

Principes d'espacement et de *layout*

L'espacement entre éléments interactifs est aussi important que leur taille individuelle pour prévenir les erreurs de sélection.

Espacement minimum : Maintenir au moins 32 DMM (3.2 cm à 1 mètre) entre les bords externes de zones interactives adjacentes. Cet espacement permet de créer une "zone morte" entre cibles, réduisant la probabilité de déclencher accidentellement l'élément adjacent lors d'un pointage imprécis.

Espacement recommandé : Un espacement de 48-64 DMM entre zones interactives est confortable et réduit significativement les erreurs. Cet espacement peut être réalisé soit par de l'espace vide, soit par des éléments non-interactifs (texte descriptif, séparateurs visuels).

Groupement spatial : Regrouper visuellement et spatialement les éléments fonctionnellement liés (par exemple, les boutons d'un même groupe de contrôle) tout en les séparant clairement des autres groupes. Cette organisation réduit la charge cognitive et facilite l'anticipation de l'action.

Éviter les grilles denses : Les *layouts* en grille dense (par exemple, des grilles 3×3 ou 4×4 de petits boutons) sont difficiles à utiliser en VR en raison de la précision limitée du pointage. Privilégier des *layouts* plus aérés, potentiellement répartis en profondeur (utilisant l'espace 3D) plutôt que compressés sur un plan 2D.

Feedback visuel des états interactifs

Étant donné l'absence de feedback tactile en RV (à moins de disposer de contrôleurs haptiques avancés), le feedback visuel devient le canal principal pour communiquer les états interactifs.

État de repos (*Default State*) : L'élément interactif doit être visuellement distinguable des éléments non-interactifs. Ceci peut être réalisé par :

- Une affordance visuelle claire (forme de bouton avec relief, ombre portée)
- Une couleur distinctive (typiquement une couleur d'accent contrastant avec les éléments statiques)
- Un léger effet de matériau (légère brillance, texture spécifique)

État de survol (*Hover State*) : Lorsque le rayon de pointage entre dans la *hit zone*, un *feedback* visuel immédiat doit indiquer que l'élément est pointé. Options efficaces :

- Changement de couleur (légère intensification ou changement de teinte)
- Émission lumineuse (*glow*) subtile autour de l'élément
- Légère animation d'échelle (grossissement de 5-10%)
- Apparition d'un contour ou d'une surbrillance
- Changement de curseur à l'extrémité du rayon de pointage

L'état de survol doit être suffisamment distinct pour être immédiatement perceptible, mais pas si intense qu'il devienne distrayant. La transition vers l'état de survol devrait être rapide (50-100ms) pour que le *feedback* soit perçu comme instantané.

État de clic/activation (*Active/Pressed State*) : Au moment du clic (pression sur le bouton du contrôleur), un *feedback* visuel bref mais distinctif confirme l'activation :

- Animation rapide (100-150ms) d'enfoncement ou de compression
- Flash ou pulse lumineux bref
- Changement de couleur momentané

Ce *feedback* est crucial car, contrairement aux interfaces physiques où le *feedback* tactile du clic est immédiat, en RV l'utilisateur dépend entièrement du *feedback* visuel (et éventuellement sonore) pour confirmer que l'action a été enregistrée.

État Désactivé (*Disabled State*) : Les éléments temporairement non interactifs doivent être clairement distingués pour éviter la frustration d'une tentative d'interaction infructueuse :

- Réduction d'opacité (typiquement 40-60% de l'opacité normale)
- Désaturation de la couleur (conversion vers niveaux de gris ou réduction de saturation à 20-30%)
- Absence de réponse au survol (pas de changement d'état *hover*)

L'état désactivé doit être suffisamment distinct pour être immédiatement reconnaissable, mais l'élément doit rester visible pour que l'utilisateur comprenne quelle fonction est temporairement indisponible.

Considérations sur le problème de scintillement et aliasing

En raison de la densité de pixels limitée des HMD actuels (typiquement 10-20 pixels par degré d'angle visuel), les éléments UI fins peuvent présenter des artefacts visuels indésirables.

Scintillement temporel (*Temporal Flickering*) : Les contours fins (lignes de 1-2 pixels d'épaisseur) peuvent scintiller ou apparaître/disparaître lors de légers mouvements de tête, en raison de l'aliasing spatial. Ce problème est particulièrement prononcé pour les lignes qui ne sont pas parfaitement alignées avec la grille de pixels.

Solution : Utiliser des éléments d'au moins 3-4 pixels d'épaisseur minimum à la distance de visualisation. Pour les contours et les lignes, privilégier une épaisseur de 4-6 DMM minimum. Appliquer un anti-aliasing (MSAA, FXAA, ou TAA) de haute qualité pour lisser les contours.

Aliasing chromatique : Les contrastes extrêmes sur des contours fins (par exemple, ligne blanche fine sur fond noir) peuvent créer des artefacts d'aliasing chromatique (franges colorées) particulièrement visibles en périphérie des lentilles du HMD.

Solution : Éviter les contrastes noir/blanc purs sur des éléments fins. Utiliser des couleurs légèrement atténuées et des épaisseurs de ligne plus généreuses. Les algorithmes de

correction d'aberration chromatique du HMD atténuent ce problème mais ne l'éliminent pas complètement.

Patterns réguliers : Les patterns réguliers fins (grilles, textures répétitives à haute fréquence spatiale) peuvent créer des effets de moiré désagréables.

Solution : Éviter les patterns réguliers fins dans les éléments d'interface. Si des patterns sont nécessaires (par exemple, pour des textures de fond), utiliser des patterns irréguliers ou organiques, ou des patterns à basse fréquence spatiale.

Interaction multimodale : son et haptique

Bien que cette section se concentre sur les aspects visuels de l'interfaçage, il est important de noter que le feedback multimodal améliore significativement l'utilisabilité en RV.

Feedback sonore : Des sons brefs et non intrusifs peuvent accompagner les interactions.

- Son de survol (*hover sound*) : court son subtil (10-20ms) lorsque le rayon entre dans une hit zone
- Son de clic : son distinctif mais pas désagréable lors de l'activation d'un bouton
- Son d'erreur : indication auditive lorsqu'une action ne peut pas être effectuée

Le *feedback* sonore est particulièrement utile car il ne requiert pas que l'utilisateur regarde l'élément interactif, permettant ainsi une interaction plus fluide où l'attention visuelle peut être dirigée ailleurs.

Feedback haptique : Les contrôleurs RV modernes intègrent des moteurs haptiques (vibration) qui peuvent fournir un *feedback* tactile :

- Vibration légère (50-100ms) lors du survol pour confirmer que le pointeur est sur un élément interactif
- Vibration plus marquée (100-150ms) lors du clic pour simuler le *feedback* tactile d'un bouton physique

B.4 Le tableau récapitulatif des avantages et inconvénients/limites de la RVI

Tableau 10

Avantages de la RVI

Catégorie	Avantages identifiés	Auteurs cités
Ingénierie pédagogique	Scénarisation des apprentissages Réduction des risques (dérivage) Économie de moyens dans certaines configurations Attrait intrinsèque de la technologie	Arnaldi et al. (2018)
Contrôle expérimental	Contrôle précis des paramètres (stimulation, avatar, configuration) Adaptabilité à divers contextes de recherche	Dewez et al. (2019)
Engagement	Engagement supérieur des utilisateurs	Schutte et Stilinović (2017)
Motivation	Influence sur l'attention Amélioration via gamification (partage et conservation des acquis) Possibilité d'explorer/modéliser des phénomènes abstraits Engagement accru et temps accru sur la tâche Persévérance développée	Long et al. (2011) Mildner et al. (2015) Cooper et al. (2019) Chen et al. (2020)
Immersion technologique	Approches pédagogiques constructivistes Apprentissages basés sur simulation	Jensen et Konradsen (2018) Edwards et al. (2019)

Catégorie	Avantages identifiés	Auteurs cités
	Amélioration compréhension concepts abstraits (multisensoriel, tactile) Création de modèles mentaux corrects Pallier déficits pensée opérationnelle concrète	Del Sol Roo (2017) ; Ucar et al. (2017) ; Yoon et al. (2017)
Sécurité	Manipulation d'objets dans environnements dangereux de manière sécurisée	Jensen et Konradsen (2018) Makransky et al. (2019)
Coûts	Réduction significative des coûts de formation (équipements coûteux)	Cook et al. (2019) Cooper et al. (2019)
Autonomie	Augmentation possibilités de gestion du processus d'apprentissage	Altinpulluk (2019)
Apprentissage naturel	Engagement sensoriel comparable au processus naturel d'apprentissage de l'enfant	Hite (2022)
Compétences cognitives	Amélioration acuité spatiale Stimulation créativité Enrichissement processus métacognitifs	Lee et Shea (2020)
Collaboration	Facilitation de l'interaction Environnements d'apprentissage dynamiques et interactifs	Demitriadou et al. (2020)

Tableau B.4.2

Inconvénients et limites de la RVI

Catégorie	Limites identifiées	Auteurs cités
Manque de données	Littérature encore insuffisante malgré expansion Besoin d'études longitudinales Nombreux pans peu documentés (rayonnements bleus, ergonomie, psychologie, neurologie) Principe de vigilance appliqué	ANSES (2021) Makransky et Petersen (2021)
Ergonomie	Poids des casques occlusifs source d'inconfort (515g Quest 3) Fatigue visuelle (résultats contradictoires) Proximité des écrans problématique	ANSES (2021) Kim et al. (2019) Gao et al. (2019)
Effets post-exposition	Perturbations habileté manuelle (doublement taux d'erreur pour 60% participants) Perturbation orientation du corps Erreur spatiale de 15% pendant 6h Besoin de ~65 essais pour retour à la normale Sous-évaluation rotations réelles après exposition RV	Champney et al. (2007) Harm et al. (2008) Day et al. (2018) Viaud-Delmon et al. (1999)
Populations à risque	Confusion RV/réel pour populations à risque Proscrire : idéation paranoïaque, troubles dissociatifs Possibles traumatismes physiologiques et changements comportementaux Modifications psychologiques ou neurologiques	Ramirez et LaBarge (2018)
Éthique et sécurité	Atteinte vie privée et cyber sécurité Questions législatives Impact environnemental production casques	ANSES (2021) Ramirez et LaBarge (2018)

Catégorie	Limites identifiées	Auteurs cités
	<p>Questions sociétales pour utilisateur et entourage</p> <p>Risque moral unique</p>	

Catégorie	Limites identifiées	Auteurs cités
Niveau d'expertise	<p>Limite majeure : Efficacité dépend du niveau préalable</p> <p>Effet de réversion de l'expertise</p> <p>Effets négatifs pour novices (surcharge cognitive extrinsèque)</p> <p>Effets positifs uniquement pour niveau intermédiaire</p> <p>Inadaptée pour apprentissage de contenus nouveaux par débutants</p> <p>Rapport coût-bénéfice questionnable</p>	Poupard et al. (2024)
Motivation	<p>Seulement 20% études montrent effet positif motivation+performance</p> <p>67% études : effets incohérents</p> <p>Confusion terminologique (motivation intrinsèque, motivation à apprendre, plaisir, intérêt situationnel)</p> <p>Motivation dirigée vers exploration technologie plutôt que contenu</p> <p>Dissociation motivation d'usage/motivation d'apprentissage</p>	Poupard et al. (2024)
Charge cognitive	<p>38% études rapportent effets négatifs vs 15% effets positifs</p> <p>Augmentation charge cognitive extrinsèque</p> <p>Performance d'apprentissage diminuée (effet moyen à large)</p> <p>3 mécanismes :</p>	<p>Poupard et al. (2024)</p> <p>Sweller (1994, 2023)</p> <p>Mayer (2005)</p> <p>Jensen et Konradsen (2018)</p> <p>Chen et al. (2020)</p> <p>Albus et al. (2021)</p>

Catégorie	Limites identifiées	Auteurs cités
	<p>1. Complexité excessive environnement (surcharge, distractions)</p> <p>2. "Détail séduisant" (richesse perceptive capte attention sur éléments non pertinents)</p> <p>3. Effet de nouveauté (double pénalité : apprendre système + perception divertissement)</p> <p>Augmentation charge de travail confirmée en chirurgie</p> <p>Influence spectateurs/examineurs (augmentation CL)</p>	
Potentialités charge cognitive	<p>Possibilité d'optimisation avec conception appropriée</p> <p>Bon dosage informations nécessaire</p> <p>Signalisation adéquate peut réduire surcharge</p> <p>Avatar actif peut alléger CL sur tâche rappel</p> <p>Prudence avec agents PNJ</p> <p>Systèmes adaptatifs possibles (EEG, ECG, EDA)</p>	<p>Albus et al. (2021)</p> <p>Steed et al. (2016)</p> <p>Collins et al. (2019)</p> <p>Gerry et al. (2018)</p>
Cinétose	<p>Frein majeur</p> <p>Incohérence sensori-motrice (visuo-vestibulaire)</p> <p>Conflit entre systèmes : visuel, vestibulaire, proprioceptif</p> <p>3 groupes symptômes (SSQ) :</p> <ul style="list-style-type: none"> - Nauséux (nausée, maux ventre, salivation) - Oculomoteurs (fatigue visuelle, vision floue, migraines) - Désorientation (vertiges, étourdissements) <p>Latence boucle d'interaction problématique (variations 10-100ms)</p> <p>Difficile prévoir utilisateurs sensibles</p>	<p>ANSES (2021)</p> <p>Poupard et al. (2024)</p> <p>Reason et Brand (1975)</p> <p>Kennedy et al. (1993)</p> <p>Arnaldi et al. (2018)</p>

Catégorie	Limites identifiées	Auteurs cités
	Populations à surveiller : femmes enceintes, troubles oculomoteurs, migraines, troubles proprioception, anxiété, mal des transports	

ÉLÉMENTS CONTEXTUELS IMPORTANTS

- La RVI présente des intérêts particuliers dans domaines à fort enjeu sécuritaire/technique (médical, industrie, chirurgie)
- L'efficacité dépend fondamentalement du contexte d'usage et du public cible
- Le principe de précaution est appliqué sur plusieurs aspects par manque de recul
- Le principe d'équivalence recommandé : éviter expériences mauvaises en RV qui le seraient IRL
- La conception pédagogique rigoureuse apparaît comme condition sine qua non de l'efficacité

ANNEXE C - CONCEPTION DU DISPOSITIF DE RVI

Cette section documente l'ensemble du processus de conception et de développement du dispositif de RVI utilisé dans cette recherche. Elle comprend les spécifications techniques complètes destinées aux développeurs, incluant le storyboard détaillé de l'expérience, l'inventaire des matériaux et assets (physiques, audio, scénaristiques), la description du système de résolution d'énigme avec ses règles explicites et implicites, les captures d'écran du prototype, ainsi que la conception du tutoriel d'initiation permettant aux participants de se familiariser progressivement avec l'environnement virtuel. Cette documentation vise à la fois à assurer la reproductibilité de l'étude et à fournir un guide pratique pour les développeurs souhaitant implémenter ou adapter ce dispositif.

C.1 Documentation technique pour les développeurs

Cette sous-section présente la documentation technique destinée à guider le développement du dispositif de RVI. Elle comprend le storyboard global de l'expérience et les spécifications techniques des principales phases du dispositif.

C.1.1 *Storyboard complet de l'expérience (10 Étapes)*

Le storyboard constitue la structure narrative et interactive globale de l'expérience. Il décompose le parcours utilisateur en 10 étapes séquentielles (voir tableau C.1.1), chacune ayant des objectifs spécifiques en termes de gameplay, d'apprentissage, et de sensibilisation. Cette structure permet de gérer progressivement la charge cognitive de l'utilisateur, en introduisant d'abord les mécaniques d'interaction de base (tutoriel), puis le contexte narratif (immersion), avant de présenter les défis interactifs et les situations de sensibilisation.

Le storyboard constitue la structure narrative et interactive globale de l'expérience. Il décompose le parcours utilisateur en 10 étapes séquentielles, chacune ayant des objectifs spécifiques en termes de gameplay, d'apprentissage, et de sensibilisation.

Tableau 11
Structure en 10 étapes

Étape	Environnement / Visuel	Interaction utilisateur	Effet pédagogique / sensoriel
1. <i>Tutoriel</i>	Salle vide, instructions claires	Manipulation basique RV	Familiarisation technique
2. <i>Immersion</i>	Salle de classe réaliste → fondu → sons amplifiés, couleurs saturées	L'utilisateur observe	Mise en condition sensorielle TSA
3. <i>Consigne de l'enseignant</i>	Enseignant explique l'énigme au tableau	Aucun, écoute passive	Surcharge cognitive via carte mentale
4. <i>Horloge</i>	Horloge centrale, symboles animés, aiguille trace une ligne	Fixer du regard → découverte d'un indice	Illustration de la recherche de repères
5. <i>Interaction enseignant</i>	Enseignant s'approche, voix brouillée, expressions incohérentes	Solution : détourner le regard	Stratégie d'évitement visuel
6. <i>Tableau de consignes</i>	Tableau à droite, consigne 7 surlignée	Lire et mémoriser	Stress → régulation via écouteurs
7. <i>Plantes</i>	Trois plantes qui s'illuminent	Fixer chaque plante 3s (remplir jauge)	Routine + persévérance malgré interruptions
8. <i>Résolution de l'énigme</i>	Tableau principal avec labyrinthe	Saisir craie et tracer solution	Application des indices accumulés
9. <i>Exposé final</i>	Enseignant demande de se tourner vers la classe, murmures en fond	Observer camarades quelques secondes	Illustration du stress social

Étape	Environnement / Visuel	Interaction utilisateur	Effet pédagogique / sensoriel
<i>10. Fin</i>	Fondu + voix off rappelant concepts	Retrait du casque	Débriefing pédagogique

Cette structure permet de gérer progressivement la charge cognitive de l'utilisateur, en introduisant d'abord les mécaniques d'interaction de base (tutoriel), puis le contexte narratif, avant de présenter les défis interactifs et les situations de sensibilisation.

Voici à présent la structure initiale de l'outil :

C.1.2 *Spécifications techniques du tutoriel*

Le tutoriel constitue la première interaction de l'utilisateur avec le dispositif de RV. Son objectif principal est de réduire l'anxiété technologique et de développer les compétences techniques minimales nécessaires pour naviguer dans l'expérience principale. La conception du tutoriel s'appuie sur les principes établis par les tutoriels de référence en RVI (*Meta First Steps*, *Lies Beneath*) et sur les recherches en pédagogie de l'apprentissage technologique.

Environnement visuel

L'environnement du tutoriel est volontairement minimal et non distrayant :

- Fond : Espace blanc uniforme (RGB : 245, 245, 245) sans lignes de perspective ou indices de profondeur dans les phases initiales. Ce choix délibéré vise à éliminer toute distraction visuelle et à focaliser l'attention exclusivement sur les objets d'interaction et les instructions.
- Introduction progressive de la profondeur : Après que l'utilisateur a maîtrisé les interactions de base avec les objets flottants, des lignes de construction légères (grille de sol, lignes de perspective) sont progressivement introduites (apparition en fade-in sur 2-3 secondes). Cette introduction graduelle permet au système vestibulaire de s'habituer progressivement aux indices de profondeur et à la structure spatiale 3D, réduisant ainsi le risque de désorientation ou de cinétose (Gorisse et al., 2018).

Cette approche d'introduction progressive est directement inspirée des tutoriels *Meta Quest* qui utilisent une stratégie similaire pour faciliter l'acclimatation à l'espace virtuel tridimensionnel.

Instructions vocales : séquence complète des 14 voice lines

Le tutoriel utilise une narration vocale guidée pour diriger l'utilisateur à travers les différentes étapes. La voix utilisée est calme, claire, et encourageante, avec un débit de parole modéré (environ 150 mots/minute, légèrement plus lent que la parole conversationnelle normale de 180-200 mots/minute) pour assurer une compréhension optimale.

Tableau 12*Séquence des 14 voice lines du tutoriel***N° Contenu de la voice line**

- VL1 Bonjour, bienvenu dans ce tutoriel qui va vous permettre de vous familiariser avec la réalité virtuelle
- VL2 Tout d'abord regardez autour de vous
- VL3 Vous voyez la manette sur votre droite ? C'est la même que celle que vous avez dans la main
- VL4 Vous voyez ce gros bouton dessous ? C'est avec lui que vous allez pouvoir attraper des objets
- VL5 Ouvrez totalement la main, n'ayez pas peur la manette est fixée grâce à la sangle elle ne tombera pas
- VL6 Essayez de saisir le bout de bois devant vous en mettant la main dessus et en appuyant sur le bouton
- VL7 Si vous gardez le bouton appuyé, le bâton restera dans votre main
- VL8 Si vous ouvrez de nouveau la main vous lâcherez le bâton, essayez pour voir
- VL9 Essayez de refaire cette action pour vous habituer
- VL10 Maintenant, regardez la cible sur votre gauche, si vous posez votre regard dessus elle changera de couleur
- VL11 Essayez de la faire changer de couleur
- VL12 Maintenant que vous avez les bases de la réalité virtuelle, prenez un instant pour vous familiariser avec l'environnement
- VL13 Une porte va apparaître, quand vous sentez que c'est le bon moment pour vous, attrapez la poignée et l'expérience débutera
- VL14 C'est à vous

Matériaux physiques du tutoriel

Le tutoriel utilise trois objets physiques pour l'apprentissage des interactions :

- Image bâton : cylindre virtuel (désigné comme "bout de bois" dans les instructions vocales) servant d'objet de manipulation
 - Fonction : Apprentissage de la saisie et du déplacement d'objets via le bouton de préhension du contrôleur
- Contrôleur (représentation virtuelle) : Représentation visuelle du contrôleur physique dans l'environnement virtuel
 - Fonction : Permettre à l'utilisateur de voir la manette virtuelle et d'identifier le bouton de préhension
 - Position : Visible sur la droite de l'utilisateur (VL3)
- Cible : Objet cible réagissant au regard de l'utilisateur
 - Fonction : Apprentissage de l'interaction par le regard (changement de couleur lors de la fixation visuelle)
 - Position : Sur la gauche de l'utilisateur (VL10)

C.2 Matériaux et assets du dispositif

Cette sous-section présente l'inventaire des matériaux et assets utilisés dans le dispositif de réalité virtuelle immersive. Elle s'organise en trois catégories : les matériaux physiques, les matériaux audio, et les matériaux scénaristiques.

C.2.1 Matériaux physiques : environnement de la salle de classe

Les matériaux physiques constituent les éléments visuels et spatiaux de l'environnement virtuel. Cette section documente les assets 3D, textures et éléments d'environnement, tous en *lowpoly*, utilisés pour créer la salle de classe virtuelle.

Modèles 3D et Assets

L'environnement de la salle de classe utilise des modèles 3D. Source : Unity Asset Store

Description : Modèles de salle de classe universitaire incluant le mobilier standard (tables, chaises, bureau enseignant, tableau, étagères)

Usage : Création de l'environnement principal où se déroule l'expérience

- 1 bureau étudiant - modèle identique qui sera utilisé 12 fois
- 1 chaise étudiant - modèle identique qui sera utilisé 12 fois
- 1 bureau enseignant
- 1 estrade pour le bureau
- 1 tableau noir (Layer modifiable)
- 1 Fenêtre - modèle identique qui sera utilisé X fois (Pas de visibilité sur l'extérieur - Fenêtre opaque mais lumineuse - Taille standard : 225cm sur 125cm)
- 1 Porte d'entrée
- 1 Panneau sortie verte (issue de secours)
- 1 Extincteur
- 1 feuille blanche
- 1 crayon
- 1 chauffage standard mural qui sera utilisé 6 fois
- 1 luminaire plafonnier incrusté dans le plafond qui sera utilisé 4 fois

Agents virtuels :

Description : Modèles de personnages humanoïdes représentant l'enseignante et les autres étudiants (les agents). L'utilisation du *lowpoly* est ici aussi de rigueur.

Fonction : Créer la présence sociale et les interactions dans l'environnement virtuel

Textures :

- Sol :
 - Type : Texture de carreaux
 - Description : Revêtement de sol type carrelage ou vinyle, typique des salles de classe
- Plafond :
 - Couleur : Blanc
 - Type : Faux plafond standard
 - Description : Surface supérieure de la salle avec panneaux d'éclairage intégrés
- Murs :
 - Couleur : Blanc
 - Type : Peinture mate
 - Description : Surfaces murales de la salle de classe

C.2.2 Matériaux audio : Voix off et ambiances sonores

Les matériaux audio constituent une composante essentielle de l'expérience immersive. Cette section documente des exemples issus des blocs audio de la phase d'immersion et de la phase principale

Exemple : Le décomposé des sons phase d'immersion :

- Bloc 1 : 7 mots : « Bienvenue dans la salle de classe virtuelle. »
- Bloc 2 : 29 mots : « Vous allez vivre l'expérience de vie commune d'un étudiant TSA dans une situation d'apprentissage. Pour cela, vous allez intégrer le corps virtuel de l'étudiant situé à côté de vous. »
- Bloc 3 : 3 mots : « Vous êtes prêt ? »
- Bloc 4 : Pas de mots : bande de son d'ambiance « salle de classe »

Exemple : Le décomposé des sons phase principale (introduction à l'énigme) :

- Bloc 1 : 74 mots : « Aujourd'hui, vous allez devoir résoudre cette énigme. Le principe est simple, vous devez rallier la sortie de ce labyrinthe, située en haut à droite en partant

du point sur la gauche. Mais d'abord, vous devez comprendre les règles qui vont vous obliger à passer d'une certaine manière dans le labyrinthe. Vous ne pouvez pas repasser sur un endroit où vous êtes déjà passé. Il n'y a qu'une seule solution à ce problème, vous avez une heure. »

Le but est que l'utilisateur mémorise cette dernière information qui sera utile lors de la phase de renforcement. Cette question de délai est importante et ne doit pas être oubliée.

- Bloc 2 : 7 mots : « Tout va bien ? Vous avez besoin d'aide ? »
- Bloc 3 : 7 mots : « Levez la main si vous m'avez compris »

C.2.3 Matériaux scénaristiques : Narration de l'expérience

Les matériaux scénaristiques constituent le contenu narratif qui structure l'expérience et guide l'utilisateur à travers les différentes phases. Cette section documente les éléments narratifs du dispositif.

Structure narrative globale du matériel scénaristique

L'immersion TSA : Après un bref fondu, les couleurs deviennent plus vives et les sons plus forts (saturation sonore et auditive). En tant qu'avatar d'un étudiant TSA, l'utilisateur perçoit l'environnement sous le filtre du TSA. L'enseignant prend la parole et s'adresse à toute la classe. Il présente, avec des gestes et un exemple, une énigme dessinée sur le tableau noir situé derrière lui :

« Aujourd'hui, vous allez devoir résoudre cette énigme. Le principe est simple, vous devez rallier la sortie de ce labyrinthe, située en haut à droite en partant du point sur la gauche. Mais d'abord, vous devez comprendre les règles qui vont vous obliger à passer d'une certaine manière dans le labyrinthe. Vous ne pouvez pas repasser sur un endroit où vous êtes déjà passé. Il n'y a qu'une seule solution à ce problème, vous avez une heure »

La carte mentale : Après avoir reçu la consigne, l'avatar de l'utilisateur devenu l'étudiant TSA commence à développer la carte mentale qui rend complexe la consigne avec les éléments qui y sont associées. L'avatar étudiant TSA est perturbé par la consigne, mais aussi par l'environnement en lui-même (représentation agressive).

Première tâche - L'horloge : l'avatar étudiant TSA commence à chercher les indices. L'environnement devient trouble sauf un point qui reste clair : l'horloge centrale au-dessus du tableau. Dès qu'il pose son regard dessus, le tracé de l'horloge se modifie. L'aiguille des

secondes est obligée de passer sur des symboles à chaque tour qu'elle fait. Ces symboles, à côté des chiffres, servent à montrer les 12 divisions des heures. Ils ont la même forme que des symboles dans l'énigme. Maintenant, quand l'aiguille passe sur les symboles hexagonaux, elle laisse une ligne jaune. Première solution pour l'énigme, cette idée est intégrée dans la zone 1 de la carte mentale.

Phase de stress - le rapport social : La résolution de cette tâche provoque une réaction de l'enseignant. Il s'approche de l'avatar étudiant TSA en s'adressant à lui. La voix est brouillée et inaudible tant que l'avatar étudiant TSA regarde dans la direction de l'enseignant. La solution est de détourner le regard vers une autre zone. La voix off donne une information afin de guider l'avatar étudiant TSA.

Dès que l'avatar étudiant TSA pose son regard sur le sol à côté de lui (consigne donnée par la voix off), la voix de l'enseignant devient claire et audible. L'avatar étudiant TSA comprend ce que dit l'enseignant. L'enseignant lui demande de lever la main. L'information acquise, l'avatar étudiant TSA lève le bras. Une fois la main levée, l'enseignant repart alors au tableau. Cette interaction a pour conséquence d'augmenter le niveau de stress de l'avatar étudiant TSA (représenté par un overlay de pulsation rougeâtre et un bruit de pulsation cardiaque).

Seconde tâche - Le tableau de consignes : Le comportement de l'enseignant a mis l'avatar étudiant TSA dans un état de stress qu'il faut maintenant réguler. Les écouteurs situés sur le bureau commencent à clignoter. L'avatar étudiant TSA attrape les écouteurs et les mets sur sa tête. Un bruit blanc très léger occulte les sons de l'état « stress ». Le son et la lumière reviennent à leur état initial.

Le tableau sur la droite de l'avatar étudiant TSA commence à clignoter. Sur ce tableau, 10 consignes sont écrites. La septième consigne est un des indices « le début est en bas à gauche ». Dès que l'avatar étudiant TSA regarde le tableau, la consigne passe en surbrillance et clignote. Cette donnée est ajoutée dans la zone 2 de la carte mentale. Un signal lumineux, situé sur la première plante de l'étagère sur la gauche, se déclenche à la fin de la tâche.

Troisième tâche - Les plantes : Dès que l'avatar étudiant TSA pose son regard sur la plante qui s'est allumée, un curseur s'active. Ce curseur est un chronomètre de 3 secondes. Il remplit une jauge qui devient jaune. Tant que l'avatar étudiant TSA garde les yeux sur la plante, le curseur monte. Il y a trois jauge (une par plante) qu'il faut remplir. L'enseignant interpelle l'avatar étudiant TSA à la fin de la première jauge avec une *voix line*. Si l'avatar étudiant TSA

détourne le regard pour écouter l'enseignante, le curseur redescend. Pour chaque jauge non terminée, le curseur redescend si l'avatar étudiant TSA détourne le regards (si elle est complète, elle reste complète). Une fois les trois jauges remplis, une ligne jaune les relie et la tâche se termine. La règle est validée et elle rejoint la zone 3 de la carte mentale. La quatrième zone de la carte mentale sert de total des règles afin de faciliter la compréhension globale de l'énigme.

La résolution de l'énigme : La carte mentale est désormais complète. L'enseignant invite l'avatar étudiant TSA à réaliser la solution au tableau. Avec l'aide de la carte mentale, l'avatar étudiant TSA entreprend la résolution de l'énigme. Il doit saisir la craie virtuelle fixé sur le tableau en bas à gauche et déplacer cette craie sur le tableau pour faire le tracé. Il complète le tracé en rejoignant le point en haut à droite et termine l'énigme.

L'*aftermath* : Une fois terminé, l'enseignant félicite l'avatar étudiant TSA et lui demande de se retourner pour expliquer à ses camarades comment il a fait. L'overlay de stress est activé de nouveau. L'avatar étudiant TSA regarde la salle et les autres élèves pendant quelques secondes avant que la simulation ne se termine. La voix off fait un rappel de l'expérience et invite l'avatar étudiant TSA à enlever le casque.

Après un rapide tutoriel sur l'utilisation de la RV, l'utilisateur est emmené dans une salle de classe virtuelle. Pendant quelques secondes, la salle lui est présenté de manière succincte et on lui explique qu'il va incarner un étudiant avec TSA. Il est ensuite emmené sur une chaise a un emplacement spécifique (un bureau au centre de la pièce). On lui explique ensuite que sa perception va être modifier pour correspondre à la vision d'une personne avec TSA.

Après un bref fondu, les couleurs sont plus vives et les sons plus forts. Désormais dans l'avatar d'un étudiant TSA, l'utilisateur découvre le monde sous un angle nouveau. Alors qu'il observe son environnement, l'enseignant prend la parole. Ce dernier s'exprime clairement et s'adresse à toute la classe. Il présente avec des gestes et un exemple une énigme dessinée au tableau tout en expliquant à voix haute :

« Aujourd'hui, vous allez devoir résoudre cette énigme. Le principe est simple, vous devez rallier la sortie de ce labyrinthe, située en haut à droite en partant du point sur la gauche. Mais d'abord, vous devez comprendre les règles qui vont vous obliger à passer d'une certaine manière dans le labyrinthe. Vous ne pouvez pas repasser sur un endroit où vous êtes déjà passé. Il n'y a qu'une seule solution à ce problème, vous avez une heure »

Après avoir donné la consigne, l'enseignant regarde la salle. L'avatar de l'utilisateur devenu l'étudiant TSA commence à développer la carte mentale qui rend complexe la consigne avec les éléments qui y sont associées. L'avatar étudiant TSA est perturbé par la consigne, mais aussi par l'environnement en lui-même qui est agressif pour lui. Les visages et expressions des autres personnes sont problématiques et cela ne fait que rajouter à la complexité de la situation. Pour essayer de se calmer et de reprendre le contrôle d'une situation qui lui échappe de secondes en secondes, l'avatar étudiant TSA commence alors à chercher un élément de l'environnement qui pourrait le calmer. Son choix se pose sur l'horloge, avec son rythme régulier et sans surprise, une source stable pour lui.

Il passe quelques instants à contempler l'horloge quand quelque chose dans le mouvement des aiguilles attire son attention. A côté des chiffres, le marqueur qui sert à montrer les 12 divisions des heures à la même forme que l'un des symboles dans l'énigme que l'enseignant a demandé de résoudre. L'aiguille des secondes est obligée de passer sur ces symboles à chaque tour qu'elle fait. Cette dernière passe sur les symboles hexagonaux en laissant une trace. Solution possible pour l'énigme, cette idée est intégrée dans une des zones de la carte mentale.

Ce comportement a provoqué une réaction de l'enseignant, qui s'approche de l'avatar étudiant TSA en lui demandant s'il va bien. Ce dernier s'en rend compte et regarde alors l'enseignant. Le problème est que les expressions de l'enseignant ne collent pas avec son regard et ses propos. L'enseignant est souriant, mais sa voix pourrait sembler agacée et son regard pourrait être porteur de jugement. C'est incohérent et incompréhensible pour l'avatar étudiant TSA qui n'arrive pas à comprendre ce que dit l'enseignant. La seule solution est de détourner le regard pour essayer de comprendre ce que dit l'enseignant, car plus l'étudiant le regarde, moins il comprend. Le regard sur le sol à côté de lui, l'avatar étudiant TSA se concentre alors sur ce que dit l'enseignant. Avec cette stratégie de détourner le regard, il comprend que l'enseignant lui demande de lever la main. L'information acquise, l'avatar étudiant TSA lève le bras car, il a bien compris ce que lui demandait l'enseignant. Une fois la main levée, l'enseignant repart alors au tableau. Néanmoins, cette interaction forcée n'est pas sans conséquence pour l'avatar étudiant TSA.

La pression accumulée durant l'interaction a provoqué du stress et l'avatar étudiant TSA. L'enseignant, sans le vouloir, l'a mis dans un état de stress qu'il faut maintenant réguler. L'avatar étudiant TSA attrape ses écouteurs, les mets sur ses oreilles et se lance un bruit blanc pour occulter tous les autres sons de la pièce. Le calme revient et même la lumière semble à

présent plus paisible. Son regard se promène et le tableau sur sa droite attire son attention. Immédiatement, il remarque qu'une des phrases fait référence à l'énigme, c'est la septime en partant du haut, dissimulé au milieu d'autres « le début est en bas à gauche ». Il ajoute cette donnée à la carte mentale et reprend sa recherche d'indices.

La répétition et la position des plantes verte sur l'étagère de gauche attirent son attention. Il ressent le besoin de mettre en place une routine et commence à observer les plantes en forme de « L » inversé qui attirent son regard. L'enseignant l'interpelle de nouveau. Sortie de sa routine, l'avatar étudiant TSA essaye d'y retourner mais doit la reprendre depuis le début. Cette nouvelle agression le sort de sa routine alors qu'il a besoin de celle-ci pour se concentrer de nouveau. L'enseignant l'interpelle de nouveau mais cette fois, l'avatar étudiant TSA résiste et reste focalisé sur sa tâche. Une fois les 3 points observés pendant quelques secondes, l'idée est prête et rejoint la carte mentale.

La carte mentale est désormais complète. L'enseignant, qui a remarqué que l'avatar étudiant TSA avait visiblement fini, l'invite à réaliser la solution au tableau. Avec l'aide de la carte mentale et les idées ajoutées dessus, l'avatar étudiant TSA entreprend alors la résolution de l'énigme. Avec ses idées et sa carte mentale comme guide, il complète le tracé et termine l'énigme.

Une fois terminé, l'enseignant félicite l'avatar étudiant TSA et lui demande de se retourner pour expliquer à ses camarades comment il a fait. Une pression énorme s'abat alors sur l'étudiant. Son cœur palpite fort et il sent dans ses tempes les coups de sang. Un bruit de murmure remplit alors la salle. Il regarde la salle et les autres élèves pendant quelques secondes avant que la simulation ne se termine.

C.3 Système de résolution de l'énigme

Le système de résolution d'énigme constitue une composante centrale (Étape 8 du storyboard). Il prend la forme d'un labyrinthe que l'utilisateur doit résoudre en appliquant un ensemble de règles explicites et implicites. Cette section présente les règles structurant le système d'énigme.

Structure du système de règles

Le labyrinthe est régi par six règles distinctes, organisées en deux catégories :

Règles Implicites (A et B) : Règles qui ne sont pas explicitement présentées à l'utilisateur et qui doivent être découvertes par inférence ou par essai-erreur.

Règles Explicites (1, 2, 3) : Règles présentées explicitement à l'utilisateur sous forme d'instructions ou de cartes de référence.

Cette structure mixte (règles implicites + explicites) crée un défi cognitif qui valorise à la fois la capacité d'inférence, l'analyse systématique, et l'application rigoureuse de règles formelles.

C.3.1 Règles implicites

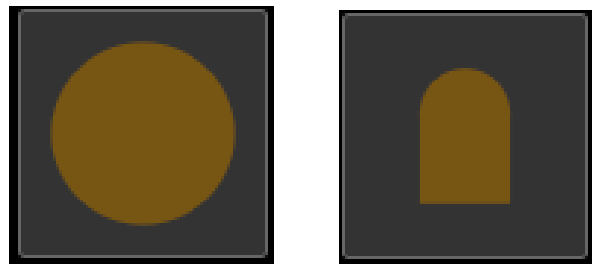
Les règles implicites définissent les contraintes fondamentales du labyrinthe mais ne sont pas explicitement communiquées à l'utilisateur.

Règle A : Points d'entrée et de sortie

Description : Le labyrinthe possède un point d'entrée unique et un point de sortie unique. Le chemin solution doit obligatoirement commencer au point d'entrée et se terminer au point de sortie.

Figure 52

Point entrée et sortie



Nature Implicite : Cette règle découle de la structure même du labyrinthe et de la présence visible des points d'entrée et de sortie. L'utilisateur doit comprendre par lui-même que le trajet doit relier ces deux points.

Règle B : Passages bloqués

Description : Certains passages du labyrinthe sont bloqués par des obstacles ou des murs. Le chemin ne peut pas traverser ces passages bloqués.

Figure 53

Passages bloqués

Nature Implicite : Les passages bloqués sont visuellement identifiables par leur apparence (murs, barrières), mais aucune instruction n'indique explicitement que ces passages sont interdits. L'utilisateur doit inférer cette contrainte de la structure visuelle du labyrinthe.



Symbolisation des règles A et B

Cette figure présente une représentation symbolique montrant l'application simultanée des règles A (entrée/sortie obligatoires) et B (évitement des passages bloqués) dans un exemple de configuration de labyrinthe.

C.3.2 Règles explicites

Les règles explicites sont présentées à l'utilisateur sous forme d'instructions formelles. Elles définissent des contraintes additionnelles que le chemin solution doit respecter.

Règle 1 : Passage obligatoire par les hexagones

Description : Le chemin solution doit obligatoirement passer par tous les hexagones noirs présents dans le labyrinthe. Ces hexagones constituent des points de passage obligatoires.

Figure 54

Les hexagones

Fonction Cognitive : Ils indiquent le chemin à suivre et facilite la perception de l'énigme afin de permettre à l'utilisateur de ne pas se perdre



Règle 2 : Séparation des carrés de couleur

Description : Les carrés de même couleur doivent être isolés dans des sections distinctes du labyrinthe. Le chemin solution sert de frontière séparant les zones contenant des carrés de couleurs différentes.

Figure 55

Les carrées de couleurs

Fonction cognitive : Cette règle introduit une contrainte de segmentation spatiale qui nécessite de considérer simultanément la position du chemin et la distribution des carrés colorés dans l'espace. Elle a été supprimée dans la version définitive du labyrinthe.



Règle 3 : Respect des divisions verticales (L Jaunes)

Figure 56

Description : Les formes en L de couleur jaune indiquent une autre direction obligatoire que l'utilisateur doit emprunter.

Le L jaune

Fonction cognitive : Cette règle impose une contrainte d'orientation qui limite les possibilités de trajet dans certaines zones du labyrinthe²².



Intégration des six règles

La résolution du labyrinthe nécessite l'application simultanée et cohérente des cinq règles (deux implicites et trois explicites). L'utilisateur doit :

1. Identifier le point d'entrée et le point de sortie (Règle A)
2. Repérer les passages bloqués et les éviter (Règle B)
3. Localiser les hexagones noirs et planifier un trajet qui les traverse tous (Règle 1)
4. Considérer la distribution des carrés colorés et planifier un trajet qui les sépare correctement (Règle 2)
5. Respecter les divisions créées par les L jaunes (Règle 3)

Cette intégration multi-contraintes crée un défi cognitif significatif qui représente certaines fonctions cognitives des personnes avec TSA.

²² Dans la version finale, cette règle a été modifiée pour être remplacée par un indice visuel textuel

C.4 Captures d'écran du moteur Unity et du prototype

Cette section présente une série de captures d'écran extraites de l'outil développé durant cette thèse. Ces captures illustrent différents aspects du dispositif, incluant les environnements, les interfaces, et les éléments de gameplay. Les figures présentées ci-dessous documentent visuellement le prototype en situation, offrant un aperçu des choix de conception visuelle, de l'agencement spatial, et des mécaniques d'interaction implémentées.

C.4.1 Extrait Unity 3D (version 20213.2.2f1) et Visual studio

Figure 57

Maquette de la salle de classe



Cette capture d'écran présente une version avant traitement. Les éléments ont été réduits au maximum pour tester les fonctionnalités de base de l'outil.

C.4.2 Environnement et perception

Figure 61

Vision neurotypique



Cette capture présente l'environnement de la salle de classe tel qu'il apparaît en mode perception neurotypique, servant de référence perceptive standard.

Figure 62

Vision TSA avec saturation visuelle et auditive



Cette capture illustre le même environnement en mode perception TSA, montrant l'effet de saturation des couleurs et des modifications visuelles simulant les particularités perceptives TSA.

Figure 63

Retour à la vision neurotypique



Cette capture montre le retour à la perception neurotypique, permettant de visualiser le contraste entre les deux modes perceptifs.

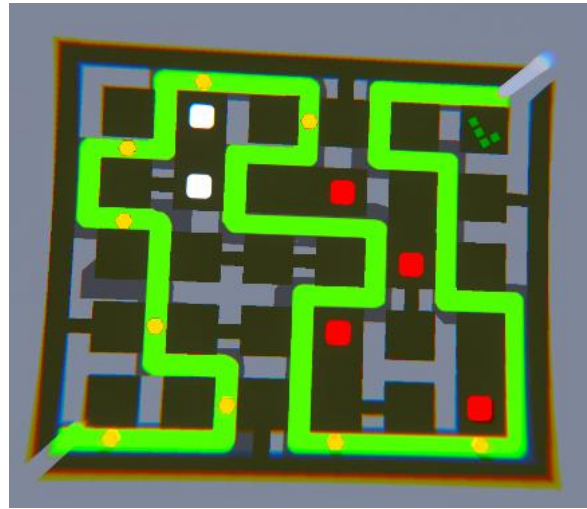
Figure 64

Upscale de l'agent enseignant



Cette capture présente l'agent virtuel enseignant en *low poly* avec un *Upscale* de 15%.

Cette capture présente l'interface du **Figure 68**
labyrinthe (avec la craie) tel qu'il apparaît *Le labyrinthe terminé*
terminé lors de la phase de résolution d'énigme.



C.4.3 Références de conception du tutoriel

Les deux figures suivantes présentent des références visuelles ayant inspiré la conception du tutoriel d'initiation à la réalité virtuelle :

Cette capture issue **Figure 69** du tutoriel officiel *Meta Tutoriel Meta First Steps* illustre les principes de conception de tutoriels RV adoptés par les développeurs de casques de réalité virtuelle grand public.



Cette capture issue **Figure 70** du jeu *Lies Beneath* *Tutoriel Lie Beneath* présente une approche alternative de tutoriel VR intégré dans un contexte narratif, ayant servi d'inspiration pour certains aspects de conception.



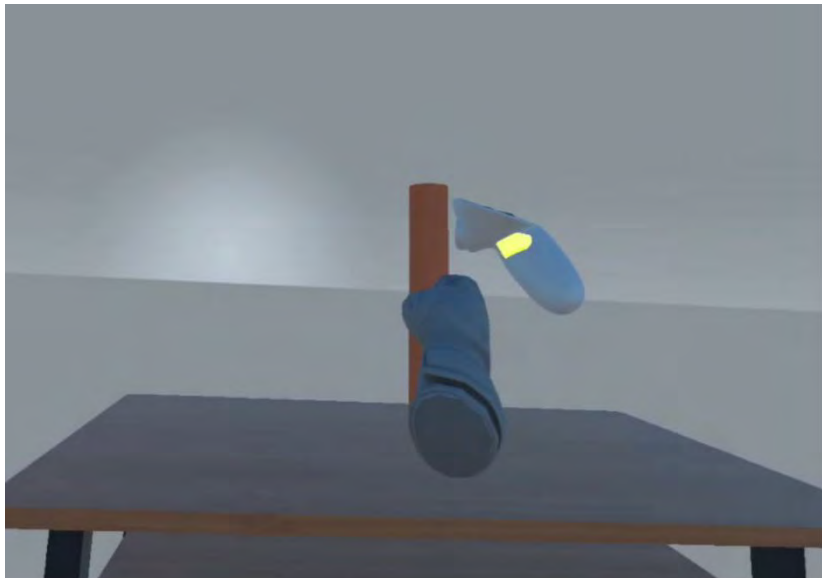
Les deux figures suivantes sont des extraits du tutoriel conçu au laboratoire et intégré à l'outil :

Cette capture montre **Figure 71**
le processus d'acquisition *Tutoriel projet TSA RVI – objet « bout de bois »*
des fondamentaux de la RVI
dans le tutoriel conçu au
laboratoire (ici les bases de
l'agentivité)



Cette capture montre
la saisie de l'objet « bout de
bois » qui permet à
l'utilisateur l'apprentissage
de l'interaction avec
l'environnement virtuel

Figure 72
Tutoriel projet TSA RVI – saisie objet « bout de bois »



ANNEXE D - INSTRUMENTS DE MESURE

Cette section rassemble l'ensemble des instruments psychométriques utilisés dans les deux expérimentations conduites durant cette recherche. Elle comprend les questionnaires de référence issus de la littérature scientifique, ainsi que les batteries complètes de mesures adaptées aux objectifs spécifiques de chaque expérimentation. Les instruments documentés incluent les échelles de mesure de la présence, de l'agentivité, de la cinétose, de l'empathie, des connaissances sur le TSA et des attitudes envers les personnes avec TSA. Cette documentation vise à assurer la reproductibilité des mesures et à fournir une référence complète pour les chercheurs souhaitant utiliser ou adapter ces instruments.

D.1 Questionnaires de référence issus de la littérature

Cette sous-section présente les questionnaires standardisés issus de la littérature scientifique utilisés pour mesurer des variables clés de l'expérience en réalité virtuelle immersive.

D.1.1 *Questionnaire Body Ownership (propriété corporelle)*

Le questionnaire *Body Ownership* mesure le sentiment de propriété corporelle, c'est-à-dire le degré auquel l'utilisateur perçoit le corps virtuel comme étant son propre corps. Cet instrument est issu des travaux sur l'illusion de la main en caoutchouc et a été adapté pour les environnements de réalité virtuelle. Il utilise une échelle de Likert en 7 points.

- It seemed as if I were feeling the touch of the paintbrush in the location where I saw the rubber hand touched.
- It seemed as though the touch I felt was caused by the paintbrush touching the rubber hand.
- I felt as if the rubber hand were my hand.
- I felt as if my (real) hand were drifting towards the right (towards the rubber hand).
- It seemed as I might have more than one left hand or arm.
- It seemed as if the touch I was feeling came from somewhere between my own hand and the rubber hand.
- I felt as if my (real) hand were turning “rubbery”.
- It appeared (visually) as if the rubber hand were drifting towards the left (towards my hand).

- The rubber hand began to resemble my own (real) hand, in terms of shape, skin tone, freckles or some other visual feature.

D.1.2 Questionnaire Embodiment (incarnation)

Le questionnaire *Embodiment* évalue le sentiment global d'incarnation dans le corps virtuel, intégrant plusieurs dimensions incluant la propriété corporelle, l'agentivité, et l'auto-localisation dans le corps virtuel.

Ownership

- 1) I felt as if the virtual body I saw when I looked down was my body.
- 2) It felt as if the virtual body I saw was someone else.
- 3) It seemed as if I might have more than one body.
- 4) I felt as if the virtual body I saw when looking in the mirror was my own body.
- 5) felt as if the virtual body I saw when looking at myself in the mirror was another person.

Agency

- 6) It felt like I could control the virtual body as if it was my own body.
- 7) The movements of the virtual body were caused by my movements.
- 8) I felt as if the movements of the virtual body were influencing my own movements.
- 9) I felt as if the virtual body was moving by itself.

Tactile Sensations

- 10) It seemed as if I felt the touch of the ____ in the location where I saw the virtual body touched.
- 11) It seemed as if the touch I felt was located somewhere between my physical body and the virtual body.
- 12) It seemed as if the touch I felt was caused by the ____ touching the virtual body.
- 13) It seemed as if my body was touching the ____.

Self-Location

- 14) I felt as if my body was located where I saw the virtual body.

15) I felt out of my body.

16) I felt as if my (real) body were drifting towards the virtual body

or as if the virtual body were drifting towards my (real) body.

Appearance

17) It felt as if my (real) body were turning into an “avatar” body.

18) At some point it felt as if my real body was starting to take on the posture or shape of the virtual body that I saw.

19) At some point it felt that the virtual body resembled my own (real) body, in terms of shape, skin tone or other visual features.

20) I felt like I was wearing different clothes from when I came to the laboratory.

Response to External Stimuli

21) I felt that my own body could be affected by ____.

22) I felt a ____ sensation in my body when I saw ____.

23) When ____ happened, I felt the instinct to ____.

24) I felt as if my body had ____.

25) I had the feeling that I might be harmed by the ____.

Presence / Ownership / Agency

7 point Likert scale

Variable / question

Presence

- Please rate your sense of being in the virtual office space, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place. I had a sense of “being there” in the virtual office space

- To what extent were there times during the experience when the virtual office space was the reality for you? There were times during the experience when the office space was the reality for me...

- When you think back about your experience, do you think of the office space more as images that you saw, or more as somewhere that you visited? The office space seems to me to be more like ...

- When you think back about your experience, do you think more as being elsewhere, or more as being in the office space? I thought more as...

- Consider your memory of being in the office space. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today? By 'structure of the memory' consider things like the extent to which you have a visual memory of the office space, whether that memory is in colour, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination, and other such structural elements. I think of the office space as a place in a way similar to other places that I've been today...

- During the time of the experience, did you often think to yourself that you were actually in the office space? During the experience I often thought that I was really seated in the office space...

Ownership

- I felt that the virtual body was my own body.

- I felt that the virtual arms were part of my body.

- I felt that the virtual arms could be harmed.

- I felt that my real arms could be harmed.

- I felt that virtual arms were not part of my body.

- I felt as if the virtual arms were from someone else's body.

Agency

- I felt as if the virtual body moved just like I wanted it to, as if it was obeying my will.

- I expected the virtual body to react in the same way as my own body.

- I felt like I controlled the virtual body as if it was my own body.

D.1.3 Questionnaire BES-A (Basic Empathy Scale - Adult)

La *Basic Empathy Scale* dans sa version pour adulte (BES-A) mesure l'empathie selon deux dimensions : l'empathie cognitive (capacité à comprendre les perspectives et états mentaux d'autrui) et l'empathie affective (capacité à ressentir émotionnellement ce qu'autrui ressent).

Version française de la BES-A non modifié – adaptation française par D'Ambrosio et al. (2009) de la BES de Jolliffe & Farrington (2006)

1. Les émotions de mes ami(e)s ne m'affectent pas
2. Après avoir été avec un(e) ami(e) qui est triste, je me sens généralement triste.
3. Je peux comprendre le bonheur que ressent un(e) ami(e) quand quelque chose de bien lui arrive.
4. Il peut m'arriver d'avoir peur quand j'écoute un film d'horreur.
5. Je me sens facilement touché(e) par les émotions des autres.
6. J'ai du mal à savoir quand mes ami(e)s ont peur.
7. Je ne suis pas triste quand je vois une personne pleurer.
8. Les sentiments des autres ne m'inquiètent pas du tout.
9. Quand quelqu'un a la sensation « d'être au plus bas », « d'avoir touché le fond », je peux comprendre ce qu'il ressent.
10. Généralement, je peux me rendre compte que mes ami(e)s sont effrayé(e)s.
11. Souvent je me sens triste lorsque je regarde des choses ou des films tristes.
12. Souvent je peux comprendre comment les gens se sentent avant même qu'ils me l'aient dit.
13. Voir une personne se mettre en colère n'a aucun effet sur moi.
14. Habituellement, je sais quand les gens sont joyeux.
15. J'ai tendance à avoir peur quand je suis avec des ami(e)s qui sont effrayé(e)s.
16. Généralement, je me rends vite compte quand un(e) ami(e) est en colère.
17. Je suis souvent envahi(e) par les sentiments de mes ami(e)s.
18. Je ne ressens rien face à la tristesse de mes ami(e)s.

19. Je ne suis généralement pas attentif (ve) aux sentiments de mes ami(e)s.

20. J'ai du mal à comprendre quand mes ami(e)s sont heureux (ses).

D.2 Batterie de mesures - expérimentation 1

Cette sous-section présente la batterie complète d'instruments de mesure utilisés lors de la première expérimentation. Cette batterie intègre des mesures de l'expérience en réalité virtuelle (présence, agentivité, cinétose) ainsi que des mesures des effets de sensibilisation (empathie, connaissances, attitudes).

D.2.1 Questionnaire complet Expérimentation 1

Version du questionnaire présentée aux utilisateurs EXP 1 :

Merci de votre participation. Vous allez à présent répondre à un certain nombre de questions portant sur l'expérience que vous venez de vivre.

Ce questionnaire prend environ 20 minutes à répondre. Les réponses sont totalement anonymes et ne seront utilisées que dans le cadre de cette recherche.

Consignes : Encercler à quel point chaque symptôme ci-dessous vous affecte en ce moment.

1. Inconfort général

Pas du tout	Un peu	Modérément	Sévèrement
-------------	--------	------------	------------

2. Fatigue

Pas du tout	Un peu	Modérément	Sévèrement
-------------	--------	------------	------------

3. Mal de tête

Pas du tout	Un peu	Modérément	Sévèrement
-------------	--------	------------	------------

4. Fatigue des yeux

Pas du tout	Un peu	Modérément	Sévèrement
-------------	--------	------------	------------

5. Nausée

Pas du tout	Un peu	Modérément	Sévèrement
-------------	--------	------------	------------

6. Difficulté de concentration

Pas du tout	Un peu	Modérément	Sévèrement
-------------	--------	------------	------------

7. Vision embrouillée

Pas du tout

Un peu

Modérément

Sévèrement

Consignes : Veuillez évaluer votre sentiment d'avoir été dans la classe virtuelle sur l'échelle suivante de 1 à 7 (1 = pas du tout et 7 = totalement)

1. Dans quelle mesure y a-t-il eu des moments, au cours de l'expérience, où la classe virtuelle était la réalité pour vous ? (Est ce que, par moment, la classe virtuelle vous a paru réelle, comme si c'était la réalité ?)

1 2 3 4 5 6 7

2. Lorsque vous repensez à ce que vous venez de vivre, pensez-vous à la classe virtuelle plutôt comme à des images que vous avez vues (comme sur un écran - 1), ou plutôt comme à un endroit que vous avez visité (comme dans un souvenir - 7) ?

1 2 3 4 5 6 7

3. Lorsque vous repensez à votre expérience, est ce que vous pensiez être vraiment dans une classe (virtuelle) ou plutôt dans un autre endroit ? (L'endroit où vous étiez avant de mettre le casque par exemple) ?

1 2 3 4 5 6 7

4. Repensez à votre souvenir d'avoir été dans la classe virtuelle. Dans quelle mesure la structure de ce souvenir est-elle similaire à celle des autres lieux où vous vous êtes rendu aujourd'hui ? (Est ce que, lorsque vous y pensez, vous pourriez comparer ce souvenir à quelque chose de réel, ou bien est-ce que pour vous, se souvenir est totalement issu d'un jeu vidéo par exemple)

1 2 3 4 5 6 7

5. Pendant l'expérience, est ce que vous avez pensé être réellement dans la salle de classe (même ponctuellement) ?

1 2 3 4 5 6 7

Consignes : Donnez votre avis sur les propositions qui suivent, sur une échelle de 1 à 7 (1 = pas du tout et 7 = totalement)

1. Les mains virtuelles bougeaient comme je le voulais, comme si elles obéissaient à ma volonté

1 2 3 4 5 6 7

2. J'ai eu l'impression de contrôler le mouvement des mains virtuelles

1 2 3 4 5 6 7

3. J'ai eu l'impression que les mains virtuelles avaient une volonté propre

1 2 3 4 5 6 7

4. J'ai eu l'impression de toucher les objets avec les mains virtuelles comme si c'était les miennes

1 2 3 4 5 6 7

5. J'ai senti que les mains virtuelles n'étaient pas les miennes

1 2 3 4 5 6 7

6. Quand l'enseignante s'est approchée de mon avatar et que l'environnement virtuel est devenu stressant, j'ai été affecté par cet évènement

1 2 3 4 5 6 7

7. A la fin de l'expérience, j'ai été affecté par la réaction de mon avatar face aux autres élèves

1 2 3 4 5 6 7

8. Le casque audio a eu un effet reposant sur moi lorsque je l'ai mis sur la tête de mon avatar

1 2 3 4 5 6 7

9. J'ai ressenti les émotions de mon avatar comme si je vivais vraiment cette expérience

1 2 3 4 5 6 7

10. Le comportement des autres agents virtuels (les élèves et l'enseignante) qui ont perturbé mon avatar m'ont également perturbé

1 2 3 4 5 6 7

11. Le rythme cardiaque que mon avatar ressentait a eu un effet sur moi

1 2 3 4 5 6 7

12. Le son amplifié que mon avatar entendait était perturbant pour moi

1 2 3 4 5 6 7

Consignes : Voici un QCM, questionnaire à choix multiples, veuillez cocher les affirmations qui vous semble juste. (Il faudrait cibler les réponses justes. S'il n'y a qu'une réponse fausse, alors inverse la question en disant : que faut-il éviter de faire, ou de quoi les personnes n'ont pas besoin. Idéalement on doit avoir 1 réponse correcte à coder pour les scores ensuite.)

Pour les personnes avec TSA SDI, la plupart du temps, la surstimulation sensorielle (son et lumière)

- a. Peut être dérangeante
- b. Peut entraîner des crises ou réactions extrêmes
- c. Est nécessaire pour qu'ils puissent travailler
- d. Je ne sais pas

Pendant qu'elles cherchent des solutions pour résoudre la tâche, les personnes avec TSA SDI ont la capacité à

- a. Trouver des réponses sans aucun élément
- b. Effacer définitivement la présence de ce qui les dérange
- c. Voir ce qui échappe aux non TSA
- d. Je ne sais pas

Après un événement stressant, les personnes avec TSA SDI n'ont surtout pas besoin

- a. D'être stimulées pour reprendre confiance en elles
- b. De brièvement s'isoler socialement
- c. De réduire les bruits de l'environnement
- d. Je ne sais pas

Lorsqu'un enseignant voit qu'une personne avec TSA SDI est concentrée sur une tâche, il est conseillé de

- a. L'interpeller, même avec bienveillance

- b. L'encourager pour qu'elle avance plus vite
- c. La laisser faire et attendre qu'elle est finie
- d. Je ne sais pas

Après une situation stressante, une personne avec TSA SDI a le plus souvent besoin

- a. D'être encouragé, voire secoué, pour se reprendre rapidement
- b. De silence ou d'écouter des sons apaisants
- c. D'être mise en confiance verbalement par la personne à l'origine du stress
- d. Je ne sais pas

Un enseignant peut interrompre une personne avec TSA plongée dans sa réflexion

- a. Pour l'aider en apportant des éléments facilitateurs
- b. Pour lui donner la réponse et éviter d'être dans l'embarras devant les autres
- c. En aucun cas
- d. Je ne sais pas

Consignes : Donnez votre avis sur les propositions qui suivent, sur une échelle de 1 à 7 (1 = pas du tout d'accord et 7 = totalement d'accord)

1. En cas d'interaction avec autrui, il arrive que les personnes avec TSA SDI n'arrivent pas à percevoir les intentions positives ou négatives de l'interlocuteur et peuvent prendre certaines interactions pour des agressions

1 2 3 4 5 6 7

2. En cas d'interaction avec autrui, il arrive que les personnes avec TSA SDI ne se rendent pas du tout compte que quelqu'un leur parle

1 2 3 4 5 6 7

3. Lors de la résolution d'une tâche ou d'un problème, une personne avec TSA SDI perçoit des informations dans l'environnement et rend la tâche plus complexe.

1 2 3 4 5 6 7

4. Lors de la résolution d'une tâche, une personne avec TSA SDI peut inventer des réponses avec très peu d'éléments

1 2 3 4 5 6 7

5. Durant la résolution de cette tâche, les personnes avec TSA SDI ont tendance à créer des cartes mentales en décomposant le problème en plusieurs blocs ou morceaux

1 2 3 4 5 6 7

6. Durant la résolution de cette tâche, les personnes avec TSA SDI sont capables de calculer mentalement des probabilités mathématiques avec des données basiques

1 2 3 4 5 6 7

7. Une personne avec TSA SDI peut prendre la parole en public si elle a été préparée préalablement

1 2 3 4 5 6 7

8. Les personnes avec TSA SDI ne sont pas limitées dans leur apprentissage et peuvent acquérir des connaissances comme les autres, mais juste différemment

1 2 3 4 5 6 7

Question ouverte : répondez librement à cette question

Selon vous, comment doit se comporter un enseignant avec une personne TSA SDI dans sa pratique pédagogique et sa communication avec la personne ?

.....
.....
.....
.....
.....
.....

D.3 Batterie de mesures expérimentation 2

D.3.1 Questionnaire complet expérimentation 2

Age – niveau scolaire – genre – profession

Cinétose :

Consignes : Encerclez à quel point chaque symptôme ci-dessous vous affecte en ce moment sur l'échelle suivante de 1 à 4 (1 = pas du tout et 4 = totalement).

1. Inconfort général
2. Fatigue
3. Mal de tête
4. Fatigue des yeux
5. Nausée
6. Difficulté de concentration
7. Vision embrouillée

Présence :

Veillez évaluer votre sentiment d'être dans la classe virtuelle sur l'échelle suivante de 1 à 7, où 7 représente votre expérience normale d'être dans un lieu. *J'ai eu l'impression « d'être là » dans la classe virtuelle.*

1. Dans quelle mesure y a-t-il eu des moments, au cours de l'expérience, où la classe virtuelle était la réalité pour vous ? (Est ce que, par moment, la classe virtuelle vous a paru réelle, comme si c'était la réalité ?)
2. Lorsque vous repensez à ce que vous venez de vivre, pensez-vous à la classe virtuelle plutôt comme à des images que vous avez vues (comme sur un écran - 1), ou plutôt comme à un endroit que vous avez visité (comme dans un souvenir - 7) ?
3. Lorsque vous repensez à votre expérience, est ce que vous pensiez être vraiment dans une classe (virtuelle) ou plutôt dans un autre endroit ? (L'endroit où vous étiez avant de mettre le casque par exemple) ?

4. Repensez à votre souvenir d'avoir été dans la classe virtuelle. Dans quelle mesure la structure de ce souvenir est-elle similaire à celle des autres lieux où vous vous êtes rendu aujourd'hui ? (Est ce que, lorsque vous y pensez, vous pourriez comparer ce souvenir à quelque chose de réel, ou bien est-ce que pour vous, se souvenir est totalement issu d'un jeu vidéo par exemple)

5. Pendant l'expérience, est ce que vous avez pensé être réellement dans la salle de classe (même ponctuellement) ?

Agentivité :

Consignes : Donnez votre avis sur les propositions qui suivent, sur une échelle de 1 à 7

(1) Pas du tout d'accord - (2) Pas d'accord - (3) Plutôt pas d'accord – (4) Ni d'accord / ni pas d'accord - (5) Plutôt d'accord - (6) D'accord - (7) Tout à fait d'accord

1. Les mains virtuelles bougeaient comme je le voulais, comme si elles obéissaient à ma volonté
2. J'ai eu l'impression de contrôler le mouvement des mains virtuelles
3. J'ai eu l'impression que les mains virtuelles avaient une volonté propre
4. J'ai eu l'impression de toucher les objets avec les mains virtuelles comme si c'était les miennes
5. J'ai senti que les mains virtuelles n'étaient pas les miennes

Empathie :

Consignes : Donnez votre avis sur les propositions qui suivent, sur une échelle de 1 à 7

(1) Pas du tout d'accord - (2) Pas d'accord - (3) Plutôt pas d'accord – (4) Ni d'accord / ni pas d'accord - (5) Plutôt d'accord - (6) D'accord - (7) Tout à fait d'accord

6. Quand l'enseignante s'est approchée de mon avatar et que l'environnement virtuel est devenu stressant, j'ai été affecté par cet événement
7. A la fin de l'expérience, j'ai été affecté par la réaction de mon avatar face aux autres élèves
8. Le casque audio a eu un effet reposant sur moi lorsque je l'ai mis sur la tête de mon avatar
9. J'ai ressenti les émotions de mon avatar comme si je vivais vraiment cette expérience
10. Le comportement des autres agents virtuels (les élèves et l'enseignante) qui ont perturbé mon avatar m'ont également perturbé

11. Le rythme cardiaque que mon avatar ressentait a eu un effet sur moi

12. Le son amplifié que mon avatar entendait était perturbant pour moi

Attitude des participants envers le TSA

Merci d'avoir accepté de répondre à ce questionnaire. Veuillez répondre aussi honnêtement que possible. Il n'y a pas de « bonnes » réponses.

(1) Pas du tout d'accord - (2) Pas d'accord - (3) Plutôt pas d'accord – (4) Plutôt d'accord - (5) D'accord - (6) Tout à fait d'accord

1. Les personnes autistes (avec Trouble du Spectre de l'Autisme Sans Déficience Intellectuelle - TSA SDI) peuvent tout à fait réussir à l'université

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6

2. Les personnes autistes (TSA SDI) devraient avoir l'opportunité d'aller à l'université.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6

3. Les personnes autistes (TSA SDI) devraient être institutionnalisées pour leur sécurité et celle des autres.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6

4. J'aurais peur en présence d'une personne autistes (TSA SDI)

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6

5. Je serais à l'aise de m'asseoir à côté d'une personne autistes (TSA SDI) dans la même classe.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6

6. Les enfants autistes (TSA SDI) devraient être pleinement intégrés dans les classes ordinaires.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6

7. Les personnes autistes (TSA SDI) ne peuvent pas comprendre les sentiments des autres.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6

8. Les élèves autistes (TSA SDI) qui sont intégrés dans des classes ordinaires constituent une distraction pour les élèves non autistes de cette classe.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6

Charges cognitives des participants :

Consignes : Donnez votre avis sur les propositions qui suivent, sur une échelle de 1 à 7

(1) Pas du tout d'accord - (2) Pas d'accord - (3) Plutôt pas d'accord – (4) Ni d'accord / ni pas d'accord - (5) Plutôt d'accord - (6) *D'accord* - (7) Tout à fait d'accord

1. J'ai trouvé qu'il était difficile de comprendre dans cette activité comment un apprenant avec Trouble du Spectre de l'Autisme Sans Déficience Intellectuelle (TSA SDI) pouvait traiter l'information et apprendre

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

2. J'ai dû investir un effort mental important dans cette activité pour apprendre comment procède un apprenant TSA SDI pour acquérir des connaissances.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Connaissances sur le TSA :

Les propositions qui vont suivre concernent la situation d'apprentissage que vous venez de vivre lors de la simulation. Elles se réfèrent à ce que peuvent vivre les apprenants avec Trouble du Spectre de l'Autisme Sans Déficience Intellectuelle (TSA SDI) dans un contexte scolaire.

Consignes : Donnez votre avis sur les propositions qui suivent, sur une échelle de 1 à 7

(1) Pas du tout d'accord - (2) Pas d'accord - (3) Plutôt pas d'accord – (4) Ni d'accord / ni pas d'accord - (5) Plutôt d'accord - (6) D'accord - (7) Tout à fait d'accord

Q1 : Les apprenants avec Trouble du Spectre de l'Autisme Sans Déficience Intellectuelle (TSA SDI) peuvent parfois mal comprendre les intentions (positives ou négatives) de leurs interlocuteurs et voir certaines interactions comme des agressions.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q2 : Les apprenants avec TSA SDI peuvent ne pas remarquer que quelqu'un leur parle.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q3 : Quand ils résolvent un problème, les apprenants avec TSA SDI remarquent beaucoup de détails autour d'eux, ce qui peut compliquer la tâche.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q4 : Pour résoudre un problème, les apprenants avec TSA SDI préfèrent souvent le diviser en plusieurs petites parties.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q5 : Les apprenants avec Trouble du Spectre de l'Autisme Sans Déficience Intellectuelle (TSA SDI) peuvent prendre la parole devant leurs camarades s'ils ont eu le temps de s'y préparer.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q6 : Les apprenants TSA SDI ont une approche différente de l'apprentissage.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q7 : Les bruits, lumières ou autres stimulations intenses peuvent gêner les apprenants avec TSA SDI.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q8 : Dans une salle de classe, la surstimulation sensorielle provoque systématiquement des crises ou des réactions intenses chez les apprenants avec TSA SDI

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q9 : Quand ils cherchent des solutions à un problème, les apprenants avec TSA SDI remarquent souvent des détails que d'autres ne voient pas.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q10 : Après une situation stressante, les apprenants avec Trouble du Spectre de l'Autisme Sans Déficiência Intellectuelle (TSA SDI) préfèrent souvent le silence ou écouter des sons apaisants pour se calmer.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q11 : Quand un apprenant avec TSA SDI est concentré sur une tâche, il est conseillé de le laisser terminer avant d'intervenir.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q12 : Un enseignant peut interrompre un apprenant avec TSA SDI pour l'accompagner en lui donnant des informations utiles.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

Q13 : Après une interaction stressante, les apprenants s avec TSA SDI ont besoin d'être rassurées par la personne a l'origine de ce stress.

Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord / ni pas d'accord	Plutôt d'accord	D'accord	Tout à fait d'accord
1	2	3	4	5	6	7

ANNEXE E - RÉSULTATS STATISTIQUES

Cette section présente l'ensemble des analyses statistiques complètes conduites sur les données des deux expérimentations. Elle comprend les tableaux détaillés des analyses de variance (ANOVA) à mesures répétées, les tests post-hoc de comparaisons multiples. Cette documentation vise à fournir une transparence maximale sur les résultats obtenus et à permettre la vérification et la réplication des analyses. Les résultats sont organisés par expérimentation, facilitant la consultation et la compréhension des effets observés.

E.1 Résultats expérimentation 1

Cette sous-section présente les résultats statistiques complets de la première expérimentation, organisés par variable dépendante.

E.1.1 Analyses de variance (ANOVA) à mesures répétées

Les tableaux suivants présentent les résultats des ANOVA à mesures répétées pour chaque variable dépendante mesurée à différents temps de mesure (pré-test, post-test immédiat, post-test différé).

Partie QCM

Anova pour mesures répétées

Anova pour mesures répétées

Effets intra-sujets

	Somme des carrés	ddl	Carrés moyens	F	p	η^2_p
Mesures répétées - facteur 1	1.21	2	0.603	2.75	0.067	0.031
Mesures répétées - facteur 1 * Condition	2.76	6	0.460	2.10	0.056	0.069
Résidu	37.27	170	0.219			

Note. Somme des carrés de type 3

Effets inter-sujets

	Somme des carrés	ddl	Carrés moyens	F	p	η^2_p
Condition	3.04	3	1.014	1.49	0.224	0.050
Résidu	57.94	85	0.682			

Note. Somme des carrés de type 3

Partie échelle de Likert

E.1.2 Tests post-hoc de comparaisons multiples

Les tableaux suivants présentent les résultats des tests post-hoc (correction de Tukey HSD) pour les comparaisons par paires entre les différents temps de mesure, lorsque l'effet principal du temps était significatif dans l'ANOVA.

Post hoc Tuckey exp 1 QCM:

Comparaisons post hoc - Mesures répétées - facteur 1 * Condition

		Comparaison		Différence moyenne	Erreur standard	ddl	t	Ptukey	
Mesures répétées - facteur 1	Condition	Mesures répétées - facteur 1	Condition						
Pré questionnaire	FS	- Pré questionnaire	PP	-0.18635	0.1560	91.0	-1.1946	0.988	
		- Pré questionnaire	PP-A	-0.66630	0.1575	91.0	-4.2312	0.003	
		- Pré questionnaire	VR	-0.37850	0.1668	91.0	-2.2687	0.506	
		- Quest immédiat	FS	-0.81292	0.1216	91.0	-6.6827	< .001	
		- Quest immédiat	PP	-0.80827	0.1305	91.0	-6.1953	< .001	
		- Quest immédiat	PP-A	-0.71430	0.1311	91.0	-5.4471	< .001	
		- Quest immédiat	VR	-0.73600	0.1354	91.0	-5.4359	< .001	
		- Quest différé	FS	-0.70875	0.1627	91.0	-4.3566	0.002	
		- Quest différé	PP	-0.70635	0.1747	91.0	-4.0437	0.006	
		- Quest différé	PP-A	-0.14790	0.1767	91.0	-0.8370	0.999	
		- Quest différé	VR	-0.74300	0.1894	91.0	-3.9230	0.009	
		PP	- Pré questionnaire	PP-A	-0.47995	0.1544	91.0	-3.1094	0.096
	- Pré questionnaire		VR	-0.19215	0.1639	91.0	-1.1724	0.990	
	- Quest immédiat		FS	-0.62657	0.1281	91.0	-4.8909	< .001	
	- Quest immédiat		PP	-0.62192	0.1169	91.0	-5.3214	< .001	
	- Quest immédiat		PP-A	-0.52795	0.1274	91.0	-4.1451	0.004	
	- Quest immédiat		VR	-0.54965	0.1318	91.0	-4.1719	0.004	
	- Quest différé		FS	-0.52240	0.1761	91.0	-2.9657	0.136	
	- Quest différé		PP	-0.52000	0.1563	91.0	-3.3269	0.054	
	- Quest différé		PP-A	0.03845	0.1739	91.0	0.2210	1.000	
	- Quest différé		VR	-0.55665	0.1868	91.0	-2.9798	0.132	
	PP-A		- Pré questionnaire	VR	0.28780	0.1653	91.0	1.7409	0.844
			- Quest immédiat	FS	-0.14662	0.1299	91.0	-1.1285	0.993
		- Quest immédiat	PP	-0.14197	0.1285	91.0	-1.1047	0.994	
		- Quest immédiat	PP-A	-0.04800	0.1192	91.0	-0.4027	1.000	
		- Quest immédiat	VR	-0.06970	0.1335	91.0	-0.5220	1.000	
		- Quest différé	FS	-0.04245	0.1775	91.0	-0.2392	1.000	
		- Quest différé	PP	-0.04005	0.1732	91.0	-0.2312	1.000	
		- Quest différé	PP-A	0.51840	0.1594	91.0	3.2522	0.066	
		- Quest différé	VR	-0.07670	0.1881	91.0	-0.4079	1.000	
VR		- Quest immédiat	FS	-0.43442	0.1411	91.0	-3.0783	0.104	
		- Quest immédiat	PP	-0.42977	0.1398	91.0	-3.0736	0.105	
		- Quest immédiat	PP-A	-0.33580	0.1405	91.0	-2.3909	0.423	
	- Quest immédiat	VR	-0.35750	0.1333	91.0	-2.6828	0.252		
	- Quest différé	FS	-0.33025	0.1858	91.0	-1.7772	0.826		
	- Quest différé	PP	-0.32785	0.1818	91.0	-1.8036	0.812		
	- Quest différé	PP-A	0.23060	0.1837	91.0	1.2551	0.983		
	- Quest différé	VR	-0.36450	0.1782	91.0	-2.0453	0.662		
	Quest immédiat	FS	- Quest immédiat	PP	0.00465	0.0954	91.0	0.0487	1.000
			- Quest immédiat	PP-A	0.09862	0.0963	91.0	1.0239	0.997
			- Quest immédiat	VR	0.07692	0.1020	91.0	0.7538	1.000

Comparaisons post hoc - Mesures répétées - facteur 1 * Condition

Comparaison				Différence moyenne	Erreur standard	ddl	t	Ptuke	
Mesures répétées - facteur 1	Condition	Mesures répétées - facteur 1	Condition						
Quest diffé	PP	- Quest diffé	FS	0.10417	0.1430	91.0	0.7282	1.000	
		- Quest diffé	PP	0.10657	0.1503	91.0	0.7090	1.000	
		- Quest diffé	PP-A	0.66502	0.1527	91.0	4.3559	0.002	
		- Quest diffé	VR	0.06992	0.1672	91.0	0.4182	1.000	
		- Quest immé	PP-A	0.09397	0.0944	91.0	0.9954	0.997	
		- Quest immé	VR	0.07227	0.1002	91.0	0.7209	1.000	
		- Quest diffé	FS	0.09952	0.1540	91.0	0.6462	1.000	
		- Quest diffé	PP	0.10192	0.1374	91.0	0.7416	1.000	
		- Quest diffé	PP-A	0.66037	0.1515	91.0	4.3597	0.002	
		- Quest diffé	VR	0.06527	0.1661	91.0	0.3930	1.000	
		- Quest immé	VR	-0.02170	0.1011	91.0	-0.2146	1.000	
		- Quest diffé	FS	0.00555	0.1546	91.0	0.0359	1.000	
	PP-A	- Quest diffé	PP	0.00795	0.1497	91.0	0.0531	1.000	
		- Quest diffé	PP-A	0.56640	0.1402	91.0	4.0414	0.006	
		- Quest diffé	VR	-0.02870	0.1666	91.0	-0.1722	1.000	
		VR	- Quest diffé	FS	0.02725	0.1582	91.0	0.1722	1.000
			- Quest diffé	PP	0.02965	0.1534	91.0	0.1933	1.000
			- Quest diffé	PP-A	0.58810	0.1557	91.0	3.7762	0.014
	Quest diffé	FS	- Quest diffé	VR	-0.00700	0.1567	91.0	-0.0447	1.000
			- Quest diffé	PP	0.00240	0.1929	91.0	0.0125	1.000
			- Quest diffé	PP-A	0.56085	0.1947	91.0	2.8800	0.166
		PP	- Quest diffé	VR	-0.03425	0.2063	91.0	-0.1660	1.000
			- Quest diffé	PP-A	0.55845	0.1909	91.0	2.9256	0.150
			- Quest diffé	VR	-0.03665	0.2027	91.0	-0.1808	1.000
PP-A		- Quest diffé	VR	-0.59510	0.2044	91.0	-2.9109	0.155	

Post hoc Tuckey exp 1 Likert

Comparaisons post hoc - Mesures répétées - facteur 1 * Condition

Comparaison								
Mesures répétées - facteur 1	Condition	Mesures répétées - facteur 1	Condition	Différence moyenne	Erreur standard	ddl	t	Ptukey
Quest immédiat	FS	- Quest immédiat	PP	0.20208	0.189	85.0	1.07029	0.995
		- Quest immédiat	PP-A	0.14926	0.206	85.0	0.72586	1.000
		- Quest immédiat	VR	0.10750	0.161	85.0	0.66576	1.000
		- Quest différé	FS	0.46650	0.173	85.0	2.68943	0.250
		- Quest différé	PP	0.23131	0.171	85.0	1.35528	0.969
		- Quest différé	PP-A	0.20716	0.183	85.0	1.13410	0.992
		- Quest différé	VR	0.03950	0.151	85.0	0.26156	1.000
		- Quest immédiat	PP	0.17474	0.194	85.0	0.90188	0.999
		- Quest immédiat	PP-A	0.12193	0.210	85.0	0.58009	1.000
		- Quest immédiat	VR	0.08017	0.207	85.0	0.38684	1.000
	- Quest différé	FS	0.43917	0.119	85.0	3.70175	0.018	
	- Quest différé	PP	0.20397	0.176	85.0	1.15808	0.991	
	- Quest différé	PP-A	0.17982	0.188	85.0	0.95766	0.998	
	- Quest différé	VR	0.01217	0.186	85.0	0.06553	1.000	
	- Quest immédiat	PP-A	-0.05281	0.207	85.0	-0.25565	1.000	
	- Quest immédiat	VR	-0.09458	0.204	85.0	-0.46457	1.000	
	- Quest différé	FS	0.26442	0.175	85.0	1.51461	0.932	
	- Quest différé	PP	0.02923	0.114	85.0	0.25645	1.000	
	- Quest différé	PP-A	0.00508	0.184	85.0	0.02765	1.000	
	- Quest différé	VR	-0.16258	0.182	85.0	-0.89535	0.999	
Quest différé	PP-A	- Quest immédiat	VR	-0.04176	0.219	85.0	-0.19045	1.000
		- Quest différé	FS	0.31724	0.193	85.0	1.64662	0.886
		- Quest différé	PP	0.08204	0.190	85.0	0.43146	1.000
		- Quest différé	PP-A	0.05789	0.133	85.0	0.43420	1.000
		- Quest différé	VR	-0.10976	0.199	85.0	-0.55151	1.000
	VR	- Quest différé	FS	0.35900	0.189	85.0	1.89513	0.759
		- Quest différé	PP	0.12381	0.187	85.0	0.66248	1.000
		- Quest différé	PP-A	0.09966	0.198	85.0	0.50358	1.000
		- Quest différé	VR	-0.06800	0.130	85.0	-0.52323	1.000
		- Quest différé	PP	-0.23519	0.155	85.0	-1.51943	0.931
FS	- Quest différé	PP-A	-0.25934	0.168	85.0	-1.54444	0.923	
	- Quest différé	VR	-0.42700	0.166	85.0	-2.57912	0.308	
	- Quest différé	PP-A	-0.02415	0.165	85.0	-0.14633	1.000	
	- Quest différé	VR	-0.19181	0.163	85.0	-1.17934	0.989	
	- Quest différé	VR	-0.16766	0.175	85.0	-0.95705	0.998	

Comparaisons post hoc - Mesures répétées - facteur 1 * Condition

		Comparaison		Différence moyenne	Erreur standard	ddl	t	Ptukey	
Mesures répétées - facteur 1	Condition	Mesures répétées - facteur 1	Condition						
Pré questionnaire	FS	- Pré questionnaire	PP	-0.05003	0.168	85.0	-0.29767	1.000	
		- Pré questionnaire	PP-A	-0.28700	0.182	85.0	-1.57400	0.914	
		- Pré questionnaire	VR	-0.48942	0.180	85.0	-2.72242	0.234	
		- Quest immédiat	FS	-0.46208	0.147	85.0	-3.13486	0.091	
		- Quest immédiat	PP	-0.28734	0.181	85.0	-1.58875	0.908	
		- Quest immédiat	PP-A	-0.34015	0.198	85.0	-1.71479	0.856	
		- Quest immédiat	VR	-0.38192	0.195	85.0	-1.95621	0.721	
		- Quest différé	FS	-0.02292	0.138	85.0	-0.16623	1.000	
		- Quest différé	PP	-0.25811	0.162	85.0	-1.59488	0.906	
		- Quest différé	PP-A	-0.28226	0.174	85.0	-1.61812	0.898	
		- Quest différé	VR	-0.44992	0.172	85.0	-2.61328	0.289	
		PP	- Pré questionnaire	PP-A	-0.23696	0.179	85.0	-1.32227	0.974
	- Pré questionnaire		VR	-0.43938	0.177	85.0	-2.48800	0.362	
	- Quest immédiat		FS	-0.41205	0.182	85.0	-2.26547	0.509	
	- Quest immédiat		PP	-0.23731	0.142	85.0	-1.67568	0.874	
	- Quest immédiat		PP-A	-0.29012	0.195	85.0	-1.48403	0.941	
	- Quest immédiat		VR	-0.33188	0.192	85.0	-1.72572	0.851	
	- Quest différé		FS	0.02712	0.161	85.0	0.16810	1.000	
	- Quest différé		PP	-0.20808	0.132	85.0	-1.57098	0.915	
	- Quest différé		PP-A	-0.23223	0.171	85.0	-1.35673	0.968	
	- Quest différé		VR	-0.39988	0.169	85.0	-2.36825	0.439	
	PP-A		- Pré questionnaire	VR	-0.20242	0.190	85.0	-1.06414	0.995
			- Quest immédiat	FS	-0.17509	0.195	85.0	-0.89727	0.999
		- Quest immédiat	PP	-3.44e-4	0.191	85.0	-0.00180	1.000	
		- Quest immédiat	PP-A	-0.05316	0.166	85.0	-0.32088	1.000	
		- Quest immédiat	VR	-0.09492	0.205	85.0	-0.46327	1.000	
		- Quest différé	FS	0.26408	0.176	85.0	1.49950	0.937	
		- Quest différé	PP	0.02889	0.173	85.0	0.16662	1.000	
		- Quest différé	PP-A	0.00474	0.155	85.0	0.03057	1.000	
	VR	- Quest différé	VR	-0.16292	0.183	85.0	-0.89004	0.999	
		- Quest immédiat	FS	0.02733	0.193	85.0	0.14181	1.000	

E.2 Résultats expérimentation 2

Cette sous-section présente les résultats statistiques complets de la seconde expérimentation, organisés par variable dépendante.

E.2.1 Analyses de variance (ANOVA) à mesures répétées

Anova pour mesures répétées - Partie connaissances

Effets intra-sujets

	Somme des carrés	ddl	Carrés moyens	F	p	η^2_p
Temps de la mesure	12.214	2	6.1070	73.613	< .001	0.564
Temps de la mesure * Condition	0.101	2	0.0506	0.609	0.545	0.011
Résidu	9.457	114	0.0830			

Note. Somme des carrés de type 3

Effets inter-sujets

	Somme des carrés	ddl	Carrés moyens	F	p	η^2_p
Condition	0.0699	1	0.0699	0.209	0.649	0.004
Résidu	19.0565	57	0.3343			

Note. Somme des carrés de type 3

Anova pour mesures répétées – Partie attitudes

Effets intra-sujets

	Somme carrés	des ddl	Carrés moyens	F	p	η^2_p
Temps de la mesure	1.390	2	0.695	6.34	0.002	0.100
Temps de la mesure * Condition	0.725	2	0.362	3.31	0.040	0.055
Résidu	12.492	114	0.110			

Note. Somme des carrés de type 3

Effets inter-sujets

	Somme des carrés	ddl	Carrés moyens	F	p	η^2_p
Condition	0.928	1	0.928	1.94	0.169	0.033
Résidu	27.297	57	0.479			

Note. Somme des carrés de type 3

E.2.2 Tests ost-Hohc de comparaisons multiples

Anova pour mesures répétées - Partie connaissances

Tests post hoc

Comparaisons post hoc - Condition

Comparaison		Différence moyenne	Erreur standard	ddl	t	ptukey
Condition	Condition					
VR	- FS	0.0398	0.0870	57.0	0.457	0.649

Comparaisons post hoc - Temps de la mesure

Comparaison		Différence moyenne	Erreur standard	ddl	t	ptukey
Temps de la mesure	Temps de la mesure					
Pré quest	- Quest immédiat	-0.5872	0.0567	57.0	-10.36	< .001
	- Quest différé	-0.5232	0.0592	57.0	-8.83	< .001
Quest immédiat	- Quest différé	0.0640	0.0417	57.0	1.53	0.283

Comparaisons post hoc - Temps de la mesure * Condition

Comparaison									
Temps de la mesure	Condition		Temps de la mesure	Condition	Différence moyenne	Erreur standard	ddl	t	p _{Tukey}
Pré quest	VR	-	Pré quest	FS	0.0987	0.1076	57.0	0.917	0.941
		-	Quest immédiat	VR	-0.5286	0.0822	57.0	-6.433	< .001
		-	Quest immédiat	FS	-0.5471	0.1068	57.0	-5.125	< .001
		-	Quest différé	VR	-0.4936	0.0859	57.0	-5.747	< .001
	-	Quest différé	FS	-0.4542	0.1068	57.0	-4.252	0.001	
	FS	-	Quest immédiat	VR	-0.6272	0.1067	57.0	-5.880	< .001
		-	Quest immédiat	FS	-0.6458	0.0781	57.0	-8.270	< .001
		-	Quest différé	VR	-0.5922	0.1068	57.0	-5.548	< .001
-		Quest différé	FS	-0.5529	0.0816	57.0	-6.774	< .001	
Quest immédiat	VR	-	Quest immédiat	FS	-0.0186	0.1058	57.0	-0.176	1.000

Comparaisons post hoc - Temps de la mesure * Condition

Comparaison								
Temps de la mesure	Condition	Temps de la mesure	Condition	Différence moyenne	Erreur standard	ddl	t	p _{tukey}
		- Quest différé	VR	0.0350	0.0604	57.0	0.579	0.992
		- Quest différé	FS	0.0743	0.1059	57.0	0.702	0.981
	FS	- Quest différé	VR	0.0536	0.1059	57.0	0.506	0.996
		- Quest différé	FS	0.0929	0.0574	57.0	1.618	0.590
Quest différé	VR	- Quest différé	FS	0.0393	0.1060	57.0	0.371	0.999

Anova pour mesures répétées – Partie attitudes

Tests post hoc

Comparaisons post hoc - Condition

Comparaison							
Condition	Condition	Différence moyenne	Erreur standard	ddl	t	p _{tukey}	
VR	- FS	0.145	0.104	57.0	1.39	0.169	

Comparaisons post hoc - Temps de la mesure

Comparaison							
Temps de la mesure	Temps de la mesure	Différence moyenne	Erreur standard	ddl	t	p _{Tukey}	
Pré quest	- Quest immédiat	-0.2015	0.0515	57.0	-3.910	<.001	
	- Quest différé	-0.0303	0.0708	57.0	-0.427	0.904	
Quest immédiat	- Quest différé	0.1712	0.0592	57.0	2.894	0.015	

Comparaisons post hoc - Temps de la mesure * Condition

Comparaison									
Temps de la mesure	Condition	Temps de la mesure	Condition	Différence moyenne	Erreur standard	ddl	t	p _{Tukey}	
Pré quest	VR	- Pré quest	FS	0.1771	0.1235	57.0	1.435	0.706	
		- Quest immédiat	VR	-0.2546	0.0747	57.0	-3.408	0.015	
		- Quest immédiat	FS	0.0287	0.1209	57.0	0.238	1.000	
	FS	- Quest différé	VR	0.0711	0.1027	57.0	0.692	0.982	
		- Quest différé	FS	0.0455	0.1291	57.0	0.352	0.999	
		- Quest immédiat	VR	-0.4318	0.1207	57.0	-3.578	0.009	

Comparaisons post hoc - Temps de la mesure * Condition

Comparaison								
Temps de la mesure	Condition	Temps de la mesure	Condition	Différence moyenne	Erreur standard	ddl	t	Ptukey
		Quest immédiat	FS	-0.1484	0.0710	57.0	-2.090	0.307
		Quest différé	VR	-0.1061	0.1297	57.0	-0.818	0.963
		Quest différé	FS	-0.1316	0.0976	57.0	-1.349	0.757
Quest immédiat	VR	Quest immédiat	FS	0.2834	0.1181	57.0	2.400	0.174
		Quest différé	VR	0.3257	0.0858	57.0	3.797	0.005
	FS	Quest différé	FS	0.3002	0.1265	57.0	2.373	0.183
		Quest différé	VR	0.0423	0.1273	57.0	0.332	0.999
		Quest différé	FS	0.0168	0.0815	57.0	0.206	1.000
Quest différé	VR	Quest différé	FS	-0.0256	0.1351	57.0	-0.189	1.000