

THESE / UNIVERSITE RENNES 2

sous le sceau de l'Université Européenne de Bretagne

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITE RENNES 2

Mention : Psychologie

Ecole doctorale Sciences Humaines et Sociales

présentée par

Lauriane POULIQUEN-LARDY

Préparée à l'Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN)

UMR CNRS 6597

Collaboration à distance : étude de la compréhension mutuelle dans les environnements virtuels collaboratifs immersifs

Le cas de la communication spatiale

Thèse soutenue le 19 mai 2016

devant le jury composé de :

Jean-Marie BURKHARDT
Directeur de recherche, IFSTTAR / *rapporteur*

Michel DENIS
Directeur de recherche émérite, CNRS / *examinateur*

Françoise DETIENNE
Directrice de recherche, CNRS / *rapporteur*

François GUILLAUME
Ingénieur de recherche, Airbus Group / *invité*

Franck MARS
Chargé de recherche, CNRS / *directeur de thèse*

Isabelle MILLEVILLE
Chargée de recherche, CNRS / *encadrant*

Pouliquen-Lardy, Lauriane. Collaboration à distance :
étude de la compréhension mutuelle dans les environnements virtuels collaboratifs immersifs : le cas de la communication spatiale - 2016

Table des matières

Préambule 7
Chapitre 1 Contexte industriel et approche ergonomique 9
Chapitre 2 Contexte théorique – Etat de l'art 13
1. Travail à distance et EVCIs 13
1.1 Principes généraux : Réalité virtuelle et EVCIs 13
1.2 Dimension technique : contraintes et possibles 16
1.3 Dimension subjective : immersion, présence et coprésence 19
2. Collaboration 21
2.1 Représentations individuelles et partagées 22
2.2 Charge mentale et principe du moindre effort collaboratif 27
3. Espace : perception, représentation et communication 30
3.1 Plusieurs espaces d'interaction 30
3.2 Référentiels spatiaux 34
3.3 Coûts cognitifs et transformations mentales 37
3.4 Les énoncés spatialisés 44
3.5 Choix des cadres de référence 47
3.6 Résumé 53
4. Synthèse et questions de recherche 54
Chapitre 3 Effet du rôle dans la communication spatiale et la gestion du moindre effort collaboratif 57
1. Contexte 57
2. Matériels et méthodes 58
2.1 Participants 58
2.2 Matériels 59
2.3 Procédure 59
2.4 Analyse des énoncés spatiaux 62
2.5 Traitements statistiques 63
3. Résultats 64
3.1 Enoncés spatialisés 64
3.2 Questionnaires 66
4. Synthèse et discussion 68
4.1 Effet du rôle sur les cadres de référence 68
4.2 Prise de perspective et transformations mentales 69
4.3 Gestion collective des efforts 69
4.4 Référentiel commun 70
Chapitre 4 Etudes sur la production et la compréhension d'énoncés spatialisés 71
1. Etude 1 : Production d'énoncés sans indice 72
1.1 Contexte 72
1.2 Matériels et méthodes 73
1.3 Résultats - Production d'énoncés sans indice 78
1.4 Synthèse et discussion – Production d'énoncés sans indice 94
2. Etude 2 : Production d'énoncés avec indices spatiaux 99
2.1 Matériels et méthodes 99

2.2 Résultats - Production d'énoncés avec indices spatiaux 103
2.3 Synthèse et discussion – Production d'énoncés avec indices spatiaux 115
3. Etude 3 : Compréhension d'énoncés spatialisés 118
3.1 Matériels et méthodes 118
3.2 Résultats - Compréhension d'énoncés spatialisés 122
3.3 Synthèse et discussion – Compréhension d'énoncés spatialisés 133
Chapitre 5 Discussion générale 137
1. La production d'énoncés spatialisés : deux niveaux de prise de perspective 137
2. Compréhension et utilisation d'énoncés spatialisés 141
3. Le référentiel commun permet la gestion collective de la charge mentale 142
3.1 Le principe du moindre effort collaboratif se manifeste sous plusieurs formes 142
3.2 Construction du référentiel commun 144
4. Limites des études 145
5. Implications techniques et perspectives 146

Remerciements

Ce travail fait partie du projet PIVIPP géré par l'IRT Jules Verne (Institut de Recherche Technologique sur les technologies avancées de production, matériaux composites, métalliques et hybrides). Je souhaite remercier les partenaires industriels et académiques de ce projet, à savoir Airbus Group, Airbus et l'IRCCyN.

Je tiens à remercier du fond du cœur Franck MARS et Isabelle MILLEVILLE pour m'avoir offert l'opportunité de réaliser ce travail avec eux et pour leur encadrement tout au long de ces trois ans et demi. Vous avez su ménager ma susceptibilité, mais surtout vous m'avez fait confiance, cela m'a permis de m'épanouir tant personnellement que professionnellement. Merci de m'avoir encouragée et d'avoir accepté mes choix personnels que j'ai ainsi pu vivre pleinement.

Je tiens également à remercier madame Françoise DETIENNE et monsieur Jean-Marie BURKHARDT pour avoir accepté de relire ce manuscrit et d'apporter un œil critique à ce travail. De même, je souhaite remercier Michel DENIS et François GUILLAUME d'avoir accepté de participer au jury.

Cette thèse n'aurait pu aboutir sans le travail précieux de Loup VUARNESSON qui a assuré toute la partie technique pour les études décrites dans le chapitre 4. Je le remercie vivement pour son implication et sa bonne humeur qui m'ont permis de réaliser la collecte des données expérimentales dans de bonnes conditions. Merci également à Mohammed RHARDA et Florent LAROCHE pour leur soutien technique. La collecte des données n'aurait d'ailleurs pas été possible sans la mobilisation de l'ensemble des participants volontaires qui a répondu présent aux moments opportuns. Je les en remercie chaleureusement. Toute la partie expérimentale décrite dans le chapitre 3 a été réalisée au Nemo Lab, un grand merci à Jérémy LE THIEC, Sidi SET, Jean-Pierre COLLET et Yann-Henri LAUDRAIN pour leur implication.

Merci également à Camilo CHARRON avec qui j'ai passé de nombreuses heures à discuter du traitement des données. Merci pour ta patience et ta disponibilité.

Je remercie l'ensemble de mes collègues de l'IRCCyN et de l'IRT pour le soutien qu'ils ont pu m'apporter. Merci au collectif de doctorants de l'IRCCyN (Marie, Sahab, Jonathan, Johan, Fippo, Bertrand, Emna, Nadine, Maïssa, Maxime, Armel, Louis-Marie, Toussain, et ceux que j'oublie) pour les pauses animées qui m'ont bien changé les idées. Merci au collectif de doctorants de l'IRT et à Serge MASSEROT pour les journées de formation qui m'ont permis de me ressourcer.

D'un point de vue plus personnel, ce travail n'aurait pu aboutir sans le soutien de mes proches. Merci à ma famille, de sang et de cœur, de m'avoir supportée et soutenue.

Marie H., merci de m'avoir aidée à traverser cette crise. Sahab, merci pour ton sourire, ta bonne humeur et ta sincérité.

Enfin, je termine ces remerciements par deux dédicaces. La première va à mon fils, à qui je vais pouvoir consacrer plus de temps. La seconde va à celui qui m'a supportée au quotidien, qui m'a soutenue, rassurée, encouragée, à qui j'en ai fait baver mais qui est resté, qui a assuré toute la logistique ces derniers mois et qui veille sur nous. Merci pour ta confiance. A toi, Ju.

Préambule

Ce travail de thèse s'inscrit dans un projet mené à l'Institut de Recherche Technologique (IRT) Jules Verne, institution créée dans le cadre des Investissements d'avenir. Ce projet collaboratif a été monté et réalisé par deux industriels, Airbus Group, Airbus, et un laboratoire académique, l'IRCCyN (UMR CNRS 6597). Ce projet nommé PIVIPP pour *Plateau d'intégration virtuelle produit/process*, vise à développer l'utilisation de la réalité virtuelle (RV) chez les industries manipulant des pièces de grandes dimensions, telles que l'aéronautique, le naval, les énergies marines renouvelables et le transport ferroviaire. En effet, de tels outils de visualisation en trois dimensions permettent l'optimisation des processus de conception et de fabrication dans une politique de réduction des coûts. L'objectif du projet PIVIPP est de faciliter, promouvoir et élargir le spectre d'utilisation de ces équipements de haute technologie. Pour cela il s'articule autour de deux axes de recherche : le travail collaboratif à distance, dans lequel s'inscrit ce travail de recherche, et le déploiement d'interfaces tangibles reconfigurables.

L'intérêt des projets collaboratifs réalisés au sein de l'IRT Jules Verne est que les questions de recherche sont directement en lien avec des problématiques industrielles. Dans le cadre applicatif qui nous intéresse, le cas d'usage est proposé par Airbus. Il concerne l'utilisation de la réalité virtuelle pour l'intégration de compétences localisées sur au moins deux sites industriels distants. Les équipes géographiquement distribuées, connectées via leur réseau sécurisé et le réseau téléphonique, sont immergées dans la même maquette numérique (généralement la maquette produite à l'issue de la conception). Ainsi, elles peuvent par exemple anticiper des changements d'ordonnancement, c'est-à-dire anticiper les conséquences d'un transfert d'activité de montage d'une pièce d'un site de production à un autre site. Nous reviendrons plus en détail sur ce cas d'usage ainsi que sur la demande portée par les industriels dans le **Chapitre 1 Contexte industriel et approche ergonomique**. Dans ce premier chapitre, nous détaillerons également les objectifs et les méthodes employées. De façon générale, l'objectif de ce travail est d'apporter des connaissances nouvelles sur la manière dont les opérateurs se représentent et communiquent des informations spatiales pour l'amélioration des outils industriels existants. Pour cela, nous avons utilisé une approche centrée utilisateur avec le déploiement d'une méthodologie de psychologie expérimentale. Nous avons combiné cette démarche avec une approche ergonomique, notamment pour les observations initiales en situations réelles. Ce contexte industriel est un support aux travaux de recherche réalisés au cours de ce travail de thèse.

Dans le **Chapitre 2 Contexte théorique – Etat de l'art**, nous situerons notre travail par rapport aux solutions technologiques actuellement disponibles. Nous nous positionnerons également par rapport aux courants théoriques et connaissances existants dans le domaine des environnements virtuels collaboratifs. Dans la section 1 de ce chapitre, nous présenterons les contraintes liées au travail à distance, les limites imposées mais également les possibilités offertes par la technologie de la réalité virtuelle. Sans être exhaustifs sur la dimension technique, il est nécessaire de comprendre comment fonctionne la réalité virtuelle pour voir les atouts que ces dispositifs apportent comparativement à des dispositifs non immersifs, mais également pour comprendre les différences avec des situations de collaboration en présentiel. Nous présenterons ensuite les concepts clés développés autour de la collaboration (section 2). Nous détaillerons notamment le concept de référentiel commun et le principe du moindre effort collaboratif. Par rapport aux problématiques industrielles autour de la compréhension mutuelle, nous verrons comment ces concepts se traduisent au sein de la communication verbale. La section suivante (3) s'appuiera sur les recherches sur la cognition spatiale pour comprendre comment nous percevons et nous représentons l'environnement qui nous entoure. Enfin, la section 4 fera le lien entre la représentation mentale de l'espace, la façon dont nous l'exprimons verbalement et les processus

Pouliquen-Lardy, Lauriane. Collaboration à distance :
étude de la compréhension mutuelle dans les environnements virtuels collaboratifs immersifs : le cas de la communication spatiale - 2016

cognitifs associés. Ainsi, nous présenterons la notion de cadre de référence et verrons les facteurs influençant le choix de tel ou tel cadre de référence. Pour conclure ce chapitre théorique, récapitulant les connaissances théoriques qui nous serons utiles pour répondre à la question industrielle, nous pourrons formuler les questions de recherche auxquelles le travail expérimental cherchera à répondre.

Les **Chapitre 3 et 4** seront consacrés à la présentation de nos propres recherches réparties en deux parties expérimentales. Dans la première partie, notre objectif était de rester très proche des situations industrielles, et de regarder la construction, par la communication verbale, d'un référentiel spatial commun lors d'une tâche de collaboration à distance. La seconde partie expérimentale, composée de trois études complémentaires, s'est focalisée sur la production, d'une part, et la compréhension, d'autre part, d'instructions spatialisées.

Dans le **Chapitre 5 Discussion générale** les contributions apportées par les résultats de ces études sont discutées, par rapport, dans un premier temps, au cadre théorique, et dans un second temps, aux attentes industrielles.

Chapitre 1 Contexte industriel et approche ergonomique

La motivation du travail mené dans cette thèse est de répondre à des besoins réels d'opérateurs en situations de travail. Les retombées visent donc un double intérêt : participer aux connaissances scientifiques et contribuer au développement des outils industriels existants.

Les besoins ont été formulés par Airbus Group, grand groupe international dans le secteur de l'aéronautique, au travers de l'implication du NemoLab (Airbus Saint-Nazaire) et de l'équipe de réalité virtuelle d'Airbus Group Innovation (Suresnes). Ils concernent l'amélioration des situations de travail associant des équipes géographiquement distribuées. En effet, l'organisation sur plusieurs sites sépare physiquement des équipes devant travailler en étroite collaboration les unes avec les autres. Un des exemples de collaboration inter-sites est la modification d'ordonnancement dans le processus de fabrication d'un avion. Le processus industriel évoluant au fil du temps et les avions n'étant pas produits sur un site unique, une modification de tâche dans le processus de production peut avoir des répercussions sur plusieurs sites. De telles modifications peuvent entraîner des perturbations fortes dans le travail des opérateurs (appelés compagnons dans ces entreprises), induisant des retards et l'augmentation des coûts de production. Afin de réduire ces coûts, les perturbations peuvent être anticipées en amont, grâce à des outils de simulation (numériques ou physiques), en intégrant les expertises des différents sites concernés.

Aujourd'hui les analyses de faisabilité, et l'anticipation des répercussions sur la production, sont réalisées à distance grâce à un réseau de plateformes de réalité virtuelle (RV) et la création d'environnements virtuels collaboratifs immersifs (EVCIs). Ces outils numériques ont l'avantage d'offrir une visualisation à l'échelle 1 des ateliers et des avions et permettent de modifier les données rapidement et à moindre coût. Ces avantages sont à comparer aux méthodes de simulation alternatives utilisées avant le développement des EVCs, notamment le passage par une maquette physique qui est extrêmement long et coûteux. Avec ces outils numériques, les équipes peuvent ainsi tester plusieurs scénarios avant de choisir la solution optimale à déployer dans les ateliers de production.

Dans ce contexte, la RV et les EVCIs sont des outils dédiés à la collaboration entre différents sites industriels. Bien que ces outils soient déjà disponibles, leur utilisation n'est pas optimale et les équipes font encore trop souvent appel à des réunions en présentiel, c'est-à-dire lorsque les équipes se rejoignent physiquement et résolvent les problèmes de manière co-localisée. La faible utilisation de ces outils peut s'expliquer notamment par la nécessité d'une période d'apprentissage, coûteuse en temps et en énergie. Ainsi l'objectif d'Airbus et d'Airbus Group est de promouvoir l'utilisation des plateaux de RV, pour minimiser les réunions en présentiel. Une dimension importante de leur questionnement est d'atteindre des performances similaires sans augmenter la charge cognitive des opérateurs.

Les situations proposées comme cas d'usage ont les caractéristiques suivantes, illustrées Figure 1 :

- il s'agit de situations collaboratives, c'est-à-dire qu'elles impliquent au moins deux entités (sites ou personnes) ayant un but commun ;
- ces entités géographiquement distribuées ont des expertises singulières (notion d'asymétrie) ;
- les tâches à réaliser comprennent une dimension spatiale importante ;
- ces situations sont techniquement contraintes : les échanges verbaux passent par un canal téléphonique et les opérateurs sont immergés dans un environnement virtuel collaboratif.

Pouliquen-Lardy, Lauriane. Collaboration à distance :
étude de la compréhension mutuelle dans les environnements virtuels collaboratifs immersifs : le cas de la communication spatiale - 2016

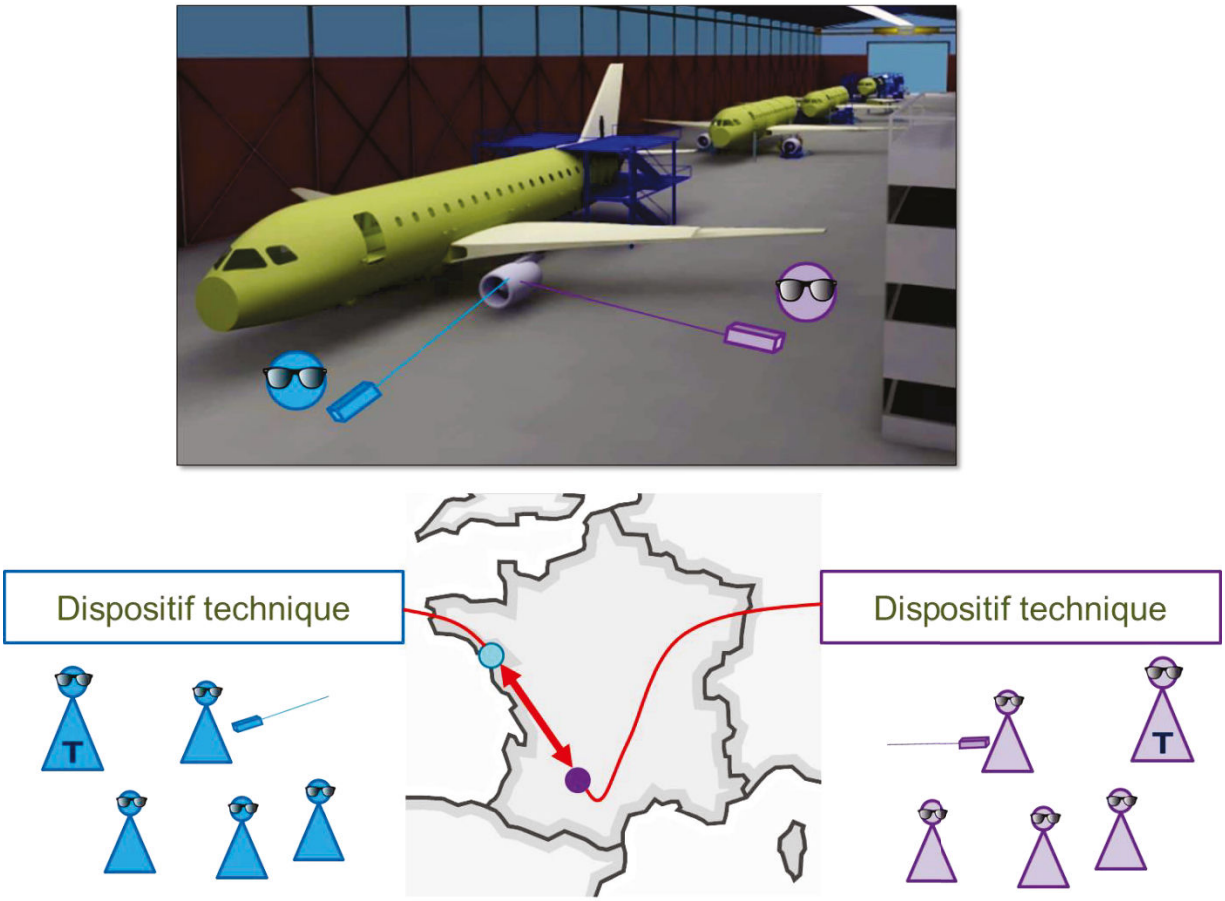


Figure 1 : Schéma des situations collaboratives distantes. T : responsables des dispositifs techniques sur chaque site.

Comme évoqué en préambule, pour répondre à ces questions, nous nous sommes appuyés sur les méthodes de la psychologie expérimentale et de l'ergonomie de l'activité. Notre objet d'étude est l'Homme au travail (nous utiliserons également le terme d'opérateur) : nous nous intéressons non seulement à son comportement et à ses performances (qui sont observables et objectivables), mais nous les inscrivons dans ce que l'ergonomie appelle l'activité de travail, c'est-à-dire ce que l'opérateur fait de manière visible et invisible (Cazamian, Hubault, & Noulin, 1996; Darses & Hoc, 2004). « Elle [l'analyse de l'activité] peut dès lors s'interpréter comme suit : le concept d'activité désigne un *fait virtuel* qui permet de comprendre la formation, et la relation entre eux, des *faits actuels* que constituent le Comportement, et la Performance, lesquels interviennent aussi, sous la forme de témoignages et de traces, comme *indices* du *fait* historique qu'est l'Activité-Vécu du Sujet. » (Hubault, Noulin, & Rabit, 1996).

Nous nous intéressons particulièrement aux processus cognitifs qui sous-tendent la réalisation de la tâche et à leurs coûts. Nous développerons particulièrement les concepts de processus cognitifs liés aux représentations et de charge mentale dans les sections 2.1 et 2.2 du chapitre suivant. Par ailleurs, puisque nous n'avons pas directement accès aux processus cognitifs (Denis, 2012), nous nous sommes intéressés aux résultats de ces processus dans le domaine particulier de la communication verbale. L'analyse des traces de l'activité a donc été possible par l'analyse des énoncés, de mesures objectives telles que les temps de préparation, de production ou de compréhension de ces énoncés, et de mesures et subjectives telles que l'évaluation de la charge mentale.

Pour ne pas rester sur un modèle purement théorique de l'opérateur et des processus cognitifs nécessaires pour accomplir la tâche, il nous semble important d'avoir une idée très précise de l'activité des opérateurs en situations réelles de travail. Pour cela nous avons réalisé des observations préalables

de situations collaboratives distantes au sein d'Airbus. Nous présentons ci-après un exemple détaillé d'observation réelle. Sans entrer dans le détail des analyses, nous rapporterons le contexte et observations majeures ainsi que les conclusions qui ont guidé les approches théorique et expérimentale.

La session collaborative s'est déroulée entre le site de Saint-Nazaire et celui de Toulouse. Elle portait sur la modification du processus industriel d'un avion en phase série. Il s'agissait d'anticiper le transfert d'une étape dans le processus de fabrication d'un site de production à l'autre. Cette étape était initialement réalisée sur le site de Toulouse, la session de travail consistait à évaluer différentes solutions pour que cette étape soit, *in fine*, réalisée sur le site de Saint-Nazaire. D'après les informations qu'ils avaient reçues pour ce transfert d'activité, les opérateurs de Saint-Nazaire avaient proposé trois scenarios. Les ingénieurs responsables de l'équipement de réalité virtuelle avaient au préalable implémenté le système afin de simuler chacun de ces scenarios. L'objectif de la session collaborative était de valider l'un des scénarios grâce à l'expertise des opérateurs réalisant actuellement la tâche à Toulouse et des opérateurs qui devront la réaliser à Saint-Nazaire, en intégrant les contraintes techniques, humaines et financières. Pour cela, une dizaine d'acteurs étaient présents sur le site de Saint-Nazaire, représentant différents corps de métiers de l'entreprise : opérateurs de lignes, préparateurs, service qualité notamment. Etaient également présents les deux ingénieurs responsables de la salle de réalité virtuelle. Un nombre plus restreint d'acteurs était présent sur le site de Toulouse (environ 4 personnes). Les deux sites étaient mis en relation par téléphone, avec un système de haut-parleurs pour que tout le monde puisse entendre.

Les trois scénarios ont d'abord été présentés sous forme de diapositives (power point) diffusées sur les murs immersifs des deux sites, en deux dimensions et de manière synchronisée entre les deux sites. Les équipes ont ensuite joué les trois scénarios, immergées dans l'environnement virtuel de la maquette numérique.

D'un point de vue technique, toutes les personnes avaient à leur disposition les lunettes pour voir l'environnement en 3D, mais tous ne les ont pas utilisées. Seule une personne était équipée de la manette et des lunettes portant les marqueurs associés au système de capture, ce qui lui permettait d'incarner l'avatar et d'interagir avec l'environnement. Chaque site disposait d'un avatar représenté par une tête et un laser (orientation de la manette). Grâce à cet avatar, chaque site avait son propre point de vue dans l'environnement, en fonction de leur localisation respective.

Seules quelques informations de l'EV transitaient d'un site à l'autre (position de l'avatar, objet sélectionné, pièce manipulée). Toutes les applications n'étaient pas forcément partagées entre les sites (notamment les mesures de cotes réalisées à Saint-Nazaire n'étaient pas visibles par Toulouse).

Le premier constat que nous avons pu faire est l'importance du canal auditif pour la réussite de ces sessions collaboratives distantes. Nous avons même pu constater la nécessiter de différentier plusieurs niveaux d'échanges liés à la présence des différents corps de métiers : les ingénieurs responsables du plateau technique d'une part (pour assurer une bonne coordination technique entre les deux sites distants) et les compagnons d'autre part, eux-mêmes organisés en métiers. Nous avons pu constater que les prises de décisions dans ces situations requerraient une bonne **compréhension mutuelle**, basée principalement sur les échanges verbaux.

Le second constat est le rôle central de la manipulation de l'environnement virtuel par un des ingénieurs responsables de la réalité virtuelle. Le succès de la session a nécessité de nombreux échanges entre cet ingénieur et les opérateurs industriels pour qu'il puisse interpréter correctement les besoins de visualisation et de manipulation. Outre la collaboration inter-sites, nous avons observé un intense travail collaboratif sur un même site.

Le troisième constat est la prégnance de la dimension spatiale dans ces situations de travail. En effet, les objets sur lesquels les opérateurs prennent des décisions sont de grande, voire très grande taille (tronçons d'avions) dans lesquels il est nécessaire de se déplacer pour en appréhender la totalité. De plus, ce sont des objets complexes, avec plusieurs composants et des particularités techniques multiples. Il semble donc difficile de se faire une représentation mentale détaillée de l'ensemble du problème. Le recours à la visualisation de la maquette numérique échelle 1 comme support des représentations et des échanges est un atout pour ce type de situations. Par ailleurs, nous avons pu constater des situations d'ambiguïté sur l'objet des échanges : les opérateurs distants n'arrivant pas à se comprendre mutuellement sur la pièce d'intérêt. Ainsi, la compréhension mutuelle dans ces environnements passe par des échanges verbaux contenant un grand nombre d'informations spatiales. C'est sur cette dimension que nous avons choisi d'axer les travaux réalisés au cours de ce travail de thèse. Il nous semble important de comprendre comment les opérateurs perçoivent et se représentent leur environnement de travail au travers des environnements virtuels dans lesquels ils évoluent. Et à partir de ces représentations comment ils échangent les informations pour se comprendre mutuellement afin de réaliser leur travail collectif.

Les observations préalables nous ont permis de reformuler les problématiques industrielles de la manière suivante :

- Comment les opérateurs se comprennent-ils mutuellement à distance, via un EVCI ?
- De quelle manière les opérateurs échangent-ils des informations spatialisées ?
- Quels coûts ou surcoûts cognitifs ces situations impliquent-elles ?
- Les EVCIs peuvent-ils offrir des aides pour faciliter ces nouvelles situations de travail ?

Ces problématiques peuvent se décliner en questions plus fondamentales sur la nature des processus cognitifs impliqués dans la communication spatiale :

- Comment les représentations individuelles intègrent-elles la dimension collective pour réaliser l'activité ? Autrement dit, comment les opérateurs construisent-ils à distance leur référentiel commun, et plus particulièrement le référentiel spatial commun ?
- Quels sont les cadres de référence utilisés dans les communications verbales ?
- Quels sont les impacts sur la charge mentale des cadres de référence utilisés lors de ces dialogues ?
- Quels types d'indices visuels peuvent faciliter la production ou la compréhension d'énoncés spatialisés ?

Pour répondre à ces questions, nous proposons dans le chapitre suivant une synthèse des recherches en lien avec les problématiques soulevées ici.

Chapitre 2 Contexte théorique – Etat de l'art

Une des particularités des situations de travail qui nous intéressent est qu'elles concernent des équipes géographiquement éloignées et qu'elles utilisent des moyens techniques très spécifiques : les environnements virtuels collaboratifs immersifs (EVCIs). En effet, il ne s'agit pas de "simples" situations de travail à distance. Avec l'utilisation d'EVCIs, les opérateurs sont totalement immergés dans l'environnement dans lequel ils travaillent. La réalité virtuelle (RV) permet non seulement une visualisation 3D des environnements de travail à l'échelle 1, mais de par sa nature numérique, elle offre également d'autres possibilités de visualisation et d'interaction aux utilisateurs (Vacherand-Revel, 2007). Un des objectifs de ce travail de recherche étant de pouvoir proposer des pistes de développement d'aides à la collaboration, il est nécessaire de bien comprendre les contraintes et les limites qu'imposent les outils utilisés, et le champ des possibles existant et à inventer.

1. Travail à distance et EVCIs

Travailler à distance nécessite l'utilisation d'outils. L'ensemble de l'activité est donc médiatisé (ou médié selon les auteurs (Gronier, 2006; Navarro, 2001) et dépend des caractéristiques techniques des outils à disposition et des réseaux permettant à ces outils d'être connectés entre eux.
Un domaine de recherche s'est développé autour de ces nouveaux modes de travail : le *travail collaboratif assisté par ordinateur* (TCAO), surtout connu sous le sigle anglo-saxon de *CSCW* pour Computer-Supported Cooperative Work (Benali, Bourguin, David, Derycke, & Ferraris, 2002; Newlands, Anderson, Thomson, Ion, & Dickson, 2002; Benford, Brown, Reynard, & Greenhalgh, 1996; Benford, Greenhalgh, Rodden, & Pycock, 2001). Les activités réalisées à distance dépendent des limites techniques des systèmes choisis (et de l'état des avancées technologiques). Il faut donc garder en mémoire qu'une situation de travail, et les observations qui y sont associées, sont vraies à un moment donné, et qu'elles peuvent se transformer rapidement avec les évolutions technologiques.

1.1 Principes généraux : Réalité virtuelle et EVCIs

Il existe plusieurs définitions de la réalité virtuelle, selon que l'on s'intéresse à la technologie ou à l'expérience vécue. Notre travail s'appuie sur la définition fonctionnelle proposée dans le traité de la réalité virtuelle : « La réalité virtuelle va lui permettre de s'extraire de la réalité physique pour changer virtuellement de **temps**, de **lieu** et(ou) de type d'**interaction** : interaction avec un environnement simulant la réalité ou interaction avec un monde imaginaire ou symbolique » (Burkhardt & Fuchs, 2006). Dans les situations industrielles qui nous intéressent, les utilisateurs sont immergés visuellement dans l'environnement virtuel, il s'agit donc d'Environnement Virtuel Collaboratif Immersif (EVCI) - *Immersive Collaborative Virtual Environment* (ICVE). Voici à quoi renvoient chacun de ces termes :

Environnements virtuels : Il s'agit d'environnements informatiques, numériques, dont le contenu peut varier considérablement, selon le but de leur utilisation. Ils peuvent viser à reproduire fidèlement la réalité, mais peuvent tout autant relever de l'imaginaire ou du symbolique (p. 5, Fuchs, Moreau, & Papin, 2006).

Collaboratifs : Ces environnements sont mis en réseau, c'est-à-dire qu'ils sont accessibles via différents sites et dans lesquels plusieurs utilisateurs peuvent interagir simultanément (Churchill, Snowdon, & Munro, 2001). D'autres auteurs emploient le terme *distribué* (*distributed*, Benford, Bowers, Fahlén, & Greenhalgh, 1994; Rodriguez, Jessel, & Torguet, 2002) ou *partagé* (s*hared*, Basdogan, Ho,

Srinivasan, & Slater, 2000; Durlach & Slater, 2000; Ott & Dillenbourg, 2001; Steed, Slater, Sadagic, Bullock, & Tromp, 1999) pour désigner cette particularité. Dans un ouvrage dédié aux environnements virtuels collaboratifs, Churchill et ses collaborateurs (2001) en donnent la définition suivante : « Un EVC est un espace (ou ensemble de lieux) virtuel informatisé et distribué. Dans ces espaces, les utilisateurs peuvent se rencontrer et interagir les uns avec les autres, avec des agents ou des objets virtuels. Les EVCs peuvent varier dans la richesse de leurs représentations depuis des espaces graphiques 3D, 2,5D à des environnements 2D, ou des environnements composés uniquement de texte. L'accès à ces EVCs n'est nullement limité aux appareils de bureau, mais pourrait également inclure les appareils mobiles ou portables, les kiosques publics, etc. »[1] (p. 4, Churchill et al., 2001).
Du point de vue technique, cela signifie que des informations transitent, mais l'architecture peut varier d'un environnement à l'autre (Benford et al., 2001).

Immersifs : Par rapport à la définition des EVCs donnée par Churchill et ses collaborateurs (2001), le fait de préciser immersif désigne le type d'interface utilisé. Une interface dite immersive va chercher à tromper les sens de l'utilisateur pour lui donner l'illusion que l'environnement numérique est là (p.17, Wang & Tsai, 2011). Pour Burkhardt (2003b), « l'ambition affichée est de faire « disparaître » l'interface par l'intermédiaire des technologies de réalité virtuelle ». Ainsi, l'immersion est caractérisée objectivement par le nombre de sens stimulés et le taux de saturation de ces sens et la résolution des dispositifs de stimulation utilisés.

Dans les situations d'usage qui font référence dans ce travail, le contenu est très proche du réel sur certains aspects (localisation, taille et forme des pièces notamment), mais s'en éloigne fortement sur d'autres (couleurs, propriétés). Le caractère distribué est géré par le partage et la transmission d'un nombre limité d'informations via le réseau sécurisé de l'entreprise. En effet, chaque site dispose de la même simulation (maquette numérique). Seules les informations sur la localisation des avatars, et des pièces manipulées par l'un ou l'autre des sites transitent entre les sites. Cela permet de réduire les risques de latence. Enfin, l'immersion visuelle des utilisateurs est assurée par des dispositifs de vidéo projection stéréoscopiques, type mur immersif ou CAVE (Cave Automatic Virtual Environment). Les sites sont également équipés de systèmes de capture de mouvement permettant de localiser des marqueurs associés aux lunettes stéréoscopiques et à la manette gérant les interactions avec l'environnement. La Figure 2 illustre le type de dispositif immersif utilisé par le site d'Airbus de Saint-Nazaire.

[1] *A CVE is a computer-based, distributed, virtual space or set of places. In such places, people can meet and interact with others, with agents or with virtual objects. CVEs might vary in their representational richness from 3D graphical spaces, 2.5D and 2D environments, to text-based environments. Access to CVEs is by no means limited to desktop devices, but might well include mobile or wearable devices, public kiosks, etc.*



Figure 2 : Mur immersif du Technocampus Smartfactory de Montoir de Bretagne. On peut voir un opérateur de chez Airbus équipé des lunettes stéréoscopiques sur lesquelles sont disposés des marqueurs (boules réfléchissantes) dont la localisation est capturée via les caméras dont trois sont visibles sur la photo (aux coins du mur). Un système à double rétroprojection assure l'affichage des deux images sur l'écran.

Le principe commun à toutes les définitions de réalité virtuelle (que les interfaces soient immersives ou non) est **l'interactivité** entre l'environnement et l'utilisateur (Fuchs et al., 2006). En effet, l'objectif est de fournir aux utilisateurs des outils pour interagir avec les données numériques. Le principe fondamental est que « l'utilisateur agit sur l'environnement virtuel grâce à l'usage d'interfaces motrices qui captent ses actions (gestes, déplacements, voix, etc.). Ces activités sont transmises au calculateur qui l'interprète comme une demande de modification de l'environnement. Conformément à cette sollicitation de modification, le calculateur évalue les transformations à apporter à l'environnement virtuel et les restitutions sensorielles (images, sons, efforts, etc.) à transmettre aux interfaces sensorielles. Cette boucle en environnement virtuel interactif n'est que la transposition de la boucle « perception, cognition, action » du comportement de l'homme dans un monde réel. » (Fuchs et al., 2006). Cette boucle est illustrée Figure 3.

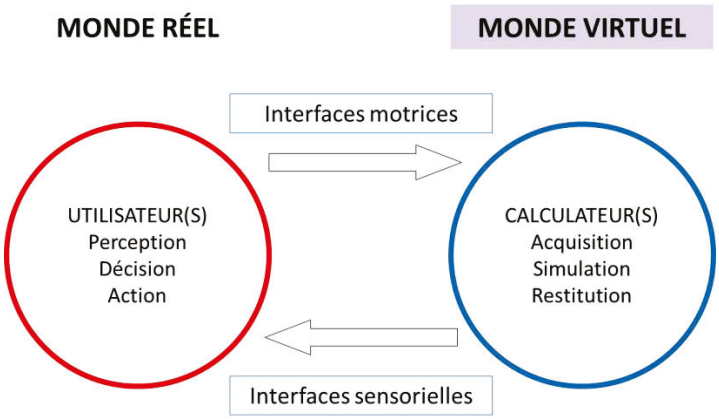


Figure 3 : Boucle « perception, cognition, action » sur laquelle s'appuie le principe de réalité virtuelle - tiré du Traité de la réalité virtuelle

1.2 Dimension technique : contraintes et possibles

La réalité virtuelle et les environnements virtuels collaboratifs immersifs sont limités techniquement par les différentes briques technologiques qui les composent. Burkhardt (2003b) distingue trois parties fonctionnelles des EV : les dispositifs de présentation d'information, d'entrée d'information et le moteur de réalité virtuelle. A ces trois composantes s'ajoutent celles permettant l'organisation et la liaison entre sites distants, à savoir le réseau et l'architecture. On notera ici que les limites viennent également de l'imagination et du mode de développement des applications. En effet, les possibilités offertes par ces outils sont très vastes, les choix guidant le développement des technologies (soft et hardware) sont donc déterminants dans l'orientation des possibilités offertes aux utilisateurs.

Il ne s'agit pas ici d'un travail sur la dimension technique, mais bien sur l'utilisation de ces outils. Nous aborderons donc les questions des limitations et capacités techniques sous l'angle des différentes composantes permettant l'utilisation : présentation sensorielle de l'environnement, interaction avec l'EV et fonctionnalités collaboratives.

1.2.1 Présentation sensorielle de l'environnement virtuel

La façon dont le dispositif fournit l'environnement à l'utilisateur est déterminante pour l'usage qu'il pourra en faire. On peut ainsi distinguer le mode de présentation de l'environnement selon le canal sensoriel visé.

Présentation des données dans la modalité visuelle

Le rendu graphique d'un environnement virtuel est fonction de plusieurs paramètres : le contenu de l'environnement (les données, ce qu'elles représentent), le logiciel qui permet l'affichage, la puissance de calcul et l'interface utilisée.

Pour afficher un environnement en trois dimensions, il faut tromper le système visuel sur la notion de profondeur. Pour cela, chacun des yeux reçoit une image plane, c'est la combinaison des deux qui trompe l'utilisateur et donne l'illusion de profondeur. C'est le principe de la stéréoscopie.

Concernant l'affichage, il existe plusieurs types d'interfaces visuelles. On va distinguer les dispositifs portatifs, type visiocasques (HMD pour Head Mounted Displays), des dispositifs à vidéo projection. Il existe des dispositifs de la taille d'un bureau pour visualiser et manipuler des objets de petites tailles (bureau virtuel) et des dispositifs de très grande taille, avec un ou plusieurs écrans. Un CAVE est un dispositif immersif de la forme d'un cube, dont la totalité des faces peuvent servir d'écran (dans la

majorité des cas au moins trois faces, puis sol et plafond). La Figure 4 montre trois types de dispositifs immersifs visuels.







Figure 4 : Interfaces visuelles immersives de haut en bas : HMD[2], CAVE et Mur immersif, bureau virtuel[3]

[2] https://commons.wikimedia.org/wiki/File:NVIS_nVisor_ST50.png

[3] Le dispositif Toucheo. © Inria / Photo H. Raguet. https://interstices.info/jcms/int_66813/quand-la-realite-virtuelle-rencontre-les-surfaces-tactiles

Présentation des données dans la modalité sonore

Tout comme pour la modalité visuelle, il existe plusieurs types de dispositifs techniques permettant la présentation des données dans la modalité auditive. Pour avoir un rendu en trois dimensions, le principe est le même que pour la stéréoscopie : donner l'illusion au système auditif que les sources sonores sont réparties dans l'espace. La précision des casques est meilleure que celle des haut-parleurs, mais, comme pour les visiocasques, ils isolent les utilisateurs les uns des autres ce qui peut poser problème lorsque plusieurs opérateurs doivent travailler ensemble sur un même site.

Présentation des données dans la modalité haptique

Plusieurs dispositifs permettent de donner une illusion de toucher aux utilisateurs. Il existe par exemple des gants haptiques pour les sensations au niveau de la main, des bras à retour d'effort pour la manipulation d'objet. D'autres solutions, telles que les interfaces tangibles reconfigurables étudiées dans l'autre axe de travail du projet PIVIPP entrent dans cette catégorie. L'apport de telles interfaces est de permettre d'intégrer une dimension physique aux simulations dans les EVCIs. Certains usages requièrent de simuler plus que la dimension visuelle, et nécessitent l'intégration de données sensorielles physiques pour tester différents contacts avec l'environnement : appuis, encombrement, accessibilité. Par exemple, lorsque la réalité virtuelle est utilisée pour la formation des compagnons, ils doivent pouvoir prendre appui sur des éléments, comme ils auront à le faire dans l'environnement réel. Les interfaces tactiles ont pour but d'assurer ces appuis. Enfin la modalité tactile peut être présentée par d'autres moyens tels que des métaphores (sous forme de vibration par exemple).

1.2.2 Interaction

Il est également important d'examiner la manière dont l'utilisateur va pouvoir interagir avec l'environnement. On retiendra ici deux modes d'interaction : les interactions basées sur des comportements existants (schèmes comportementaux importés) et celles proposant de nouvelles modalités d'interaction (métaphores, Fuchs et al., 2006). Ainsi dans certains EV, il est possible de voler, traverser des objets, modifier leur apparence, visualiser des phénomènes non visibles (chaleur, radiations…). Ces modalités d'interaction doivent être définies en fonction des objectifs à atteindre (la tâche) et des dispositifs à disposition. C'est notamment à ce niveau que les possibilités d'interaction avec le monde virtuel dépassent les possibilités d'interaction avec le monde réel, limitées alors par les choix et l'imagination des concepteurs.

1.2.3 Mode collaboratif

Les EVCIs proposent différents outils permettant la collaboration. Un des outils les plus importants est la représentation des personnes dans les environnements : ce sont les avatars (Burkhardt, 2007; Goebbels, Lalioti, & Göbel, 2003). Leur apparence peut prendre plusieurs formes (humaine, objets) qui peuvent être plus ou moins détaillées. Cependant, il semble que la vraisemblance comportementale soit un facteur à privilégier par rapport à l'apparence physique (Churchill et al., 2001). D'autres fonctionnalités permettent la collaboration, tels que les outils de pointage, retour haptique, références exocentriques, la transcription des regards, etc… (Basdogan et al., 2000; Carletta et al., 2010; Chellali, Milleville-Pennel, & Dumas, 2013; Churchill et al., 2001; Müller, Helmert, Pannasch, & Velichkovsky, 2013). Enfin, les modes de visualisation de l'EV et de partage de l'information sont à prendre en compte (Churchill et al., 2001; Convertino, Ganoe, Schafer, Yost, & Carroll, 2005; Dumazeau & Karsenty, 2008; Gutwin & Greenberg, 2004; Schafer & Bowman, 2004). Schafer et Bowman (2004) ont étudié différents modes de visualisation (perspective normale ou monde miniature) pour une tâche collaborative de

guidage. Ils ont mis en évidence que selon la finesse de l'action à réaliser et le rôle des collaborateurs (guide ou manipulateur) les modes de visualisation facilitaient plus ou moins la performance.

1.3 Dimension subjective : immersion, présence et coprésence

Le développement d'interfaces immersives a pour objectif de procurer à l'utilisateur une certaine expérience dans l'EV. On aborde ici la dimension subjective d'une expérience vécue. Pour décrire cette expérience vécue, les auteurs travaillant sur la RV ont proposé plusieurs concepts. Il est alors question d'immersion, de réalisme, de présence et de coprésence (Bulu, 2012; Burkhardt, 2003a; Nowak, 2001). Certains auteurs pensent que la performance dans un environnement virtuel serait liée au sentiment de présence que l'utilisateur ressent. Cependant, ce lien n'a pas été clairement établi. Il nous a semblé intéressant de vérifier si ces dimensions subjectives avaient un effet sur les performances. Ces notions n'étant pas au centre du débat, nous les présenterons rapidement.

Comme nous l'avons vu précédemment, l'immersion est fortement liée à la dimension technique et la stimulation sensorielle.

La notion de réalisme est liée au contenu même de l'environnement et à son rendu graphique. Le réalisme est également lié au comportement de l'environnement, c'est-à-dire au degré avec lequel l'environnement réagit de manière plus ou moins semblable à ce qui se passe dans la réalité. C'est ce que recouvre le concept de *plausibility illusion* développé par Slater (2009).

Le sentiment de présence s'est développé avec les situations de téléopération (Steuer, 1992; Zahorik & Jenison, 1998) pour décrire l'expérience singulière ressentie par les téléopérateurs de se sentir dans un autre lieu que l'environnement réel dans lequel ils étaient. Initialement, il était question de téléprésence (Slater, 2009; Steuer, 1992; Zhao, 2003). Puis le concept s'est élargi au sentiment de présence dans les environnements non réels, et notamment les EVs. Le concept de présence est défini par Slater (2009) comme étant un phénomène subjectif lié à la sensation qu'a l'utilisateur d'*être là* dans l'EV. La présence est donc liée à un état psychologique de l'utilisateur qui caractérise son expérience relative au lieu dans l'EV. Pour Burkhardt (2003b) le concept de présence vise « à caractériser l'expérience perceptive et cognitive de l'utilisateur dans le cours de son activité avec un EV. ».

Mais le sentiment de présence n'est pas le seul éprouvé lors de l'utilisation d'EV. Avec l'usage des EVCIs, la présence de collaborateurs distants dans le même EV va induire le sentiment de coprésence. La présence sociale ou coprésence est définie comme étant le sentiment d'être ensemble avec une autre personne (Riva, Davide, & IJsselsteijn, 2003; Sallnäs, Rassmus-Gröhn, & Sjöström, 2000). Pour le sentiment de coprésence, qui nous intéresse plus particulièrement ici, la notion de réciprocité est importante (Nowak, 2001; Zhao, 2003) : en effet, le sentiment de coprésence est notamment permis par l'échange et la manifestation de la prise en compte de l'autre. Si un des interlocuteurs ignore l'autre, ou s'il ne peut lui manifester sa présence, alors le sentiment de coprésence pourra être altéré.

Différents facteurs de variabilité ont été identifiés comme pouvant impacter les sentiments de présence (Gerhard, Moore, & Hobbs, 2001; Slater, 1999; Slater, Usoh, & Chrysanthou, 1995; Slater & Usoh, 1993) et coprésence (Axelsson, Abelin, Heldal, Schroeder, & Wideström, 2001; Bailenson et al., 2005; Basdogan et al., 2000; Chellali et al., 2013; Garau et al., 2003; Goebbels et al., 2003; Linebarger & Kessler, 2002; Schroeder et al., 2001) : tant au niveau technique (la résolution, la latence, les caractéristiques des avatars, la symétrie des dispositifs utilisés, les modalités sensorielles, etc.) que personnel (connaissances préalables, l'état de fatigue, le niveau de stress, l'intérêt pour la tâche, la nouveauté). Il faut également noter que l'influence de ces facteurs va varier en fonction de la tâche réalisée. En effet, les besoins ne sont pas les mêmes si les utilisateurs doivent réaliser une tâche ensemble, s'ils échangent simplement leurs avis, s'ils sont en situation d'apprentissage, ou en

compétition. Par ailleurs, les études ne s'accordent pas sur l'impact des sentiments de présence et coprésence sur la performance (Burkhardt, 2007; Schuemie, Van Der Straaten, Krijn, & Van Der Mast, 2001).

Le terme réalité virtuel est un terme générique pour parler des systèmes informatiques qui permettent **une interaction en temps réels** des utilisateurs avec les données numériques que constituent l'environnement virtuel.

Les EVCIs sont des systèmes de RV mis en réseau qui permettent à plusieurs utilisateurs d'interagir avec et au sein d'un même EV. La dimension immersive est assurée par les interfaces visant à donner l'illusion aux utilisateurs que l'environnement numérique est présent autour d'eux, se substituant à l'environnement réel.

La réalité virtuelle, de par son caractère informatique, n'a de limites que celles de l'état de la technique (notamment de développement des logiciels et des dispositifs de visualisation) et de l'imagination. Pour reprendre les termes de Burkhardt (2003b), « l'activité de l'ergonome consiste alors moins à valider de façon définitive des choix de dialogue et de présentation d'information qu'à enrichir ces nouveaux paradigmes à base d'EV par une méthodologie efficace de prise en compte des utilisateurs et de leur activité au niveau de l'interaction et de la présentation d'information. ». Pour cela des allers-retours sont toujours nécessaires entre les possibilités techniques et les besoins des utilisateurs.

2. Collaboration

Dans les situations qui nous intéressent, l'objectif n'est atteignable qu'en combinant les compétences de plusieurs compagnons, c'est-à-dire que la tâche ne peut pas être réalisée par un seul opérateur. Il existe une littérature dense sur les activités collectives. Les chercheurs dans ce domaine différencient plusieurs formes d'activités collectives, dont la collaboration, la coaction, la coordination, ou la coopération distribuée, avec des définitions différentes selon les auteurs (Gronier, 2006; Kvan, 2000). Ces définitions renvoient souvent à des classes de situations particulières, dont les caractéristiques et définitions sont non transposables d'une situation à l'autre (Zacklad, 2003). Ces situations varient notamment selon la taille des collectifs impliqués (entreprise, service, équipe, binôme), la temporalité de la réalisation de la tâche (une réunion de quelques heures à des projets courant sur plusieurs jours), et le niveau de partage du but et sous-buts (Hoc, 2003).

Au sein de ces activités collectives, nous positionnons notre réflexion dans le cadre des situations de *collaboration*, qui font partie de la classe des activités coopératives, que Hoc (2003) définit de la manière suivante : « *Nous considérons que deux agents* [définition généralisable à plus de deux agents] *sont en situation de coopération aux deux conditions minimales suivantes :*

- *Ils poursuivent chacun des buts qui peuvent entrer en* interférence*, au niveau des buts, des ressources, des procédures, etc.*
- *Ils font en sorte de traiter ces interférences pour* faciliter *les activités individuelles ou(et) la tâche commune quand elle existe.* »

Les activités collaboratives sont donc des situations de coopération dans lesquelles la tâche commune joue un rôle central.

De plus, dans les situations d'intérêt, les compétences sont distribuées ce qui induit la notion de rôle, ou de répartition des tâches (Détienne, Barcellini, & Burkhardt, 2007). Les rôles peuvent être définis au préalable, ou se construire au cours de l'activité (Roberts, Wolff, Otto, & Steed, 2003). Lund et collaborateurs (2006) utilisent la notion de degré de symétrie pour décrire la répartition des responsabilités et des contributions de chacun dans la tâche.

Par ailleurs, la notion de tâche commune est à mettre en relation avec celle de but commun, ou de but supposé commun pour reprendre la nuance introduite par Pavard et Karsenty (1997). En effet, un but, pour Tomasello et ses collaborateurs (2005) est « la représentation mentale d'un état désiré ». Pour les situations qui nous intéressent, les équipes distantes, lorsqu'elles se réunissent lors des sessions de travail dans les EVCIs, ont pour but de résoudre un problème, de prendre des décisions qui seront appliquées de manière opérationnelle sur les chaines de production. Le but commun de ces équipes est donc une idée générale d'une solution optimale, dont la réalisation finale est hypothétique au moins au début de la session collaborative. Le but global commun s'accompagne d'un ensemble de buts "personnels", qui diffèrent d'un opérateur à l'autre, ou d'une équipe à l'autre, selon son métier. En effet, chaque métier connait un certain nombre de contraintes propres. Le but commun et les buts personnels évoluent au cours de la réalisation de la tâche. Pour passer de cette projection très théorique qu'est le but, à l'état réel finalisé souhaité, il faut que les acteurs agissent. Tomasello et ses collaborateurs (2005) parlent alors d'*intentions*, qui comprennent le ou les plans d'action en direction de ce but. En effet, la collaboration fait partie d'un processus actif visant à transformer une situation initiale, vers un état final souhaité par plusieurs personnes. Il y a donc déploiement d'un certain nombre de plans d'action, ce qui nécessite un certain degré de planification.

Comme nous l'avons vu dans la définition des activités coopératives (et donc collaboratives), la collaboration se manifeste par la présence d'interférences, c'est-à-dire d'interactions entre le cours des

évènements gérés par les différentes personnes (Hoc, 2001, 2003). Les actions de l'un modifient les actions de l'autre. Ainsi, « les activités individuelles ne sont pas indépendantes les unes des autres » (Hoc, 2001). Pour Rabardel et Béguin (2005) il s'agit d'un processus cyclique « où le résultat de l'activité d'une personne constitue la source d'activité d'une autre ». Dans le modèle proposé par Hoc, nous retiendrons que les interférences peuvent être gérées à plusieurs niveaux :

- Au minimum, quels que soient les niveaux supérieurs atteints, il y aura toujours une adaptation aux interférences sur le moment, c'est-à-dire ce que Hoc (Hoc, 1998) appelle une collaboration dans l'action : à l'instant T les collaborateurs s'adaptent aux actions des autres, mais également aux aléas de l'environnement.
- Dans les situations qui nous intéressent, la collaboration est prévue, il y a donc une anticipation par rapport au moment d'exécution. Cette anticipation peut prendre différentes formes : planification, répartition des rôles, allocation de moyens, mais aussi retour d'expérience, formations communes… Ce sont les niveaux que Hoc (1998, 2001, 2003) appelle coopération dans la planification et métacoopération. A ces niveaux, l'objectif est de faciliter la collaboration dans l'action.

Par ailleurs, dans son modèle, Hoc fait également référence aux activités de *coordination* et de *communication*. Comme nous l'avons vu, la communication est particulièrement importante lorsque les équipes travaillent à distance.

Pour répondre aux exigences de la collaboration dans l'action, les collaborateurs développent un ensemble de ressources pour améliorer et faciliter l'atteinte du but commun, dans les meilleures conditions possibles et notamment à moindre coût cognitif. Pour une collaboration efficace, les collaborateurs doivent pouvoir prendre en compte les contraintes et besoins de leurs collaborateurs (Convertino et al., 2005). Pour cela, chaque opérateur doit avoir une représentation occurrente de la situation dans laquelle un certain nombre d'éléments sont partagés entre les collaborateurs.

2.1 Représentations individuelles et partagées

2.1.1 Représentation occurrente de la situation

Quelle que soit l'activité réalisée, pour être opérationnel l'opérateur construit une représentation occurrente de la situation, basée sur un modèle de son environnement (Hoc, 2003). On s'inscrit ici dans un modèle proposé et largement décrit par Endsley (1995) qui est celui de la conscience de la situation (*situation awareness*). Ce concept a été développé dans le contexte des activités de pilotage d'avion et sert notamment à modéliser les activités de prise de décision en situation dynamique. En effet, les situations d'aviation (civile ou militaire) nécessitent d'avoir une représentation de la situation toujours à jour, précise et complète, pour supporter les décisions prises sous contrainte, notamment temporelle. La conscience de la situation est définie par Endsley (1995) de la manière suivante : « La *conscience de la situation* est la perception des éléments dans l'environnement au sein d'un volume de temps et d'espace, la compréhension de leur signification, et la projection de leur état dans un avenir proche »[4]. On voit ainsi que le concept de conscience de la situation va au-delà de "simplement" percevoir la situation. Il comprend trois niveaux, que nous allons détailler ci-après. Pour bien comprendre et illustrer chacun de ces niveaux, nous reprendrons l'exemple du cas de référence.

[4] "Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"

Le premier niveau correspond à la perception de l'environnement : quelles sont les informations présentes détectées par l'opérateur ? Dans le cas d'étude, ce premier niveau est celui de la perception de l'environnement de travail, qu'il soit réel ou virtuel, ou une combinaison des deux.

Au deuxième niveau, cette perception est mise en relation avec d'autres éléments pour lui donner un sens. C'est la façon dont l'opérateur interprète les signaux. Hoc (2003) appelle ce niveau diagnostic (au sens strict). Ce diagnostic est possible en mettant en relation les signaux détectés avec des connaissances préalables (Denis, 1994; Hoc, 2003; Richard, 1990). Si l'opérateur est immergé dans une maquette numérique, alors il peut en déduire de quel avion il s'agit et l'étape du processus de production. Il le sait parce qu'il l'a appris par son expérience de terrain, par sa formation, ou parce qu'il l'a déduit de ses expériences préalables (Burkhardt, Détienne, & Wiedenbeck, 1997). Sans entrer dans le détail des différents types de connaissances, nous définissons une connaissance comme un élément stocké en mémoire, qui peut être une connaissance avérée (considérée comme vraie) ou une croyance (non prouvée, Hoc 2003; Richard 1990). L'accès aux connaissances peut se faire plus ou moins facilement. Ce que nous retiendrons ici, c'est que l'opérateur donne un sens aux éléments perçus de manière isolée. Tout comme pour le niveau perceptif, la compréhension de la situation peut avoir plusieurs états. Le sens donné à la situation peut être correcte ou erroné, et cela peut avoir plusieurs incidences. En effet, l'interprétation de la situation peut être fausse et mener à des actions qui vont dégrader la performance. Mais elle peut tout aussi bien être fausse et amener des actions performantes. L'opérateur sera alors conforté dans son interprétation erronée. Dans le cas où les actions menées dégradent la performance, l'opérateur pourra faire évoluer sa compréhension (le modèle ou représentation opérant) induisant de nouvelles actions et ainsi de suite.

Cet ajustement des actions est décrit au troisième niveau composant la conscience de la situation : celui de la projection de l'évolution de la situation au cours du temps. Pour Hoc (2003), ce niveau de la représentation est celui du pronostic. Cette projection de la situation telle qu'elle est comprise, permet notamment d'avoir un point de comparaison avec l'évolution réelle de la situation. S'il y a un écart, l'opérateur doit alors ajuster son modèle (diagnostic et pronostic associé) ce qui lui permettra de mettre en place de nouvelles actions. A cette étape de construction de la conscience de la situation, l'opérateur sait quelles étapes ont été réalisées et quelles étapes sont encore à réaliser. Ces connaissances sont très spécifiques au métier, ainsi un électricien saura que l'avion à ce moment de la production doit être équipé de telles pièces, il sait aussi qu'elles doivent être montées avant que d'autres composants soient montés, sous peine de devoir démonter certaines structures. La détection des écarts au pronostic permet de faire évoluer d'une part la représentation pour optimiser l'ajustement des actions en situation, et d'autre part les représentations/connaissances mises en mémoire à long terme. Ce stockage permettra de réutiliser l'expérience acquise lors de prochaines situations similaires.

La Figure 5 reprend le schéma proposé par Endsley (1995) pour décrire ces différents niveaux et les liens avec les déterminants externes (Task/System Factors) et internes (Individuals Factors). Malgré ce que suggère cette figure, les trois niveaux décrits ci-dessus ne se succèdent pas. Pour revenir et compléter la description du premier niveau, celui de la perception, l'opérateur peut chercher activement des informations dans son environnement, sur la base de sa compréhension de la situation (diagnostic) et de la projection (pronostic) qu'il en a. De plus, un opérateur peut ne pas percevoir des stimulus pourtant présents dans l'environnement (Berthoz, 2005) ou percevoir des évènements qui ne le sont pas (Ameisen, 2012).

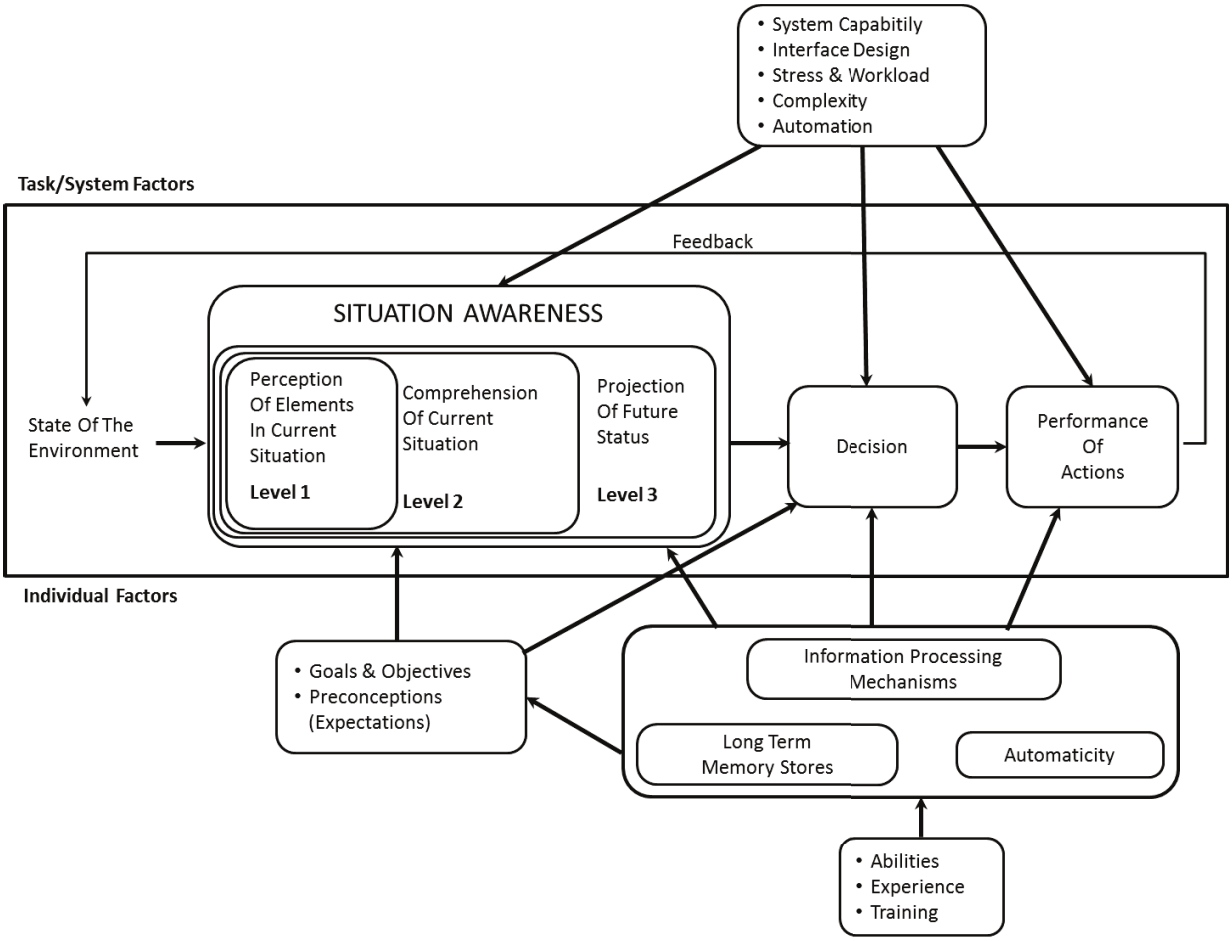


Figure 5 : Modèle de la conscience de la situation d'après Endsley (1995)

Comme nous l'avons vu, la représentation de la situation pour être opérationnelle est circonstancielle ou située, c'est-à-dire valable uniquement dans le contexte particulier dans lequel elle est construite et pour une finalité spécifique liée à l'activité en cours (Richard, 1990). Elle se différencie des connaissances qui sont plus générales et stockées en mémoire à long terme, alors que les représentations occurrentes sont caractéristiques de la mémoire de travail (Figure 6).

Du point de vue de la performance, avoir une bonne *conscience de la situation* est un des facteurs permettant d'atteindre une bonne performance, mais elle n'est ni indispensable, ni suffisante. En effet, si la conscience de la situation est bonne mais que les compétences techniques ou les moyens à disposition ne sont pas bons, alors la performance ne peut être optimale (Wickens, 2008). Inversement, dans certaines situations, notamment très automatisées, il est possible que l'opérateur ait une très mauvaise conscience de la situation, mais que la performance soit élevée (Wickens, 2008). Les problèmes graves interviennent alors lorsque le système tombe en panne : puisque l'opérateur n'a pas une bonne conscience de la situation, il ne peut la maîtriser ou intervenir de manière pertinente.

Enfin, Hoc (2003) ajoute une précision sur la notion de *situation*, qui ne doit pas être comprise comme uniquement l'environnement de la tâche. En effet, les opérateurs intègrent à leur représentation de l'environnement, des données concernant leur état propre et des caractéristiques personnelles. Par exemple, un novice, sachant qu'il ne maîtrise pas bien un procédé aura tendance à redoubler d'attention dans la réalisation de cette étape de l'activité. « La situation doit donc être envisagée, non pas au sens restreint de l'état et de l'évolution de l'environnement externe aux opérateurs, mais au sens plus large de l'état et de l'évolution d'un système opérateur-tâche » (Hoc, 2003). Dans le contexte d'une activité

collaborative, nous parlerons de situation pour désigner un système comprenant plusieurs opérateurs dans un contexte de réalisation d'une tâche commune. Dans la partie qui suit, nous détaillerons comment les opérateurs intègrent leurs collaborateurs dans la représentation de la situation.

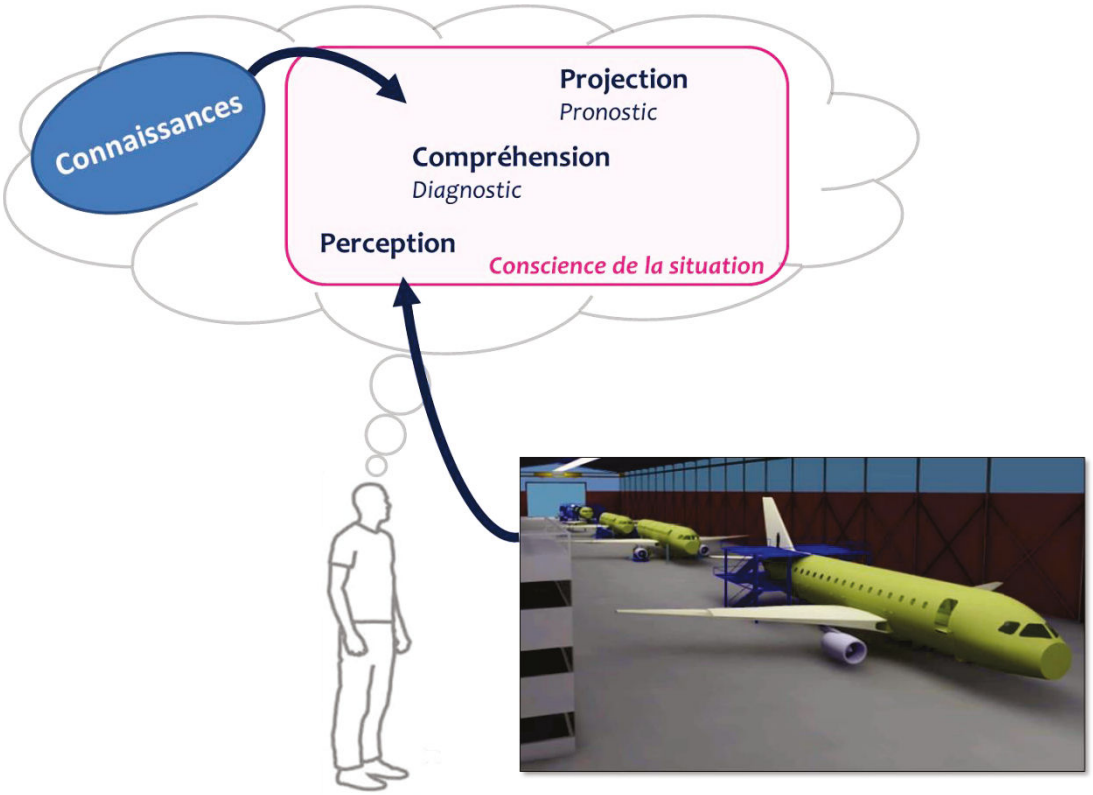


Figure 6 : Conscience de la situation

2.1.2 Le référentiel commun

Comme nous l'avons vu dans la partie précédente, pour réaliser une tâche, l'opérateur s'appuie sur une représentation fonctionnelle de l'environnement et du contexte. Pour réaliser une tâche en collaboration, les opérateurs ont besoin d'avoir des informations partagées. Le *référentiel commun* (RC) est la partie qui intègre les informations partagées entre les collaborateurs dans la représentation fonctionnelle ou *conscience de la situation*. Nous reprenons ici le concept proposé par Leplat (1991) et de Terssac et Chabaud (1990), également repris par Hoc (1998, 2001, 2003). Initialement, le référentiel opératif commun est défini comme « représentation fonctionnelle commune aux opérateurs, qui oriente et contrôle l'activité que ceux-ci exécutent collectivement » (Leplat 1991). Le référentiel commun comprend « à la fois des connaissances et des représentations partagées pour pouvoir tenir compte des autres agents dans ses décisions » (Hoc 2003). On retrouve chez d'autres auteurs des synonymes ou des concepts très proches dont Giboin (2004) récapitule les différentes formulations. Puisque les représentations sont individuelles, leur partage est relatif. Cela conduit à plusieurs niveaux de mutualisation, illustrés Figure 7. Hoc (2003) distingue ainsi :

- les représentations communes avérées non conscientes : les représentations sont effectivement communes, mais l'opérateur n'a pas conscience que l'autre opérateur a également cette représentation ;

- les représentations communes avérées conscientes : les représentations sont effectivement communes, et l'opérateur le sait ;
- les représentations conscientes supposées communes : l'opérateur croit qu'elles sont communes ;
- les représentations privées : propres à l'opérateur et identifiées comme telles.

Les représentations ou connaissances sont donc associées à des informations complémentaires, que sont les métainformations ou méta-connaissances (Keysar, Barr, Balin, & Paek, 1998). Nous venons de voir que ces méta-informations peuvent concerner le degré de mutualisation des représentations, mais elles peuvent également renseigner sur le degré de certitude (Gronier & Sagot, 2007). Ainsi, le fait que le collaborateur évoque la procédure informe qu'il en a connaissance, mais son degré de maitrise n'est pas renseigné. Par ailleurs, dans les méta-informations constituant le référentiel commun, les opérateurs intègrent les contraintes de leurs collaborateurs. En effet, Spante et collaborateurs (2004), lors d'une tâche collaborative de réalisation d'un puzzle en trois dimensions à distance avec des dispositifs asymétriques (immersif et non immersif), ont demandé aux participants d'échanger leurs dispositifs au cours de la tâche. Après l'échange la collaboration a été améliorée. Les participants ont ainsi pu composer avec les contraintes du dispositif non immersif, assignant des tâches de contrôle et de supervision, avec les possibilités d'action offertes par le dispositif immersif, plus pratique pour sélectionner et déplacer les objets virtuels (Spante, Schroeder, Axelsson, & Christie, 2004). On notera que les représentations que deux opérateurs ont d'une même situation, un même évènement, ou un même objet peuvent avoir différents degrés de similitudes : elles peuvent être soit identiques, soit différentes mais compatibles, soit différentes et incompatibles (Hoc 2003, Figure 7). A priori, plus les opérateurs ont de représentations comprises dans leur référentiel commun, plus leurs *conscience de la situation* seront similaires et compatibles, plus leur activité collective sera facilitée. Ainsi, c'est au cours de la situation de collaboration que les opérateurs vont ajuster mutuellement le référentiel commun, c'est-à-dire le degré de mutualité des représentations et la conscience qu'ils ont de ce partage. C'est ce que les auteurs appellent « grounding » ou « processus d'établissement (et d'entretien) du terrain commun » : le fait de faire rentrer dans le référentiel commun (*common ground*) une information par échange d'indices de compréhension (Chellali, 2009; Clark & Brennan, 1991; Giboin, 2004; Nova, 2004). Ces indices peuvent prendre plusieurs formes (action validant la compréhension, questionnement, acquiescement, reformulation…), dont l'usage est influencé par les moyens techniques à disposition (Gergle, Kraut, & Fussell, 2004). De plus, plus une référence est considérée comme partagée, plus selle sera utilisée par la suite et plus les opérateurs s'appuieront dessus pour la suite de leurs actions (Brown-Schmidt, 2012). Des représentations communes avérées non conscientes peuvent ainsi passer à l'état de représentation commune avérées conscientes si le cours des évènements rend explicite le partage de l'information. Par exemple, deux collaborateurs peuvent connaitre individuellement une procédure (connaissance procédurale) et penser tous deux qu'ils sont les seuls à la connaitre (méta-connaissance). Elle ne fait donc pas partie du référentiel commun conscient. Au cours de l'activité, si l'un des collaborateurs fait référence à cette procédure, l'autre pourra également évoquer qu'il la connait. Cette connaissance partagée va alors faire partie des représentations communes avérées conscientes. Garrod et Pickering (2009) parlent de l'alignement des représentations : plus les collaborateurs partagent et ont conscience de partager les connaissances, représentations, informations, plus leurs représentations sont alignées. Ce passage d'une catégorie à l'autre n'est pas anodin. En effet, le fait de savoir que l'autre sait permet d'élargir le socle commun sur lequel les collaborateurs vont s'appuyer pour réaliser leur tâche. Par ailleurs, cela peut modifier le comportement, notamment langagier, nous y reviendrons dans la section suivante.

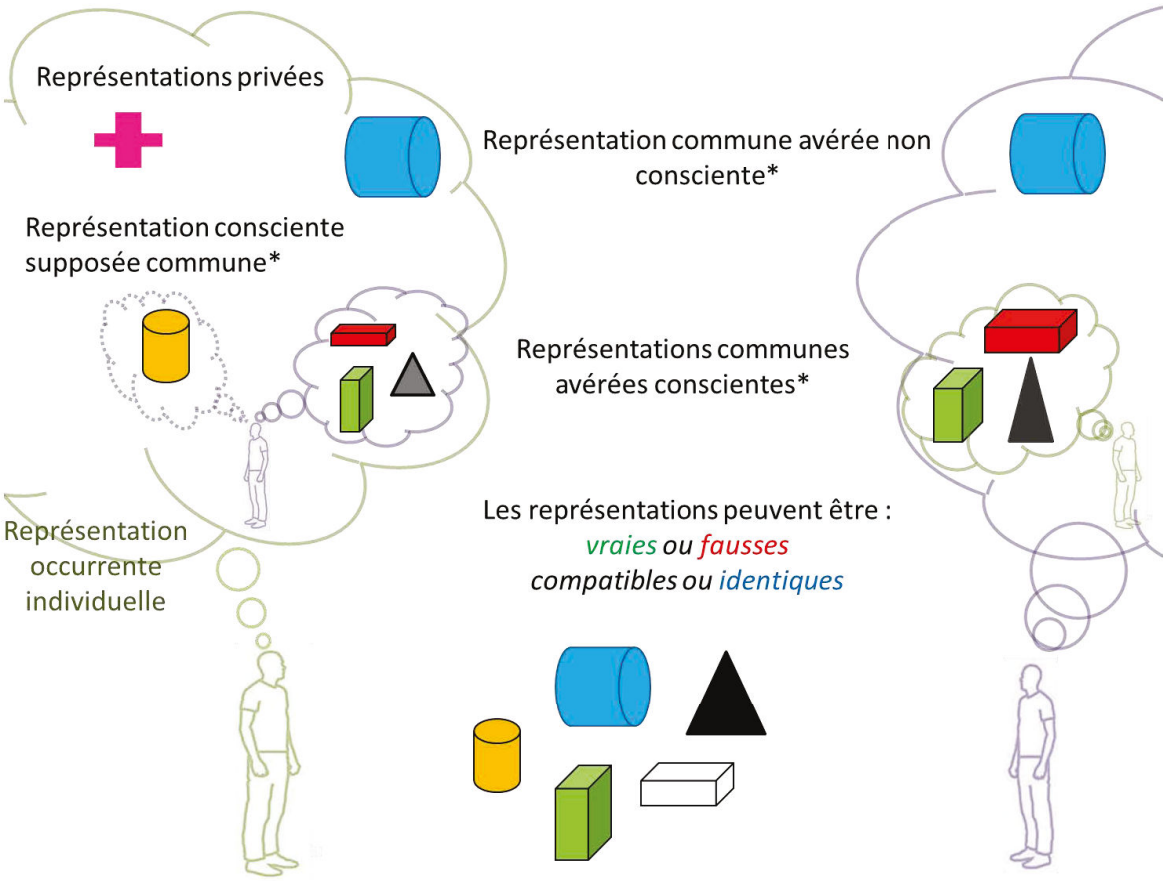


Figure 7 : Différents types de représentations, selon leur niveau de partage, … * représentations constituant le référentiel commun. Chaque forme représente le contenu d'une représentation.

Par ailleurs, le contenu et la nature du RC dépendent de la tâche à réaliser, de la situation et de la familiarité entre les opérateurs. En effet, si les opérateurs n'ont jamais travaillé ensemble au préalable, ils ne peuvent pas « s'appuyer sur des normes, des conventions, voire des routines de travail collectif préexistantes » (Vacherand-Revel, 2007). L'enjeu sera alors de « se constituer une histoire commune ». Inversement, des opérateurs se connaissant et ayant déjà travaillé ensemble, auront déjà un terrain favorable avec un ensemble de références communes (un bagage commun) et des « habitudes communes de travail » (Vacherand-Revel, 2007). Cependant, à chaque situation de travail, les opérateurs ont besoin de réactiver, de réactualiser, de contextualiser leur RC. Pour être opérationnel il faut que leur référentiel commun soit d'actualité.

Comme nous avons pu le mettre en évidence dans cette partie, plus les collaborateurs ont une représentation partagée de la situation, plus la performance sera meilleure (Giboin, 2004). Cependant, ce référentiel commun a un coût individuel et collectif (Convertino et al., 2005; MacMillan, Entin, & Serfaty, 2004).

2.2 Charge mentale et principe du moindre effort collaboratif

2.2.1 Charge mentale

Du point de vue cognitif, la réalisation de toute activité a un coût. En effet, l'Homme a des capacités cognitives limitées. Il ne peut traiter simultanément qu'un certain nombre d'informations, ne peut faire de traitements complexes sur de longues durées sans que la performance soit dégradée et

l'énergie investie entraine de la fatigue (Hart, 2006; Kuan et al., 2007). L'Homme a également des capacités mnésiques limitées. Il est donc nécessaire de gérer ses ressources en fonction des contraintes de la tâche et de ses capacités propres (astreinte). Pour décrire ce phénomène de limitations et de coût cognitif, les auteurs ont proposé le concept de charge mentale (Raufaste, Daurat, Mélan, & Ribert-van de Weerdt, 2004). Il désigne les coûts ou efforts qu'un opérateur doit fournir pour soutenir les exigences de l'activité (Hart, 2006). D'autres auteurs abordent la question de la charge mentale en termes de capacité mentale disponible pour réaliser d'autres tâches concurrentes (Sperandio, 1996).

Plusieurs facteurs influencent la charge mentale : des facteurs de contrainte, tels que la quantité d'information à traiter, la pression temporelle, les objectifs de performance, et des facteurs d'astreinte, tels que la motivation, l'expertise, la fatigue (Raufaste et al., 2004). Dans le domaine de l'aviation, Kuan et collaborateurs (2007) ont observé dans des situations à fort niveau de stress une augmentation de la charge mentale et une diminution de la qualité de la conscience de la situation entrainant des erreurs de jugements et de prédiction. Hoc (2003) rapporte que les contrôleurs aériens, lors de phases d'activité moins intenses, ne diminuaient pas leur charge de travail : ils en profitaient pour anticiper d'avantage les évènements à venir. Cette stratégie permettait une meilleure gestion des conflits, moments fortement sollicitants où les décisions doivent être prises vite, en intégrant un grand nombre de facteurs. L'élaboration du diagnostic et du pronostic complets et fiables impose un coût cognitif supplémentaire, mais cette énergie investie permet d'anticiper les problèmes laissant plus de temps pour les gérer s'ils surviennent.

Dans l'idée d'évaluer cette charge mentale, les auteurs ont proposé plusieurs techniques pour l'appréhender (Raufaste et al., 2004). Certaines s'appuient sur des mesures objectives, telles que la mesure du rythme cardiaque, les réponses électrodermales ou les doubles tâches (Leplat & Sperandio, 1967) ; alors que d'autres l'évaluent par des mesures subjectives. Dans ce travail de recherche nous utiliserons le test développé par la NASA, le NASA Task Load Index (NASA TLX, Hart, 2006). Ce test est basé sur la décomposition de la charge mentale en six composantes : demandes mentale, physique, temporelle, frustration, effort et performance. La définition de chacune des composantes est donnée dans l'Annexe 1. La force du test réside en la pondération de chaque composante. Il existe d'autres tests basés sur des mesures objectives, tel que l'échelle de Cooper-Harper (Cooper & Harper, 1969).

2.2.2 Principe du moindre effort collaboratif

Ainsi, une partie de l'activité de l'opérateur consiste à gérer sa charge de travail, en termes de ressources externes et internes (Endsley, 1995; Giboin, 2004). Lors d'activités collaboratives, Clark et Brennan (1991) ont proposé l'hypothèse d'une gestion collective de la charge mentale : c'est le principe du moindre effort collaboratif. Cette hypothèse s'est développée en opposition au principe du moindre effort, qui stipule que les stratégies de minimisation de la charge mentale sont toujours privilégiées. Or ce principe individuel ne semblait pas expliquer toutes les situations.

Dans la construction du référentiel commun, les opérateurs intègrent des méta-informations sur les contraintes de leur collaborateur, en évaluant notamment la charge mentale de leur collaborateur. De telles évaluations au cours de l'activité peuvent conduire à des modifications dans la répartition des rôles. Par exemple, dans l'étude proposée par Spante (2004), les participants amélioraient leur activité après avoir testé les deux systèmes techniques. Cette prise en compte des contraintes techniques du collaborateur leur a permis une meilleure gestion des ressources globales et une meilleure répartition des rôles. Dans l'étude menée par Gergle et collaborateurs (2004), le phénomène de grounding change selon les modalités techniques mises à disposition des participants. Ainsi les mots sont remplacés par des gestes

lorsqu'ils permettent une communication directe. Ce changement de stratégie peut être la manifestation du choix du meilleur compromis entre coût de production et fiabilité de l'échange.

D'autres études ont mis en évidence l'influence de caractéristiques personnelles sur le comportement des opérateurs. Hoc (2003) précise que, même si les opérateurs partagent un certain nombre de choses, ils ne partagent pas tout, sinon la charge de travail serait trop élevée.

Le principe du moindre effort collaboratif propose que les opérateurs assument parfois une charge cognitive plus importante pour diminuer la charge cognitive globale liée à l'ensemble de l'activité. Cette stratégie peut permettre par exemple de diminuer la charge cognitive d'un collaborateur, ou de diminuer le risque d'erreur qui nécessiterait des couts supplémentaires pour revenir sur ce qui a été fait (Clark et Brennan 1991).

---

3. Espace : perception, représentation et communication

Dans les situations qui nous intéressent, la question de la collaboration est liée à celle de l'espace. En effet, les objets sur lesquels les opérateurs travaillent sont des objets 3D et ont la particularité d'être de grande taille. La manière dont les opérateurs se représentent leur environnement de travail et la manière dont ils échangent les informations spatiales sont donc importantes.

Dans cette partie et la suivante, nous allons chercher à répondre aux questions suivantes :

- Comment l'espace est-il perçu, codé et représenté ?
- Comment l'espace est-il partagé, surtout au travers de la communication verbale ?
- Est-ce que le langage peut renseigner sur la représentation spatiale sous-jacente ?

Quel que soit la discipline s'intéressant à l'espace, les auteurs ont recours à la notion de cadre de référence (*reference frame*, *frame of reference*). Il s'agit d'un système de coordonnées (repère et direction) permettant de localiser une cible par rapport à d'autres éléments de l'environnement (Chellali, 2009; Luyat, 2009; Majid, Bowerman, Kita, Haun, & Levinson, 2004; McNamara, 2003; Shelton & McNamara, 2001). Dans la littérature, les termes de référentiel spatial, système ou cadre de référence sont utilisés de manière interchangeable, il nous semble important de dissocier ce qui relève de la représentation de ce qui relève du langage. En effet, même si les deux ne sont pas totalement indépendants l'un de l'autre (Hayward & Tarr, 1995), les mots sont conditionnés par la dimension linéaire des énoncés et par le vocabulaire à disposition du locuteur (Daniel et Denis, 1998). Plusieurs études ont d'ailleurs mis en évidence des effets de la langue sur les performances de tâches spatialisées (Levinson, 1996; Majid et al., 2004; Mishra, Dasen, & Niraula, 2003). De plus, Verjat (1994) souligne l'effet des modalités dans lesquelles se déroule la tâche (visuelle, linguistique, motrice) sur les performances. Dans son ouvrage, Levinson (2003) propose de différencier trois niveaux de traitement de l'information spatiale : perceptuel, conceptuel et linguistique. L'auteur argumente que, bien qu'il s'agisse de niveaux différents, des mécanismes existent pour passer de l'un à l'autre afin d'assurer une cohérence globale. Il précise d'ailleurs dans quelle modalité se situe les définitions des « trois cadres de référence linguistiques ». Pour reprendre cette distinction, nous avons fait le choix d'utiliser le terme *référentiel spatial* pour désigner les structures cognitives, et *cadre de référence* pour évoquer les systèmes de coordonnées utilisés dans le langage. Cette distinction ne nie pas les processus partagés entre les niveaux cognitif et langagier, elle permet d'envisager qu'il puisse exister des processus spécifiques à l'un ou l'autres de ces niveaux.

3.1 Plusieurs espaces d'interaction

Les auteurs distinguent plusieurs espaces perceptifs selon les modalités sensorielles impliquées pour percevoir l'environnement et le type d'interactions qui y ont lieu (Figure 8). Cette catégorisation montre la dépendance forte de l'espace au corps. Tversky (2005a) parle de cognition *incarnée*. C'est par ses capacités (sensorimotrices) que le corps perçoit et interagit avec l'environnement. Mais le corps n'est pas indépendant de notre esprit et celui-ci influence notre manière de percevoir ce corps et les possibilités d'interaction avec l'environnement (Berthoz, 2005, p. 129). Witt (2011) conclut que « la perception n'est pas une représentation objective de l'environnement mais reflète plutôt la relation entre l'environnement et les capacités d'action de celui qui le perçoit au sein de cet environnement »[5].

[5] « Perception is not an objective representation of the environment but instead reflects the relationship between the environment and the perceiver's ability to act within it. » (Witt, 2011).

Le premier niveau, local ou *figural*, correspond à la spatialité de chaque partie du corps, définies par leurs capacités physiques (sensorielles et motrices) propres (Tenbrink, Coventry, & Andonova, 2011). Il y a donc un espace sensorimoteur de la main, du bras, de la tête et ainsi pour toutes les parties du corps (Paillard, 1991).

Le niveau suivant correspond à l'espace du corps, qui est délimité par la peau et qui prend en compte la position relative des segments corporels les uns par rapport aux autres (Honoré, Richard, & Mars, 2002; Iachini, Coello, Frassinetti, & Ruggiero, 2014; Tversky, 2005b, 2008). A ce niveau, les auteurs différentient le schéma corporel de l'image du corps (Mars, Honoré, Richard, & Coquery, 1998). Le premier, objectif, est « un standard auquel sont rapportés les changements de posture, avant qu'ils n'entrent dans le champ de la conscience » (Head & Holmes, 1911). Le second, subjectif, est une représentation liée à l'expérience personnelle de la personne (Berthoz, 2005; Honoré et al., 2002). Ainsi, « l'anorexique se « vit » gros, même si dans le miroir il pourrait se percevoir maigre » (Berthoz, 2005, p. 129).

Au niveau suivant le corps entier entre en interaction avec son environnement mais sans déplacement de celui-ci, c'est-à-dire sans qu'il y ait navigation. C'est l'espace proche, péricorporel ou péripersonnel (Honoré et al., 2002; Iachini et al., 2014; Tversky, 2008) « dont la limite est donnée par les points que l'organisme peut atteindre sans activité locomotrice » (Honoré et al., 2002). La représentation de cet espace est influencée par les capacités d'interaction avec l'environnement (Honoré et al., 2002). En effet, plusieurs études ont montré l'effet des capacités d'actions, qu'elles soient liées à des outils à disposition (Bourgeois, Farnè, & Coello, 2014; Farnè, Serino, & Làdavas, 2007; Osiurak, Morgado, & Palluel-Germain, 2012) ou à des capacités physiques personnelles (Witt, 2011) sur les capacités d'interaction avec l'environnement (distance suffisante pour atteindre un objet par exemple). Cet espace est également nommé *micro-espace* qui est le « lieu de la manipulation de petits objets où il est facile pour le sujet de changer de points de vue par rapport à l'objet » (Pallascio, Talbot, Allaire, & Mongeau, 1990).

Dans un ensemble d'études Iachini et Coello se sont intéressés aux liens entre l'espace péri-personnel et l'espace interpersonnel, en mesurant respectivement les distances d'atteinte et de confort interpersonnel (Iachini et al., 2014). Ils ont montré que ces deux espaces partageaient certains mécanismes mais différaient sur d'autres (Iachini et al., 2014). Il semble donc que l'espace d'interaction varie selon l'objet de l'interaction (simple objet ou objet social) et qu'il soit modulé par des facteurs sociologiques.

Tenbrink et collaborateurs (2011) proposent un espace intermédiaire entre l'espace péricorporel et l'espace de navigation : l'espace *vista*. Il s'agit de l'espace visuellement perceptible, donc plus vaste que l'espace atteignable sans déplacement.

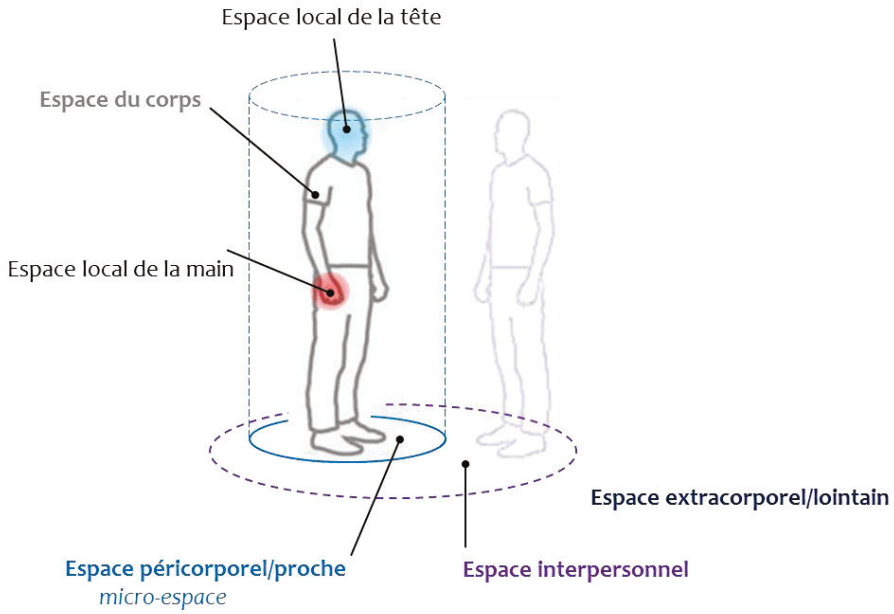
Au-delà de l'espace proche, c'est l'espace extracorporel ou lointain (Honoré et al., 2002; Tversky, 2005b, 2008), c'est-à-dire « hors d'atteinte sans activité locomotrice » (Honoré, Richard, & Mars, 2002). Certains auteurs marquent des distinctions selon la relation entre navigation et contrôle visuel. Pour Tversky (2008), l'*espace de navigation*, ne peut pas être perçu d'un seul endroit (Tversky, 2008). Pallascio et ses collaborateurs (Pallascio et al., 1990) distinguent le *méso-espace* qui « est l'espace des déplacements du sujet dans un domaine contrôlé par la vue et qui s'obtient par le recollement de micro-espaces connexes » du *macro-espace* qui « est celui qui nécessite une représentation implicite des mouvements relatifs à plusieurs systèmes de références, que l'on pourrait imager par un «recollement de cartes», selon l'expression de Brousseau (1986). »

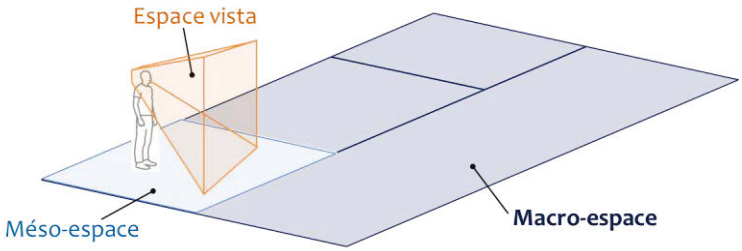
Tenbrink et ses collaborateurs (2011) distinguent l'*espace environnemental* de l'*espace géographique* sur le critère de l'échelle et de la représentativité. Selon eux, il est possible de se construire

une représentation d'un espace environnemental (quartier, campus, trajet maison-travail) grâce à une exposition sensorimotrice suffisante à l'environnement. Inversement, un espace géographique (ville, région, pays…) ne pourrait être représenté mentalement sur la seule base de l'expérience sensorimotrice (distances trop grandes). Il ne pourrait être appréhendé que par des représentations symboliques telles que des cartes, réduisant l'espace géographique à l'espace figuratif[6].

Tous ces espaces sont basés sur une expérience immédiate du monde réel (perception directe ou symbolique), mais l'espace peut également être pensé : se rappeler de comment aller quelque part, se rappeler le lieu de ses vacances, ou déduit : nous sommes capables d'imaginer un lieu à partir de sa description verbale (Loomis, Klatzky, Avraamides, Lippa, & Golledge, 2007). Les représentations sont alors basées sur des processus de rappel des informations stockées en mémoire à long terme et de traitement de l'information. Les auteurs distinguent les images mentales des modèles mentaux. Les premières sont « l'évocation mentale sous forme de représentations à caractère quasi sensoriel, des aspects physiques, des caractéristiques formelles d'objets qui ne sont plus dans le champ perceptif » (Denis, 1982, 1994; Kosslyn & Thompson, 2003), alors que les modèles mentaux sont amodaux, plus abstraits et schématiques (Bryant, Tversky, & Lanca, 2001).

[6] « *Environmental* space surrounds the body, and can't be perceived without locomotion on the part of the perceiver. It relates to spaces containing large buildings and cities. Although not directly perceivable in a short time, Montello argues that this level of space is nevertheless perceivable with enough exposure. Finally, *geographical* space is at a larger scale again, not directly perceivable through direct experience/navigation, and learnable only through symbolic representations, such as map formats that reduce geographical space to figural space. » (Tenbrink, Coventry, & Andonova, 2011).







Espace environnemental



Espace géographique

Figure 8 : Les différents types d'espaces selon l'interaction avec l'environnement.

3.2 Référentiels spatiaux

Nous venons de voir qu'il existe plusieurs espaces d'interaction entre un opérateur et son environnement. Dans cette partie, nous détaillons les différentes manières dont ces espaces sont représentés, codés et mémorisés dans différents référentiels spatiaux (Berthoz, 2005). Comme évoqué précédemment, un référentiel spatial est un système de coordonnées (repère et directions organisatrices) permettant de localiser une cible par rapport à d'autres éléments d'une scène (Chellali, 2009; Luyat, 2009; Majid et al., 2004; McNamara, 2003; Shelton & McNamara, 2001).

Au niveau local, les informations sensorimotrices sont codées dans le référentiel de l'organe ou de la partie du corps qui interagit avec l'environnement. Ainsi, les saccades oculaires sont codées dans un système de coordonnées rétino-centré, et les mouvements dans la tête codés dans un système visuo-céphalo-moteur (Paillard, 1991).

Les différents espaces du corps recoupent donc une myriade de référentiels différents, traités dans des zones spécifiques du cerveau (Berthoz, 2005 p. 153). Pour effectuer un mouvement de la main avec un retour visuel, le cerveau doit donc combiner deux référentiels distincts : l'un rétino-centré et l'autre centré sur la main. Malgré cette multiplicité de référentiels, notre perception et nos actions sont cohérentes. L'hypothèse proposée par Berthoz (2005) est que la coordination des référentiels est assurée par un système flexible, assurant les transitions d'un référentiel à l'autre. Il semble que cette flexibilité et coordination des référentiels s'acquièrent par l'expérience, nous permettant d'accumuler des connaissances sur l'espace. Ces connaissances « suscitent non seulement des comportements adéquats, mais créent également des attentes, des inférences et des prédictions pertinentes sur une grande diversité d'espaces » (Tversky, 2005, p. 162). Ce sont ces connaissances qui permettent de prédire la trajectoire d'une balle et de produire les mouvements qui permettront de l'intercepter ou de la relancer. Ce sont aussi ces connaissances qui créent les illusions perceptives. Comme déjà évoqué plus haut, l'illusion de différence de taille des personnes situées dans les « chambres de Ames » sont possibles parce que notre expérience du monde nous a convaincu que les murs d'une pièce sont perpendiculaires les uns aux autres, nous construisons la perspective associée à cette connaissance.

Les objets peuvent être localisés en utilisant le corps de celui qui les perçoit comme système de référence. C'est le référentiel égocentré (Paillard, 1991), par opposition aux référentiels exocentrés, dont le système de coordonnées n'est pas lié au corps de celui qui perçoit. Le référentiel égocentré est composé de trois axes, trois plans orthogonaux illustrés Figure 9 (Bryant and Tversky, 1999). Les axes composant le référentiel égocentré dérivent de « leurs propriétés perceptives et comportementales » (Tversky, 2005b), c'est-à-dire de leurs capacités d'interaction avec le monde. Ainsi, l'axe tête/pieds, également nommé axe transversal, « est asymétrique d'un point de vue perceptif et comportemental ; pour une personne canonique, donc debout, il correspond au seul axe asymétrique dans le monde, celui qui est formé par la gravité. L'axe avant/arrière du corps est aussi asymétrique ». En effet les capacités sensorielles du corps distinguent l'espace immédiatement perceptible par la vue et sur lequel il est possible de manipuler, l'avant, de l'espace non perçu immédiatement et sur lequel il n'est pas facile d'agir, l'arrière. Cette distinction est renforcée au niveau moteur : l'avant étant le sens de déplacement privilégié. « L'axe gauche/droite du corps ne présente aucune asymétrie saillante, ni à partir du corps, ni à partir du monde » (Tversky, 2005b). Dans ce référentiel égocentré, il semble que les directions qui le composent ne se valent pas, c'est-à-dire qu'il serait plus facile de percevoir, localiser et mémoriser des objets dans certains axes plutôt que d'autres. En effet, plusieurs études, sur le postulat que « la rapidité reflète la proéminence dans l'espace cognitif » (Tversky, 2005a, 2005b), ont montré la prédominance de l'axe transversal (tête/pieds) sur les deux autres, et de l'axe frontal (avant/arrière) sur l'axe sagittal (droite/gauche). Ainsi, la saillance perceptive et fonctionnelle du corps affecterait la gestion et

l'organisation de l'espace mental (Tversky, 2005b). C'est la théorie du cadre de référence (*Spatial Framework Theory*). Cependant, cette préférence semble également dépendre des modalités cognitives impliquées. En effet, si les objets sont présents, alors ceux localisés devant la personne sont plus facilement retrouvés que les autres, et les objets derrière sont retrouvés plus lentement (Bryant et al., 2001). Pour les tâches de rappel, il semble que les informations soient organisées autrement, avec les effets suivants : haut/bas privilégié par rapport à l'axe devant/derrière, lui-même privilégié à l'axe droite/gauche (Tversky, 2008). En ce qui concerne la condition perceptive, la prédominance sensorielle semble l'emporter sur les autres caractéristiques : les objets dans le champ de vision sont plus facilement retrouvés car immédiatement perçus. Inversement, retrouver un objet situé derrière nécessite de faire appel au souvenir, induisant une étape et un coût cognitif supplémentaires.

Par ailleurs, « Il y a deux façons, ou stratégies mentales, de se rappeler un chemin parcouru ; la première est la « stratégie cognitive de route » égocentrée, topokinestésique, qui vise à se rappeler le mouvement (traverser à tel endroit, prendre à droite ou à gauche, rencontrer quelqu'un...). Il s'agit d'une mémoire séquentielle des actions. La seconde vise à se décentrer en retraçant le chemin parcouru grâce à une carte, avec une stratégie de survol allocentrée, topographique, indépendante de la perception. Pour être capable de changer de point de vue – de retrouver un chemin différent, de changer de perspective – il faut donc pouvoir changer de stratégie. » (Berthoz, 2013).

Mais les auteurs ne semblent pas faire la distinction entre un référentiel égocentré basé sur le corps physiquement présent dans l'environnement et un référentiel égocentré basé sur le souvenir de ce corps dans l'environnement. En effet, il semblerait plus simple de se représenter la situation telle qu'elle est actuellement, d'après les indices sensoriels directement accessibles, que de se représenter une scène telle qu'elle a été vécue d'après les souvenirs de ce qui a été perçu.

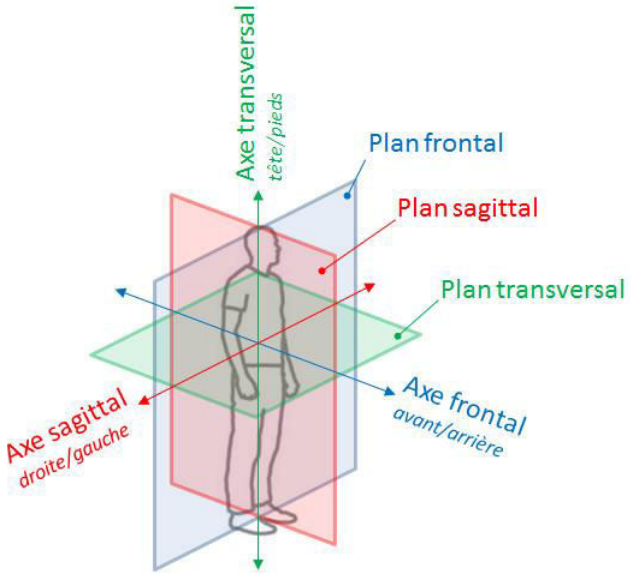


Figure 9 : Plans et axes du référentiel égocentré

Pour coder la localisation d'un objet dans une scène, un observateur peut utiliser plusieurs référentiels autres que son propre corps (Figure 10). Il s'agit alors de référentiels *exocentrés*, indépendants de toute perspective localisant les objets les uns par rapport aux autres (object-to-object). Nous distinguerons deux types de référentiels exocentrés : ceux utilisant les caractéristiques de l'environnement, que nous nommerons *allocentrés*, et ceux centrés sur des objets ou des personnes que nous nommerons *centrés sur* la référence utilisée pour situer l'objet-cible.

Dans un référentiel allocentré, les objets sont localisés par rapport à des points de repères ou des invariants de l'environnement, comme la gravité. Les référentiels centrés sur une référence peuvent avoir plusieurs types de directions organisatrices, selon qu'elles sont définies ou non par les propriétés intrinsèques de la référence. On distinguera alors les référentiels centrés intrinsèques et extrinsèques.

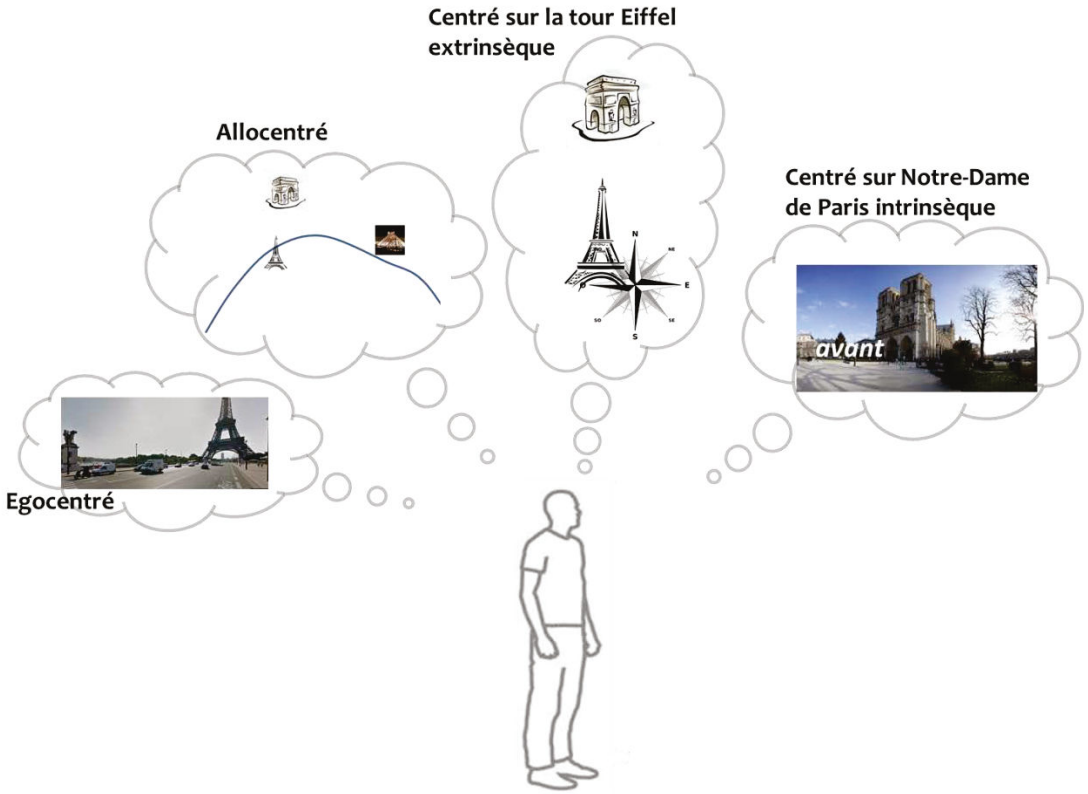


Figure 10 : Les différents types de référentiels dans lesquels un environnement peut être représenté mentalement.

Les directions des référentiels exocentrés intrinsèques sont données par les propriétés intrinsèques de l'objet ou la personne de référence. On retrouve ici l'effet de la notion de fonctionnalité : en effet les objets mobiles peuvent avoir un sens de déplacement donnant une direction organisatrice (Figure 11). Les voitures ont ainsi un *avant*, donné par le sens de déplacement le plus courant du véhicule, et par voie de conséquence elles ont également un *arrière*. L'usage permet également de distinguer le *haut* du *bas*, renforcé par le référentiel gravitaire. Ces deux directions permettent de déduire la troisième direction, et d'attribuer une *droite* et une *gauche* au véhicule (Bryant & Tversky, 1999). D'autres objets immobiles ont également des orientations, liées à leurs usages (Figure 11) : appareils photo, télévision, fourchette, arrosoir… Cependant ces orientations peuvent être ambiguës (Watson, Pickering, & Branigan, 2004). Les exemples de l'appareil photo et du poste de télévision, illustrés Figure 11, montrent que les directions liées aux usages ne sont pas liées à la localisation de l'utilisateur par rapport à l'objet, mais sont bien liées à leurs modes de fonctionnement et d'utilisation.

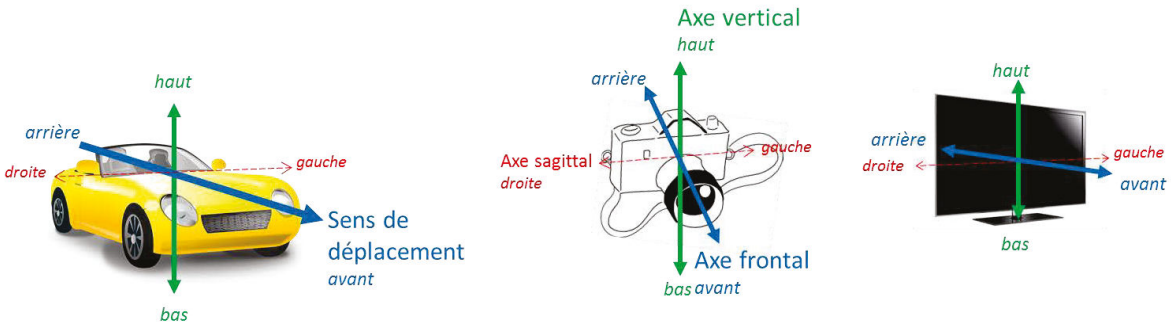


Figure 11 : systèmes de références exocentrés intrinsèques

L'Homme est également capable de reconnaitre le corps d'une autre personne et de localiser une cible par rapport à une autre personne. Cette reconnaissance passe notamment par l'identification de la position relative des segments du corps de l'autre. Il est ainsi possible de déterminer les actions ou postures impossibles chez d'autres personnes (Tversky, 2008). La localisation d'une cible par rapport à une autre personne utilise les mêmes axes et plans que ceux du référentiel égocentré : tête/pieds, avant/arrière et droite/gauche. Plusieurs études ont d'ailleurs montré que la prédominance de certains plans présents dans le référentiel égocentré est également à l'œuvre dans les référentiels exocentrés intrinsèques (Bryant & Tversky, 1999).

Le choix d'un référentiel plutôt qu'un autre semble également influencé par la tâche à réaliser. Ainsi, il semble que les référentiels spatiaux utilisés dans les représentations occurrentes pour la réalisation d'actions motrices soient égocentrés, alors que les informations seraient mémorisées dans des référentiels exocentrés (Avraamides & Kelly, 2008). Mais cette généralité est controversée, en effet il semble que la mémorisation de scènes spatiales, soient mémorisées selon des perspectives privilégiées, souvent liées à l'usage. On aura plus facilement tendance à se représenter une pièce depuis la porte par laquelle on entre. (Lackner & Dizio, 1998) ont mis en évidence ce phénomène à l'aide d'une *rotating room*. Si le point d'accès habituel à cette pièce est modifié, les utilisateurs ont du mal à reconnaitre le lieu et à retrouver les objets : la représentation de la pièce est inscrite dans une représentation plus large. Si le point de passage entre ces deux espaces (la pièce et son environnement) est modifié, alors le rappel de la connaissance a du mal à s'accorder avec les indices visuels.

Dans la partie suivante, nous allons présenter les mécanismes qui permettent de passer d'un référentiel à l'autre.

3.3 Coûts cognitifs et transformations mentales

L'Homme est en permanence en train de gérer plusieurs référentiels. Nous allons voir que cette gestion implique plusieurs coûts cognitifs possibles.

Plusieurs études ont mis en évidence qu'il était plus facile de traiter (dans l'action ou au niveau de la mémorisation) des informations spatiales lorsque plusieurs référentiels sont alignés que lorsqu'ils divergent, quelle que soit leur nature (Galati & Avraamides, 2014; Mou & McNamara, 2002; Tamborello, Sun, & Wang, 2012).

Dans l'étude de Tamborello et collaborateurs (2012) la tâche consistait à indiquer sur un clavier à l'aide des flèches droite/gauche (flèche vers le haut pour indiquer qu'il ne doit pas tourner) de quel côté tourner un canon pour atteindre une cible. La situation comprenait deux canons de couleurs différentes. Chacun des canons avait pour cible uniquement celles de sa couleur. Les canons étaient orientés soit dans la même direction, soit dans des directions divergentes (90° ou 180°). Les stimuli utilisés sont illustrés

Figure 12. Les résultats indiquent que les temps de réaction sont plus longs quand les directions divergent. Il semble plus compliqué et plus coûteux de gérer deux référentiels exocentrés intrinsèques, qu'un seul. Mais les temps de réaction augmentent également avec le degré de rotation par rapport à l'orientation du participant : ce résultat suggère que les participants opéraient une sorte de transformation mentale (rotation mentale ou prise de perspective) pour aligner leur référentiel avec celui du canon cible[7]. Ce dernier résultat est cohérent avec nombre d'études sur la prise de perspective et les rotations mentales, nous y reviendront dans la partie suivante.

L'étude de Tamborello et collaborateurs (2012) a également mis en évidence que le choix du référentiel s'opère également en prenant des informations circonstancielles. Dans leur étude ce phénomène apparait sous la forme de la prise en compte de la probabilité d'apparition d'une cible. En effet, les cibles (rouge:bleu) étaient soit en nombre égal (probabilité équivalente) ou inégal (plus forte probabilité pour l'une des cibles). Dans le cas où les cibles bleues étaient plus nombreuses, lorsque la cible désignée était effectivement bleue, alors les temps de réaction étaient plus courts que dans la situation à probabilité équivalente. Il semble donc que les participants privilégiaient le référentiel désigné par la probabilité. Si la cible désignée n'était pas celle avec la plus forte probabilité (celle en plus petit effectif), les temps de réaction étaient plus longs, suggérant la mise en place d'un effort supplémentaire pour changer de référentiel. Ce résultat est à mettre en lien avec le concept de *conscience de la situation* : il semble plus pertinent et plus opérationnel de privilégier le référentiel le plus saillant. Du point de vue de la charge mentale, cette stratégie semble être plus économique dans la majorité des cas, mais demande d'assurer un effort supplémentaire lorsque finalement le référentiel choisi n'est pas pertinent.



Figure 12 : Stimuli utilisés dans l'étude de Tamborello et collaborateurs (2012)

Dans leur étude, Galati et Avraamides (2014) étudient la pluralité des référentiels spatiaux dans deux types de tâches : mnésique (rappel et prise de perspective) et linguistique (description). Nous présentons ici la dimension mnésique et reviendrons sur la dimension linguistique dans la partie suivante. Le dispositif expérimental permet au participant de se représenter la scène dans plusieurs référentiels spatiaux : un référentiel égocentré basé sur le point de vue du participant (*director*), un référentiel intrinsèque centré sur la structure (axe de symétrie, comme illustré Figure 13 gauche) et un référentiel intrinsèque centré sur un collaborateur (*matcher*). Trois configurations sont proposées : la structure est alignée soit avec le participant, soit avec le *matcher*, soit avec aucun des deux ; et deux conditions de connaissance mutuelle : le participant sait ou non où se trouve le *matcher*. Les tests mnésiques montrent que les participants prennent en compte à la fois leur orientation et celle de leur interlocuteur par rapport au dispositif pour mémoriser et se représenter la structure. Pour la tâche de rappel, les participants devaient reconstituer le dispositif sous forme de plan (Figure 13 droite). Si le référentiel du directeur est

[7] « people perform some type of mental self-rotation and perspective taking to align themselves with the targeted cannon. »

aligné avec celui du dispositif, quelle que soit leur connaissance et l'orientation de leur collaborateur, alors c'est la configuration qu'ils préfèrent. Au contraire, si le référentiel du directeur n'est pas aligné avec celui du dispositif, les stratégies varient en fonction de l'interlocuteur. Si les directeurs ne savent pas où est leur interlocuteur, ils gardent préférentiellement leur référentiel. Si leur interlocuteur est aligné avec le référentiel du dispositif, ils prendront préférentiellement cette perspective pour organiser et mémoriser l'information. Ces stratégies d'organisation des informations spatiales en mémoire sont également visibles dans les performances de prise de perspective : les participants étaient amenés à imaginer qu'ils se situaient à l'endroit de l'un des objets situés sur le dispositif (temps pour changer de perspective) et devaient, depuis cette perspective, pointer vers un autre objet (temps pour situer l'objet). Les temps de réponse pour les deux mesures sont plus courts dans les directions orthogonales des orientations choisies pour organiser les informations en mémoire, que dans les directions obliques.

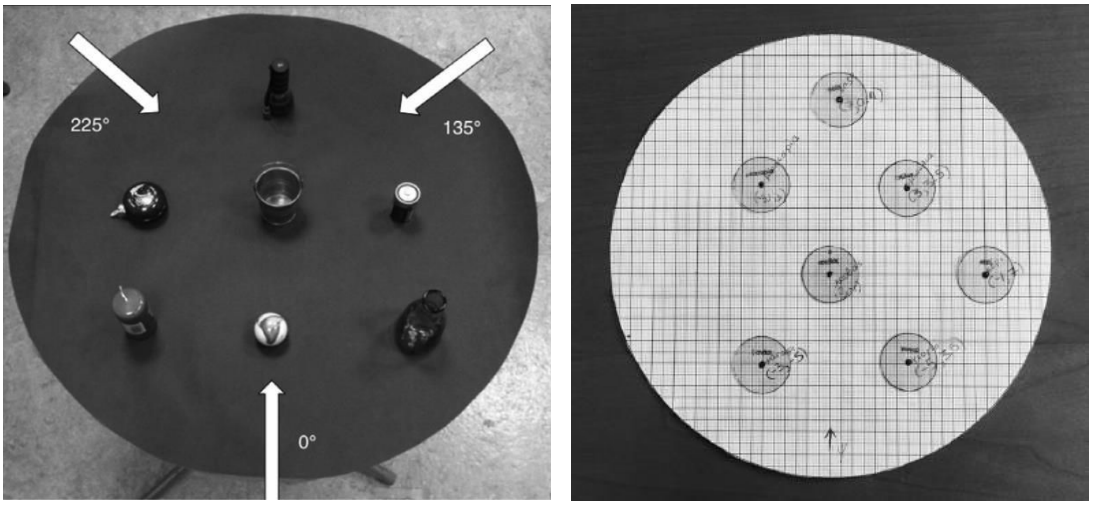


Figure 13 : D'après Galati et Avraamides (2014). Gauche : structure utilisée, Droite : test de rappel.

Ce bénéfice de l'alignement des référentiels est à mettre en relation avec des effets constatés dans l'organisation des informations spatiales en mémoire. Il semble que pour mémoriser une information spatiale, elle subit une réorganisation (construction) qui peut amener ensuite à des erreurs de jugements (Tversky, 2008). Tversky (2005) rapporte deux types d'erreur. La première est une erreur d'alignement des points de repères, qui sont mémorisés « comme étant plus alignés les uns par rapport aux autres qu'ils ne le sont réellement. ». Cette déformation est illustrée avec l'exemple de Rio et Boston ou Rome et Philadelphie (Figure 14). Rio est bien à l'est de Boston et Rome au nord de Philadelphie, alors que la majorité des personnes s'imaginent l'inverse. La seconde erreur est liée à la rotation de référentiels pour faire coïncider des directions pourtant non alignées. Cette déformation est illustrée avec l'exemple de l'Amérique latine : « il est demandé aux étudiants d'orienter un « découpage » de l'Amérique du Sud dans un cadre de référence « nord-sud » et « est-ouest ». L'Amérique du Sud est un territoire allongé, mais son axe d'élongation est incliné par rapport aux directions cardinales. Une majorité significative d'étudiants a orienté l'Amérique du Sud comme plus verticale qu'elle ne l'est réellement » (Tversky, 2005).

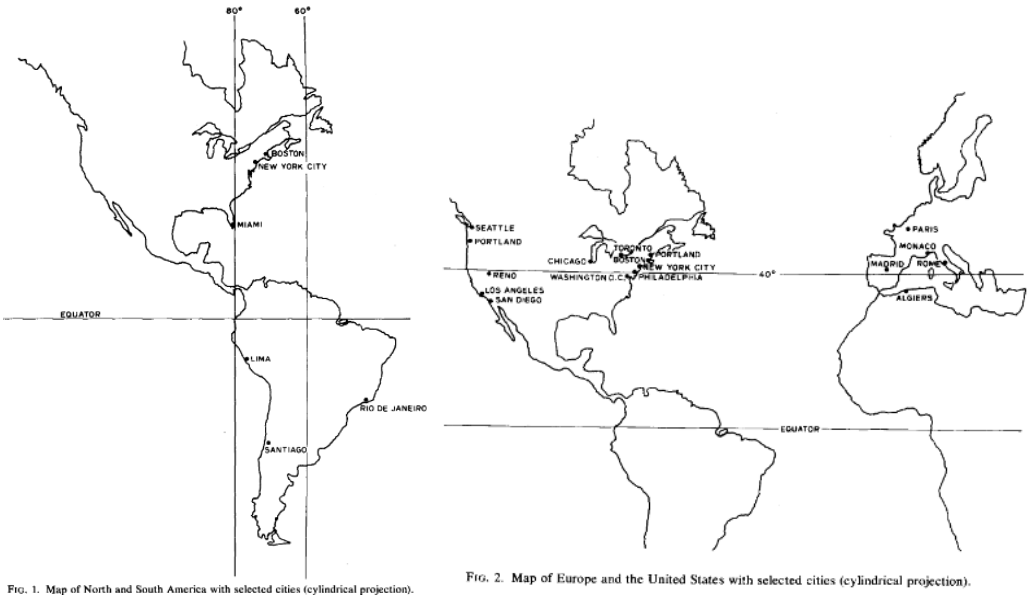


FIG. 1. Map of North and South America with selected cities (cylindrical projection).

FIG. 2. Map of Europe and the United States with selected cities (cylindrical projection).

Figure 14 : figures illustrant les déformations d'alignement, d'après Tversky, 1981

Pour revenir à la gestion des référentiels au cours de l'action, plusieurs études se sont intéressées aux changements de référentiels et aux transformations mentales associées. Les auteurs semblent s'accorder sur le fait de distinguer deux types de transformations mentales : celles où la personne fait bouger l'objet mentalement (object-based) et celles où la personne change de perspective (Zacks & Michelon, 2005). Le choix de l'une ou l'autre transformation dépend de la tâche, du jugement qui doit être fait, mais également des capacités personnelles.

Des tests pour évaluer les transformations mentales, se sont beaucoup développés (Albaret & Aubert, 1996; Kaltner, Jansen, & Riecke, 2014; Peters et al., 1995; Pylyshyn, 1979). Shepard et Metzler (1971) ont développé un test consistant à identifier si deux images correspondent à des prises de vue d'un même objet ou d'objets différents. Les images utilisées sont illustrées Figure 15. Ils ont montré que les temps de réaction augmentent avec le degré de rotation entre les deux pièces à comparer (Metzler & Shepard, 1974). Il semble que le temps de traitement par le cerveau soit proportionnel à l'écart entre les deux figures, suggérant qu'il opère une rotation mentale (de l'objet ou de la perspective).

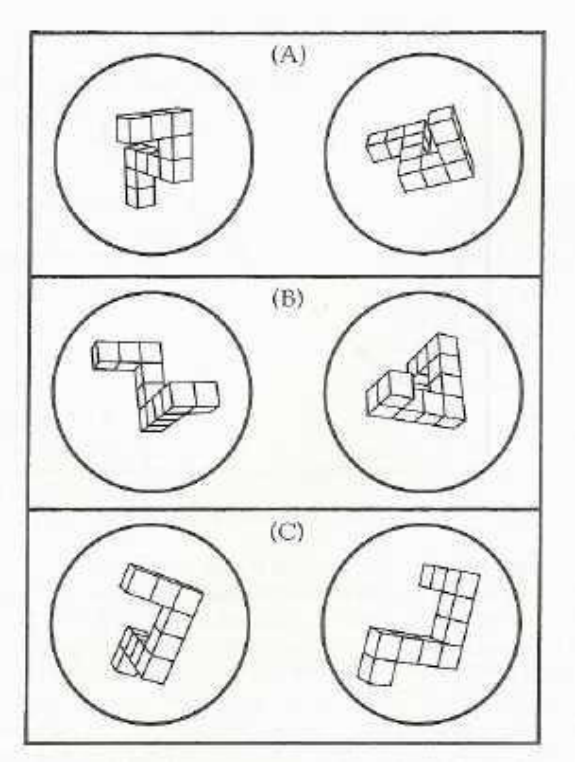


Figure 15 : Figures utilisées dans les tests de Shepard et Metzler (1971)

De plus, il semble que ce phénomène de rotation mentale soit influencé par le type de stimuli, plus les stimuli sont proches de postures humaines possibles, plus il est facile de statuer (Amorim, Isableu, & Jarraya, 2006).

Le terme rotation mentale est parfois assimilé aux transformations basées sur l'objet (rotation de l'objet), et il est alors mis en opposition aux processus de prise et changement de perspective. Alors qu'il s'agit de deux niveaux différents : les rotations mentales sont un type d'opération mentale, mais elles peuvent impliquer la rotation d'un objet ou d'une perspective. On retrouve donc des effets de rotations mentales (augmentation des temps de réaction avec les écarts d'orientation) pour des changements d'orientation de référentiels intrinsèques (rotations d'objets) ou égocentrés (changement de perspective). A noter que Kozhevnikov et Hegarty (2001) ont développé un test dédié spécifiquement à la capacité de changement de perspective : Object perspective taking test présenté dans la Figure 16.

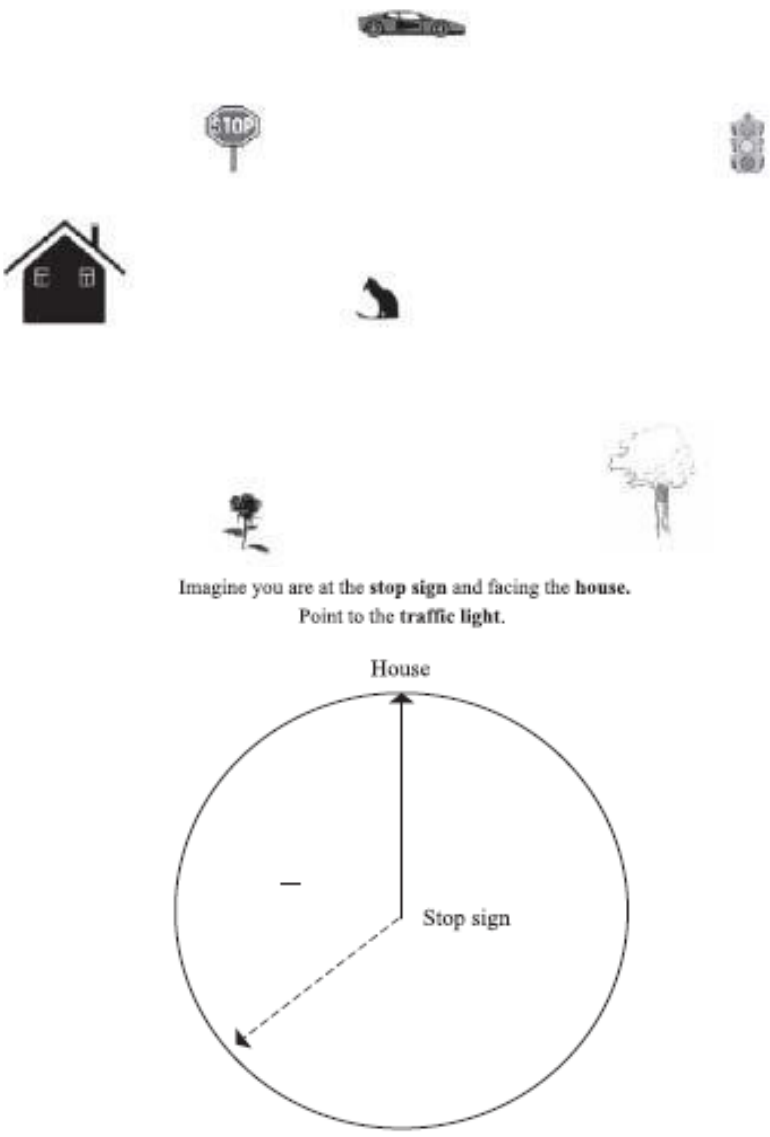


Fig. 1. Example of an item in the Object Perspective Taking Test. The dashed arrow indicates a participant's response to the item (direction to the traffic light).

Figure 16 : Principe du test sur la prise perspective développé par Kozhevnikov et Hegarty (2001) : Object perspective taking test

Dans l'étude de Tamborello et collaborateurs (2012), présentée page 38, les auteurs mettent en évidence deux phénomènes. D'une part ils constatent un processus de prise de perspective, avec l'augmentation des temps de réaction avec l'augmentation de l'angle entre le référentiel égocentré du participant et le référentiel intrinsèque du canon-cible. D'autre part, les temps de réaction augmentent également avec l'éloignement de la cible par rapport à la direction du canon, ce qui suggère un processus de perception relative par rapport à l'axe directeur (alors qu'on pourrait penser que plus c'est loin, plus c'est discriminable). Mais cette étude laisse ouverte la question de l'alignement avec la direction organisatrice, puisque dans ces conditions les modes de réponses changent par rapport aux conditions de latéralité.

Michelon et Zacks (2006) rapportent les résultats de quatre études différentes. Ils s'intéressent à deux jugements : déterminer de quel côté de l'avatar la cible se trouve (droite/gauche), déterminer si un

objet cible est visible ou non par l'avatar. La Figure 17 montre le type de stimuli sur lesquels les participants devaient se prononcer.

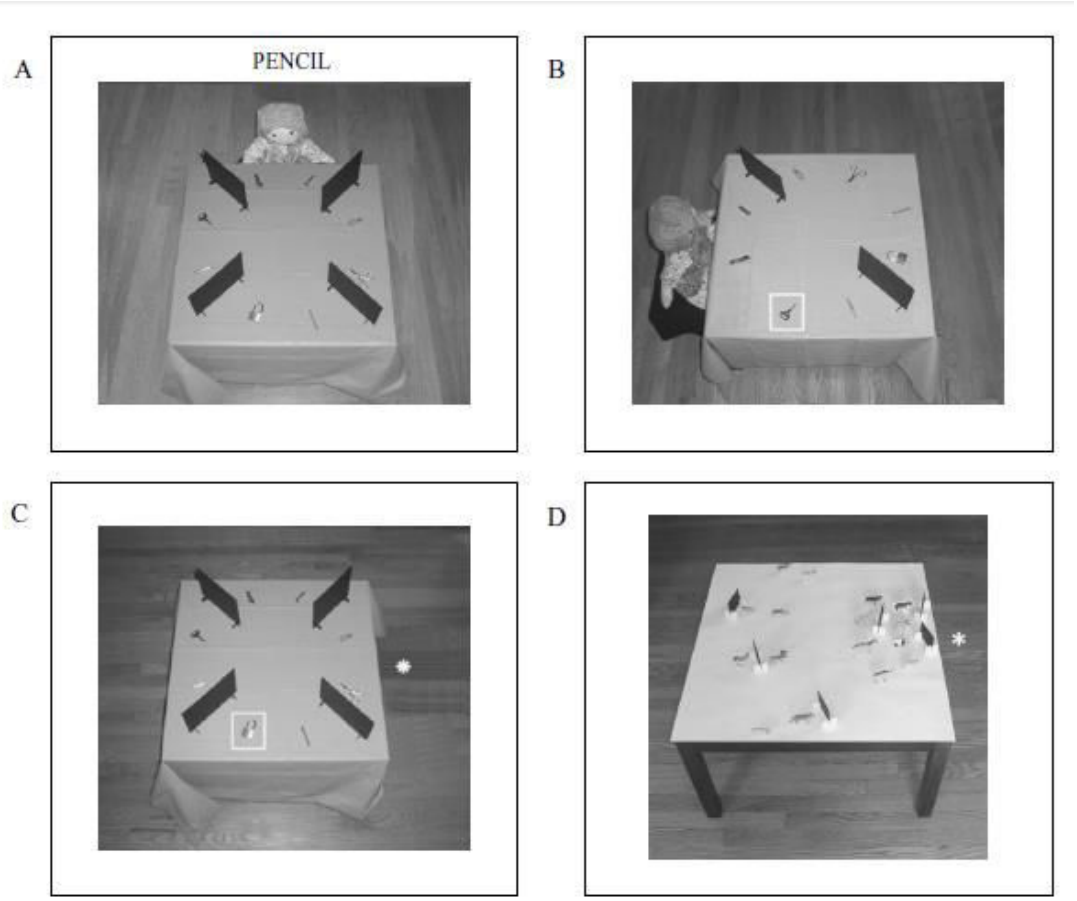


Figure 1. Example of the displays used in Experiment 1 (panel A), Experiment 2 (panel B), Experiment 3 (panel C), and Experiment 4 (panel D). In the left–right task, participants decided whether or not the target object was on the left side of the doll avatar. In the visibility task, they decided whether or not the avatar could see the target object. Photographs of the display were originally color photographs. Target objects were indicated by a yellow square (Experiments 2 and 3) or outlined in yellow (Experiment 4), and a yellow asterisk represented the avatar in Experiments 3 and 4.

Figure 17 : Stimulus utilisés dans l'étude de Michelon et Zacks (2006)

Les résultats de leur étude montrent que pour les jugements de latéralité (droite/gauche) les temps de réaction augmentent avec le degré de rotation entre le participant et l'avatar (résultats similaires à d'autres études). Ils en concluent que ce temps est symptomatique du fait que les participants font des rotations mentales : temps que les participants fassent coïncider mentalement leur perspective avec celle de l'avatar.

Dans les jugements de visibilité, les temps de réaction ne varient pas en fonction du degré de rotation entre le participant et l'avatar. Par contre, ils varient en fonction de la distance entre l'avatar et la cible. Il semble que pour ce jugement, seul la construction mentale d'une ligne de mire soit nécessaire : plus la ligne est longue, plus le temps pour émettre le jugement est long.

Les auteurs concluent que tracer une ligne de mire (line-of-sight) est plus simple que prendre la perspective de l'avatar, cela ne nécessite pas d'imaginer cette nouvelle perspective et n'engendre pas de conflit entre la perspective actuelle et la perspective imaginée.

Dans leur étude, Roberts et Aman (1993) demandent aux participants de déterminer dans quelle direction tourner un triangle pour qu'il fasse face à un point (dispositif très proche de celui de Tamborello

et al, 2012). Les auteurs constatent une augmentation des temps de réaction, uniquement à partir de 90°. Selon eux ce résultat implique qu'il n'y a pas de rotation mentale avant puisque la droite et la gauche du participant et de la référence sont identiques.

3.4 Les énoncés spatialisés

Dans le domaine linguistique, les auteurs ont également eu recours à des systèmes de référence pour catégoriser les énoncés spatialisés. Même si la verbalisation d'informations spatialisées est basée sur des mécanismes cognitifs similaires à la mémorisation et à la perception de l'espace, il n'en reste pas moins que la modalité linguistique diffère et influence l'évaluation des capacités cognitives liées à l'espace (Verjat, 1994). Ainsi, nous avons utilisé jusqu'à présent le terme ***référentiel spatial*** pour désigner le système de coordonnées utilisé pour coder et mémoriser l'information. Nous utiliserons ici le terme ***cadre de référence*** (CdR) pour désigner les systèmes de coordonnées utilisés dans le langage.

Il existe beaucoup de classifications des énoncés spatialisés, souvent créées en fonction de la tâche et des objectifs de l'étude. Par exemple, les études concernant le développement d'outils d'aides à la navigation, vont s'intéresser à la manière dont les points de repères sont introduits et à leur pertinence (Roger, Knutsen, Bonnardel, & Le Bigot, 2013), alors que pour l'étude de tâches de co-manipulation, il sera plus pertinent de s'intéresser aux choix des perspectives (Schober, 1996).

Notre objectif ici est de comprendre comment les opérateurs échangent les informations spatialisées et quelles sont les stratégies collectives ou individuelles de gestion de la charge mentale.
Ces classifications ne sont d'ailleurs pas universelles, et dépendent fortement des langues dans lesquelles elles sont construites. Les langues ne catégorisent pas les choses de la même manière (Hagège, 2005).

Les énoncés contenant des informations spatialisées ne font pas tous appel à un système de coordonnées. Certains énoncés tels que « c'est ici », « there is a bar nearby » (exemple issu de Roger et al, 2011) ou « ils sont à la Mairie » ne font pas appel à un cadre de référence. Le premier a été désigné par certains auteurs de déictiques (c'est-à-dire compréhensible uniquement dans le contexte de l'énonciation), et les deux suivants renvoient à l'utilisation de noms pour désigner des lieux (*placenames*). Diessel (2014) souligne l'importance des termes démonstratifs, qui, selon lui, sont rattachés à des référentiels égocentrés, qu'il nomme également déictiques. Nous reprenons la terminologie proposée par Schober (1995) pour qui ces énoncés sont *neutres* (neutral « that do not depend on any one view »). Ces énoncés sont également difficiles à mettre en relation avec les représentations sous-jacentes. Pour Diessel (2014), les expressions déictiques s'appuient sur un référentiel égocentré. Mais ce que nous retiendrons surtout de ces réflexions, c'est que les énoncés et les référentiels cognitifs associés ne sont dans tous les cas pas dissociables des situations d'énonciation. Mais le choix de ce type d'énoncé peut être stratégique du point de vue collaboratif : étant donné qu'aucune perspective n'est choisie, aucune perspective n'est imposée au destinataire pour comprendre l'énoncé. Nous reviendrons sur ces aspects dans la partie suivante.

Pour les énoncés comprenant un système de coordonnées, il existe de nombreuses classifications (Retz-Schmidt, 1988). Les classifications les plus courantes distinguent les énoncés selon l'origine du cadre de référence. Sur cette distinction, trois types d'énoncés sont proposés : centrés sur une personne (viewer-centered, egocentric, …), qu'il s'agisse du locuteur ou d'une autre personne ; centrés sur un objet (object-centered), avec ses orientations propres ; centré sur l'environnement (environment-centered), généralement ancré sur des repères géographiques et utilisant les directions cardinales (Carlson-Radvansky & Irwin, 1993; Diessel, 2014; Retz-Schmidt, 1988).

Par ailleurs, Levinson (2003) propose une autre classification, résumée Figure 18. Cette classification repose sur trois critères : l'origine du système de coordonnées, la nature des coordonnées et

la présence ou non d'une référence (*relatum*) par rapport à laquelle la cible est localisée. Il définit ainsi les cadres de référence absolu, intrinsèque et relatif. Le cadre de référence absolu est indépendant de tout point de vue et utilise un système de coordonnées indépendant de l'objet à localiser et de la référence. D'autres auteurs parlent d'énoncés viewer-independant (Galati, Pelle, Berthoz, & Committeri, 2010). Le cadre de référence intrinsèque utilise le système de coordonnées de la référence. Enfin, le cadre de référence relatif est le seul cadre de référence tripartite, dans lequel le système de coordonnées est celui d'un observateur et situe la cible par rapport à un autre objet.

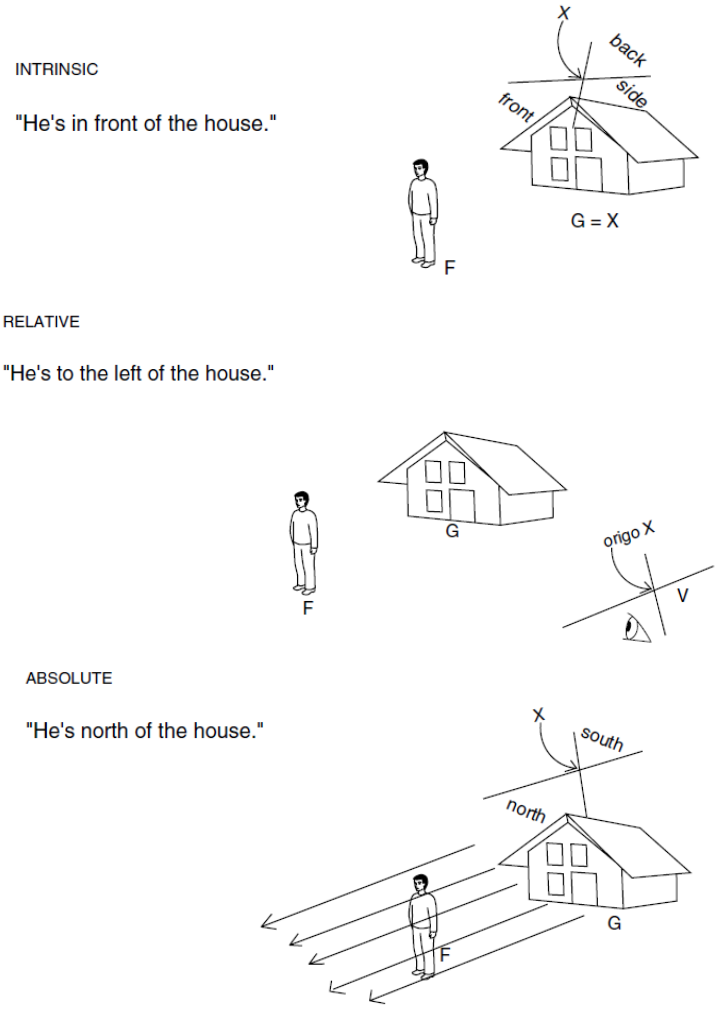


Figure 18 : Les trois cadres de référence proposés par Levinson, d'après Levinson (2003)

Les deux classifications exposées ci-avant sont en fait très générales. Elles servent bien à l'analyse des énoncés, sous l'angle linguistique de leur construction syntaxique, mais reflètent peu les usages et ne tiennent pas compte des processus cognitifs sous-jacents. En effet, d'un point de vue cognitif, il semble que les mécanismes mentaux pour produire un énoncé centré sur soi ou centré sur quelqu'un d'autre ne soient pas les mêmes (nous y reviendrons dans la partie suivante).

Les classifications utilisées lors d'études sur la verbalisation au cours de tâches spatialisées, ont permis de mettre en évidence d'autres catégories d'énoncés, souvent basés sur la perspective depuis laquelle ils sont construits.

Les études portant sur des tâches de navigation ont mis en évidence/s'appuient sur deux types de descriptions : *route descriptions* dont le point de vue change avec la déambulation, les énoncés sont centrés sur la ou les personnes qui naviguent (égocentré), et *survey perspective* dont le point de vue est

situé au-dessus de l'environnement, les énoncés utilisent un Cadre de référence absolu, allocentré ou environment centered (Roger, Bonnardel, & Le Bigot, 2011).

Lors de tâches sur des environnements plus petits, Schober (1995) utilise la distinction entre des énoncés centrés sur le locuteur, le destinataire, sur les deux (c'est-à-dire vrai depuis les deux perspectives, qu'elles soient alignées ou non), centrés sur l'environnement ou neutres.
Cette distinction entre les énoncés centrés sur le locuteur et le destinataire est également présente dans d'autres études (Duran, Dale, & Kreuz, 2011).

Nous proposons donc une classification des énoncés spatiaux à plusieurs niveaux :

Les énoncés *neutres* sont les énoncés sans système de coordonnées. Ces énoncés regroupent les énoncés utilisant les termes déictiques, les noms de lieu, les relations de proximité. Il est difficile d'après ces énoncés de définir les représentations/référentiels sous-jacents. Plusieurs auteurs pensent que ces énoncés sont plus simples à produire et à comprendre. Cependant, pour les énoncés utilisant les noms de lieu, la compréhension mutuelle suppose que le locuteur et le destinataire aient la même association nom/lieu, c'est-à-dire que cette connaissance fasse partie de leur référentiel commun. De même, pour les énoncés déictiques (utilisant le corps pour désigner quelque chose), leur compréhension mutuelle dépend du référentiel commun, qu'il s'agisse de l'espace visuellement partagé par les deux interlocuteurs ou relatif à leur expérience passée.

Les énoncés utilisant un cadre de référence *égocentré* utilisent la perspective du locuteur (speaker-centered, Schober, 1995). Ces énoncés sont basés sur un référentiel égocentré du point de vue du locuteur, mais sur un référentiel exocentré centré sur le locuteur pour les destinataires. Il semble que du point de vue de la production des énoncés spatialisés, ce cadre de référence soit le plus simple à utiliser. Schober (1996) a notamment montré que les énoncés égocentrés étaient produits/initiés plus rapidement que des énoncés exocentrés centrés sur un objet ou une autre personne. Cependant, peu voire pas d'étude n'a mesuré la dimension subjective de cette facilité pour la production.

En opposition aux énoncés centrés sur le locuteur, nous qualifierons les énoncés dont la perspective n'est pas celle du locuteur d'*exocentrés*, c'est-à-dire qui demandent une décentration pour le locuteur, c'est-à-dire un changement de référentiel par rapport au référentiel égocentré du locuteur. Nous distinguons plusieurs cadres de référence exocentrés.

Certains énoncés exocentrés utilisent le point de vue et le système de coordonnées du destinataire de l'énoncé, nommés *centré sur le destinataire* (addressee-centered, Schober, 1995). Le référentiel associé est celui du destinataire. Il semble que ces énoncés soient différents des énoncés égocentrés et centrés sur un objet (Schober, 1996).

D'autres énoncés exocentrés utilisent un objet et ses propriétés intrinsèques pour système de coordonnées, il s'agit des énoncés *centrés sur un objet* (object-centered, Schober, 1995). Le référentiel associé est celui de l'objet.

Enfin, certains énoncés sont centrés sur une personne ou un objet, mais n'utilisent pas leurs propriétés intrinsèques pour définir les directions organisatrices. Il s'agit alors d'énoncés (égo ou exocentrés) *extrinsèques* : le locuteur utilise un système de coordonnées indépendant de l'objet ou de la personne pris comme référence (Table 1). De même que pour les énoncés neutres, il est difficile de dire quel type de représentation sous-tend ce type d'énoncés.

Table 1 : Résumé des cadres de références utilisant un système de coordonnées (référence et directions organisatrices).

		Directions organisatrices	
		Intrinsèques à la référence	Extrinsèques à la référence
Référence	Locuteur	Egocentré intrinsèque	Egocentré extrinsèque
	Autre	Exocentré intrinsèque	Exocentré extrinsèque

Les énoncés spatialisés peuvent renseigner sur la manière dont les collaborateurs se représentent l'espace et échangent les informations spatialisées. Ils donnent également des pistes de réflexion sur les processus à l'œuvre, tant du point de vue individuel que collectif.

3.5 Choix des cadres de référence

De nombreuses études se sont intéressées aux énoncés produits lors de tâches collaboratives réelles ou fictives, dans des environnements d'échelles différentes (Schober, 1993; Tenbrink et al., 2011). Nous présentons ici certains de ces travaux qui nous permettrons, en lien avec les réflexions développées dans les parties précédentes, de dégager des questions de recherche propres aux situations particulières de collaboration à distance.

Les études réalisées par Schober sont très structurantes sur la compréhension des dialogues spatialisés (Schober, 1993, 1995, 1996, 2009a). En 1993, il étudie les perspectives utilisées pour décrire la localisation de cible dans un dispositif présenté Figure 19. Le participant (*director*) devait décrire quel cercle était pointé d'une croix (cible), pour que le destinataire (*matcher*) puisse inscrire une croix sur le cercle compris comme étant la cible. Son protocole comprenait plusieurs conditions : avec ou sans collaborateur. Lorsqu'un collaborateur était présent, les deux participants pouvaient échanger pour réaliser la tâche. Les rôles étaient échangés de deux manières, soit à la moitié des essais (32/64), soit à chaque essai. Chacun disposait d'un exemplaire papier représentant une table circulaire sur laquelle sont placés deux cercles, étaient également représentés par des flèches les points de vue de chacun (à 0°, 90°, 180° ou 270°) (Figure 19 gauche). Les résultats montrent d'une part que les énoncés diffèrent s'il s'agit d'un dialogue réel ou d'un monologue. Les énoncés contiennent plus de mots lorsque le collaborateur est virtuel. Ce résultat peut s'expliquer par l'absence de référentiel commun sur lequel s'appuyer et qu'il n'y a pas de phénomène de grounding possible.

D'autre part, les résultats montrent que les interlocuteurs parlent jusqu'à ce qu'ils soient surs qu'ils se sont compris (grounding). De plus, contrairement au principe du moindre effort (individuel), les *directors* n'ont pas toujours gardé leur perspective même lorsque l'interlocuteur est désaxé par rapport à eux. Schober (1993) suggère que cette stratégie demande un effort cognitif supplémentaire. Il conclut que « la prise de perspective dans une conversation est une chose compliquée. Les locuteurs ne font pas que prendre en compte leur propre biais égocentrique, la position relative du destinataire, et la nature des objets qu'ils décrivent. Ils prennent également en compte ce que leur destinataire renvoie de leur compréhension, et la manière dont certains endroits ont déjà été décrits. »[8].

[8] « spatial perspective-taking in conversation is a complicated affair. Speakers do more than take into account their own egocentric bias, their addressee's relative position, and the nature of the objects they are describing. They also take into account what their addressee gives as evidence of having understood, and how their addressees have already described locations. »

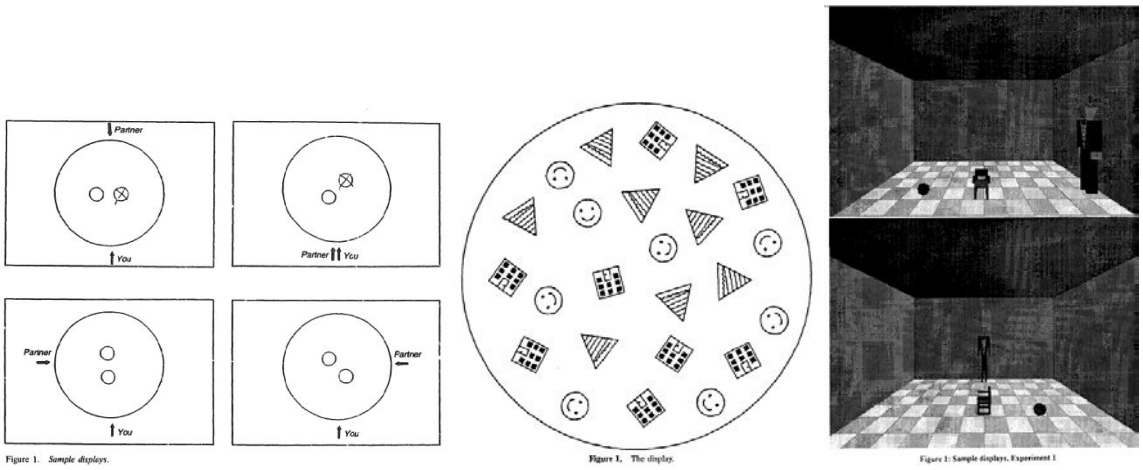


Figure 1. *Sample displays.*

Figure 1. The display.

Figure 1: Sample displays. Experiment 1

Figure 19 : images des dispositifs utilisés par Schober, en 1993, 1995 et 1996

Schober (1995) publie les résultats d'une autre étude dans laquelle il s'intéresse à la répartition de l'effort lors de conversations spatialisées. Un participant (*director*) devait expliquer à un autre participant (*matcher*) comment disposer des jetons sur un plateau de jeu comprenant trois types de figures (Figure 19 milieu). Les participants ne se voyaient pas, et disposaient chacun d'un plateau (identique), qui était soit orienté de la même manière (0°), soit à 90° soit à 180°. Chacun connaissait le point de vue de l'autre sur le plateau. Pour réaliser la tâche, les collaborateurs pouvaient échanger autant qu'ils le souhaitaient. L'analyse a porté sur le choix des Cadre de référence, avec les catégories suivantes : centrés sur le locuteur, sur le destinataire, sur les deux (c'est-à-dire vrai depuis les deux perspectives, qu'elles soient alignées ou non), sur un objet, sur l'environnement ou neutres. Les résultats de cette étude montrent que c'est l'effort de celui qui n'a pas l'information spatiale, donc du *matcher*, qui est minimisé puisque les deux collaborateurs utilisent plus la perspective du *matcher* que celle du *director*. Il semble également que l'effort collectif soit minimisé, puisque l'utilisation de descriptions neutres augmente au cours du temps. Par contre les conclusions restent floues concernant les processus cognitifs à l'œuvre pour gérer les situations les plus difficiles. Schober replace son étude dans trois modèles explicatifs possibles : le modèle des rotations mentales, lesquelles impliquent une difficulté croissante avec le degré de rotation de l'orientation de l'interlocuteur, le modèle de l'inversion des mots (*word-reversal*), selon ce modèle la situation dans laquelle les interlocuteurs se font face serait plus simple que pour lorsqu'ils sont orientés à 90°, et le modèle égocentrique/non égocentrique, qui prédit une difficulté identique quelle que soit le changement d'orientation. Schober a récapitulé les effets attendus selon les trois modèles, chez le locuteur et le destinataire, en fonction de la situation (degré de rotation entre l'orientation du locuteur et du destinataire) et des cadres de référence centré sur le locuteur et l'interlocuteur (Figure 20).

TABLE 1
Three Models of Relative Effort for Speaker- and Addressee-Centered Descriptions

	When Speaker Uses a Speaker-Centered (Egocentric) Description		When Speaker Uses an Addressee-Centered Description	
Offset	**Speaker's *Production***	**Addressee's *Comprehension***	**Speaker's Production**	**Addressee's Comprehension**
Mental-Rotation Model				
Same (0°)	easiest	easiest	easiest	easiest
90°	easiest	harder	harder	easiest
180°	easiest	hardest	hardest	easiest
Word-Reversal Model				
Same (0°)	easiest	easiest	easiest	easiest
90°	easiest	hardest	hardest	easiest
180°	easiest	harder	harder	easiest
Egocentric/Nonegocentric Model				
Same (0°)	easiest	easiest	easiest	easiest
90°	easiest	hard	hard	easiest
180°	easiest	hard	hard	easiest

Figure 20 : Tableau repris de Schober (1995)

Les résultats de cette étude confirment que lors de situations collaboratives, les locuteurs ne construisent pas toujours leurs énoncés dans leur référentiel (égocentré) bien que celui-ci soit supposé être le plus simple. Il semble que dans certaines circonstances, le point de vue du destinataire soit privilégié, notamment, comme c'est le cas dans cette étude, lorsqu'il doit réaliser une action ou prendre une décision à partir de l'énoncé proposé. Bien que l'auteur relie ces résultats à des théories plus générales (modèles des processus cognitifs et principe du moindre effort collaboratif), l'étude se base uniquement sur les énoncés utilisés. L'auteur conclut lui-même que des mesures telles que les temps de latence pourraient compléter l'étude et apporter des arguments sur la répartition de la charge mentale.

Plusieurs autres études ont confirmé ces résultats sur la prise de perspective spontanée de l'interlocuteur, notamment dans des tâches de guidage (Goschler, Andonova, & Ross, 2008; Roger et al., 2011, 2013). Roger et ses collaborateurs (2013) ont même mis en évidence l'effet du rôle, pressentit dans l'étude de Schober de 1995, sur l'introduction des points de repères dans une tâche de guidage en temps réel. Les résultats montrent que le guide introduit plus de points de repères que la personne guidée, même s'ils notent que cette dernière participe activement au dialogue spatial. Ils constatent également que les guides prennent plus facilement la perspective de la personne guidée. Ils discutent leurs résultats en termes de principe du moindre effort collaboratif. Selon eux, la nature des stratégies de réduction des efforts collectifs dépend du rôle joué par la personne au sein du binôme.

Ainsi, plusieurs études ont mis en évidence que la perspective de la personne guidée/de la personne qui réalise l'action est privilégiée, notamment par le guide/celui qui donne les instructions (*director*). Cette stratégie semble faire partie du principe du moindre effort collaboratif pour minimiser la charge mentale de la personne devant interpréter et utiliser l'information spatialisée pour réaliser la tâche (personne guidée, *matcher*). Cette prise de perspective aurait un surcoût pour la personne devant changer de référentiel spatial (guide, *director*). Pour compléter les données sur les efforts liés à ce phénomène de prise de perspective, plusieurs études ont ajouté des mesures de temps de réaction.

Comme évoqué précédemment, Schober (1996) mesure les temps de réaction pour la production d'énoncés soit égocentrés, soit centrés sur un objet latéralisé (une chaise), soit centrés sur le destinataire

(addressee-centered) représenté par un avatar dans la scène (Figure 19 droite). Les résultats montrent que les énoncés égocentrés sont produits plus rapidement que les autres, et que les énoncés centrés sur le destinataire sont produits plus lentement que les autres. Schober en conclut qu'il est plus facile de garder sa propre perspective que de prendre celle du destinataire. Les énoncés centrés sur des objets semblent être d'une difficulté intermédiaire. Cependant, dans cette étude, Schober ne laisse pas le choix aux participants du cadre de référence utilisé, et ne prend pas en compte la difficulté subjective de la tâche.

Dans l'étude de Galati et Avraamides (2014) – déjà évoquée dans la partie précédente - le dispositif très proche de celui proposé par Schober en 1995. La tâche collaborative consistait pour un participant (*director*) à décrire à un autre participant (*matcher*) un dispositif spatial comprenant un axe de symétrie pour qu'il (le *matcher*) le reproduise (Figure 13). Dans l'étude proposée par Galati et Avraamides, les auteurs ont rajouté des tâches de rappel pour voir comment les informations sont stockées en mémoire (résultats présentés dans la partie cognition spatiale). Les auteurs ont également ajouté trois conditions d'alignement du dispositif (avec soit le *director* seul, soit le *matcher* seul, soit avec aucun des deux interlocuteurs) et deux conditions de connaissance de la situation (le *director* sait ou non où se trouve le *matcher*). Les résultats sur les énoncés montrent que selon les conditions, les directeurs vont privilégier leur point de vue, ou l'alignement du point de vue de l'autre avec le dispositif. Ces résultats montrent que la structure intrinsèque et la position relative du collaborateur entrent dans les critères pris en compte dans le choix des cadres de référence. De plus, lorsque plusieurs référentiels sont orientés dans la même direction, il semble que cette direction soit privilégiée, même si elle n'est pas alignée avec le locuteur (*director*).

Outre la prise en compte de la position relative de l'interlocuteur comme facteurs influençant le choix des cadres de référence dans la production d'énoncés spatialisés, Schober (2009) a montré que les locuteurs prenaient également en compte les aptitudes de leurs collaborateurs, notamment spatiales. Cette prise en compte permettrait l'optimisation des capacités de chacun. Ainsi, un locuteur avec des capacités spatiales élevées assumerait plus de prise de perspectives centrées sur le destinataire qu'un locuteur aux capacités spatiales moindres.

Par ailleurs, Carlson-Radvansky et Radvansky (1996) ont montré que le choix du cadre de référence est influencé par la dimension fonctionnelle qui relie la cible (l'objet à localiser) et la référence (l'objet par rapport auquel la cible est localisée). Ainsi, il semble que les énoncés utilisant les caractéristiques intrinsèques de la référence soient préférés lorsqu'il y a un lien logique/fonctionnel entre les deux entités, alors que les énoncés utilisant la perspective du locuteur ou d'autres stratégies sont préférées lorsqu'il n'y a pas le lien fonctionnel est bafoué, comme sur l'image présentée Figure 21.

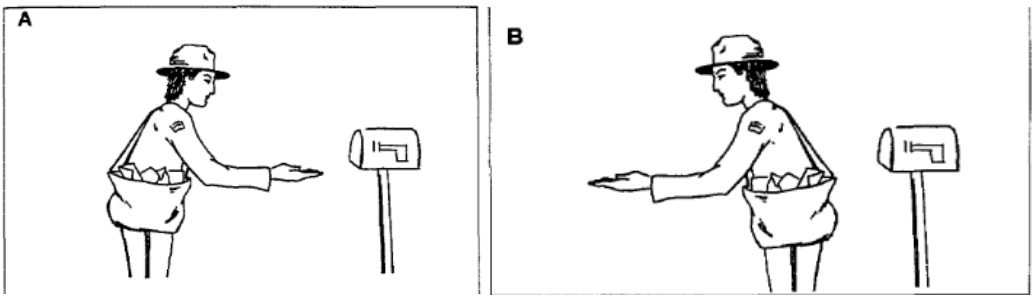


Figure 21 : Stimulus utilisés pour l'étude de Carlons-Radvansky et Radvansky (1996).

Dans une étude sur une tâche de co-manipulation dans un environnement virtuel, Chellali et ses collaborateurs (2013) ont mis en évidence les effets positifs de la présence d'une référence spatiale fixe dans l'environnement virtuel de travail.

Bien que le dialogue soit une activité coopérative, peu d'études s'intéressent à la fois aux mécanismes de production et de compréhension des énoncés spatialisés.

Burigo et Sacchi (2013) se sont intéressés aux mécanismes de compréhension et de production d'énoncés spatialisés dans deux études distinctes. L'étude sur la production d'énoncés spatialisés consistait à décrire la localisation d'objets dans des scènes simples comprenant deux objets, l'objet à localiser étant désigné par un point. Les images utilisées sont celles également utilisées dans l'étude sur la compréhension, Figure 22. Les auteurs ont mesuré à la fois le temps de préparation de l'énoncé et le temps de production. Les résultats confortent l'idée que les énoncés ne sont pas planifiés entièrement avant de commencer la phrase. Ils ont également montré des effets de l'axe sur lequel est positionnée la cible et de l'orientation respective des objets composant la scène.

La partie de l'étude sur la compréhension des énoncés spatialisés consistait à placer la cible ou la référence d'après un énoncé. Les résultats ont mis en évidence l'importance de l'orientation fonctionnelle des objets. Les auteurs ont également mis en évidence qu'il est plus long de placer la référence que la cible car le lien logique doit être reconstruit. Enfin, les auteurs ont mis en évidence que les temps d'action sont plus long lorsque le plan frontal est impliqué que les autres plans.

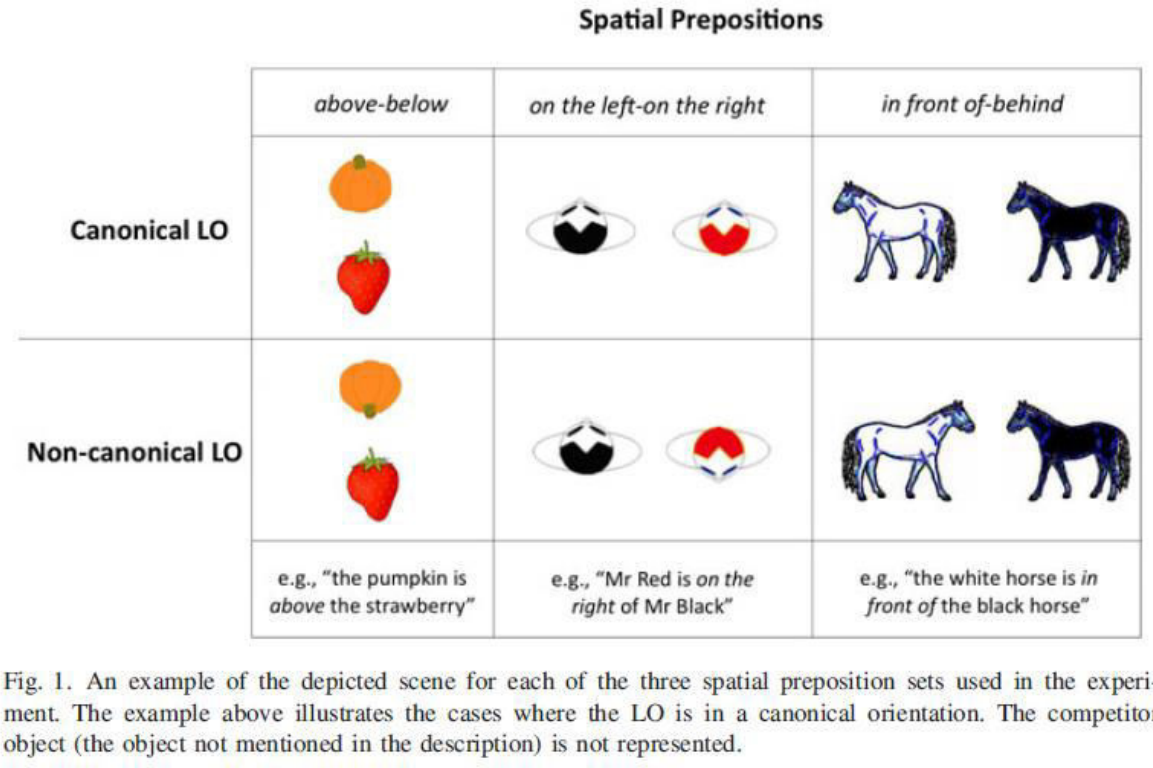


Fig. 1. An example of the depicted scene for each of the three spatial preposition sets used in the experiment. The example above illustrates the cases where the LO is in a canonical orientation. The competitor object (the object not mentioned in the description) is not represented.

Figure 22 : Stimulus utilisés dans l'étude de Burigo et Sacchi (2013)

La compréhension des énoncés demande une interprétation, c'est-à-dire donner du sens dans un cadre de référence, surtout si l'énoncé est ambigu ou non précisé (Tversky & Hard, 2009). Dans leur étude, Duran et ses collaborateurs (2011) ont testé la compréhension d'énoncés spatialisés produits par une personne plus ou moins présente (*worker*, le locuteur). Le dispositif est présenté Figure 23. Cette étude montre plusieurs résultats intéressants. Il semble que les participants ne puissent pas être tous considérés de la même manière. En effet, certaines personnes gardent leur perspective (égocentriques), alors que d'autres prennent la perspective de celui qui donne l'instruction (other-centriques / centré sur le locuteur). Pour les participants qui privilégient la perspective du *worker* (other-centrique), les temps de réaction sont plus longs pour répondre même s'ils sont alignés avec le *worker*. Ce résultat suggère que

cela leur demande plus de temps pour prendre la perspective de l'autre. D'autre part, les résultats montrent l'effet d'informations apportées sur le locuteur : les participants sont informés que le *worker* ne sait pas où se trouve le participant, ce qui sous-entend qu'il ne peut pas prendre la perspective du destinataire, donc que les énoncés seront centrés sur le locuteur. Cette information induit une augmentation du nombre de participant prenant la perspective du locuteur pour comprendre l'énoncé (other-centriques, Duran et al., 2011).

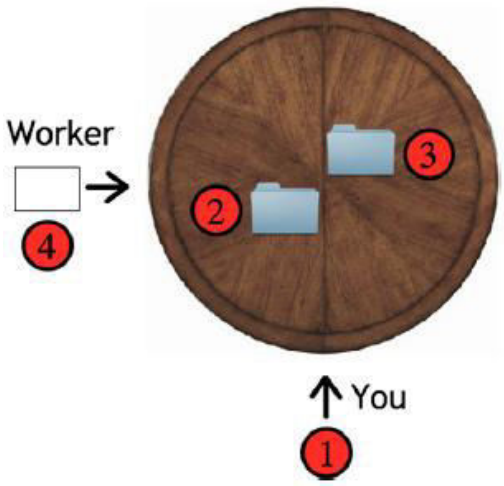


Figure 23 : Stimulus utilisé dans l'étude de Duran et al (2011)

Cette étude donne de plus amples informations sur le fait qu'il serait plus simple de garder sa propre perspective que d'en changer, notamment pour comprendre les messages. Ainsi, on trouve ici un argument pour les situations où les locuteurs prennent la perspective du destinataire. Bien que la prise de perspective du destinataire par le locuteur demande un effort supplémentaire, cette prise de perspective offre un vrai bénéfice du point de vue de la compréhension.

Carlson et Van Deman (2008) ont testé quatre hypothèses concernant la manière dont les cadres de référence sont sélectionnés pour interpréter un énoncé lorsque plusieurs référentiels sont disponibles Ils ont demandé aux participants de déterminer si un énoncé est vrai ou faux, quel que soit le Cadre de référence utilisé. Par exemple l'énoncé « La sphère verte est au-dessus de la voiture » est vrai dans un cadre de référence relatif (du point de vue du locuteur), mais l'énoncé « La sphère verte est devant la voiture » est également vrai, cette fois-ci dans un cadre de référence intrinsèque à la voiture.

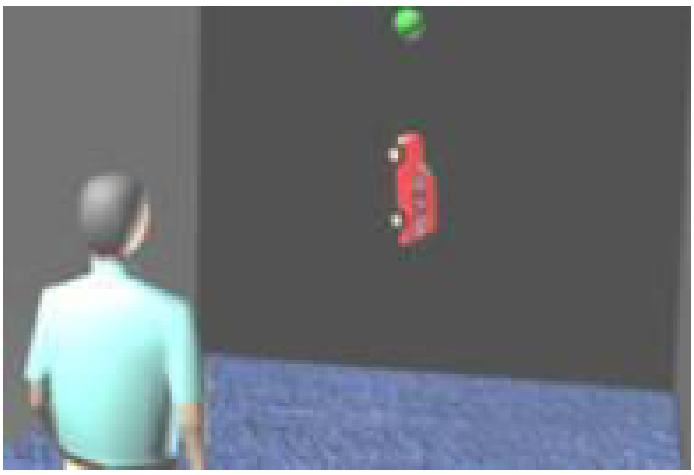


Figure 24 : Stimulus utilisé dans l'étude de Carlson et Van Deman (2008)

Ils ont ainsi regardé l'effet des cadres de référence et directions utilisés dans les énoncés des essais précédents sur les temps de réaction. Des temps de réaction plus longs suggèrent un processus de levée d'inhibition, inhibition induite par l'essai précédent. C'est-à-dire que la compréhension d'un énoncé dans un cadre, axe ou direction de référence induirait une inhibition des autres cadres, axes, ou directions de référence. Les résultats confirment l'hypothèse de préférence des axes. En effet, les effets sont différents selon l'axe et le terme employé.

Dans leur étude, Ishikawa et Kiyomoto (2008) montrent que, dans une tâche de compréhension d'énoncés spatialisés, passer d'instructions données dans un cadre de référence relatif à des instructions données dans un Cadre de référence absolu ne demande pas le même effort que de passer d'instructions données dans un Cadre de référence absolu à des instructions données dans un Cadre de référence relatif. Alors que les résultats ne montrent a priori, pas de différence de difficulté initiale. Ainsi, on peut supposer que pour traiter l'information du Cadre de référence absolu, les participants ont déployé un effort cognitif plus important pour atteindre la même performance. Le fait de passer à un Cadre de référence demandant une CM moindre, leur permet d'atteindre des performances meilleures. En effet, en mobilisant le même effort, mais si la tâche est plus simple, alors les performances sont meilleures. Ils concluent que la construction d'images mentales sont similaires dans les deux Cadre de référence, la différence ne viendrait pas de la représentation (image mentale) mais plutôt du temps requis (lié à la facilité) de traiter l'information dans ces deux cadres de référence. Ils concluent qu'il vaut mieux éviter de mixer les Cadre de référence dans les systèmes d'aide à la navigation.

3.6 Résumé

Ainsi, il semble que pour se construire une représentation spatiale d'un environnement plusieurs processus entrent en jeu. Ces processus varient selon la source de cette représentation. En effet elle peut soit s'appuyer sur la perception de l'environnement présent, soit s'appuyer sur des connaissances mémorisées. D'autre part, cette représentation peut se faire dans différents référentiels spatiaux, selon la source des informations, les caractéristiques de l'environnement et l'utilisation qui en sera faite. Dans les référentiels égocentrés, des mécanismes de prédominance perceptive et fonctionnelle organisent l'espace en zones plus ou moins privilégiées. Dans les référentiels indépendants de la perspective égocentrée, plusieurs niveaux d'informations peuvent être codés : la localisation respective des objets les uns par rapport aux autres, l'orientation des objets qui en ont une, et le codage des objets relativement à d'autres objets selon leurs directions organisatrices propres. De nouveau des mécanismes de prédominance perceptive et fonctionnelle sont à l'œuvre pour organiser l'espace en zones plus ou moins saillantes. Les coûts associés à ces représentations dépendent notamment des efforts liés aux changements de perspectives ou aux rotations mentales nécessaires.

---

4. Synthèse et questions de recherche

Comme nous l'avons vu dans cette revue de la littérature, les environnements virtuels collaboratifs immersifs offrent de nouvelles situations de travail à distance. Ces environnements offrent un support aux activités collaboratives dont la réussite est conditionnée par la mise en place et le maintien d'un référentiel commun opérationnel. Les collaborateurs distants s'appuient sur le dialogue pour échanger les informations. La communication verbale est donc un élément clé de la performance dans ces situations collaboratives distantes.

De par la nature spatiale des informations sur lesquelles les opérateurs de l'industrie aéronautique travaillent, il nous semble particulièrement pertinent de s'intéresser aux communications verbales portant des informations spatialisées. Ces communications verbales s'appuient sur des représentations spatiales. Nous avons mis en évidence, d'après différentes études, que ces représentations sont construites dans des référentiels spatiaux, qui peuvent être égo- ou exocentrés. La source des informations (perceptives ou mnésiques), les caractéristiques de la scène (organisation interne, localisation des interlocuteurs…), mais également la tâche à effectuer influencent la manière dont la scène est représentée et mémorisée. Pour être échangées, les informations spatiales sont communiquées dans des énoncés utilisant différents cadres de référence. Tout comme les représentations, plusieurs études ont montré que le choix des cadres de référence utilisés dans la production d'énoncés spatialisés est conditionné par la situation. Enfin, les énoncés spatialisés doivent être interprétés, et transposés dans le référentiel dans lequel l'action est réalisée.

L'objectif des études présentées ci-après est de mieux comprendre la manière dont les opérateurs construisent le référentiel spatial commun à distance, comment ils gèrent collectivement la charge mentale lors de ce type d'activité collective, et quels processus cognitifs sont à l'œuvre lors de dialogues spatiaux.

D'un point de vue pratique, comme Benford et ses collaborateurs l'ont formulé en 1996, « l'essence des environnements virtuels collaboratifs est que l'espace partagé définit un référentiel spatial commun et cohérent. »[9] L'objectif est de voir comment les EVCIs peuvent soutenir au mieux la collaboration à distance et diminuer la charge mentale induite par l'absence des informations disponibles en présentiel.

---

[9] « The essence of CVEs is that the shared space defines a consistent and common spatial frame of reference. In other words, there is a well established co-ordinate system in which the relative positions and orientations of different objects can be measured. »(Benford, Brown, Reynard, & Greenhalgh, 1996)

Partie expérimentale

Pouliquen-Lardy, Lauriane. Collaboration à distance :
étude de la compréhension mutuelle dans les environnements virtuels collaboratifs immersifs : le cas de la communication spatiale - 2016

Chapitre 3 Effet du rôle dans la communication spatiale et la gestion du moindre effort collaboratif

1. Contexte

Dans cette première partie expérimentale, notre objectif était de proposer un protocole proche des situations réelles observées chez Airbus, et d'étudier l'effet du rôle sur la communication spatiale. Le rôle étant principalement défini par les connaissances préalables, nous avons choisi de prendre des participants novices par rapport aux compétences techniques que des opérateurs de chez Airbus ont et que nous n'aurions pas contrôlées. L'absence de connaissance préalable sur la tâche à réaliser permet également de maitriser les informations constituant le référentiel commun. Ainsi en choisissant les informations fournies à chacun des membres de l'équipe cela permet de définir les rôles de chacun. Nous avons ainsi proposé deux rôles : guide et manipulateur. Ce type de rôle a déjà été utilisé dans d'autres études (Galati & Avraamides, 2014; Goschler et al., 2008; Roger et al., 2013; Schober, 1995). Ces études ont d'ailleurs mis en évidence des effets du rôle sur les cadres de référence utilisés, et ont notamment montré que les guides prenaient la perspective de la personne guidée pour donner leurs instructions. L'objectif ici est d'étudier l'effet du rôle sur la communication spatiale et sur la construction du référentiel commun, avec un regard particulier sur la gestion de l'effort dans une activité collaborative.

Pour réaliser une tâche, un opérateur doit avoir une représentation occurrente de la situation. Cette représentation est basée à la fois sur sa perception de l'environnement, mais également sur ses connaissances préalables. Ainsi l'opérateur a une compréhension de la situation (diagnostic), mais peut également prévoir son évolution (pronostic). Cette conscience de la situation permet notamment d'anticiper les évènements et de gérer les ressources pour optimiser la réalisation de la tâche. De plus, lors d'une tâche collaborative, les opérateurs ont besoin d'informations mutualisées pour coordonner leurs actions et atteindre leur objectif commun. Ces informations mutualisées constituent le référentiel commun. Ce référentiel est mis à jour au cours de la collaboration. Une des sources majeures de ces mises à jour est la communication verbale. Cependant, il ne suffit pas qu'une information soit dite pour qu'elle fasse partie du référentiel commun, le partage de cette information doit être confirmé par les deux parties. C'est ce que Clark et Brennan (1991) appellent le phénomène de *grounding*. Ainsi la communication verbale est une activité dans laquelle les interlocuteurs sont mutuellement investis, c'est-à-dire que l'effort de compréhension mutuelle est supporté par l'ensemble des collaborateurs (Clark & Brennan, 1991). Plusieurs études ont d'ailleurs mis en évidence que cet effort était minimisé sur l'ensemble de la tâche et pas uniquement aux niveaux individuels : c'est le principe du moindre effort collaboratif (Clark & Brennan, 1991; Gergle et al., 2004; Roger et al., 2013; Spante et al., 2004).

Dans les situations qui nous intéressent, plusieurs facteurs peuvent demander des efforts aux collaborateurs. D'une part, ils sont immergés dans un environnement virtuel, leur relation à leur environnement de travail est donc modifiée par rapport à des situations dans un environnement réel. D'autre part ils sont éloignés géographiquement, ce qui implique une communication verbale par téléphone. Nous insistons sur cette dimension technique car elle peut contraindre fortement la collaboration. Pour les situations qui nous concernent, nous nous intéressons particulièrement aux coûts cognitifs liés à la production et à la compréhension. Comme nous l'avons montré dans la partie précédente, plusieurs facteurs peuvent influencer la communication. Par exemple, plusieurs études ont montré que les temps pour prendre la perspective d'un interlocuteur augmentaient avec le degré de rotation de cet interlocuteur (Duran et al., 2011; Michelon & Zacks, 2006; Roberts & Aman, 1993; Schober, 1995). Selon ces auteurs, cette variation dans les temps de réaction correspond aux efforts

demandés pour opérer des transformations mentales. Ainsi, dans certaines situations, notamment lorsque le locuteur et le destinataire regardent dans la même direction, ni l'un ni l'autre n'a besoin d'opérer de transformation mentale pour dire et comprendre l'énoncé suivant : « regarde à droite, il y a un problème ». Inversement, si les interlocuteurs ne sont pas alignés, le locuteur devra opérer une transformation mentale pour produire l'énoncé suivant : « attention sur ta droite ». Le locuteur a également la possibilité d'imposer à son interlocuteur de faire cet effort de prise de perspective s'il dit : « attention, sur ma droite ». De plus, certains énoncés ne demandent pas de transformation mentale, notamment lorsqu'aucune perspective n'est imposée. C'est le cas des énoncés neutres, dans lesquels le locuteur laisse le destinataire choisir la perspective qui lui convient le mieux. Par exemple, un opérateur situé dans la soute qui indique à un collègue situé dans le cockpit : « attention, il y a une fuite dans la soute », se base sur ce qu'il voit pour donner l'information mais n'utilise pas de perspective dans son énoncé. Le destinataire du message n'a alors pas besoin d'imaginer la soute depuis un endroit précis pour comprendre la localisation de la fuite. Ainsi, l'étude des cadres de référence informe sur la manière dont le locuteur choisit de partager l'information à son interlocuteur, mais peut être insuffisant pour comprendre les efforts demandés à chacun pour se comprendre mutuellement. Dans cette étude, nous complèterons l'étude des cadres de référence par l'étude des transformations mentales requises pour produire et comprendre chaque énoncé spatialisé. Ces données renseigneront si les collaborateurs suivent le principe du moindre effort collaboratif en assumant un coût cognitif supplémentaire pour diminuer l'effort global. Mais l'étude du dialogue spatial n'est pas une mesure directe de la charge mentale, et à ce jour aucune étude n'a fait le lien entre prise de perspective, transformations mentales et charge mentale.

Cette première étude propose donc d'étudier la communication spatialisée lors d'une tâche collaborative réalisée à distance dans un environnement virtuel collaboratif immersif, sous différents aspects. D'une part, nous allons regarder les cadres de référence utilisés pour voir quelles perspectives les guides et les manipulateurs choisissent pour partager les informations spatiales. A ce sujet, notre hypothèse est que le rôle affecte les perspectives choisies. Etant donné que c'est le guide qui est censé donner les informations au manipulateur, nous supposons qu'il sera plus enclin à prendre la perspective du manipulateur que l'inverse. Le second aspect sous lequel nous abordons la communication spatiale porte sur les transformations mentales requises pour produire et comprendre les énoncés spatialisés. Notre hypothèse est que plus un locuteur change de perspective plus il devra opérer de transformations mentales, dont des rotations mentales qui ont un coût cognitif. Enfin, la dernière hypothèse concernant cette étude, est que les deux locuteurs vont tenter de minimiser les coûts liés à l'échange d'informations spatialisées. Cela peut par exemple se traduire par l'utilisation privilégiée d'énoncés neutres puisqu'ils n'imposent pas de perspective ni pour le locuteur ni pour le destinataire.

2. Matériels et méthodes

Nous avons proposé une tâche collaborative de guidage-manipulation réalisée à distance en binôme (guide-manipulateur) dans un environnement virtuel collaboratif immersif.

2.1 Participants

Six femmes et vingt-deux hommes, ayant le français pour langue maternelle, ont participé à cette étude. Les participants étaient âgés de 20 à 54 ans, et la moyenne était de 24 ans. Ils ont travaillé en binômes. Pour contrôler un possible effet du genre sur les résultats, aucun binôme n'était exclusivement féminin, et trois participants féminins ont pris le rôle du guide, et trois autres ont pris le rôle du manipulateur.

2.2 Matériels

Les sessions se sont déroulées au NEMO Lab, sur le site d'Airbus Saint-Nazaire. L'environnement virtuel était issu d'une maquette numérique représentant un atelier de l'usine Airbus Saint-Nazaire comprenant deux avions en cours de montage. Les participants étaient immergés dans l'environnement virtuel grâce à deux murs immersifs à double rétroprojection, de 3,5 x 2,2m et de 2 x 3 m. Les murs étant de taille et de qualité d'image différente, pour éviter un effet dû à l'asymétrie technique, chacun des rôles a été associé autant de fois à chacun des murs. Les participants étaient équipés de lunettes stéréoscopiques et d'une manette (wiimote) pour les déplacements (dans les trois directions) et la sélection d'objets. Une fiche technique était fournie aux participants au tout début de la session (avant la phase d'entrainement) expliquant comment procéder pour se déplacer et accéder au menu (Annexe 2). Les lunettes et la manette étaient équipées de marqueurs, associés à un système de tracking pour la parallaxe, la localisation et l'orientation du pointeur (associé à la manette). Chaque participant était représenté par un avatar composé d'une tête et d'un laser, comme illustré Figure 25.
Pour les sessions collaboratives, les participants étaient équipés d'oreillettes téléphoniques. L'enregistrement des conversations était assuré par un dictaphone relié à un troisième téléphone (audioconférence). Les sessions étaient également filmées.



Figure 25 : Un des participants devant un des murs immersifs. La tête rose et le laser sont l'avatar du collaborateur. Les participants devaient déplacer le cube bleu transparent sur la droite de l'écran.

2.3 Procédure

A leur arrivée, les participants étaient présentés l'un à l'autre et à l'équipe présente sur place (personnel technique et chercheurs). Après une explication du déroulement de la séance, les participants étaient répartis dans les salles associées aux deux dispositifs immersifs. Ils prenaient alors connaissance du scénario et de leur rôle. Le scénario proposé était la préparation de la venue d'un expert conformité,

basé à Toulouse, pour le contrôle de plusieurs pièces dans un des avions de l'atelier de Saint-Nazaire. Le contrôle des pièces en question devait immobiliser tout l'atelier, et la direction souhaitait que cette immobilisation dure le moins longtemps possible. Pour anticiper la venue de l'expert, une session collaborative distante a été organisée entre l'expert conformité (manipulateur) et le responsable de l'atelier de Saint-Nazaire (guide).

Les rôles étaient expliqués dans des notices imprimées (Annexe 3 et Annexe 4). Après en avoir pris connaissance, les participants avaient un entrainement spécifique à leur rôle. Pour des raisons pratiques de lisibilité dans les sections résultats et discussion, nous garderons les termes guides et manipulateurs, plutôt que responsable atelier et expert conformité.

Les informations données au guide portaient sur l'atelier, ses contraintes techniques et financières. L'entrainement se déroulait dans le même environnement que la session collaborative, à savoir la représentation de l'atelier et des avions. Les participants devaient visiter l'ensemble de l'environnement (les trois niveaux), et apprendre les modes de déplacement des différents obstacles présents dans l'environnement. Pour cela, une personne support venait manipuler les objets pour faire une démonstration de leurs modes de déplacement (rotation ou translation). Pour nous assurer de leur niveau de connaissance de l'environnement, les guides devaient réaliser un plan de l'environnement après leur séance d'entrainement.

Le manipulateur recevait quant à lui des informations sur l'équipement (contraintes techniques de manipulation), et l'emplacement des pièces à contrôler (sans détail de l'environnement). La phase d'entrainement se déroulait dans un environnement virtuel représentant un atelier vide, dans lequel était présent le cube représentant l'équipement à déplacer (Figure 25) et différents objets à déplacer. L'objectif de cette session d'entrainement était que le participant soit à l'aise dans la manipulation des objets virtuels (sélection d'objets, déplacement seul ou avec objet) et plus particulièrement de son équipement.

Une fois les phases d'entrainement finies, les participants étaient immergés dans le même environnement et mis en relation par téléphone. Ils avaient alors 45 minutes pour réaliser la tâche.
Une fois le début de la session collaborative lancé, les participants étaient libres de parler et de se déplacer. Leur tâche était de déplacer le cube transparent depuis le point de départ (niveau 0, Figure 26) à deux autres points dans un des avions de l'atelier (A et B, respectivement aux niveaux 1 et 2). Il n'y avait qu'une seule et unique solution possible, illustrée Figure 26 droite. Pour accéder au point A, le chemin le plus court était celui passant par l'arrière de l'avion. L'encombrement était tel que l'équipement du manipulateur ne passait pas. Les collaborateurs devaient alors passer par le côté droit de l'avion, ce qui nécessitait l'ouverture d'un portique et l'installation d'une rampe (puisque le matériel disposé sur roulettes ne pouvait pas être porté). Pour accéder au point B, l'arrière de l'avion était obstrué par des équipements dont les coûts de déplacement prohibitifs engageaient à tester une autre solution. Cependant, l'accès par la porte latérale de l'avion donnait sur une impasse étant donné que la cloison intérieure de l'avion ne permettait pas le passage de l'équipement de contrôle (à quelques centimètres près ; cela n'était donc pas détectable visuellement : les participants devaient le constater sur place). Les participants devaient donc déplacer les équipements obstruant l'accès par l'arrière de l'avion.

Une fois les 45 minutes passées ou la mission accomplie, un questionnaire était proposé aux participants (Annexe 5). Ce questionnaire comprenait des questions sur l'état physique de la personne (fatigue, nausées), la phase d'entrainement, les sentiments de présence et coprésence, la répartition des tâches et le contrôle de la situation, les éléments sur lesquels ils se sont basés pour réaliser la tâche, le niveau de compréhension, la satisfaction, les besoins techniques identifiés, et un questionnaire général (âge, genre, latéralité).

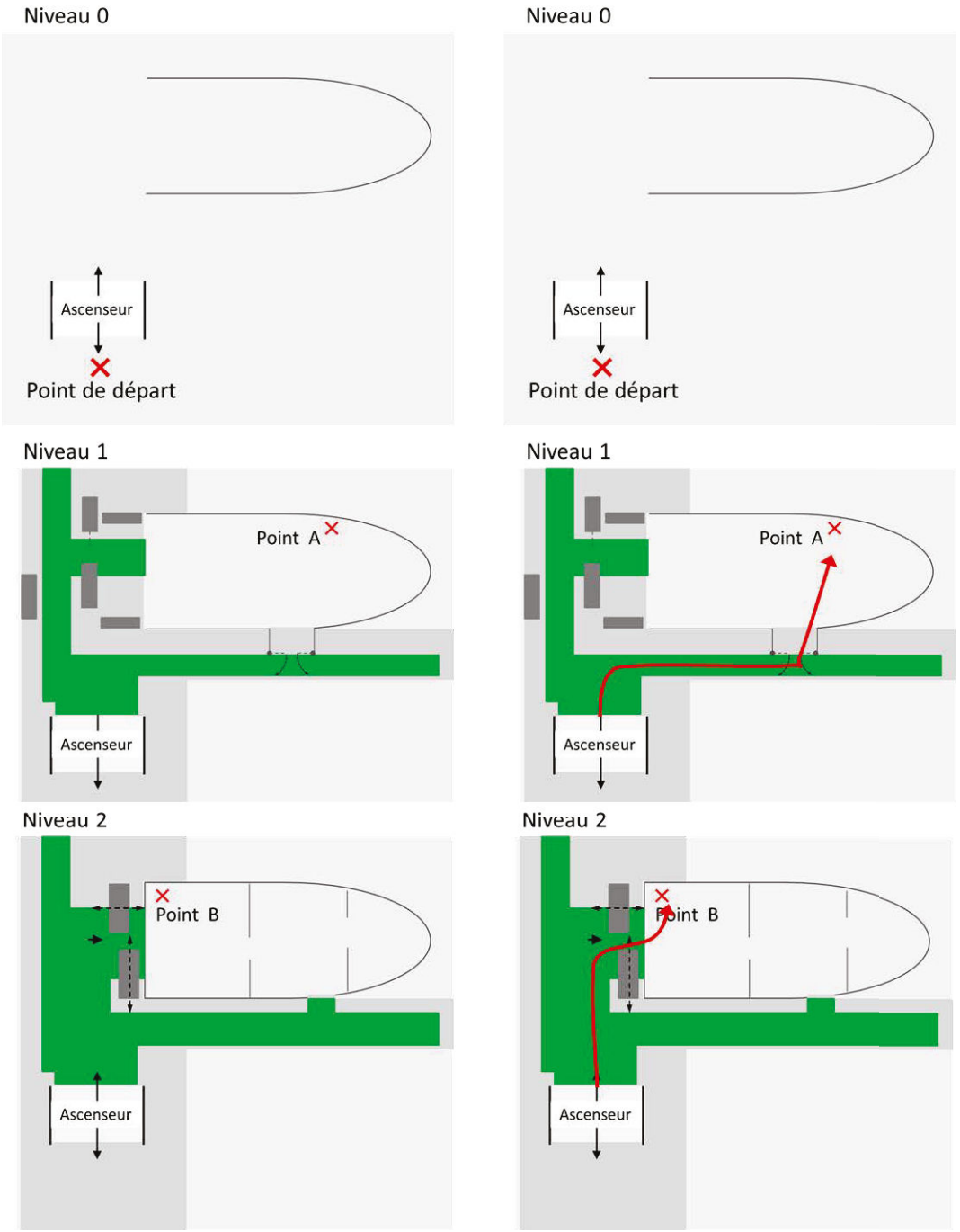


Figure 26 : plan de l'environnement virtuel. La partie de droite illustre la seule et unique solution possible.

Sept binômes ont également répondu au NASA TLX (Annexe 1) pour mesurer la charge mentale de la tâche, et ont passé deux tests pour mesurer leurs aptitudes spatiales : le Santa Barbara Sense of Direction scale (Annexe 6), et le test de rotations mentales MRT-A ©, présenté en Annexe 7. Le NASA TLX est un questionnaire en deux étapes, où les participants doivent d'abord noter chacune des composantes (exigence mentale, exigence physique, exigence temporelle, effort, performance, frustration) puis à pondérer le poids de chaque composante par rapport aux autres. Le Santa Barbara Sense of Direction Scale est une auto-évaluation basée sur quinze questions, du type « Je suis doué pour indiquer des directions », auxquelles les participants répondent sur une échelle continue bornée entre « Entièrement d'accord » et « Désaccord complet ». Le score global est calculé grâce à la moyenne des réponses aux

quinze questions. Le Mental Rotation Test consiste en une succession de 24 lignes de test. Sur chaque ligne, le participant doit trouver, parmi quatre images proposées, les deux correspondant à des vues d'un objet cible ayant subi une rotation sur l'axe vertical. Le score global est la somme des tests réussis.

2.4 Analyse des énoncés spatiaux

Les données audio et vidéo ont été traitées a posteriori. Les conversations ont été retranscrites à l'aide du logiciel Adobe Premier Elements® puis analysées l'aide d'Actogram Kronos®, un logiciel d'horodatage et de codage d'observations.

Les énoncés ont été codés selon le cadre de référence utilisé. Pour cela nous avions défini cinq catégories :

- neutres : les énoncés n'utilisent pas de perspective pour localiser l'objet de l'énoncé. ;
- égocentrés : les énoncés utilisent la perspective du locuteur ;
- centrés sur le destinataire : les énoncés utilisent la perspective du destinataire ;
- centrés sur un objet : les énoncés utilisent un objet, avec ses propriétés intrinsèques et son orientation, comme référence ;
- centrés sur autre chose : énoncés utilisant une perspective non définie par un interlocuteur ou un objet présent à l'instant de l'énoncé (perspective utilisée par le passé ou localisation future de l'un des interlocuteurs par exemple).

De plus, les énoncés étaient également codés en fonction des transformations mentales requises pour être produites et comprises. Ce codage ne pouvait être réalisé qu'avec les informations issues des vidéos. Nous avons alors distingué trois catégories d'énoncés :

- pas de transformation mentale requise (sans TM) ;
- transformation mentale sans perspective imposée (SP) ;
- rotation mentale (RM).

Etant donné que les énoncés ont été horodatés, il a été possible de regarder l'évolution des énoncés produits au cours du temps. Pour cela, nous avons choisi de diviser chaque session en quatre quartiles de durée égale. Nous avons ensuite regroupé le nombre d'énoncés par quartiles pour tous les binômes.

Afin d'illustrer la manière dont les énoncés ont été codés, nous proposons un exemple issu de l'une des sessions collaboratives. Les participants sont tous les deux au niveau 1, ils viennent d'essayer de passer par le portique latéral et se dirigent vers l'arrière de l'avion. Le guide a devancé le manipulateur et est arrivé à sa position décrite sur la Figure 27. Le manipulateur a suivi le guide, mais a été ralenti par des problèmes de manipulation. Il n'a pas vu le guide tourner à droite. Les photos de la Figure 28 montrent ce que chacun des participants voit au moment de l'échange reporté ci-dessous.

1- Guide : « Alors tu suis le tapis vert. »

2- Manipulateur : « J'avance ou je tourne à droite ? »

3- Guide : « Ouais tu essaies de tourner à droite. »

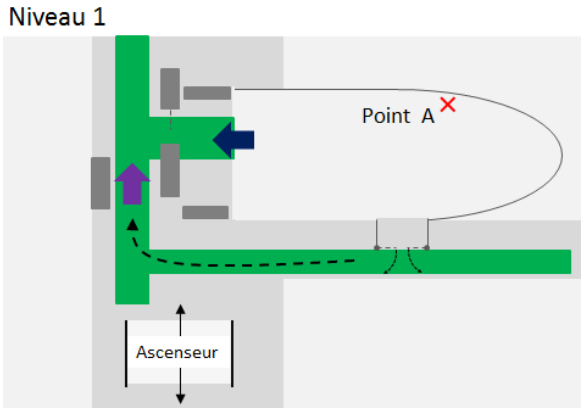


Figure 27 : Localisation et orientation du guide et du manipulateur au moment de l'échange reporté ci-avant.

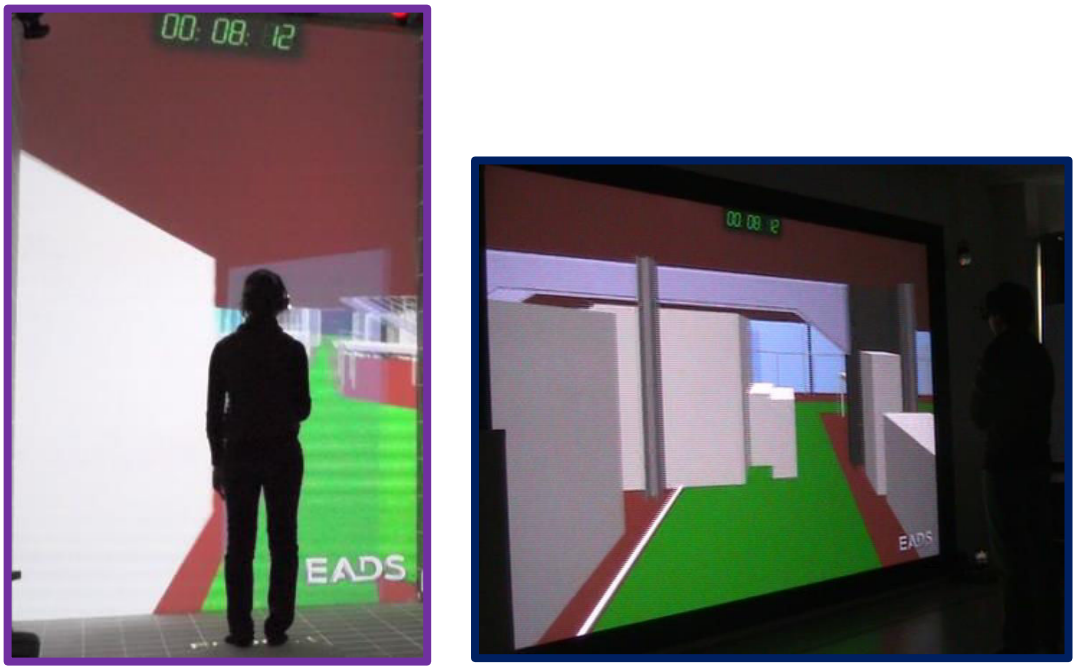


Figure 28 : Champs de vision des participants au moment de l'échange. A gauche le manipulateur, à droite le guide.

L'énoncé 1 est classé comme Neutre, Sans transformation mentale pour le locuteur, Sans transformation mentale pour le destinataire. L'énoncé 2 est classé comme Egocentré, Sans transformation mentale pour le locuteur, Rotation mentale pour le destinataire. Enfin, l'énoncé 3 est classé comme Centré sur le destinataire, Rotation mentale pour le locuteur, Sans transformation mentale pour le destinataire.

2.5 Traitements statistiques

Les analyses de fréquences, nombre d'énoncés pour les trois catégories (cadres de référence, transformations mentales pour la production et la compréhension), ont été faites grâce à des tests du Khi2. Les moyennes étaient comparées par des t-tests bilatéraux.

Pour tous les tests, le niveau de significativité utilisé était $p<.05$.

3. Résultats

3.1 Enoncés spatialisés

De manière générale les guides ont produits plus d'énoncés que les manipulateurs, en moyenne respectivement 93,2 (écart-type : 50,5) et 67 (écart-type : 25). Cette différence n'est pas significative, (t(18)=1.46, p=.16), à cause de la grande variabilité interindividuelle.

Les comparaisons du nombre d'énoncés pour chacune des trois catégories (cadres de référence, transformations mentales requises pour la production et la compréhension) ont montré un effet significatif du rôle (guide vs manipulateur) dans chacune des catégories (respectivement χ^2 (5, 1602) = 116.08, p<.001 ; χ^2(2, 1602) = 116.08, *p*<.001 et χ^2(2, 1602) = 45.12, *p*<.001). Les comparaisons détaillées entre le guide et le manipulateur sont présentées ci-après.

3.1.1 Cadres de référence

Une analyse plus détaillée pour chaque cadre de référence, illustrée Figure 29, a permis de montrer des différences significatives pour trois des cinq catégories. Les guides ont utilisé significativement plus d'énoncés **neutres** et **centrés sur le destinataire** que les manipulateurs (respectivement χ^2 (1, 1602) = 8.09, p<.01 and χ^2 (1, 1602) = 76.2, *p*<.001). Inversement, les manipulateurs ont utilisé significativement plus d'énoncés égocentrés que les guides (χ^2 (1, 1602) = 53.59, *p*<.001).

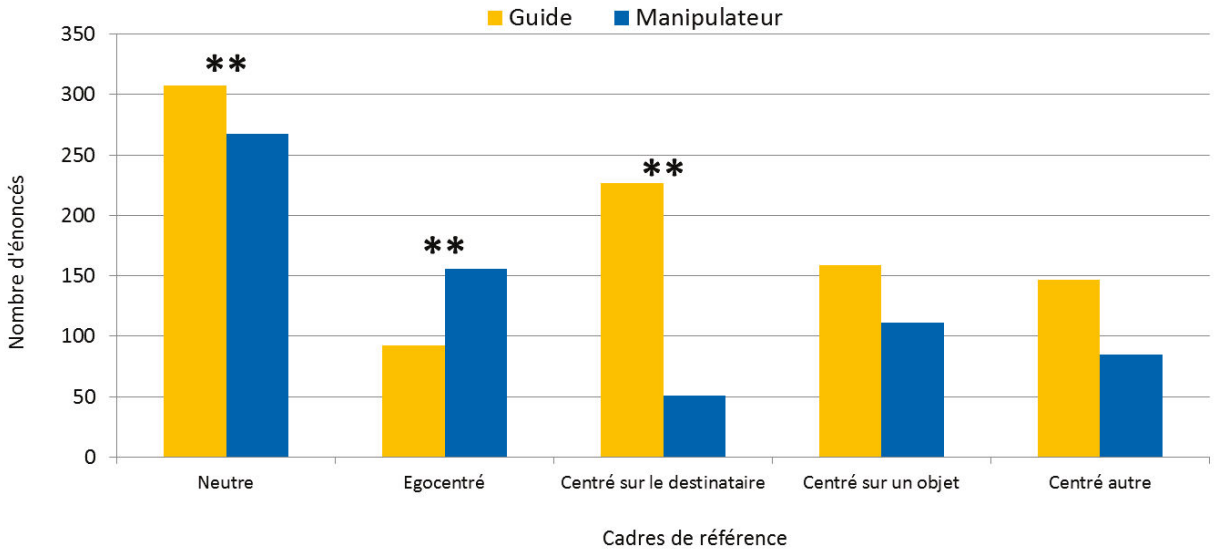


Figure 29 : Nombre d'énoncés produits par les guides et les manipulateurs selon les cadres de référence utilisés. ** *p*<.001

Nous nous sommes également intéressés à la répartition des énoncés au cours du temps. Pour étudier un éventuel effet du temps, nous avons divisé chaque session collaborative en quatre quartiles de durées équivalentes et comptabilisé le nombre d'énoncés de chaque cadre de référence pour chacun des quartiles. La répartition est significativement différente chez les guides (χ^2 (12, 831) = 22.85, p<.05) et les manipulateurs (χ^2 (12, 556) = 21.99, p<.05).

Nous avons regardé plus particulièrement la différence entre le premier et le dernier quartile pour les énoncés centrés sur le destinataire et égocentrés, illustrée par la Figure 30. Cette différence est significative pour les guides (χ^2 (1, 151) = 8.9, p<.01) et les manipulateurs (χ^2 (1, 88) = 14.31, p<.001).

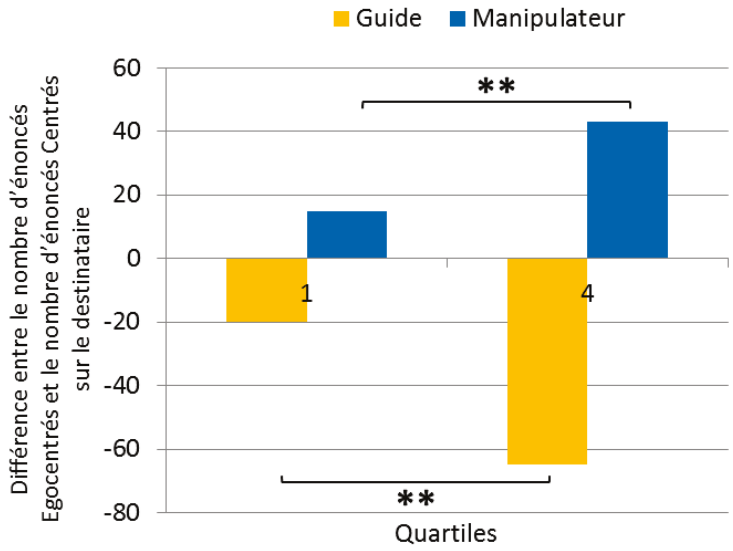


Figure 30 : Différence d'énoncés égocentrés et centrés sur le destinataire pour les guides et manipulateur entre les quartiles 1 et 4. ** *p*<.01

*3.1.2 Transformations mentales pour la **production** des énoncés*

Une analyse plus détaillée pour chaque type de transformations, illustrée Figure 31, a permis de montrer des différences significatives pour deux des trois catégories (Table 2). Les guides ont utilisé significativement plus d'énoncés ne requérant pas de transformation mentale, ainsi que plus d'énoncés requérant des rotations mentales (respectivement $\chi^2(1, 1602) = 19.44$, $p<.001$ et $\chi^2(1, 1602) = 34.35$, $p<.001$).

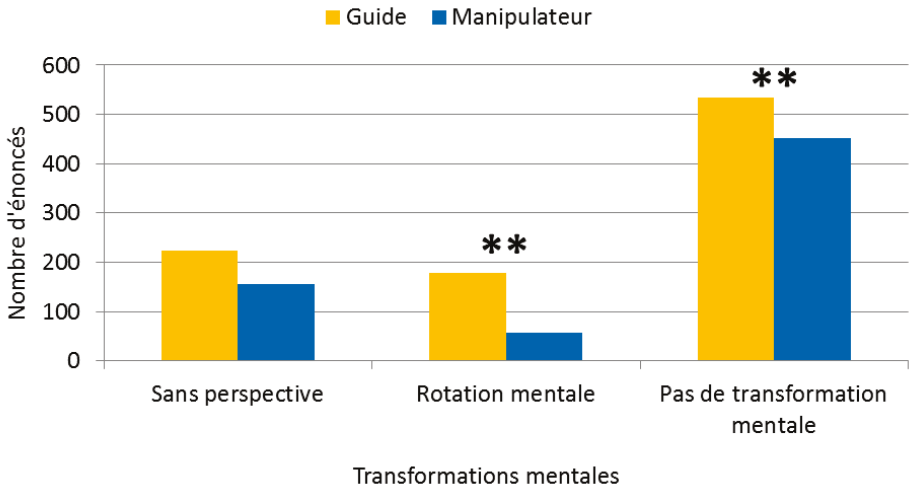


Figure 31 : nombre d'énoncés produits par les guides et les manipulateurs selon les transformations mentales requises pour leur production. ** *p*<.001

Table 2 : Nombre d'énoncés produits par les guides et les manipulateurs, et leurs pourcentages associés, selon le type de transformation mentale requis pour la production des énoncés

	Guide	Manipulateur
Sans perspective	223 (24%)	156 (23%)
Rotation mentale	179 (19%)	57 (9%)
Pas de transformation mentale	535 (57%)	452 (68%)

*3.1.3 Transformations mentales pour la **compréhension** des énoncés*

Une analyse plus détaillée pour chaque type de transformations, illustrée Figure 32, a permis de montrer des différences significatives pour deux des trois catégories (Table 3). Les guides ont utilisé significativement plus d'énoncés ne requérant pas de transformation mentale pour être compris par le destinataire que les manipulateurs ($\chi^2(1, 1602) = 20.34$, $p<.001$). Inversement, significativement plus d'énoncés requérant des rotations mentales pour être compris par les destinataires ont été produits par les manipulateurs que par les guides ($\chi^2(1, 1602) = 40.16$, $p<.001$).

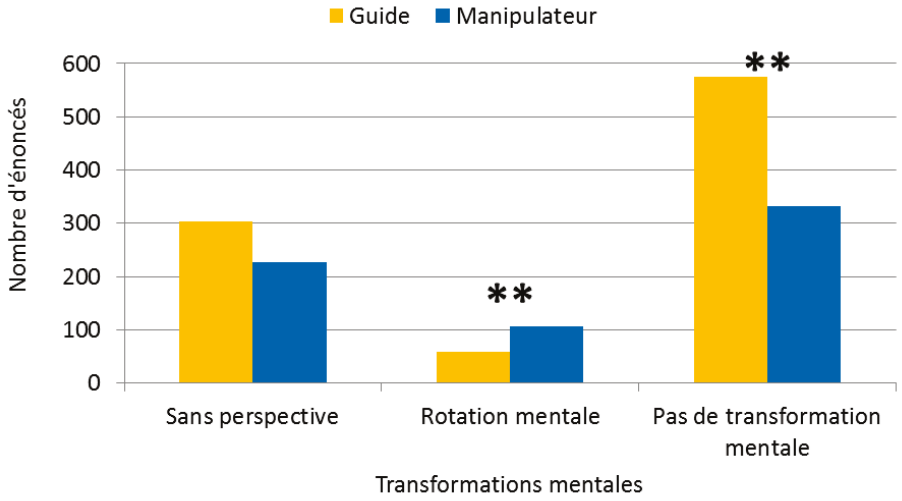


Figure 32 : nombre d'énoncés produits par les guides et les manipulateurs selon les transformations mentales requises pour leur compréhension. ** *p*<.001

Table 3 : Nombre d'énoncés produits par les guides et les manipulateurs, et leurs pourcentages associés, selon le type de transformation mentale requis pour la compréhension des énoncés

	Guide	Manipulateur
Sans perspective	304 (32%)	226 (34%)
Rotation mentale	59 (6%)	107 (16%)
Pas de transformation mentale	574 (61%)	332 (50%)

3.2 Questionnaires

3.2.1 Charge mentale

L'analyse de la charge mentale a montré un score global plus élevé pour les manipulateurs que pour les guides ($t(12)=2.52$, $p<.05$). La comparaison des composantes pour chacun des rôles n'a révélé aucune différence pour les manipulateurs : toutes les sous-échelles ont obtenus des scores similaires. Cependant, pour les guides, l'exigence mentale est significativement plus forte que les cinq autres composantes ($p<.05$ pour chaque comparaison).

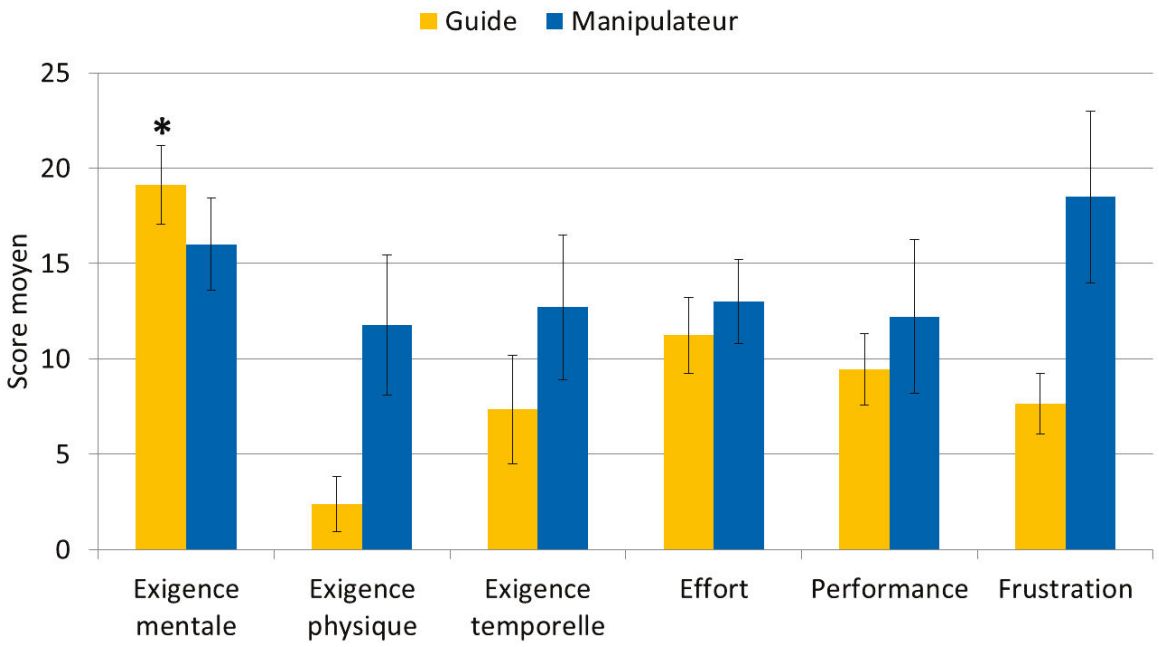


Figure 33 : Evaluations pondérées moyennes de chacune des six composantes de la charge mentale du NASA TLX. L'astérisque illustre la différence significative entre l'exigence mentale avec les cinq autres composantes pour les guides uniquement (*p<.05).

3.2.2 Présence – coprésence

Les participants devaient estimer leurs sentiments de présence et de coprésence sur une échelle bornée (nul / très fort) de 10 centimètres. En moyenne, les participants ont estimé leur sentiment de présence à 6.5 cm (*SD* = 1.9) et de coprésence à 8 cm (*SD* = 1.97). L'analyse des sentiments de présence et coprésence n'a montré aucune différence significative entre guides et manipulateurs (présence: t(26)=0.63, p=.53 ; coprésence: t(26)=0.54, p=.59).

3.2.3 Test d'habiletés spatiales

Pour les sept binômes qui ont passé les tests de capacités spatiales, il n'a pas été observé de lien entre la gestion de la collaboration et les aptitudes spatiales mesurées.

4. Synthèse et discussion

Cette étude s'est intéressée à la communication verbale véhiculant des informations de nature spatiale dans des tâches collaboratives réalisées à distance dans des environnements virtuels collaboratifs immersifs. Elle nous a permis de dégager un ensemble de résultats que nous allons maintenant résumer et discuter, notamment par rapport aux hypothèses et à l'état de la littérature.

En ce qui concerne les cadres de référence utilisés pour le partage des informations spatialisées, les guides ont produit plus d'énoncés neutres et centrés sur le destinataire que les manipulateurs. Inversement, les manipulateurs ont utilisé plus souvent le cadre de référence égocentré que les guides. De plus, les écarts entre les cadres de référence égocentrés et centrés sur le destinataire, déjà significatifs au début des sessions collaboratives, sont plus importants à la fin des sessions, et ce pour les deux rôles, mais dans des sens opposés.

En ce qui concerne les transformations mentales requises pour la production et la compréhension des énoncés, les guides ont dû opérer plus de rotations mentales que les manipulateurs. Il est également intéressant de noter que tous les participants ont principalement produit des énoncés ne nécessitant pas de transformation mentale dans la phase de production. Ce constat est différent du point de vue de la compréhension. Les guides ont en effet produit principalement des énoncés ne nécessitant pas de transformation mentale pour être compris (environ les deux tiers), alors que les manipulateurs ont produit presque autant d'énoncé nécessitant des transformations mentales que n'en nécessitant pas.

Les manipulateurs ont évalué la charge mentale globale de la tâche comme plus importante que les guides. D'autre part, la répartition de cette charge mentale n'est pas la même selon le rôle des participants : les manipulateurs n'ont pas mis en évidence de différences entre les six composantes de la charge mentale, alors que chez les guides, l'exigence mentale a été considérée comme plus importante que toutes les autres composantes.

4.1 Effet du rôle sur les cadres de référence

La première hypothèse concernait l'effet du rôle des collaborateurs sur les cadres de référence choisis pour partager les informations spatialisées, et plus précisément le fait que les guides prendraient plus facilement la perspective des manipulateurs que l'inverse. Les résultats présentés ici valident cette première hypothèse. En effet, lorsqu'il s'agissait de prendre la perspective de l'un des collaborateurs, la perspective du manipulateur a été privilégiée par les deux participants. Comme évoqué dans la partie 4.2 Choix des cadres de référence (page 47), Schober (1995) et Roger et collaborateurs (2013) avaient fait des constats similaires sur l'influence du rôle sur les cadres de référence. Dans l'étude de Schober (1995) la tâche de guidage/manipulation était à une échelle plus petite que celle de notre étude puisque les participants devaient disposer des jetons sur un plateau de jeu (Figure 19 milieu). Cependant, il avait observé que la perspective du *matcher* (manipulateur) était privilégiée. Dans une tâche de guidage à plus grande échelle, Roger et collaborateurs (2013) font également le constat que les guides prennent facilement la perspective de la personne guidée.

De plus, les résultats de notre étude montrent que ces choix sont plus marqués à la fin de la session qu'au début. Ce renforcement peut être dû au bénéfice apporté par l'utilisation des cadres de référence centrés sur le manipulateur, confortant ainsi les collaborateurs dans leurs manières d'échanger les informations spatiales. Ainsi, lorsqu'une personne guide une autre personne, l'utilisation de la perspective de la personne guidée pour donner les instructions est un phénomène robuste qui ne doit pas être sans lien avec la construction du référentiel commun et la gestion de l'effort collectif. Nous reviendrons sur ces deux points dans la troisième et quatrième partie de cette discussion. Schober (1995) avait quant à lui observé une augmentation du nombre d'énoncés neutres. Son hypothèse était que les

énoncés neutres minimisaient l'effort des deux collaborateurs. Il semble donc que selon la tâche, les cadres de références privilégiés varient.

4.2 Prise de perspective et transformations mentales

La seconde hypothèse concernait les transformations mentales requises pour produire et comprendre les énoncés spatialisés. Elle stipule que plus un opérateur change de perspective, plus il doit opérer de transformations mentales pour produire ses énoncés. Etant donné que les guides ont plus souvent changé de perspective que les manipulateurs, ils devraient avoir produit moins d'énoncés ne requérant pas de transformation mentale. Hors les résultats montrent deux choses, d'une part les guides ont produit proportionnellement plus d'énoncés ne requérant pas de transformation mentale que les manipulateurs, mais ils ont également produit plus d'énoncés nécessitant des rotations mentales. De ces résultats il faut surtout retenir que la relation entre cadres de référence et transformations mentales n'est pas si directe qu'elle semble l'être. En effet, selon la manière dont les opérateurs se positionnent et naviguent dans l'environnement, il est possible que certains changements de perspectives ne demandent pas d'opérer de transformation mentale pour produire un énoncé spatialisé, et par conséquent soient moins coûteux qu'il pourrait y paraître. Ce résultat est en opposition avec les modèles proposés par Schober (1995) pour qui un changement de perspective s'accompagne nécessairement de rotations mentales.

Le second résultat montre que lorsque des transformations mentales étaient nécessaires pour changer de perspective, les guides ont opéré plus de rotations mentales que les manipulateurs. Ce résultat confirme que les guides assument plus d'effort cognitif que les manipulateurs pour échanger les informations spatiales, mais uniquement lorsque cela est nécessaire. Ainsi, les analyses réalisées dans cette étude ne permettent pas de conclure sur le lien entre prise de perspective et transformation mentale, et encore moins sur les coûts cognitifs de ces prises de perspectives, notamment parce que le positionnement relatif des deux interlocuteurs n'était pas contrôlé. Les études sur les énoncés spatialisés n'ont jusqu'à présent pas remis cette idée en cause, et il semble communément admis que changer de perspective induise systématiquement un surcoût cognitif (Goschler et al., 2008; Schober, 1995). Les résultats présentés ici suggèrent une autre alternative : changer de perspective nécessiterait d'opérer des transformations mentales uniquement dans certaines conditions (Kessler & Rutherford, 2010). Ce sont des questions auxquelles nous tenterons de répondre dans la seconde partie expérimentale.

4.3 Gestion collective des efforts

La troisième hypothèse de cette étude portait sur la gestion collective des efforts dans la réalisation collective de la tâche, dont nous pensions observer des manifestations du principe du moindre effort collaboratif. Comme nous l'avons déjà évoqué, les guides ont minimisé l'effort de compréhension des manipulateurs tout en ne s'imposant pas toujours de surcoût individuel pour donner leurs indications. Les résultats des transformations mentales requises pour la compréhension des énoncés montrent que les manipulateurs n'ont quant à eux pas tenté de diminuer l'effort de compréhension des guides. Les manipulateurs n'ont donc pas participé autant que les guides au principe du moindre effort collaboratif au niveau de la communication des éléments spatialisés. Cependant, ce résultat est à mettre en relation avec l'analyse globale de la charge mentale. Les résultats sur la charge mentale ont mis en évidence que ce sont les manipulateurs qui ont assumé la charge mentale la plus importante pour la réalisation de la tâche. Ce résultat se comprend aisément puisque ce sont les manipulateurs qui ont dû prendre en charge tous les aspects techniques de la tâche : sélection des objets, changement de menu, déplacement de l'objet dans l'environnement, gestion des collisions, notamment. Ainsi, afin d'optimiser les performances

techniques, ils ont minimisé leurs efforts dans la communication spatiale pour allouer plus de ressources aux tâches dont ils avaient la responsabilité. Du point de vue des guides, les analyses de leur charge mentale ont montré une exigence mentale plus importante que toutes les autres composantes. L'exigence mentale correspondant « aux opérations mentales et perceptives nécessaires à la réalisation de la tâche » peut être associée aux efforts liés aux prises de perspective. Il est donc possible que l'importance de l'exigence mentale pour les guides soit liée aux choix qu'ils ont opérés pour partager les informations spatiales avec les manipulateurs. Du point de vue du moindre effort collaboratif, l'asymétrie dans les cadres de référence et les transformations mentales peut donc refléter la prise en charge par les guides d'une partie des sources d'efforts nécessaires à la réalisation de la tâche. Comme ils ne pouvaient pas faire l'action à la place du manipulateur, les guides ont minimisé l'effort là où ils le pouvaient, à savoir la communication verbale. Les guides ont donc cherché à faciliter la compréhension des énoncés qu'ils produisaient. De plus, les guides ont soulagé les manipulateurs d'une partie de la charge mentale globale tout en minimisant l'effort nécessaire à cette prise en charge. De leur côté, les manipulateurs ont orientés leurs ressources vers l'exécution de la tâche et non vers la minimisation des efforts de compréhension des guides.

4.4 Référentiel commun

Tous ces résultats indiquent que les collaborateurs ont géré collectivement les contributions de chacun afin de minimiser l'effort du manipulateur. Cette gestion collective est permise par la prise en compte du manipulateur dans la représentation occurrente de la situation sur laquelle s'appuie le guide pour collaborer. C'est le référentiel commun. En effet, si les guides avaient voulu minimiser uniquement leurs propres efforts, ils auraient utilisé des énoncés simples à produire, ce qui n'est pas le cas. Le référentiel commun sur lequel se sont appuyés les guides contient donc des informations sur la charge mentale individuelle des manipulateurs, leurs besoins et leurs contraintes. De plus, le fait que les choix de cadres de références se soient renforcés au cours du temps montre que les représentations sur les besoins ont été confirmées, et que les efforts mis en œuvre pour prendre la perspective du manipulateur se sont avérés pertinents et efficaces. Cette étude a donc permis de mettre en évidence que le référentiel spatial commun évolue en fonction des contraintes de la tâche. Cette prise en compte a également été observée par Spante et collaborateurs (2004), qui ont observé des changements de répartition des rôles après que les participants aient échangé de dispositif technique. Par leur protocole, Spante et ses collaborateurs (2004) ont rendu explicites les contraintes, alors que dans notre étude, la prise en compte des contraintes du collaborateur s'est faite spontanément.

Pour conclure sur cette première étude, nous avons pu confirmer que les guides prenaient plus souvent la perspective de la personne guidée pour échanger des informations de nature spatiale. Nous avons également pointé du doigt que les liens entre cadre de référence, transformation mentale et charge mentale ne sont pas directs. Cette étude n'apporte pas d'information sur ces liens, si ce n'est que le choix d'un cadre de référence centré sur une autre référence que soi-même n'induit pas nécessairement de transformation mentale. Ainsi, afin de mieux comprendre les coûts associés aux changements de perspective, nous avons mis en place deux études complémentaires sur la production d'énoncés spatialisés. Nous avons complété ces approches avec une étude dédiée aux coûts de compréhension des énoncés. Ces trois études sont présentées dans le chapitre suivant.

Chapitre 4 Etudes sur la production et la compréhension d'énoncés spatialisés

Dans cette seconde partie expérimentale, nous avons cherché à mieux comprendre les facteurs influençant la communication spatiale, notamment du point de vue de la charge mentale requise à la fois pour la production et la compréhension des énoncés spatialisés. Comme nous l'avons souligné dans la partie bibliographique et l'étude présentée ci-avant, la gestion des cadres de référence peut engendrer des processus mentaux plus ou moins coûteux, notamment ceux liés à la prise de perspective. Bien entendu le dialogue est un processus plus complexe, dont d'autres composantes, telles que le grounding, devraient être étudiées (Bezuidenhout, 2013; Clark & Brennan, 1991). Cependant, les phases de production et de compréhension sont déterminantes. En effet, il y aura toujours production et par conséquent compréhension des énoncés. Il est d'ailleurs possible que le principe du moindre effort collaboratif vise à minimiser ces coûts de production et de compréhension. Dans ce contexte, nous avons proposé trois études complémentaires. Les deux premières études s'intéressent à la production d'énoncés spatialisés. Dans un premier temps, l'objectif est d'identifier les configurations les plus difficiles pour la production d'énoncés spatialisés. Par configurations nous entendons l'écart d'orientation des plans sagittaux entre un locuteur et un destinataire, c'est-à-dire l'orientation du destinataire, ainsi que la position relative d'un objet à localiser (cible) par rapport au locuteur et au destinataire. Nous avons traité cette question dans la première étude, intitulée **1. Production d'énoncés sans indice**. Dans un second temps, l'objectif est de tester différentes aides visuelles qui pourraient faciliter la production des énoncés spatialisés (**2. Production d'énoncés avec indices**). Ces deux études sont l'occasion d'étudier les cadres de référence utilisés par les locuteurs pour produire leurs énoncés. La troisième étude s'intéresse à la compréhension de ces énoncés. L'objectif est alors de comparer la difficulté de compréhension des énoncés les plus utilisés lors de la production d'énoncés spatialisés (**3. Compréhension d'énoncés**).

1. Etude 1 : Production d'énoncés sans indice

1.1 Contexte

Comme nous venons de l'évoquer, l'objectif de cette étude est de mieux comprendre les processus à l'œuvre lors de la production d'énoncés spatialisés. Cette phase du dialogue comprend plusieurs étapes (Burigo & Sacchi, 2013; Daniel & Denis, 1998). En effet, les locuteurs doivent activer une représentation interne, basée sur les éléments pertinents de la scène et du but à atteindre. Ils choisissent la perspective dans laquelle ils vont produire leur énoncé, préparent et expriment leur énoncé. Burigo et Sacchi (2013), ont d'ailleurs mis en évidence que ces étapes ne sont pas successives, mais peuvent se chevaucher. Chacune de ces étapes peut être influencée par la situation et avoir des coûts plus ou moins élevés.

En ce qui concerne la construction de la représentation, elle doit s'appuyer sur la perception de l'environnement. On peut donc penser que la dimension égocentrée a une importance, au moins dans cette première étape du processus de production des énoncés spatialisés. Par ailleurs comme évoqué dans la partie théorique, la représentation de la localisation des objets se construit selon un cadre de référence mental (*mental spatial framework*, (Tversky, 2005a, 2005b). Selon cette théorie, l'axe tête/pieds serait prédominant par rapport aux axes avant/arrière et droite/gauche, et l'axe avant/arrière serait prédominant sur l'axe droite/gauche. Par ailleurs, la représentation d'une scène peut se faire en utilisant un autre système de coordonnées en privilégiant un autre référentiel, celui du destinataire par exemple. Tout comme pour le référentiel égocentré, la position de la cible par rapport aux directions intrinsèques de la référence influence la représentation et les performances associées (Bryant & Tversky, 1999; Michelon & Zacks, 2006; Tamborello et al., 2012). De plus, certains auteurs ont mis en évidence une influence de l'alignement des interlocuteurs sur la manière dont les locuteurs se représentent l'espace et sur les cadres de référence utilisés dans le dialogue spatial (Galati & Avraamides, 2014). Enfin, on peut imaginer que selon le cadre de référence qui sera choisi pour produire l'énoncé, la représentation égocentrée ou centrée sur le destinataire sera privilégiée, l'autre étant probablement inhibée.

En ce qui concerne le choix du cadre de référence, comme nous l'avons vu dans les chapitres précédents, la prise de perspective a été beaucoup étudiée (Brown-Schmidt & Hanna, 2011; Galati & Avraamides, 2014; Krauss & Fussell, 1991; Michelon & Zacks, 2006; Schober, 1993), mais aucune de ces études n'a mesuré directement la charge cognitive subjective liée à ce processus. La plupart des études comptent le nombre d'énoncés dans tel ou tel cadre de référence, avec l'hypothèse que les interlocuteurs tentent de minimiser leur charge mentale et la charge mentale collective, en choisissant les énoncés les plus simples et les plus efficaces (Bezuidenhout, 2013; Brown-Schmidt & Hanna, 2011; Galati & Avraamides, 2014; Roger et al., 2013; Schober, 1993, 1995). D'autres études s'intéressent aux temps de réaction, avec l'hypothèse que plus les réponses sont rapides, moins les processus mis en œuvre sont coûteux (Burigo & Sacchi, 2013; Schober, 1996). Bien que ces deux approches soient cohérentes et intuitives, l'étude de la charge mentale ne peut se réduire à la mesure de temps de réaction ou à la catégorisation des énoncés produits. Par ailleurs, comme expliqué dans la partie théorique, il existe plusieurs modèles explicatifs sur la manière dont la prise de perspective fonctionne et ses conséquences sur la charge mentale (Schober, 1995). On distingue alors les modèles des rotations mentales, de l'inversion des mots dans la configuration face à face et le modèle égocentrique/non égocentrique.

Enfin, en ce qui concerne la formulation des énoncés, cette étape doit être influencée par le cadre de référence utilisé, en ce sens que les mots choisis dépendent de la référence et des caractéristiques utilisées pour localiser la cible.

Dans ce contexte, nous avons choisi d'étudier les effets de l'orientation du destinataire et de la position de la cible sur l'exigence mentale, le temps de préparation des énoncés, le temps de production des énoncés et le nombre de mots par énoncés. Pour cela nous avons développé un environnement virtuel, présenté Figure 34, dans lequel la position et l'orientation du locuteur sont fixes. L'avatar du destinataire est situé au milieu de la pièce, entouré par huit cibles potentielles organisées selon des écarts de 45 degrés. Ainsi l'avatar peut faire face aux huit tables (8 orientations de l'avatar) et la cible peut être chacune de ces huit tables (8 positions de cible).



Figure 34 : Vue du participant sur l'environnement virtuel. L'avatar situé au milieu des tables représente le collaborateur du participant. Il est entouré de huit tables positionnées tous les 45°. La bouteille à décrire est celle posée sur la table verte.

D'après le contexte théorique, nos hypothèses sont les suivantes :

- Si les participants suivent le principe du moindre effort collaboratif, ils devraient utiliser l'avatar comme référence pour donner leurs instructions.
- Si les participants utilisent bien l'avatar comme référence pour produire leurs énoncés et si la localisation relative de la cible suit la théorie du cadre de référence, les interactions devraient montrer la prédominance du plan sagittal de l'avatar sur le plan frontal de l'avatar.
- Si la prise de perspective du destinataire suit le modèle des rotations mentales, alors les mesures de l'exigence mentale et les temps de préparation des énoncés doivent augmenter de manière linéaire avec le degré de rotation de l'avatar.
- Les cibles étant réparties de manière différente dans le champ de vision du participant, la détection des cibles devrait avoir une influence sur les temps de préparation des énoncés dans un référentiel égocentré.

1.2 Matériels et méthodes

1.2.1 Participants

Vingt-deux personnes ont participé à cette étude : cinq femmes et dix-sept hommes. Ils étaient âgés de 20 à 42 ans, et la moyenne d'âge était de 25,5 ans.

1.2.2 Matériels

Cette étude, ainsi que les deux suivantes, ont été réalisées au sein de la plateforme de Réalité Virtuelle VIVRE.lab de l'Ecole Centrale de Nantes. L'environnement virtuel, décrit ci-avant, a été développé spécialement pour cette étude par un ingénieur de l'équipe PsyCoTec.

Les participants étaient immergés dans l'environnement grâce à un mur immersif à double rétroprojection de 1.8m par 2.4m. Pour cela, ils étaient équipés de lunettes stéréoscopiques sur lesquelles étaient placées des marqueurs, permettant d'être repérés grâce à un système de huit caméras infra rouge, pour assurer la parallaxe et le positionnement de la tête dans l'environnement.

Les participants étaient équipés de casque avec micro intégré, et avaient accès à un clavier disposé sur une table devant eux.

Afin d'évaluer les aptitudes spatiales des participants, ils ont réalisé le Santa Barbara Sense of Direction Scale et le test des rotations mentales MRT-A © utilisés lors dans la première partie expérimentale. Ces tests sont présentés dans les Annexe 6 et Annexe 7. Les participants ont également répondu à un questionnaire, présenté en Annexe 13.

1.2.3 Procédure

Le déroulement de l'étude était présenté aux participants à leur arrivée. Ensuite ils devaient prendre connaissance des consignes et du scénario (Annexe 9). Les participants-locuteurs devaient donner des instructions à un collaborateur virtuel pour que ce collaborateur puisse sélectionner un objet-cible parmi huit cibles potentielles. Le collaborateur-destinataire de l'instruction était représenté par un avatar. Seul le participant-locuteur recevait l'information sur l'objet-cible à sélectionner. Seul son collaborateur-destinataire avait la possibilité de sélectionner l'objet-cible. Les instructions devaient être données de manière verbale, sans qu'aucun échange ne soit possible : il fallait donc que la consigne soit la plus précise possible. Par ailleurs, une contrainte de temps était imposée : il fallait que les participants-locuteurs aillent le plus vite possible une fois que l'objet-cible était identifié.

Avant de commencer les tests, les participants étaient invités à naviguer dans l'environnement, dans lequel l'avatar de leur collaborateur n'était pas présent. L'environnement était une simple pièce, avec huit tables, sur chacune des tables était positionnée une bouteille, comme illustré Figure 34 et Figure 35. Les tables et bouteilles ne pouvaient être différenciées autrement que par leur localisation (même taille, même couleur, même forme).

Les participants disposaient de sept essais d'entrainement. A chaque essai, l'orientation et la position de la cible pouvaient changer. Le plan expérimental comprenait deux facteurs : orientation de l'avatar {huit modalités} * position de la cible {huit modalités}, ce qui donnait 64 configurations. Chacune de ces configurations était répétée trois fois. A la dernière répétition les participants devaient évaluer la charge mentale en répondant à la question suivante : « Comment noteriez-vous l'exigence mentale pour cet exercice ? ». Pour répondre, les participants devaient déplacer un curseur le long d'un segment borné entre 0 et 100, comme illustré sur le bas de la Figure 36. La distance du curseur à la borne 0 était enregistrée. Afin d'expliquer la question aux participants, la dernière répétition était précédée d'une explication (Figure 36, haut).

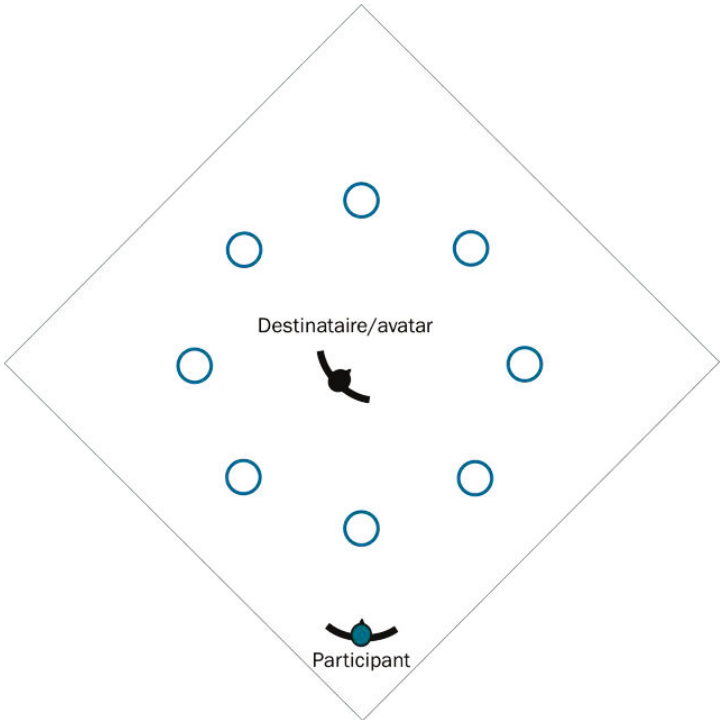


Figure 35 : Schéma de l'agencement de l'environnement virtuel, vue du dessus.

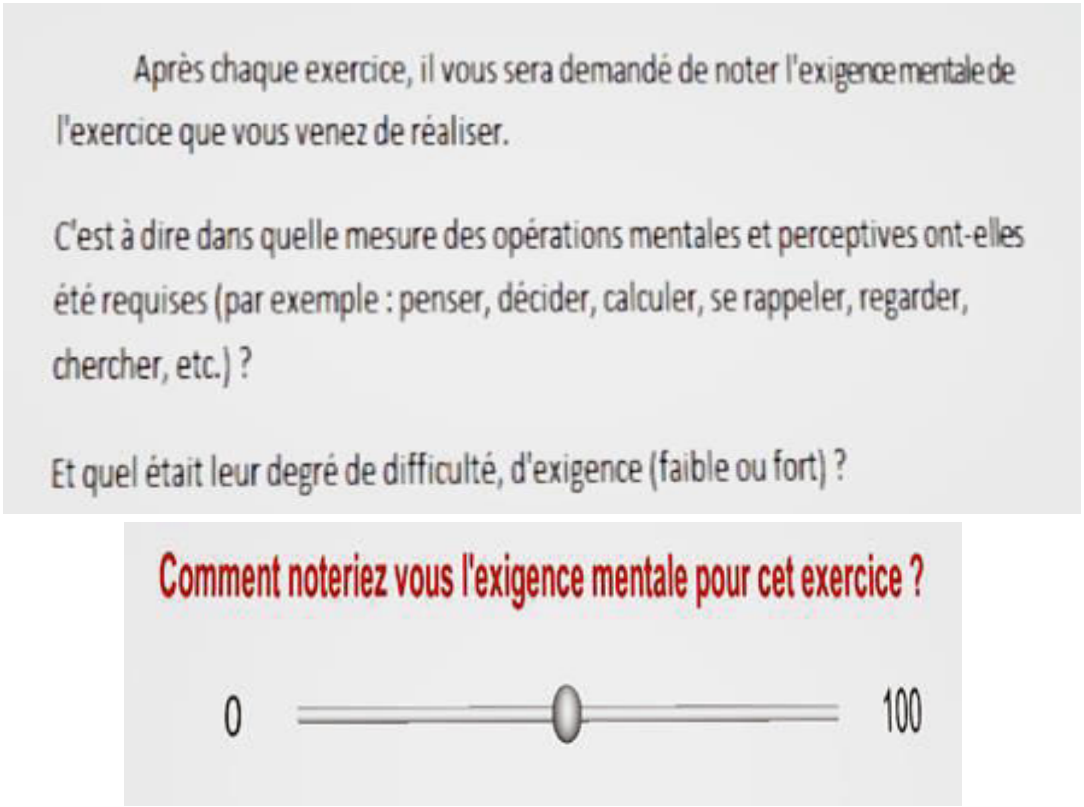


Figure 36 : Copie d'écran de la question permettant de mesurer la charge mentale après chaque essai.

Un essai se déroulait de la manière suivante :
Les participants devaient appuyer sur la touche 0 du pavé numérique pour faire apparaître la scène (avec les tables et l'avatar). Ils pouvaient prendre autant de temps qu'ils le souhaitaient pour en prendre connaissance. Ensuite, ils devaient appuyer de nouveau sur la touche 0 pour faire apparaître la bouteille

cible (la table changeait alors de couleur). Cette action déclenchait un premier chronomètre, mesurant le temps de préparation de l'énoncé, c'est-à-dire le temps entre l'apparition de la cible et le début de la production de l'énoncé. Une fois la cible désignée, les participants devaient donner leur instruction le plus vite et de la manière la plus précise possible. Les participants devaient appuyer sur la barre espace, et la maintenir enfoncée tant qu'ils parlaient. Il était bien précisé aux participants qu'ils ne devaient pas appuyer sur la barre espace dès que la cible apparaissait, mais bien quand ils commençaient à parler. Le début de la pression sur la barre espace stoppait le premier chronomètre et déclenchait le second, mesurant le temps de production de l'énoncé. Une fois que le participant relâchait la barre espace le second chronomètre s'arrêtait et la scène disparaissait. Ils pouvaient ainsi lancer un nouvel essai.
Pendant l'expérience, si le participant se trompait dans les touches, l'essai était annulé par l'expérimentateur et reporté à la fin de la série.

1.2.4 Codage des énoncés spatialisés

Chaque instruction donnée par les participants a été codée selon le ou les cadres de référence utilisés pour localiser la cible. Pour cela nous avions défini six catégories :

- **centrés sur le destinataire intrinsèque** : énoncés utilisant la perspective du destinataire, ainsi que ses propriétés intrinsèques – « la bouteille piégée est sur la table devant toi sur la droite », « la bouteille piégée est sur la table sur ta gauche derrière », « devant toi » ;
- **centrés sur le destinataire extrinsèque** : énoncés utilisant la perspective du destinataire, ainsi que des caractéristiques extrinsèques à l'avatar (directions cardinales, heures) – « la bouteille piégée est sur la table à trois heures », « au nord-est » (comme nous le verrons plus tard, les participants utilisant ce genre d'instructions explicitaient le fait que la table faisant face à leur collaborateur (avatar) était celle à midi, ou au nord) ;
- **centrés sur le destinataire combiné** : énoncés utilisant la perspective du destinataire, ainsi qu'une combinaison des caractéristiques intrinsèques et extrinsèques (le plus souvent degrés et droite/gauche) – « la bouteille piégée est sur la table à quarante-cinq degrés à ta droite » ;
- **égocentrés** : énoncés utilisant la perspective du participant – « la bouteille piégée est sur la table directement devant moi » ;
- **multiples** : énoncés utilisant plusieurs cadres de référence, c'est-à-dire donnant une information redondante – « la bouteille piégée est sur la table directement devant moi, c'est-à-dire directement à ta droite » ;
- **autres** : énoncés utilisant un autre moyen de référencement, par exemple le fait de faire référence à la situation précédente – « même table ».

1.2.5 Traitements statistiques

Le nombre d'énoncés par cadre de référence, a été comparé grâce à des tests du Khi2. Pour les temps de préparation et de production, les valeurs des trois essais ont été moyennées. Cette répétition n'entre donc pas dans le plan d'analyse. Pour ces deux variables et l'exigence mentale, les données ont été comparées par des analyses de variance à mesures répétées (ANOVA), dont les facteurs seront précisés lors de la présentation des résultats. Lorsque l'analyse de variance se révélait significative, des analyses post-hoc (Tukey HSD) ont été réalisées. Pour tous les tests, le niveau de significativité utilisé était $p<.05$.

Pour identifier les situations les plus complexes, nous avons dans un premier temps testé les 64 configurations : 8 orientations de l'avatar * 8 positions de la cible par rapport au participant. Puis, plusieurs raisons nous ont incitées à réorganiser les données : le caractère symétrique de la situation par rapport au participant, l'usage massif de l'avatar comme référence et une interaction positive entre la

position de la cible et l'orientation de l'avatar. Nous avons ainsi regroupé les huit positions de la cible par rapport au participant en cinq positions (de A à E), comme illustré sur le schéma gauche de la Figure 37. Dans cette même logique, nous avons regroupé les huit orientations de l'avatar en cinq orientations (0°, 45°, 90°, 135° et 180°), comme illustré sur le schéma du milieu de la Figure 37. Pour étudier l'effet de la position de la cible par rapport à l'avatar, nous avons regroupé les données selon que la cible est alignée avec le plan sagittal ou frontal de l'avatar, ou qu'elle se situe en diagonale devant ou derrière lui. Ce regroupement est illustré sur le schéma à droite de la Figure 37.

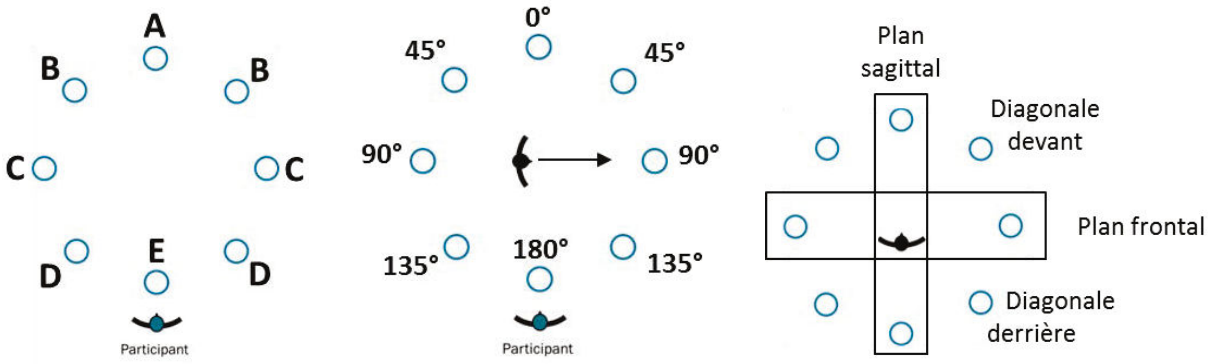


Figure 37 : Principes organisant le regroupement des données. Schéma de gauche : selon la position de la cible par rapport au participant, schéma du milieu : selon l'orientation de l'avatar par rapport au participant, schéma de droite : selon l'alignement de la cible avec les plans intrinsèques de l'avatar.

1.3 Résultats - Production d'énoncés sans indice

Avant de présenter les résultats sur les différentes variables mesurées au cours de cette étude, nous tenons à rapporter des observations sur le comportement des participants. Dans certains cas, les participants utilisaient les essais d'entrainement pour parfaire leur stratégie verbale et optimiser leurs énoncés. Notamment, lorsqu'ils choisissaient de donner leurs consignes avec un système extrinsèque, telles que les heures ou les directions cardinales, ils informaient leur collaborateur de cette stratégie. Cela leur permettait ensuite de réduire considérablement la longueur de leurs instructions.

1.3.1 Cadres de référence

Les participants ont utilisé différents cadres de référence pour donner leurs instructions. La répartition générale est présentée Figure 38. Ils ont utilisé en grande majorité (80%) des énoncés utilisant la perspective du destinataire, 7 % ont été produits dans le cadre de référence égocentré, c'est-à-dire centré sur le locuteur, dans 11% des énoncés l'information est formulée dans plusieurs cadres de référence (Multiple), et 2% sont codés dans un autre système (référence à l'essai précédent par exemple).

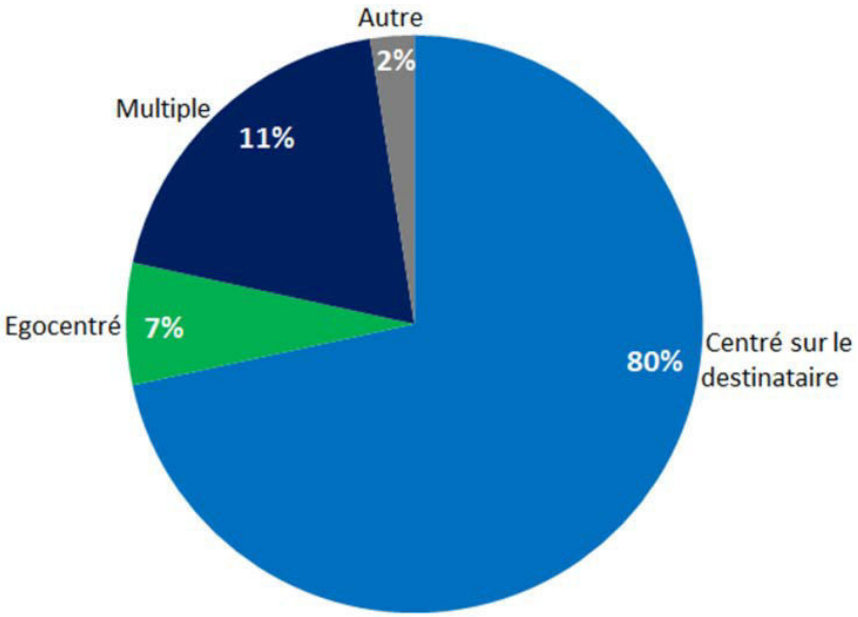


Figure 38 : Cadres de référence utilisés par les participants pour donner leurs instructions.

Les analyses sur l'ensemble des énoncés n'ont mis en évidence ni d'effet de l'orientation de l'avatar par rapport au participant (χ^2 (35, 4224) = 5, p=1), ni d'effet de la cible par rapport au participant (χ^2 (35, 4224) = 42.1, p=.19). Les seules analyses ayant montré des effets sont celles regardant la position de la cible par rapport à l'avatar (χ^2 (35, 4224) = 854, p<.001). De par la nature symétrique des situations et des résultats, nous avons regroupé les données selon que la cible se situe dans le plan frontal ou sagittal de l'avatar, ou en diagonale devant ou derrière lui comme illustré sur la partie droite de la Figure 37. Les analyses selon ce regroupement montrent que les participants ont produits des énoncés différents selon la position de la cible par rapport aux axes intrinsèques de l'avatar (χ^2 (15, 4224) = 850.75, p<.001). Ces résultats sont présentés Figure 39. Nous soulignons ici plusieurs résultats intéressants (en rose dans la Table 4). D'une part, les participants n'ont pas utilisé de cadre de référence centré sur le destinataire **combinés** lorsque la cible est alignée avec le plan sagittal de l'avatar, contrairement aux autres configurations (24% en diagonale, 16% dans le plan frontal et 23% en diagonale derrière). D'autre part, les participants ont produit significativement plus d'énoncés centrés sur le destinataire **intrinsèques** lorsque

la cible est alignée avec le plan sagittal de l'avatar (66%) que dans toutes les autres configurations, notamment lorsque la cible est alignée avec le plan frontal de l'avatar (41%). Par ailleurs, les participants ont utilisé plus d'énoncés centrés sur le destinataire **extrinsèques** dans les conditions en diagonale (devant, 22%, et derrière, 21%) que lorsque la cible est alignée avec l'un des plans intrinsèques de l'avatar (sagittal, 16%, et frontal, 17%). Enfin, les participants ont plus doublé leurs énoncés (**Multiple**) lorsque la cible est en diagonale derrière l'avatar (28%) que dans toutes les autres configurations, notamment lorsque la cible est en diagonale devant l'avatar (17%).

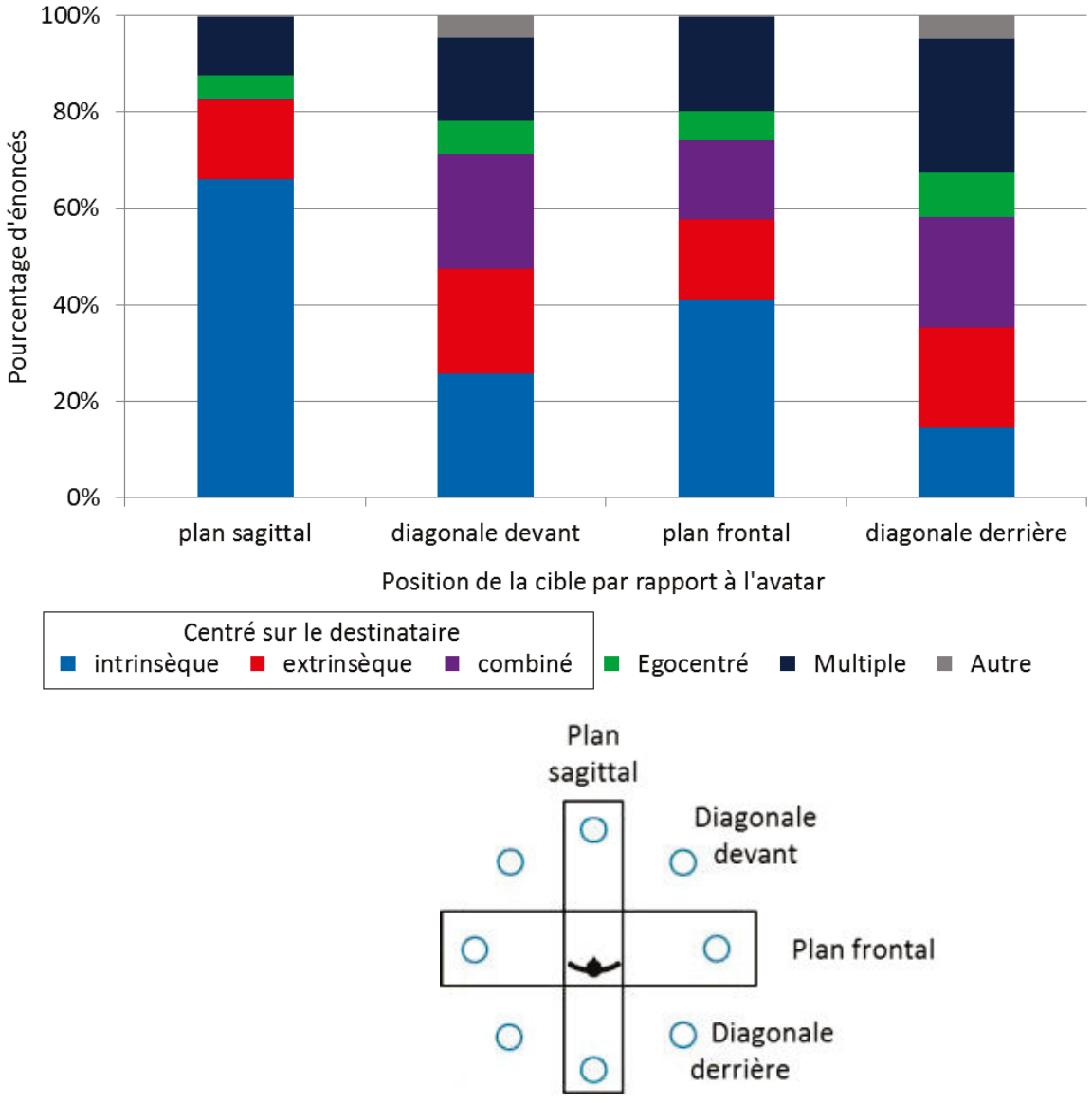


Figure 39 : Pourcentage d'énoncés produits dans chaque cadre de référence selon la position de cible par rapport à l'avatar.

Table 4 : Résultats des tests du Khi2 pour les comparaisons du nombre d'énoncés par cadre de référence. * p<.05, **p<.001.

Centré sur le destinataire intrinsèque

	diagonale devant	plan frontal	diagonale derrière
plan sagittal	344.4**	132.6**	585**
diagonale devant		55.2**	42.5**
plan frontal			186.7**

Centré sur le destinataire extrinsèque

	diagonale devant	plan frontal	diagonale derrière
plan sagittal	9.3**	.05	7.2**
diagonale devant		7.9**	.1
plan frontal			6*

Centré sur le destinataire combiné

	diagonale devant	plan frontal	diagonale derrière
plan sagittal	282.9**	185.3**	268.8**
diagonale devant		18.4**	.3
plan frontal			13.9**

Egocentré

	diagonale devant	plan frontal	diagonale derrière
plan sagittal	2.7	.9	13.7**
diagonale devant		.5	4.3*
plan frontal			7.7**

Multiple

	diagonale devant	plan frontal	diagonale derrière
plan sagittal	12.3**	22.9**	83**
diagonale devant		1.7	32.8**
plan frontal			19.9**

Synthèse des résultats sur les cadres de référence utilisés pour donner les instructions

Les participants ont principalement donné leurs instructions en utilisant l'avatar comme référence.

Les caractéristiques intrinsèques de l'avatar ont été plus souvent utilisées lorsque la cible était alignée avec l'un des plans intrinsèques de l'avatar, et de manière plus marquée dans le plan sagittal que dans le plan frontal.

Les participants ont plus souvent utilisé plusieurs cadres de référence dans un même énoncé pour localiser la cible lorsque celle-ci était en diagonale derrière l'avatar, notamment par rapport à la configuration diagonale devant.

Ainsi, les quatre positions possibles de la cible par rapport à l'avatar ont suscité des comportements verbaux singuliers.

1.3.2 Exigence mentale

La première analyse de variance (ANOVA) à deux facteurs (8 orientations de l'avatar * 8 positions de la cible) sur les mesures de l'exigence mentale a mis en évidence un effet de la position de la cible codée par rapport au participant ($F(7, 147)=3.15$, $p<.01$), un effet de l'orientation de l'avatar ($F(7, 147)=3.89$, $p<.01$) et une interaction significative entre les deux variables ($F(49, 1029)=12.03$, $p<.001$). De nouvelles analyses ont été réalisées à partir de plusieurs regroupements, comme évoqué dans la section **Matériels et méthodes**, p. 74. Les premiers regroupements s'appuient sur la dimension symétrique de la situation par rapport au participant, nous avons donc regroupé les cibles et les orientations de l'avatar selon les schémas présentés Figure 37. Le dernier regroupement, déjà utilisé pour l'analyse des cadres de référence, s'appuie sur la symétrie des caractéristiques intrinsèques de l'avatar.

Effets de la position de la cible codée par rapport au participant sur l'exigence mentale

L'ANOVA réalisée avec les données reconfigurées selon l'axe de symétrie du plan sagittal du participant confirme l'effet de la position de la cible ($F(4, 84)=3.77$, $p<.01$). Les analyses post-hoc ont mis en évidence que la cible la plus proche du participant (E) induit une exigence mentale significativement plus faible que les cibles C et D ($p<.05$), comme illustré Figure 40.

Effets de l'orientation de l'avatar sur l'exigence mentale

L'ANOVA réalisée sur les mesures de l'exigence mentale reconfigurées selon l'axe de symétrie du plan sagittal du participant confirme l'effet de l'orientation de l'avatar sur l'exigence mentale ($F(4, 84)=5.55$, $p<.001$). De plus, les analyses post-hoc montrent que l'exigence mentale est plus faible lorsque les interlocuteurs sont dans la même orientation (0°) par rapport aux situations où l'avatar est tourné de 90° ($p<.05$), 135° ($p<.01$) et de 180° ($p<.001$). L'exigence mentale est également plus faible lorsque les interlocuteurs sont décalés de 45° par rapport à la situation où ils se font face (180°, $p<.01$). Ces résultats sont représentés Figure 41. Par ailleurs, une corrélation montre une augmentation linéaire de l'exigence mentale avec la rotation de l'avatar par rapport au participant ($R^2 = .97$).

Effets de la position de la cible codée par rapport à l'avatar sur l'exigence mentale

L'ANOVA à deux facteurs réalisée après avoir regroupé les données selon l'alignement de la cible par rapport à l'avatar : 4 positions de la cible par rapport à l'avatar * 5 orientations de l'avatar a mis en évidence un effet de la position de la cible par rapport à l'avatar ($F(3, 63)=21.81$, $p<.001$), un effet de l'orientation de l'avatar ($F(4, 84)=5.55$, $p<.001$) et une interaction significative entre les deux variables ($F(12, 252)=3.41$, $p<.001$). L'effet de l'orientation de l'avatar reste inchangé par rapport à ce qui a été présenté ci-avant, nous n'y reviendrons donc pas.

Concernant l'effet de la cible par rapport à l'avatar, les analyses post-hoc ont mis en évidence que lorsque la cible est alignée avec le plan sagittal de l'avatar (9.5), c'est-à-dire lorsque la cible est située devant ou derrière lui, l'exigence mentale est significativement plus faible que lorsque la cible est dans n'importe quelle autre position relative (25.8 pour diagonale devant, $p<.001$, 20.3 pour plan frontal, $p<.01$ et 35.9 pour diagonale derrière, $p<.001$). De même, lorsque la cible est située en diagonale derrière l'avatar, l'exigence mentale est significativement plus importante que pour toutes les autres situations (diagonale devant, $p<.05$, plan frontal, $p<.001$). De plus, il est intéressant de noter que la différence d'exigence mentale moyenne entre les situations où la cible est en diagonale devant ou sur le plan frontal de l'avatar n'est pas significative ($p>.3$).

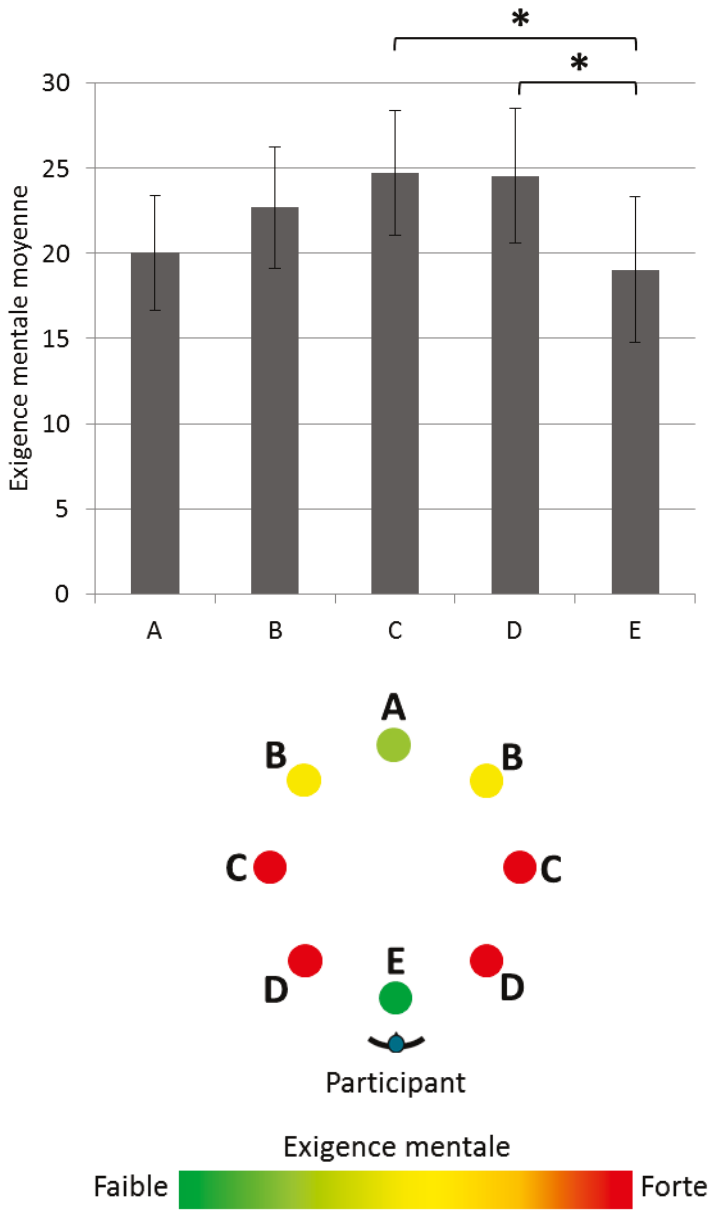


Figure 40 : Exigence mentale moyenne en fonction de la position de la cible, codée par rapport au participant. * p<.05

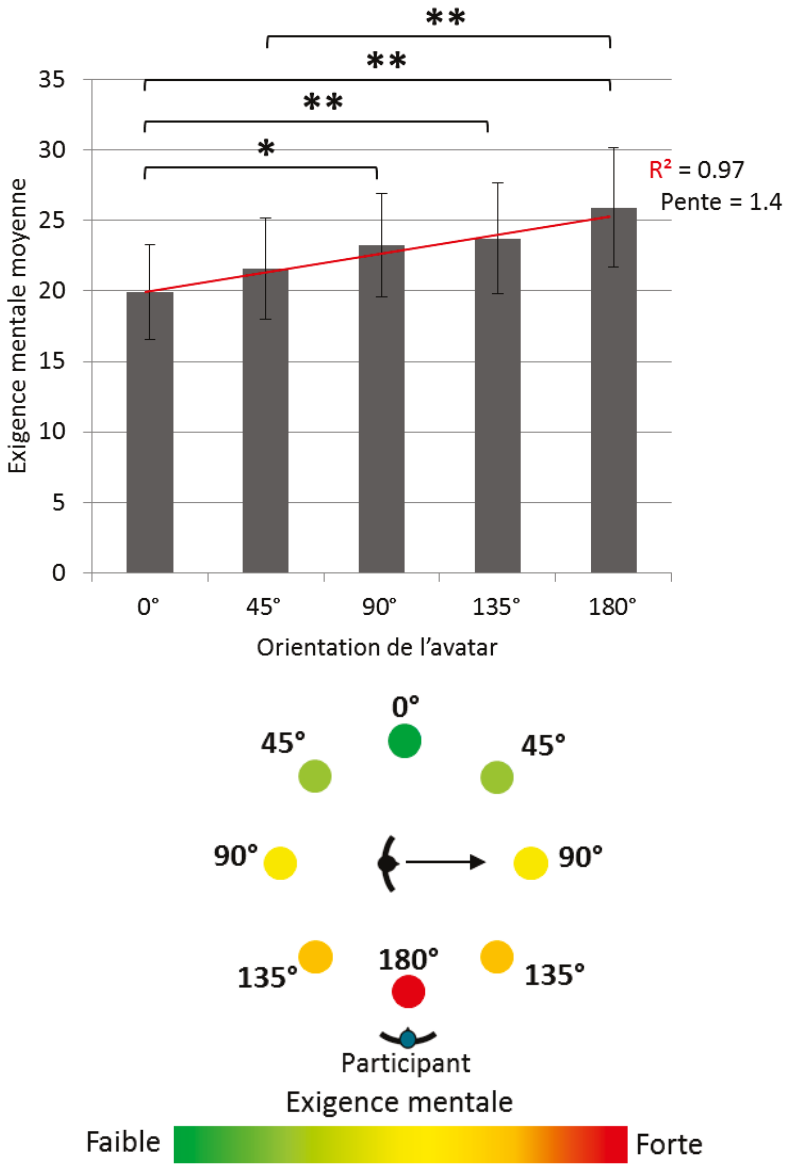


Figure 41 : Exigence mentale moyenne en fonction de l'orientation de l'avatar. * p<.05 ; ** p<.01

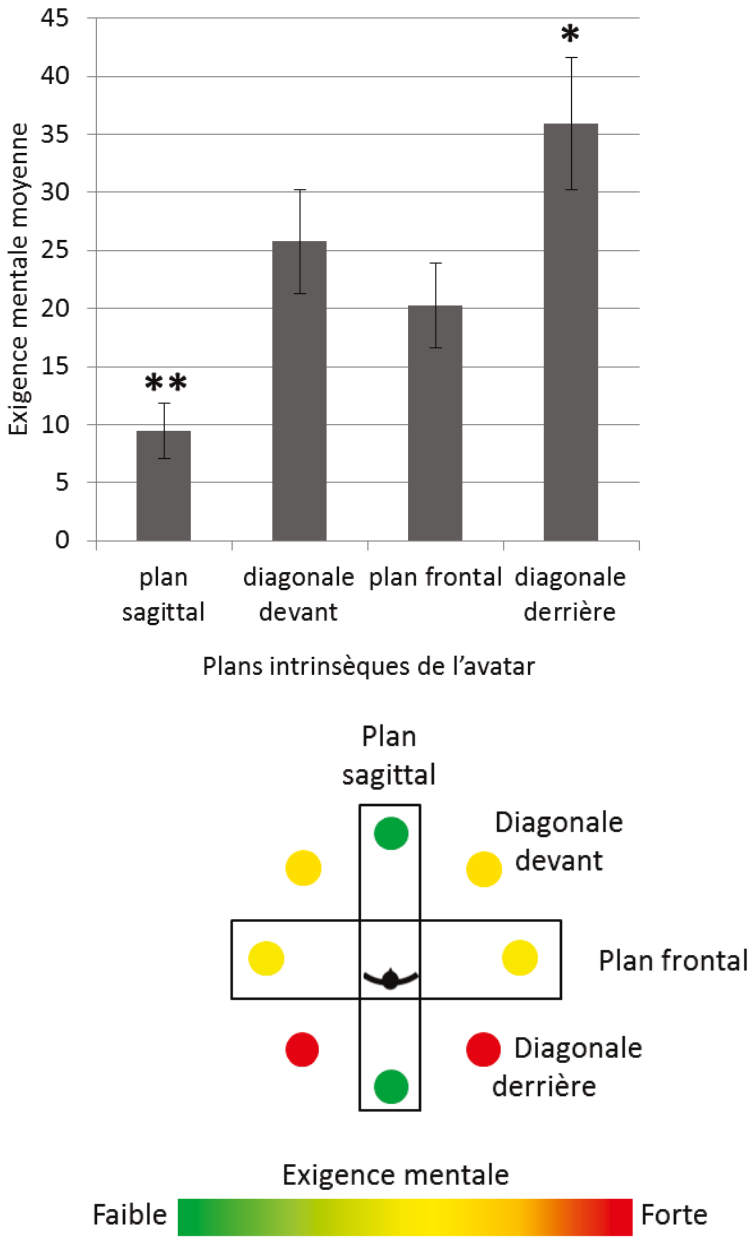


Figure 42 : Exigence mentale moyenne en fonction de la position de la cible, codée par rapport à l'avatar. * p<.05, ** p<.01

Concernant l'interaction entre la position de la cible codée par rapport à l'avatar et l'orientation de l'avatar, les analyses post-hoc ont permis de montrer que l'interaction est due à deux effets principaux. Le premier effet concerne les situations où l'avatar est perpendiculaire au participant (90°, Figure 43). En effet, lorsque l'avatar est dans cette orientation, il est intéressant de constater que l'exigence mentale n'est pas statistiquement différente que la cible soit alignée avec le plan sagittal ou avec le plan frontal de l'avatar (p=.99). Par ailleurs, l'exigence mentale est plus forte lorsque la cible est en diagonale devant que lorsqu'elle est alignée avec l'un des plans de l'avatar (sagittal p<.001, frontal p<.001), mais moins élevée que lorsque la cible se situe en diagonale derrière l'avatar (p<.001).

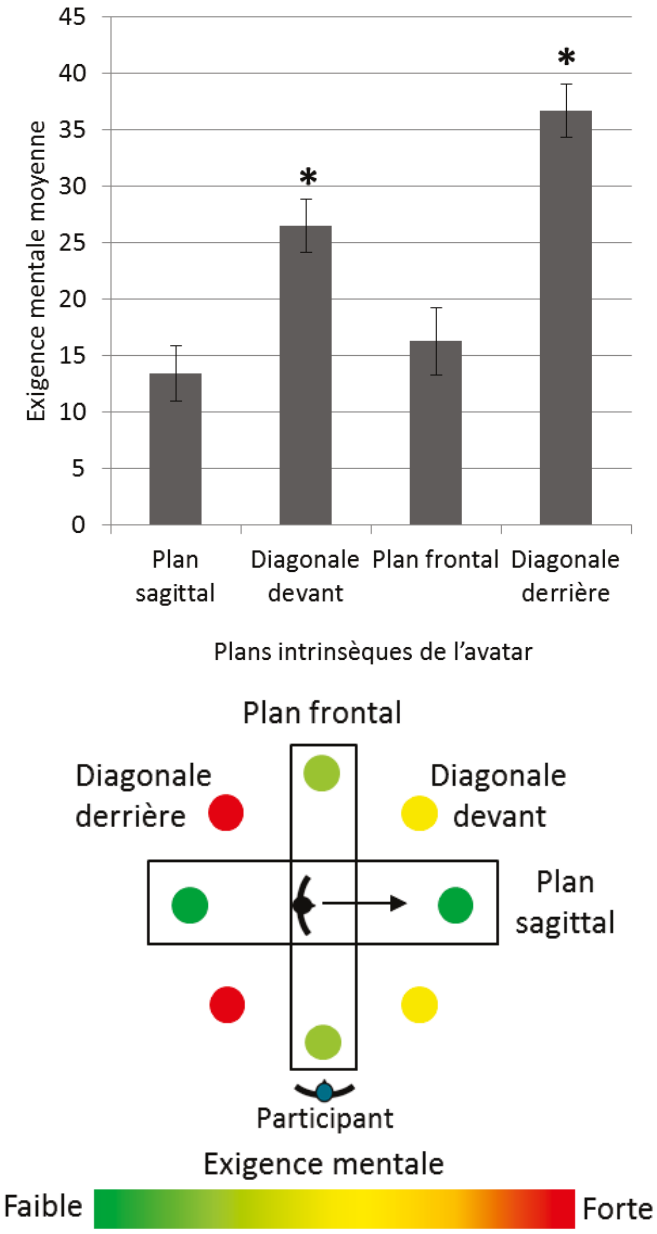


Figure 43 : Exigence mentale en fonction de la position de la cible par rapport aux plans intrinsèques de l'avatar lorsque l'avatar est tourné de 90°. * p<.05

Le second effet concerne les situations où la cible est alignée avec le plan sagittal de l'avatar. En effet, dans cette configuration, l'exigence mentale ne présente aucune différence selon l'orientation de l'avatar (pente = .2). Il n'y a d'ailleurs pas de corrélation entre l'exigence mentale et le degré de rotation de l'avatar (r=.01). L'ensemble des analyses post-hoc est présenté dans l'annexe 10.

Synthèse des résultats sur l'exigence mentale requise pour donner les instructions

L'exigence mentale est modulée par plusieurs facteurs qui interagissent :

- **La localisation de la cible par rapport au participant.** L'exigence mentale la plus faible est induite par la cible la plus proche du participant.
- **L'orientation de l'avatar.** Nous avons constaté que plus l'écart entre l'orientation du participant et l'orientation de l'avatar augmente, plus l'exigence mentale augmente aussi. Cependant cet effet n'est pas observé lorsque la cible est alignée avec le plan sagittal de l'avatar.
- **La position de la cible par rapport à l'avatar.** Les caractéristiques intrinsèques de l'avatar organisent l'espace en zones induisant une exigence mentale plus ou moins forte : l'exigence mentale la plus faible est induite par l'alignement de la cible avec le plan sagittal de l'avatar, et l'exigence mentale est la plus forte lorsque la cible est située en diagonale derrière l'avatar.

1.3.3 Temps de préparation des énoncés

Cette variable dépendante correspond au délai entre l'apparition de la cible et le début de l'énoncé. L'ANOVA à deux facteurs (8 orientations de l'avatar * 8 positions de la cible) a mis en évidence un effet de la position de la cible codée par rapport au participant ($F(7, 147)=6.07$, $p<.0001$), un effet de l'orientation de l'avatar ($F(7, 147)=3.49$, $p<.001$) et une interaction significative entre les deux variables ($F(49, 1029)=8.75$, $p<.0001$). Les mêmes regroupements de données ont été effectués que pour l'exigence mentale.

Effets de la position de la cible codée par rapport au participant sur le temps de préparation des énoncés

L'ANOVA réalisée avec les données reconfigurées confirme l'effet de la position de la cible par rapport au participant ($F(4, 84)=9.28$, $p<.0001$). Les analyses post-hoc ont mis en évidence les différences présentées Figure 44, dont l'effet principal peut se résumer de la manière suivante : plus la cible est éloignée du plan sagittal du participant, plus le temps de préparation augmente. En effet, lorsque la cible est située directement devant le participant (E), le temps de préparation est en moyenne de 1.1 secondes, ce temps est significativement plus court que lorsque la cible est située en position D (1.3s, $p<.05$), C (1.4, $p<.01$) et B (1.3s, $p<.0.5$). De plus, lorsque la cible est situé dans le plan sagittal du participant, mais au plus loin (A), le temps de préparation de l'énoncé est en moyenne de 1.2 secondes, ce qui est significativement plus court que lorsqu'elle est au plus loin du plan sagittal (position C : 1.4s, $p<.01$).

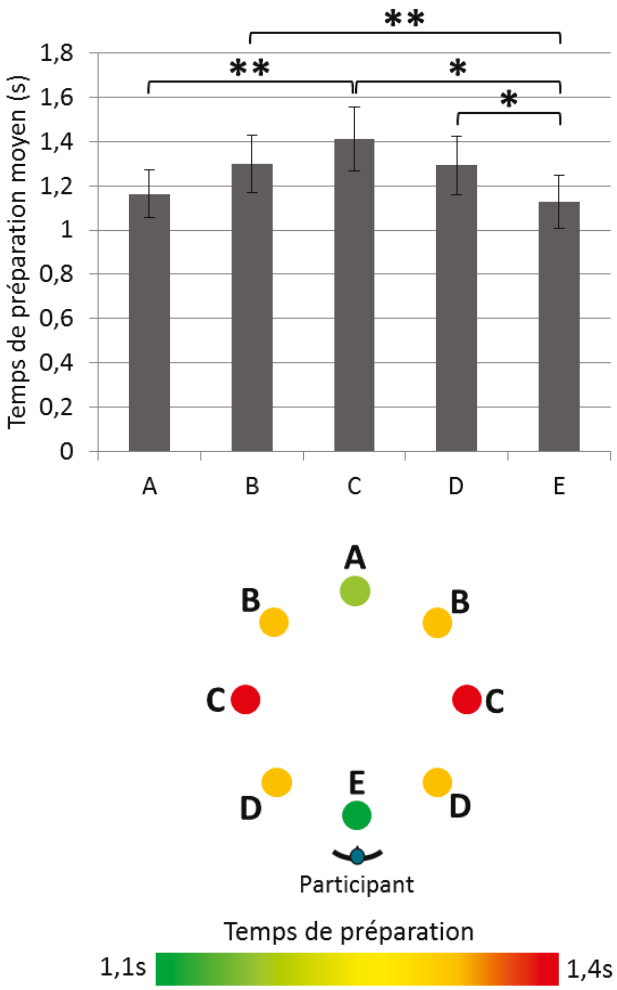


Figure 44 : Temps de préparation moyen en fonction de la position de la cible, codée par rapport au participant. * p<.05, ** p<.01

Effets de l'orientation de l'avatar sur le temps de préparation des énoncés

La seconde ANOVA réalisée avec les données reconfigurées confirme également l'effet de l'orientation de l'avatar sur le temps de préparation des énoncés ($F(4, 84)=5.81$, $p<.001$). Les analyses post-hoc ont mis en évidence les différences présentées Figure 45. Les temps de préparation des énoncés sont plus courts lorsque l'avatar est dans la même orientation que le participant (0°), ou légèrement incliné (45°), que lorsqu'ils se font face (180°) ou presque (135°). Les résultats sont résumés dans la Table 5.

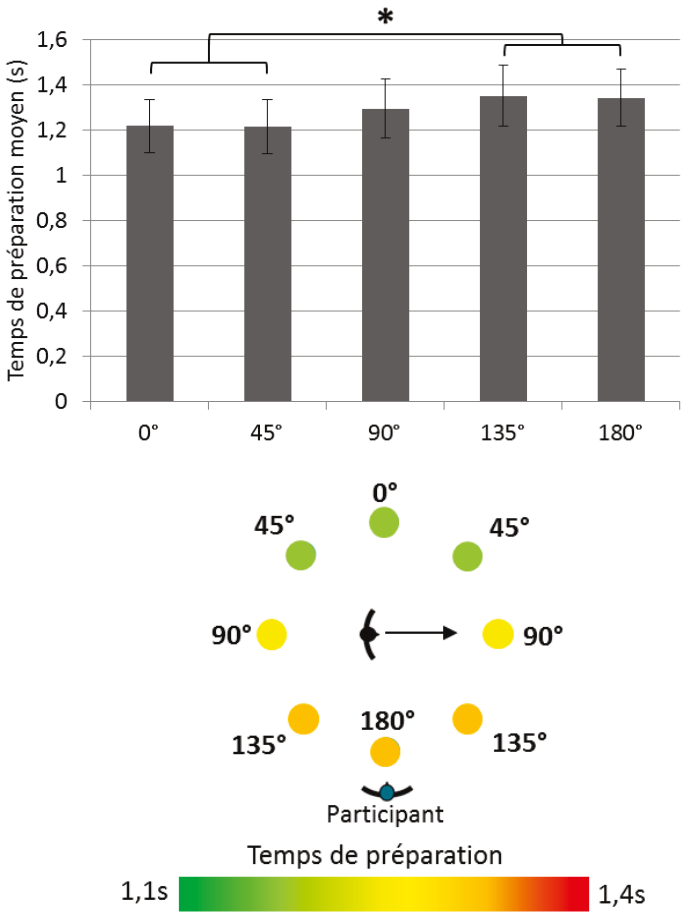


Figure 45 : Temps de préparation moyen en fonction de l'orientation de l'avatar. * p<.05

Table 5 : Résultats des analyses post-hoc (Tukey HSD) réalisées sur les temps de préparation des énoncés en fonction de l'orientation de l'avatar. Case rouge p<.05

	0°	45°	90°	135°	180°
0°		1,00	0,27	0,01	0,01
45°	1,00		0,24	0,01	0,01
90°	0,27	0,24		0,58	0,73
135°	0,01	0,01	0,58		1,00
180°	0,01	0,01	0,73	1,00	

Effets de la position de la cible codée par rapport à l'avatar sur le temps de préparation des énoncés

L'ANOVA réalisée après avoir regroupé les données selon l'alignement de la cible par rapport à l'avatar (4 positions de la cible par rapport à l'avatar * 5 orientations de l'avatar) a mis en évidence un effet de la position de la cible par rapport à l'avatar ($F(3, 63)=24.2$, $p<.0001$), un effet de l'orientation de l'avatar ($F(4, 84)=5.81$, $p<.001$) et une interaction significative entre les deux variables ($F(12, 252)=3.67$, $p=.0001$). L'effet de l'orientation de l'avatar reste inchangé par rapport à la première analyse, nous ne présenterons donc que l'effet de la position de la cible par rapport à l'avatar et son interaction avec l'orientation de l'avatar.

En ce qui concerne la position de la cible codée par rapport à l'avatar, les analyses post-hoc ont mis en évidence les mêmes effets d'organisation de l'espace par les caractéristiques intrinsèques de l'avatar que pour l'exigence mentale. On constate d'une part que les temps de préparation sont plus

courts lorsque la cible est alignée avec le plan sagittal de l'avatar (0.98s), que lorsque la cible est dans n'importe quel autre plan (1.3s pour diagonale devant, p<.01 ; 1.2s pour plan frontal, p<.05 et 1.7s pour diagonale derrière, p<.001). D'autre part, les temps de préparation sont plus longs lorsque la cible est située en diagonale derrière l'avatar que pour toutes les autres situations (diagonale devant, p<.001, plan frontal, p<.001). La différence entre les situations où la cible est en diagonale devant et sur le plan frontal de l'avatar n'est pas significative (p>.8).

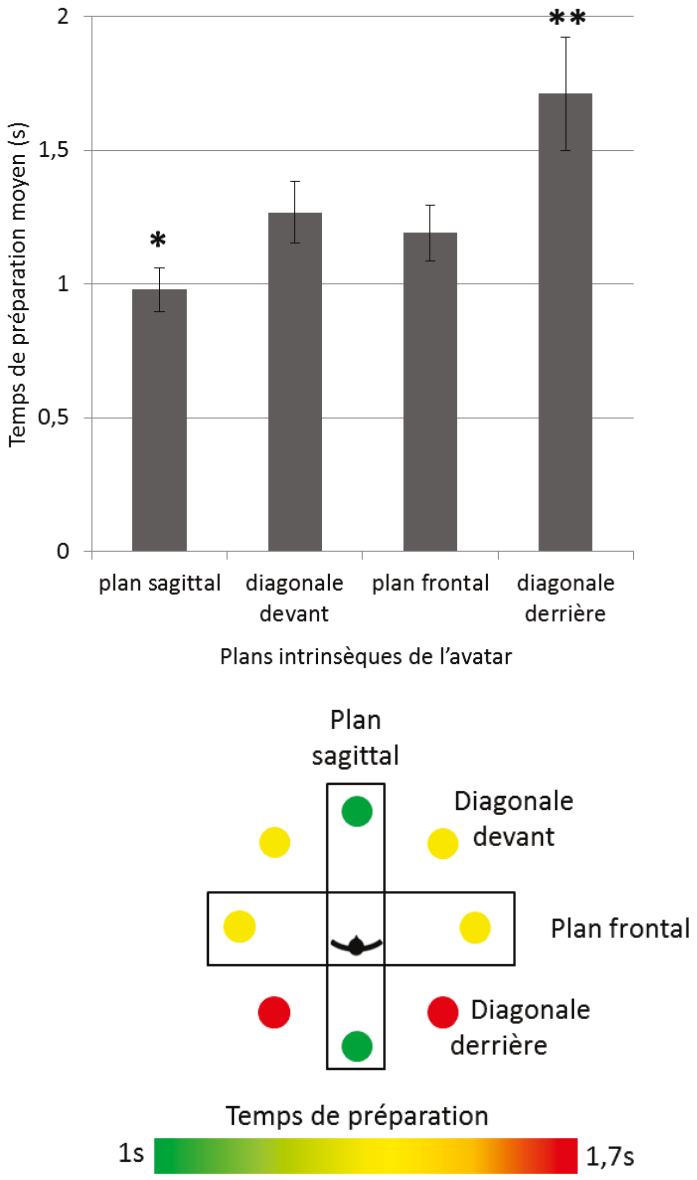


Figure 46 : Temps de préparation moyen en fonction de la position de la cible, codée par rapport à l'avatar. * p<.05, ** p<.01

En ce qui concerne l'interaction entre la position de la cible codée par rapport à l'avatar et l'orientation de l'avatar, les analyses post-hoc ont mis en évidence que cette interaction est due au fait que le pattern général présenté Figure 46 est vrai uniquement pour les orientations 0° et 180°. Pour toutes les autres orientations (45°, 90° et 135°), le temps de préparation est toujours significativement plus long lorsque la cible est en diagonale derrière l'avatar que pour les autres positions de la cible par rapport à l'avatar. De plus, lorsque l'avatar est à 135° par rapport au participant, le temps de préparation

de l'énoncé est plus court lorsque la cible est alignée avec le plan sagittal de l'avatar, que lorsqu'elle se situe en diagonale devant l'avatar.

Les analyses post-hoc révèlent deux autres interactions entre l'orientation de l'avatar et la position de la cible par rapport à celui-ci. D'une part, lorsque la cible est alignée avec le plan sagittal de l'avatar, le temps de préparation des énoncés est plus long lorsque l'avatar est orienté perpendiculairement (90° : 1.1s) que lorsqu'il fait face au participant (180° : 0.8s, $p<.05$). D'autre part, lorsque la cible est située en diagonale derrière l'avatar, le temps de préparation de l'énoncé est plus long lorsque l'avatar est tourné à 135° (1.9s) que lorsqu'il est tourné à 45° (1.5s, $p<.01$). L'ensemble des analyses post-hoc est présenté dans l'annexe 11.

Synthèse des résultats sur le temps de préparation des énoncés

Le temps de préparation des énoncés est influencé par la position de la cible par rapport au participant et à l'avatar.

L'influence de la position de la cible par rapport au participant se caractérise par une augmentation des temps de préparation avec la distance au plan sagittal du participant.

Les cibles situées dans le plan sagittal de l'avatar induisent les temps de préparation les plus courts et les cibles situées en diagonale derrière l'avatar les temps de préparation les plus longs.

L'orientation de l'avatar n'a qu'une faible influence sur le temps de préparation des énoncés.

1.3.4 Temps de production des énoncés

L'ANOVA à deux facteurs (8 orientations de l'avatar * 8 positions de la cible) sur les temps de production a mis en évidence un effet de l'orientation de l'avatar ($F(7, 147)=2.23$, $p<.05$), un effet de la position de la cible codée par rapport au participant ($F(7, 147)=2.58$, $p<.05$), et une interaction significative entre les deux variables ($F(49, 1029)=13.64$, $p<.0001$). Les analyses post-hoc ne montrent aucun effet pour l'orientation de l'avatar ou la position de la cible. Nous présentons donc directement les résultats issus du recodage des données.

L'ANOVA à deux facteurs a été réalisée après avoir regroupé les données selon l'alignement de la cible par rapport à l'avatar : 4 positions de la cible par rapport à l'avatar * 5 orientations de l'avatar. Cette analyse a mis en évidence un effet de la position de la cible par rapport à l'avatar ($F(3, 63)=22.88$, $p<.0001$), un effet de l'orientation de l'avatar ($F(4, 84)=2.95$, $p<.05$) et une interaction significative entre les deux variables ($F(12, 252)=2.25$, $p<.05$).

En ce qui concerne la position de la cible codée par rapport à l'avatar, les analyses post-hoc ont mis en évidence les mêmes types d'effets que pour l'exigence mentale et le temps de préparation des énoncés (Figure 48). Le temps de production est significativement plus court lorsque la cible est alignée avec le plan sagittal de l'avatar (1.8s) que lorsque la cible est dans n'importe quelle autre position (2.5s pour diagonale devant, $p<.0001$; 2.3s pour plan frontal, $p<.001$ et 2.9s pour diagonale derrière, $p<.0001$). De même, lorsque la cible est située en diagonale derrière l'avatar, le temps de production est significativement plus long que pour toutes les autres situations (diagonale devant, $p<.05$, plan frontal, $p<.001$). La différence de temps de production moyen entre les situations où la cible est en diagonale devant ou sur le plan frontal de l'avatar n'est pas significative.

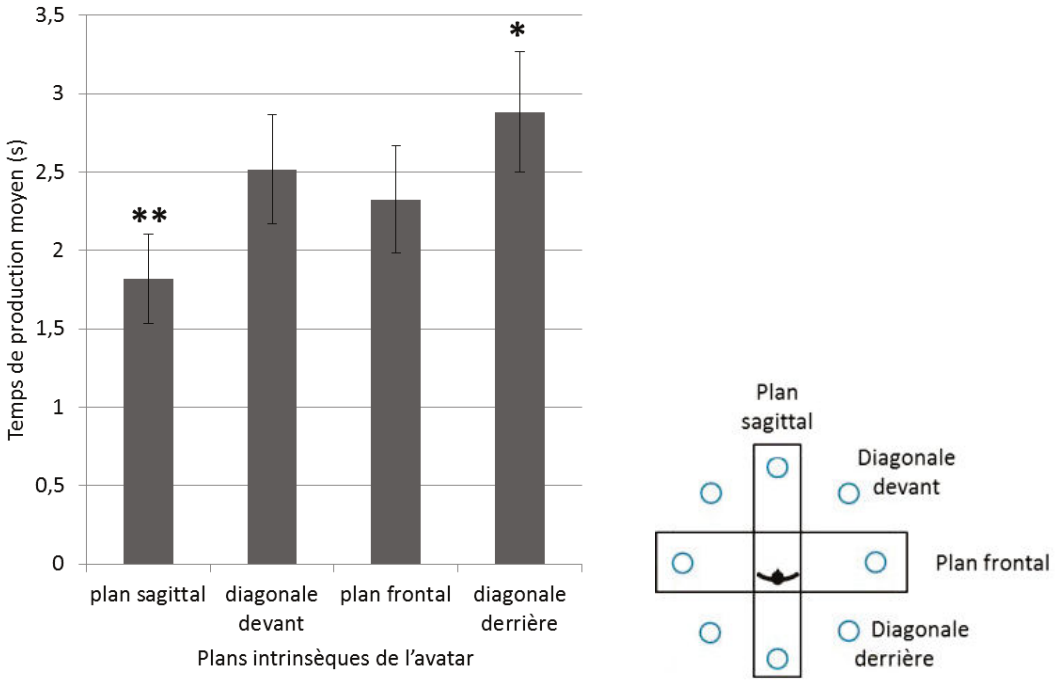


Figure 47 : Temps de production moyen en fonction de la position de la cible, codée par rapport à l'avatar. * p<.05, ** p<.01

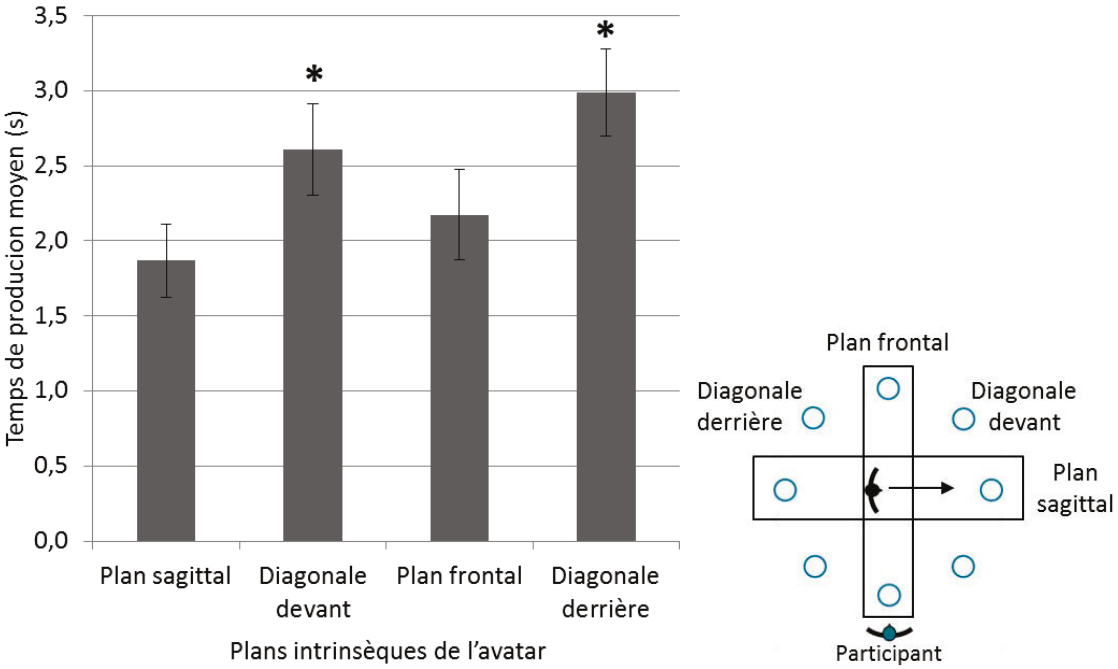


Figure 48 : Temps de production moyens en fonction de la position de la cible par rapport aux plans intrinsèques de l'avatar lorsque l'avatar est tourné de 90°. * p<.05

En ce qui concerne l'interaction entre la position de la cible codée par rapport à l'avatar et l'orientation de l'avatar, les analyses post-hoc montrent que l'interaction est principalement due au changement de pattern lorsque l'avatar est perpendiculaire au participant (90°, Figure 48). En effet, dans cette orientation, il n'y a pas de différence entre les temps de production lorsque la cible est dans le plan sagittal de l'avatar (1.9s), et dans le plan frontal (2.2s). Par contre, lorsque la cible est en diagonale devant

l'avatar (2.6s), le temps de production est plus long que lorsque la cible est alignée avec le plan sagittal (p<.01) et frontal (p<.01), mais moins long que lorsqu'elle se situe en diagonale derrière l'avatar (3s, p<.05). L'ensemble des analyses post-hoc est présenté dans l'annexe 12.

Synthèse des résultats sur le temps de production des énoncés

Tout comme les cadres de référence, l'exigence mentale et le temps de préparation des énoncés, le temps de production des énoncés est modulé par la position de la cible par rapport à l'avatar. On retrouve le même effet global des caractéristiques intrinsèques sur l'organisation de l'espace : les temps de production sont plus courts lorsque la cible est située dans le plan sagittal de l'avatar, et plus longs lorsque la cible est située en diagonale dernière l'avatar.

Cet effet est également modulé par l'orientation de l'avatar.

1.3.5 Nombre de mots

L'ANOVA à deux facteurs (8 orientations de l'avatar * 8 positions de la cible par rapport au participant) sur le nombre de mots utilisé dans les instructions a mis en évidence d'une part un effet de la position de la cible codée par rapport au participant ($F(7, 140)=2.33$, $p<.05$), et d'autre part une interaction entre les deux variables ($F(49, 980)=11.02$, $p<.0001$). Cette analyse n'a pas mis en évidence d'effet direct de l'orientation de l'avatar ($F(7, 140)=.97$, $p=.46$) sur le nombre de mots par énoncé. Les analyses post-hoc ne montrent pas d'effet de la position de la cible par rapport au participant. Nous présentons donc directement les résultats issus du recodage des données. L'ANOVA à deux facteurs réalisée après avoir regroupé les données selon l'alignement de la cible par rapport à l'avatar (4 positions de la cible par rapport à l'avatar * 5 orientations de l'avatar) n'a mis en évidence qu'un effet de la position de la cible par rapport à l'avatar ($F(3, 63)=16.05$, $p<.0001$), présenté Figure 49.

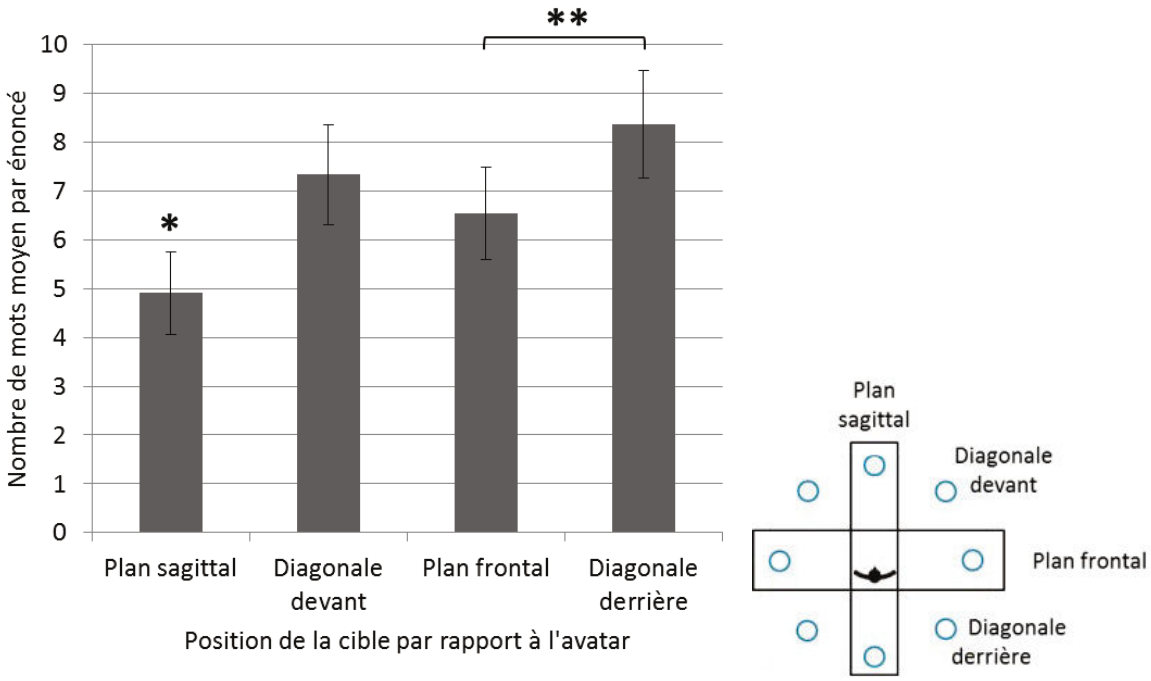


Figure 49 : Nombre de mots moyen en fonction de la position de la cible, codée par rapport à l'avatar. * p<.05, ** p<.01

Les analyses post-hoc ont montré que les participants ont utilisé en moyenne 4.9 mots pour localiser une cible située dans le plan sagittal de l'avatar, ce qui est significativement moins que dans les autres configurations (7.3 pour diagonale devant, $p<.001$; 6.5 pour le plan frontal, $p<.05$ et 8.4 pour diagonale derrière, $p<.001$). De plus, la différence entre le nombre de mots moyen lorsque la cible est dans le plan frontal de l'avatar est significative par rapport aux situations où la cible est en diagonale derrière l'avatar ($p<.01$).

Synthèse des résultats sur le nombre de mots moyen utilisé par énoncé
Le nombre de mots est modulé principalement par la position de la cible par rapport à l'avatar, et plus précisément par ses caractéristiques intrinsèques. Les participants ont utilisé moins de mots pour décrire les cibles alignées avec le plan sagittal de l'avatar.

1.3.6 Corrélations

La comparaison entre les moyennes des variables pour chaque condition a mis en évidence une corrélation entre le temps de production des énoncés et le nombre de mots ($r=.91$, $p<.001$), et entre le temps de production et le temps de préparation ($r=.49$, $p<.05$).

1.3.7 Questionnaire et autres comparaisons

Les participants ont estimé se sentir présents dans l'environnement virtuel à 3.1 sur une échelle de 10 (ET=.5) et à 6.5 (ET =.7) dans l'environnement réel. Ils ont estimé la présence de leur collaborateur avec eux dans l'environnement virtuel à 5.3 (ET=.8). Il n'y a pas de corrélation entre ces trois variables.
L'analyse des données du questionnaire a mis en évidence une corrélation négative entre l'âge des participants et leur résultat au test des rotations mentales (MRT-A ®). On constate ainsi que plus l'âge augmente, moins les performances à ce test sont bonnes ($r=-.44$, $p<.05$).

1.4 Synthèse et discussion – Production d'énoncés sans indice

L'étude sur la production d'énoncés sans indice a fourni un grand nombre de résultats dont nous allons résumer et discuter les plus importants ici.

Le premier constat est que les participants ont principalement donné leurs instructions en utilisant l'avatar comme référence. Il est intéressant de constater que toutes les variables sont d'ailleurs influencées par la position de la cible par rapport à l'avatar, et plus précisément par les caractéristiques intrinsèques de l'avatar. En effet, que ce soit pour l'exigence mentale, les temps de préparation et production des énoncés et le nombre de mots, les cibles alignées avec le plan sagittal de l'avatar induisent les mesures les plus faibles et les cibles en diagonale derrière l'avatar induisent les mesures les plus fortes.

Le second constat est l'influence de l'orientation de l'avatar sur l'exigence mentale et le temps de préparation des énoncés. L'effet global est une augmentation de l'exigence mentale avec le degré de rotation de l'avatar. Ce résultat est cependant à moduler selon la position de la cible par rapport à l'avatar. En effet, l'augmentation décrite n'est pas constatée lorsque la cible est alignée avec le plan sagittal de l'avatar.

Nous avons également constaté que l'exigence mentale et le temps de préparation des énoncés étaient influencés par la position de la cible par rapport au participant. En effet, les mesures de ces deux variables augmentent plus la cible est éloignée du plan sagittal du participant.

1.4.1 Influence de la position de la cible dans le champ de vision du locuteur

L'exigence mentale est modulée par la position de la cible par rapport au participant. Ce résultat est à mettre en lien avec les résultats observés pour les temps de préparation des énoncés. En effet, d'après la taille des effets et la variable impliquée, ces résultats indiquent que les cibles les plus proches du participant sont détectées plus vite que les cibles les plus éloignées. Ainsi, on retrouve ici l'effet largement décrit dans la littérature de l'excentricité de la cible sur sa détection (Carrasco, Evert, Chang, & Katz, 1995; Holmes, Cohen, Haith, & Morrison, 1977; Latham & Whitaker, 1996). Cet effet de l'excentricité de la cible n'affecte ni les cadres de référence choisis, ni le temps de production des énoncés, ni le nombre de mots composant ces énoncés, mais uniquement le temps de préparation et l'exigence mentale. Ce premier effet suggère que cette perception égocentrée de la scène fait partie de la représentation occurrente de la situation.

1.4.2 Influence de la position de la cible par rapport au destinataire

Le facteur influençant le plus largement la production des énoncés spatialisés est la position de la cible par rapport au destinataire. En effet, l'exigence mentale, les temps de préparation et de production des énoncés et le nombre de mots varient en fonction de la position de la cible par rapport à l'avatar. Ce résultat est cohérent avec les cadres de référence utilisés puisque 80% des énoncés sont centrés sur le destinataire. L'avatar est donc utilisé majoritairement comme référence pour localiser la cible. Par ailleurs les résultats confirment la *Spatial Framework Theory* (Tversky, 2005b) : une cible alignée avec le plan sagittal de l'avatar induit une exigence mentale plus faible qu'une cible alignée avec le plan frontal de l'avatar. Cette étude apporte des compléments par rapport à ce modèle : il y a prédominance de l'axe avant/arrière sur tous les autres axes de l'avatar, et prédominance des directions diagonale devant et plan frontal sur la direction diagonale derrière. Autrement dit, localiser une cible et décrire sa position est plus simple si elle se situe devant ou derrière l'avatar, et plus compliqué si elle se situe en diagonale derrière lui. Ce phénomène est résumé et illustré sur la Figure 50.

La prédominance de l'axe avant/arrière sur l'exigence mentale peut être attribuée à plusieurs facteurs. D'une part, les descriptions en elles-mêmes sont différentes : les énoncés décrivant les cibles alignées avec le plan sagittal sont principalement centrés sur le destinataire *intrinsèques* et contiennent moins de mots que les énoncés décrivant les autres cibles. Ainsi, la faible exigence mentale pour la description de ces cibles peut être attribuée au fait que ces cibles peuvent être décrites en peu de mots. En effet, un seul mot suffit : "devant" ou "derrière". Mais si le langage s'est doté de mots pour signifier ces emplacements, c'est qu'ils sont importants. On rejoint ici les arguments de la *Spatial Framework Theory* (Tversky, 2005b) : l'axe avant/arrière prédomine sur les autres parce qu'il est asymétrique et organise notre rapport au monde. L'avant du corps, en opposition à l'arrière, correspond à ce qui est directement perceptible et manipulable ainsi qu'au sens de la marche. Certains participants l'ont d'ailleurs souligné dans leurs commentaires, les cibles situées à l'avant et à l'arrière de l'avatar ont un statut particulier car elles sont explicables plus facilement. Par ailleurs, le fait que les temps de production des énoncés pour décrire les cibles alignées avec le plan sagittal soient plus courts que les autres est cohérent avec le faible nombre de mots de ces énoncés : moins il y a de mots, moins il faut de temps pour les dire. Enfin, les temps de préparation des énoncés localisant les cibles situées dans le plan sagittal de l'avatar sont également plus courts que pour les autres positions de cible. Ce fait peut être attribué d'une part à une détection rapide, mais également à l'association linguistique entre la position de la cible par rapport à l'avatar et le vocabulaire pour la décrire. En effet, nous disposons de mots précis pour décrire ces directions. Elles sont d'ailleurs acquises plus tôt que les autres (Verjat, 1994). L'effet de prédominance du plan sagittal observé ici est donc lié à la fois à l'organisation de l'espace et au langage décrivant cet espace. Les deux dimensions se renforçant l'une l'autre.

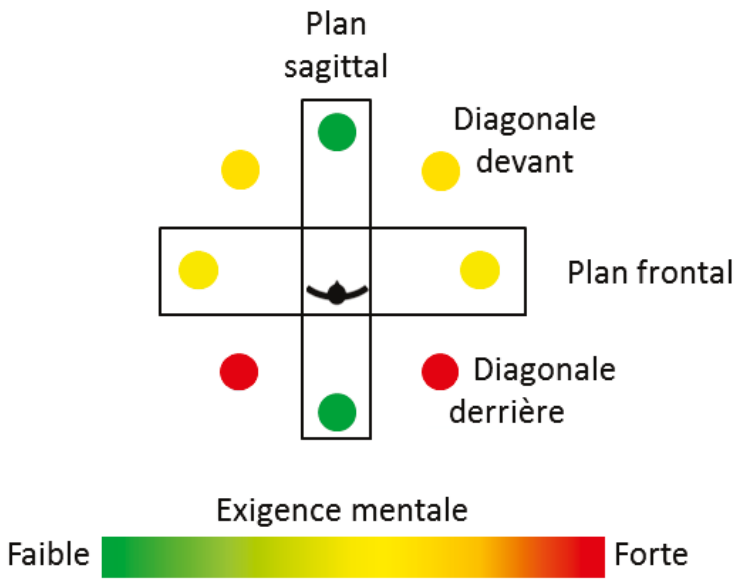


Figure 50 : Modèle de la *Spatial Framework Theory* complété avec les résultats de cette étude. La position de la cible par rapport au destinataire influence l'exigence mentale sollicitée pour indiquer la position de celle-ci.

Les situations induisant l'exigence mentale la plus forte sont celles où la cible est située en diagonale derrière l'avatar. Dans ces situations, les participants ont utilisés des cadres de référence spécifiques par rapport aux autres configurations. Par rapport aux cibles situées en diagonale devant, les participants ont produit plus d'énoncés mixtes, c'est-à-dire qu'ils ont plus souvent eu recours à plusieurs cadres de référence dans un même énoncé pour localiser la cible. Cependant cette différence de choix de

cadre de référence ne transparaît pas dans le nombre de mots, puisque la différence n'est pas significative entre les énoncés produits pour décrire les cibles situées en diagonale devant et en diagonale derrière l'avatar. On constate cependant que les temps de préparation et de production des énoncés sont significativement plus longs lorsqu'il s'agit de décrire les cibles situées en diagonale derrière qu'en diagonale devant. Ainsi les mesures des temps de préparation et de production donnent une information supplémentaire : bien qu'il n'y ait pas de différence significative dans le nombre de mots, les participants mettent significativement plus de temps pour préparer et produire les énoncés pour décrire les cibles situées en diagonale derrière que les cibles situées en diagonale devant l'avatar. Il n'est donc pas uniquement question des mots, mais du choix de ces mots et de l'hésitation que ce choix peut entrainer. Par ailleurs l'évaluation de l'exigence mentale est bien plus marquée que les autres variables. En effet, l'exigence mentale moyenne est trois fois supérieure pour les cibles situées en diagonales derrière par rapport aux cibles situées dans le plan sagittal. Ce rapport n'est pas du tout le même pour les temps de préparation (1.7/1 seconde), de production (2.9/1.8 seconde) ou le nombre de mots (8.4/4.9 mots). Ainsi, l'exigence mentale serait bien le reflet des activités cognitives (choix des cadres de référence, opérations mentales) que le participant doit effectuer pour assurer une bonne performance (Cazamian et al., 1996). De plus, cette direction n'étant pas primordiale dans notre rapport au monde, nous n'avons pas de mot la désignant spécifiquement. Le locuteur doit donc choisir entre plusieurs alternatives possibles. Cependant, ce constat d'absence de mots pour désigner la direction est également valable pour la direction diagonale devant. De ce fait, la difficulté supplémentaire constatée pour les cibles situées en diagonale derrière l'avatar viendrait du fait que pour percevoir la cible le destinataire devra se tourner, ce qui modifiera l'orientation de ses directions organisatrices.

En ce qui concerne les cibles situées en diagonale devant et sur le plan frontal de l'avatar, elles induisent des exigences mentales similaires alors que les cadres de référence utilisés pour produire les énoncés ne sont pas tout à fait les mêmes. Les cibles alignées avec le plan frontal devraient bénéficier de l'alignement par rapport aux cibles situées en diagonale. C'est effectivement le cas par rapport à la situation diagonale derrière, mais pas pour la diagonale devant. Selon nous, cet effet de prédominance des plans sur les diagonales est compensé par le fait que la cible située en diagonale devant est dans la partie de l'espace *vista*, c'est-à-dire visible et atteignable rapidement par l'avatar, contrairement à la cible située en diagonale derrière.

1.4.3 Influence de l'orientation de l'avatar : modèle des rotations mentales ?

L'exigence mentale est également influencée par l'orientation de l'avatar, elle augmente de manière linéaire avec le degré d'orientation de l'avatar. Ce résultat valide l'hypothèse que la prise de perspective est un processus similaire à une rotation mentale (Schober, 1995). Cependant, ce résultat est à nuancer en fonction de la position de la cible par rapport au destinataire. En effet, l'interaction entre l'orientation de l'avatar et la position de la cible par rapport à lui montre que le modèle des rotations mentales ne s'applique pas lorsque la cible est dans le plan sagittal de l'avatar. Ainsi, comme nous l'avons évoqué plus tôt, le plan sagittal de la référence utilisée est prédominant : il est perçu plus vite et peut être décrit par un seul mot. Il semble donc que les locuteurs n'aient pas besoin d'opérer de transformation mentale pour produire un énoncé décrivant une cible alignée avec le plan sagittal de l'avatar, malgré le fait qu'ils utilisent un cadre de référence centré sur le destinataire. Ce résultat est à rapprocher des études sur la prise de perspective (Kessler & Rutherford, 2010; Michelon & Zacks, 2006). Dans ces études, les auteurs proposaient deux tâches : déterminer la latéralité d'une cible par rapport à une autre perspective, ou déterminer la visibilité d'une cible depuis cette perspective. Les résultats indiquent que les temps de réaction augmentent avec l'augmentation de l'écart d'orientation uniquement pour la tâche

impliquant la latéralité, mais pas la visibilité. Les auteurs concluent qu'il existe deux niveaux de prise de perspective. Le niveau 1 correspond au fait d'imaginer la ligne de visibilité depuis un point de vue donné, alors que le niveau 2 correspond à la simulation mentale du mouvement du corps permettant de se mettre dans la perspective donnée. Par rapport à notre étude où une seule tâche est demandée aux participants, les résultats nous incitent à faire la distinction entre prise de perspective verbale et prise de perspective cognitive. La prise de perspective verbale est l'utilisation d'un cadre de référence centré sur une référence, ici le destinataire. La prise de perspective cognitive est le processus mental qui demande au locuteur de se mettre à la place de la référence utilisée (transformation mentale), dont l'effort augmente avec l'écart d'orientation. Cette distinction incite donc à compléter les études sur les dialogues spatialisés utilisant les cadres de référence, soit avec des mesures de la charge mentale, mais ce n'est pas toujours possible lors d'une tâche complexe, soit de prendre en compte la position de la cible par rapport à la référence utilisée. En tout cas, ce résultat permet d'appréhender d'une autre manière la gestion de la charge mentale au cours d'activités collaboratives.

Les résultats ont mis en évidence un autre effet de l'orientation de l'avatar sur l'exigence mentale. Lorsque l'avatar est orienté à 90°, le schéma général de la *Spatial Framework Theory* est modifié. Cette modification est due au fait que dans cette orientation de l'avatar, le plan frontal de l'avatar est aligné avec le plan sagittal du participant, ce qui induit une diminution de l'exigence mentale pour les cibles situées dans le plan frontal de l'avatar. La diminution est telle qu'il n'y a plus de différence d'exigence mentale entre les cibles situées dans le plan sagittal et frontal de l'avatar. Ce résultat conforte le fait que lorsque deux directions organisatrices, de deux cadres de référence différents, sont alignées, cela diminue l'exigence mentale. Cela renforce également l'idée soulevée plus haut de la prédominance des plans intrinsèques sur les directions diagonales. Ces différences d'exigence mentale sont également retrouvées pour les temps de préparation et de production des énoncés.

1.4.4 Principe du moindre effort collaboratif

Nous avons évoqué dans les paragraphes précédents les différents facteurs influençant l'exigence mentale. Nous allons discuter ici de la gestion de cette composante de la charge mentale. L'hypothèse que nous avions posée était que si les participants suivaient le principe du moindre effort collaboratif, ils prendraient la perspective de l'avatar pour minimiser l'effort de leur collaborateur. C'est ce que nous avons observé, 80% des énoncés ont été produit dans un cadre de référence centré sur le destinataire et seule une personne a choisi d'utiliser un cadre de référence égocentré pour donner ses indications. Cependant, cette hypothèse est discutable, car il est possible qu'ils aient pris la perspective de leur collaborateur parce que cela était plus simple pour eux. En effet, sauf pour les cibles situées directement devant le participant, les autres cibles sont difficilement distinguables les unes des autres dans un référentiel égocentré. La personne ayant utilisé uniquement le cadre de référence égocentré a dû faire appel à un système extrinsèque : elle a numéroté les tables par rapport à sa position dans l'environnement. L'utilisation de ce système supposait qu'elle se rappelle des numéros associés à chaque table ou qu'elle recompte les tables à chaque fois, ce qui est cognitivement coûteux. Inversement, la présence de l'avatar donne une référence par rapport à laquelle il est possible de situer les cibles. Comme nous l'avons vu, selon la position de la cible par rapport à l'avatar l'exigence mentale requise pour la localiser n'est pas la même, mais cette étude ne permet pas de conclure sur le fait qu'il est plus ou moins facile d'utiliser un cadre de référence égocentré ou centré sur le destinataire. Bien que la présence d'énoncés centrés sur le destinataire ne permette pas de conclure sur une gestion collective de la charge mentale, d'autres résultats permettent de le faire.

D'une part, certains participants ont justifié leur choix d'avoir pris la perspective de leur interlocuteur parce qu'ils considéraient que cela lui faciliterait la tâche (« plus facile à appliquer pour lui », « pour lui simplifier la tâche »). Il y a donc bien une notion de gestion collective dans ce choix. Un participant a même indiqué qu'il a privilégié la facilité de compréhension de son collaborateur au détriment de l'effort qu'il devait fournir pour donner les instructions, suggérant qu'il aurait été plus facile d'utiliser un cadre de référence égocentré.

D'autre part, nous avons vu que la position de la cible par rapport au participant affectait l'exigence mentale. Certains participants ont d'ailleurs soulevé le fait qu'ils se sont utilisés comme référence. Bien que la répartition des cadres de référence ne soit pas affectée par la position de la cible par rapport au participant, le cadre de référence égocentré a été utilisé plus souvent lorsque la cible était située dans le plan sagittal du participant. Il est alors possible qu'il soit plus facile de produire un énoncé égocentré lorsque la cible est dans le plan sagittal du locuteur qu'un énoncé centré sur le destinataire. La question à se poser est : est-ce que l'alignement de la cible dans le plan sagittal du locuteur est prédominant sur toutes les positions de cible par rapport à l'avatar, ou uniquement sur les situations les plus difficiles ? Le très faible taux d'utilisation d'énoncés égocentrés ne permet pas de conclure sur cette question. Notre hypothèse est que le plan sagittal du locuteur est prédominant sur les autres directions, mais que les participants ont privilégié la stabilité de leurs énoncés pour ne pas imposer de surcharge cognitive liée à des changements de cadres de référence. Etant donné que dans la plupart des cas ils ont donné leurs instructions dans un cadre de référence centré sur le destinataire, ils ont conservé ce choix pour toutes les situations, même celles où un autre cadre de référence était plus simple à utiliser. Certains participants ont d'ailleurs souligné le fait que cette stabilité était un choix collaboratif. Les participants font ainsi l'hypothèse que recevoir des informations dans un nouveau cadre de référence demande un effort supplémentaire. Nous reviendrons sur cette question dans la discussion générale.

Par ailleurs le fait que l'exigence mentale soit plus faible pour les cibles alignées dans le plan sagittal du participant et le plan frontal de l'avatar, renforce l'importance de l'alignement de la cible par rapport au plan sagittal du locuteur. Ce résultat renvoie aux résultats de l'étude de Galati et Avraamides (2014) dans laquelle ils ont mis en évidence le bénéfice de l'alignement de plusieurs systèmes de référence.

Enfin, par rapport aux modèles proposés par Schober (1995), un des participants explicite le fait que lorsque l'avatar lui fait face, sa stratégie est d'inverser droite et gauche. Cependant, cette stratégie ne semble pas être appliquée par tous et les résultats ne permettent pas de dire si elle diminue effectivement l'exigence mentale.

Pour résumer les éléments apportés par cette étude, nous avons tout d'abord apporté des éléments sur le modèle des cadres de référence : la prédominance du plan sagittal sur les autres directions centrées sur le destinataire est liée à l'importance de cette direction dans notre rapport au monde. Cette importance est telle que, non seulement nous la détectons plus vite, sans avoir à recourir à des transformations mentales coûteuses, mais nous nous sommes dotés de mots pour décrire rapidement cette direction.

L'objectif de l'étude suivante est de voir si les configurations les plus complexes, notamment celles où la cible est situé dans la diagonale arrière du participant, peuvent être rendues plus faciles en présence d'indices spatiaux.

2. Etude 2 : Production d'énoncés avec indices spatiaux

Les résultats exposés dans la partie précédente ont mis en évidence que, lors d'une tâche de production d'instructions spatialisées dans un environnement virtuel, la localisation de la cible par rapport au destinataire était un facteur important. Un premier constat, cohérent avec les résultats exposés dans la littérature, est que l'alignement de la cible avec le plan sagittal de l'avatar est un facteur facilitateur. En effet, nous avons pu constater une exigence mentale moindre et des temps de préparation et de production d'énoncés plus courts lorsque la cible est dans le plan sagittal de l'avatar, c'est-à-dire directement devant ou derrière lui, par rapport aux autres positions de la cible. Le second constat, est que la situation la plus complexe pour les locuteurs est lorsque la cible se trouve en diagonale derrière leur collaborateur. Un troisième constat est que lorsque la cible est alignée avec plusieurs références (le participant et l'avatar) la situation est plus simple que lorsqu'elle n'est alignée avec aucune. Ces constats faits, l'objectif est d'évaluer comment certains indices visuels peuvent influencer le choix des cadres de référence et l'exigence mentale associée, ce qui pourrait amener à moduler les conclusions de l'étude précédente.

L'avantage des environnements virtuels est qu'ils sont modulables, notamment dans les indices visuels mis à disposition des utilisateurs. Nous avons choisi de comparer deux types d'indices spatiaux : incarnés, i.e. directement présent sur l'avatar, et distaux (Creem-Regehr, Willemsen, Gooch, & Thompson, 2005). L'indice visuel incarné est un avatar latéralisé dont la droite et la gauche sont différentiées par des manches de couleur (Figure 51 – gauche). Notre hypothèse est que l'indice incarné, en rendant asymétrique le plan frontal qui est naturellement symétrique, faciliterait la perception de la latéralisation de l'avatar. Ainsi, pour les situations où l'avatar est désaxé par rapport au locuteur et où la cible est sur la droite ou la gauche de l'avatar, nous faisons l'hypothèse que la présence de l'avatar latéralisé diminuera l'exigence mentale. Les indices distaux sont matérialisés par quatre panneaux de couleurs différentes positionnés en périphérie de la zone de travail (Figure 51 - droite). Notre hypothèse est que le fait d'ajouter des indices distaux fixes, en donnant de nouvelles références potentielles, faciliterait un codage exocentré non centré sur le destinataire. Ces indices seraient plus fiables car indépendants de l'orientation de l'avatar et ne nécessiteraient pas de prise de perspective. Leur utilisation limiterait ainsi les ambiguïtés liées aux situations complexes (notamment lorsque la cible est située en diagonale derrière l'avatar).

L'objectif de cette étude est donc d'apporter des données sur l'effet des indices spatiaux sur la charge mentale requise pour la production d'énoncés spatialisés.

2.1 Matériels et méthodes

2.1.1 Participants

Vingt-deux personnes ont participé à cette étude : cinq femmes et dix-sept hommes. Ils étaient âgés de 19 à 67 ans, dont la moyenne d'âge était de 29,6 ans.

2.1.2 Matériels

La majorité du matériel utilisé pour cette étude est le même que celui utilisé dans l'étude précédente décrit page 73, le participant étant placé devant le même mur immersif. L'environnement virtuel est identique : une pièce carrée dans laquelle huit tables sont disposées en cercle et de manière régulière. L'avatar du collaborateur est placé au centre de ces tables et à chaque fois il fait face à une des tables. La tâche est la même : le participant doit indiquer la position de la cible à son collaborateur

représenté par l'avatar visible dans la scène. Le scénario utilisé est également le même que celui de l'étude précédente (Annexe 9).

Afin d'apprécier l'effet des indices spatiaux, quatre combinaisons d'indiçages ont été utilisées :

- l'environnement virtuel de l'expérience précédente a été repris tel quel. Cette condition est nommée **Sans indice** ;
- dans ce même environnement l'avatar simple utilisé pour la condition Sans indice est remplacé par l'avatar latéralisé. Cette condition est nommée **Avatar latéralisé** ;
- la même pièce virtuelle utilisée dans la condition Sans indice est agrémentée de quatre panneaux de couleurs différentes. Cette condition est nommée **Panneaux** ;
- enfin dans la dernière condition les deux types d'indices sont regroupés, cette condition est nommée **Panneaux + Avatar latéralisé**.

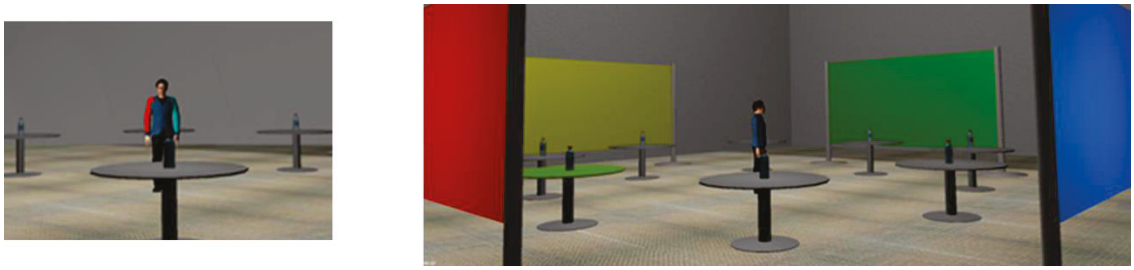


Figure 51 : Indices visuels testés dans l'étude 2. A gauche : l'avatar latéralisé, à droite : les panneaux de couleur.

A la fin de l'expérience, les participants devaient répondre à un questionnaire, dont une question a été ajoutée par rapport au questionnaire proposé lors de l'étude précédente (Annexe 13) et réaliser deux tests d'aptitudes spatiales : le test d'autoévaluation Santa Barbara et le test des rotations mentales MRT-A ®, présentés dans la partie matériels de l'étude précédente (page 74 et en annexes 6 et 7).

2.1.1 Procédure

Le plan expérimental de cette étude comprenait trois facteurs : l'orientation de l'avatar {quatre modalités (0°, 90°, 180°, 270° Figure 52)} * la position de la cible {huit modalités} * les types d'indices spatiaux {quatre modalités}, ce qui donnait 128 conditions expérimentales, chacune répétée deux fois. A la deuxième répétition, le participant devait évaluer l'exigence mentale de la situation (de la même manière que dans l'étude précédente, page 74). Les essais ont été répartis de la manière suivante : toutes les conditions expérimentales pour une même condition d'indiçage étaient présentées successivement et de manière aléatoire. Le contrebalancement des conditions d'indiçage a été contrôlé entre les participants. Nous avons choisi de procéder de cette manière pour faciliter l'immersion des participants dans l'environnement virtuel. Pour permettre aux participants de mettre en place de nouvelles stratégies s'ils le souhaitaient, une phase de visite de l'environnement et huit essais d'entrainement ont eu lieu à chaque changement d'indices spatiaux. Ce plan expérimental est résumé Figure 53. Pendant l'expérience, si le participant se trompait dans les touches, l'essai était annulé et reporté à la fin de la série.

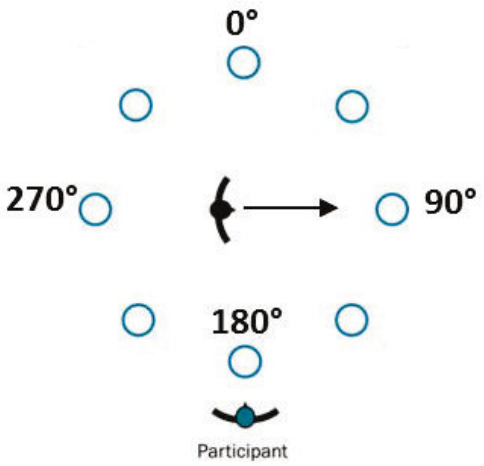


Figure 52 : Les quatre orientations de l'avatar utilisées dans l'étude Production d'énoncés avec indices spatiaux : 0°, 90°, 180° et 270°.

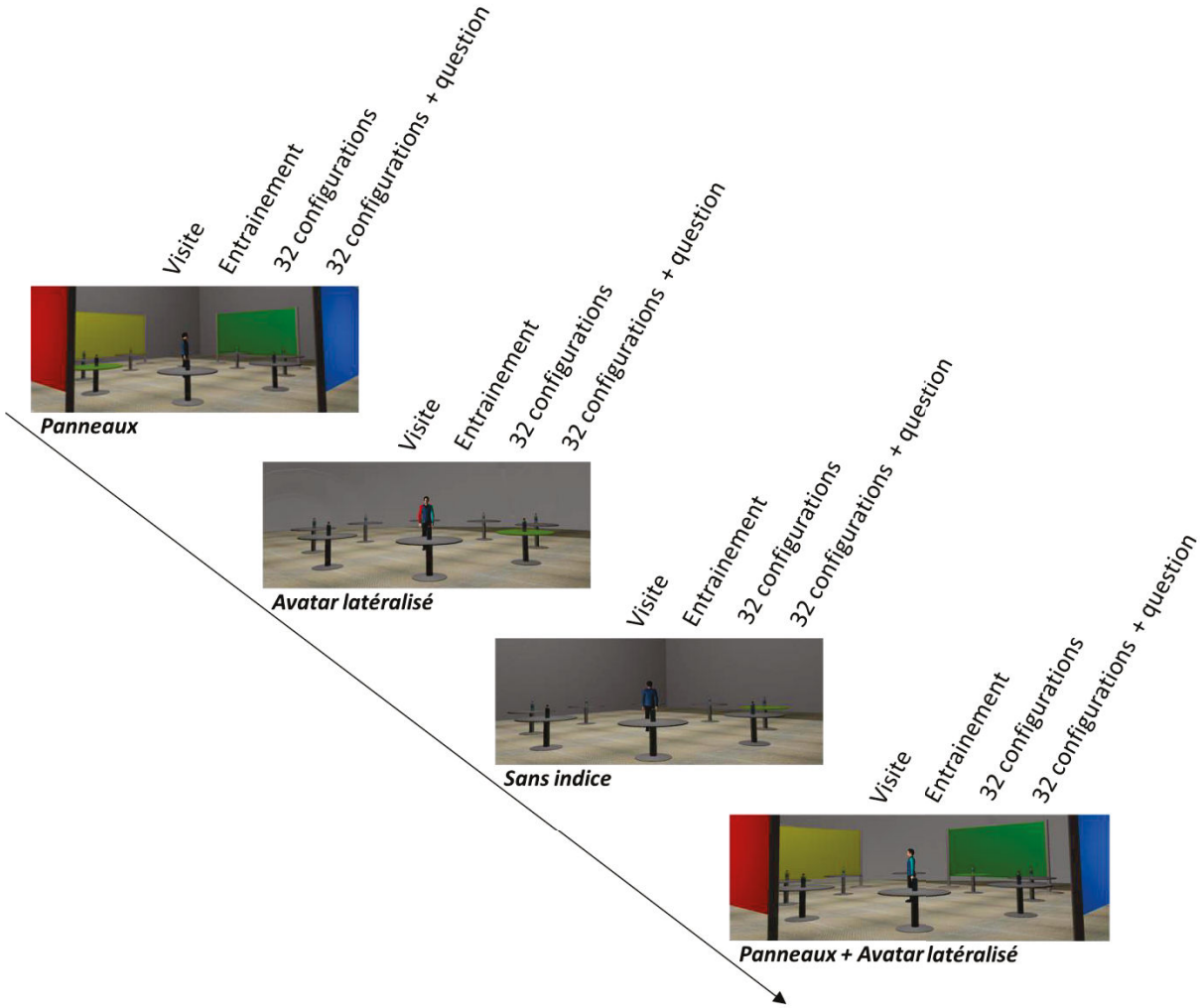


Figure 53 : Résumé du déroulement de l'étude 2. Pour chaque condition d'indiçage, les participants visitaient l'environnement, faisaient huit essais d'entrainement puis les essais tests. Lors de la deuxième répétition, les participants répondaient à la question sur l'exigence mentale. L'ordre des conditions d'indiçage est contrebalancé sur l'ensemble des participants.

2.1.2 Codage des énoncés spatialisés

Chaque instruction a été codée selon le ou les cadres de référence utilisés pour localiser la cible. Pour cela nous avons repris les catégories définies pour l'étude précédente, auxquelles nous avons ajouté les énoncés exocentrés. Nous avons également regroupé tous les énoncés centrés sur le destinataire (sans faire la distinction selon le système de coordonnées utilisé) :

- **centrés sur le destinataire** : nous avons regroupé dans cette catégorie tous les énoncés utilisant la perspective du destinataire/avatar, quelles que soient la nature des caractéristiques utilisées (intrinsèques ou non) ;
- **égocentrés** : énoncés utilisant la perspective du participant ;
- **exocentrés** : nous avons regroupé sous cette catégorie toutes les instructions utilisant les indices environnementaux (panneaux) pour localiser la cible – « la bouteille piégée est à côté du panneau bleu », « entre les panneaux bleu et rouge » ;
- **multiples** : nous avons regroupé les énoncés utilisant plusieurs cadres de référence, comme dans l'étude précédente, et les énoncés utilisant plusieurs références pour localiser la cible (l'avatar et/ou le participant et/ou les panneaux de couleur) comme dans cet exemple « la bouteille piégée est sur la table sur ta droite devant le panneau bleu » ;
- **autres** : énoncés utilisant un autre moyen de référencement, par exemple le fait de faire référence à la situation précédente – « même table ».

2.1.3 Traitements statistiques

Le nombre d'énoncés par cadre de référence, ont été comparé grâce à des tests du Khi2.

Pour les temps de préparation, de production, les valeurs des deux essais ont été moyennées. Cette répétition n'entre donc pas dans le plan d'analyse. Pour ces deux variables et l'exigence mentale, les données ont été comparées par des analyses de variance à mesures répétées (ANOVA), dont les facteurs seront précisés lors de la présentation des résultats. Lorsque l'analyse de variance se révélait significative, des analyses post-hoc (Tukey HSD) ont été réalisées. Pour tous les tests, le niveau de significativité utilisé était $p<.05$.

Pour plus de lisibilité, seuls les résultats significatifs concernant la présence des indices seront présentés.

2.2 Résultats - Production d'énoncés avec indices spatiaux

Le même constat a été fait que lors de la première étude, les participants ont souvent explicité leur stratégie au cours des premiers essais afin de réduire leurs instructions aux informations pertinentes uniquement.

2.2.1 Cadres de références

Pour les conditions *Sans indice* et *Avatar latéralisé*, les participants ont produit 99% d'énoncés centrés sur le destinataire, c'est-à-dire centrés sur l'avatar (Figure 54). Les rares énoncés non centrés sur l'avatar étaient soit égocentrés (« en face de moi »), soit faisaient référence à l'essai précédent (« même table »). En présence des indices distaux (conditions *Panneaux* et *Panneaux + Avatar latéralisé*), respectivement 45,8% et 50,6% des énoncés étaient uniquement centrés sur le destinataire, 27% et 22,3% ont fait référence uniquement aux panneaux (Exocentré) et 26% et 25% comprenaient des informations utilisant plusieurs références (Multiple).

Une observation plus détaillée des énoncés a permis d'identifier deux groupes de participants : ceux qui n'ont jamais utilisé les panneaux (7 participants) et ceux qui, lorsque les panneaux étaient présents, les ont utilisés au moins une fois (15 participants).

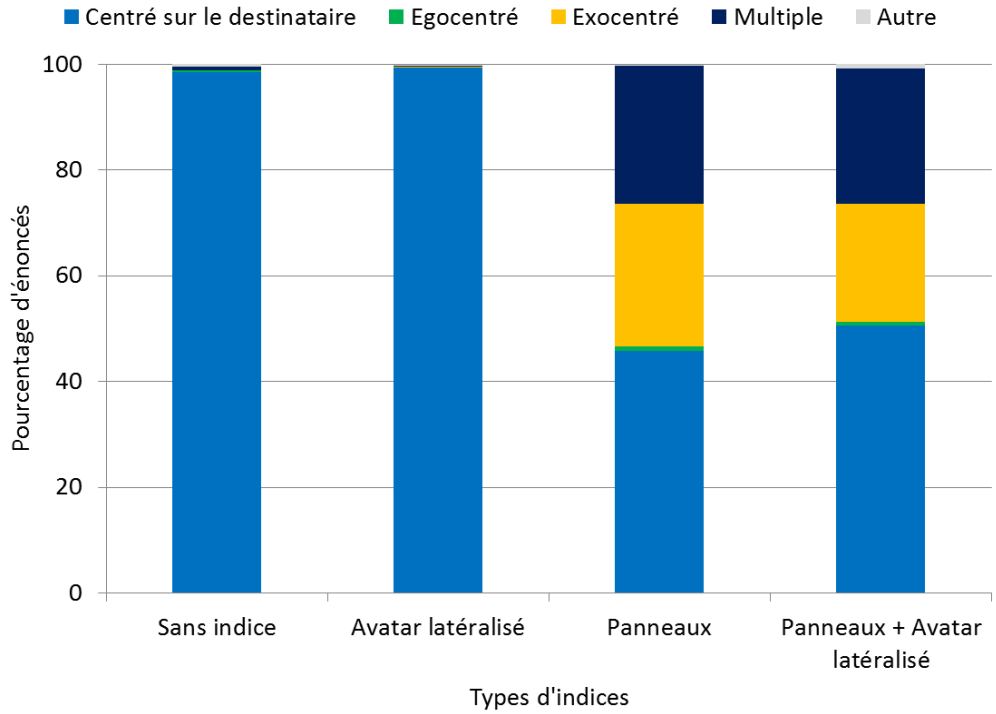


Figure 54 : Pourcentage des cadres de références utilisés pour chaque condition d'indiçage spatial.

Cette observation a suscité plusieurs interrogations :

- l'ordre d'apparition des indices a-t-il un effet sur les cadres de référence utilisés ?
- les panneaux ont-ils tout de même eu un effet (bénéfique ou néfaste) sur les performances et la charge mentale des personnes ayant gardé un cadre de référence centré sur le destinataire ?
- pour les participants ayant utilisé les indices distaux au moins une fois (Groupe 2), même en présence des panneaux, certains énoncés sont toujours centrés sur le destinataire. Ce constat amène la question suivante : est-ce que l'utilisation des panneaux est aléatoire ou est-elle préférée dans certaines configurations ?

Nous avons regardé l'ordre dans lequel les participants ont été confrontés aux différents indices spatiaux. Nous avons ainsi constaté que six des sept participants n'ayant pas utilisé les panneaux ont commencé par un environnement sans panneau (*Sans indice* ou *Avatar latéralisé*). Cependant, cette donnée n'est pas suffisante pour conclure sur un effet général de l'ordre de présentation des indices. Nous avons également regardé si cette population correspondait à tous les participants ayant commencé sans les indices distaux. Il s'avère qu'au total douze participants ont commencé par des conditions sans panneau : donc dans l'échantillon que nous avons, la moitié des participants ayant commencé sans les panneaux ne les ont pas utilisés lorsqu'ils se sont présentés. Ce choix peut s'expliquer par différentes raisons que nous discuterons dans la section suivante.

Etant donné que pour les essais sans les indices distaux (*Sans indice* et *Avatar latéralisé*), plus de 98% des énoncés sont produits dans le cadre de référence centré sur le destinataire, l'étude de la répartition des cadres de référence dans ces deux conditions d'indiçage n'est pas nécessaire. Nous nous intéresserons donc uniquement aux essais pour lesquels les indices distaux étaient présents (*Panneaux* et *Panneaux + Avatar latéralisé*). Pour les deux conditions avec indices distaux, l'analyse des résultats a mis en évidence deux effets : un effet de la position de la cible par rapport au locuteur (χ^2 (28, 2686) = 407, p<.001), et un effet de la position de la cible par rapport à l'avatar (χ^2 (28, 2686) = 257, p<.001). Par ailleurs les résultats montrent qu'il n'y a pas d'effet de l'orientation de l'avatar sur la répartition des cadres de référence (χ^2 (12, 2686) = 3.14, p=.99).

Effets de la position de la cible par rapport au participant sur les cadres de référence

Comme lors de l'étude précédente, nous avons regroupé les données selon la symétrie du plan sagittal du participant pour étudier l'effet de la position de la cible par rapport à lui. Ces résultats sont présentés Figure 55. Plusieurs résultats sont à souligner. D'une part, les seuls énoncés égocentrés ont été produits lorsque la cible est juste en face du participant (E). Les analyses montrent également que les cadres de références utilisés pour les cibles A et C ne sont pas différents (χ^2 (3, 672) = 4.12, p=.2). Le même constat de similitude peut être fait entre les cibles B et D (χ^2 (2, 672) = .02, p=.99). De plus, les participants ont utilisé plus d'énoncés centrés sur le destinataire lorsque la cible est en position A, C et E que dans les positions B et D. De même que les participants ont utilisé plus d'énoncés neutres lorsque la cible est dans les positions B et D que A, C et E. Ces résultats sont repris dans la Table 6.

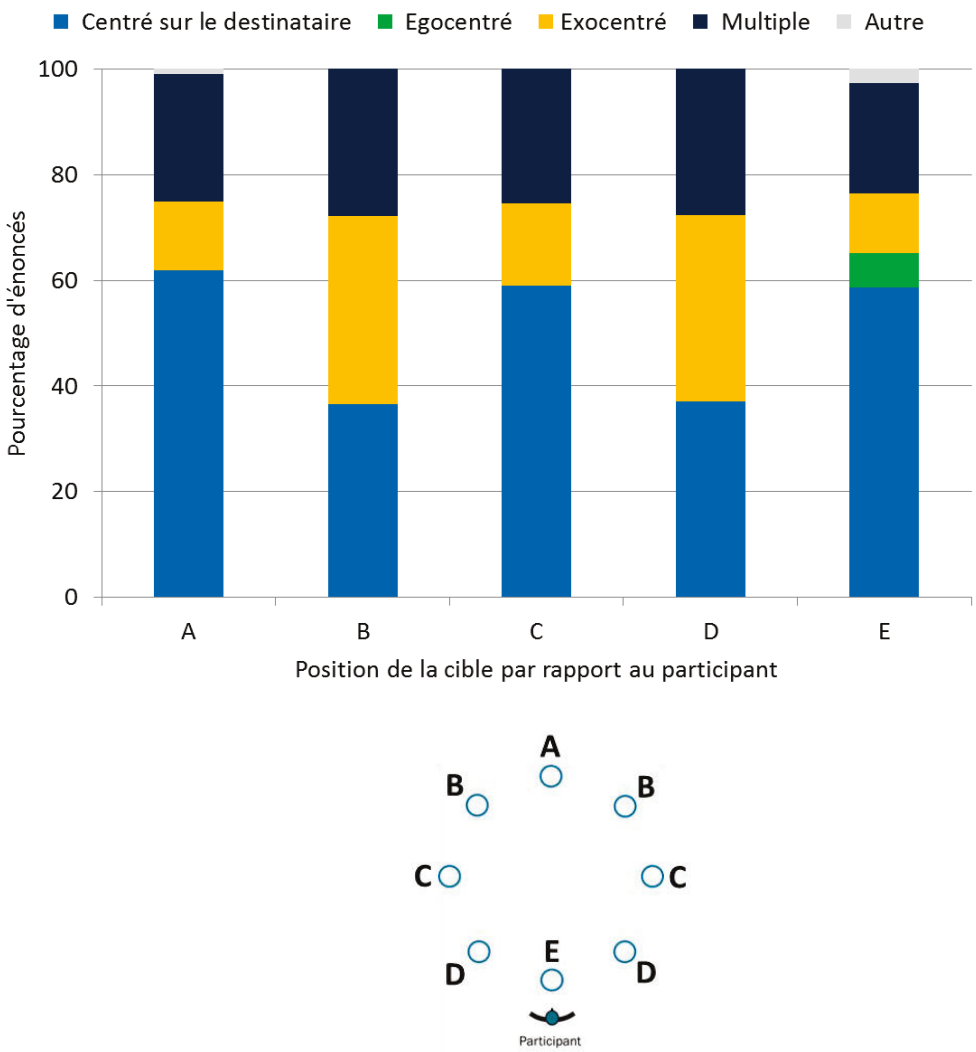


Figure 55 : Pourcentage d'énoncés produits dans chaque cadre de référence selon la position de la cible par rapport au participant (les résultats montrent les conditions *Panneaux* et *Panneaux + Avatar latéralisé* regroupées).

Table 6 :Résultats des tests du Khi-2 pour les énoncés centrés sur le destinataire et exocentrés. ** p<.01.

	Centré sur le destinataire				Exocentré		
	A	C	E		A	C	E
B	43**	33.5**	32.7**		46.1**	35.1**	55.1**
D	41.5**	32.2**	31.3**		44.5**	33.7**	53.4**

Effets de la position de la cible par rapport à l'avatar sur les cadres de référence

Concernant la position de la cible par rapport à l'avatar, l'analyse menée sur les données regroupées a soulevé des différences selon l'alignement de la cible avec les plans intrinsèques de l'avatar (*Panneaux* : χ^2 (12, 1344) = 118 ; p<.001 ; *Panneaux + Avatar latéralisé* : χ^2 (12, 1344) = 138 ; p<.001), comme illustré Figure 56.

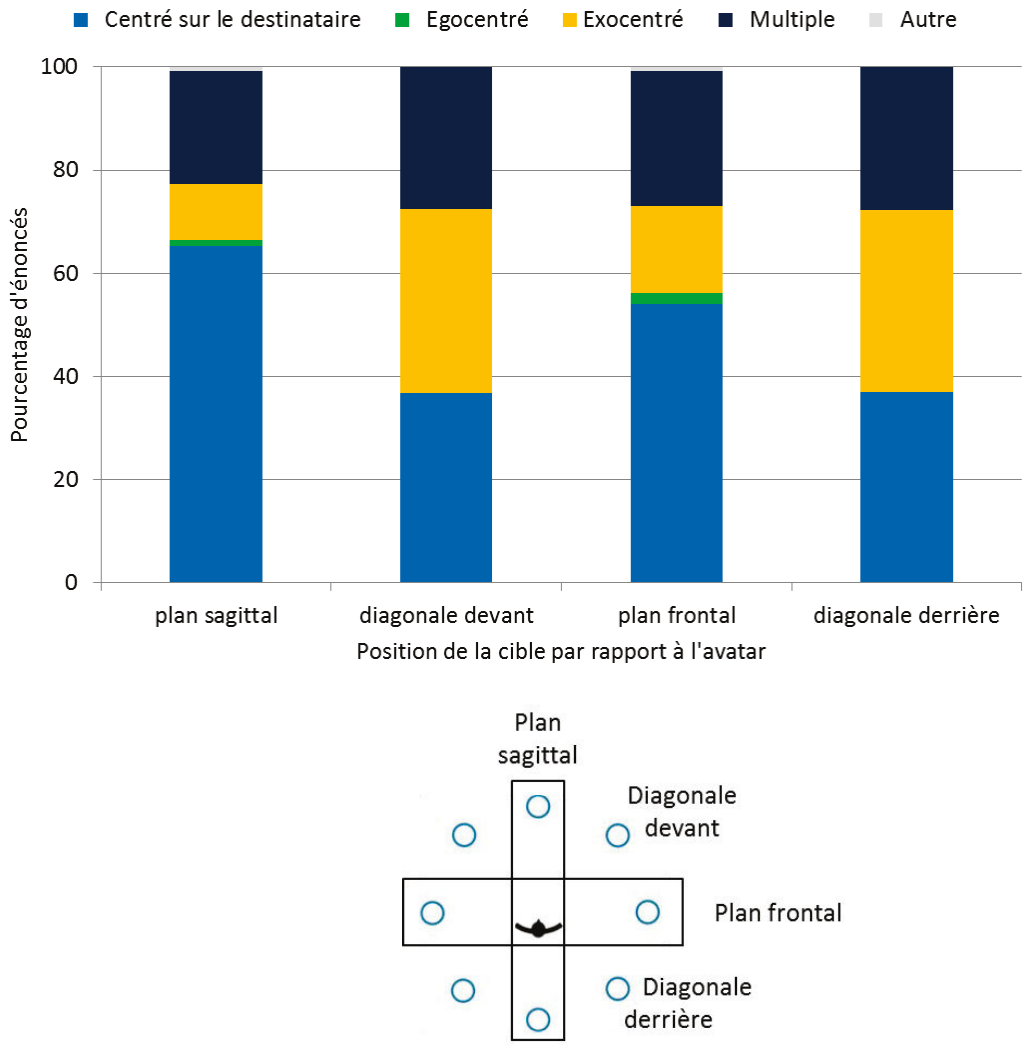


Figure 56 : Pourcentage d'énoncés pour chacun des cadres de référence en fonction de la position de la cible par rapport à l'avatar (conditions *Panneaux* et *Panneaux + Avatar latéralisé*).

Les participants ont utilisé le cadre de référence centré sur le destinataire majoritairement lorsque la cible était alignée avec un des plans intrinsèques (sagittal et frontal) de l'avatar (χ^2 (1, 2686) = 139 ; p<.001). Inversement, ils ont utilisé plus d'énoncés exocentrés lorsque la cible est en diagonale par rapport à l'avatar, que lorsqu'elle est alignée avec l'un des plans intrinsèques (χ^2 (1, 2686) = 167 ; p<.001). Alors que les configurations en diagonale ne présentent pas de différence dans les cadres de référence utilisés (χ^2 (2, 1344) = 0 ; p=.9), les cadres de références utilisés pour décrire les cibles alignées avec les plans intrinsèques de l'avatar diffèrent selon le plan concerné (χ^2 (4, 1344) = 20.1 ; p<.001). En effet, lorsque la cible est alignée avec le plan sagittal de l'avatar, 65% des énoncés sont centrés sur le destinataire, 11% sont produits dans un cadre de référence exocentré et 22% utilisent plusieurs cadres de référence (Multiple). Lorsque la cible est alignée avec le plan frontal de l'avatar, 54% des énoncés sont centrés sur le destinataire, 17% sont exocentrés et 26% utilisent plusieurs cadres de référence (Multiple). Seules les différences entre les énoncés centrés sur le destinataire (χ^2 (1, 1344) = 17.4 ; p<.001) et exocentrés (χ^2 (1, 1344) = 10.4 ; p<.001) sont significatives (Multiple : χ^2 (1, 1344) = 3.2 ; p>.05).

Synthèse des résultats sur les cadres de référence utilisés pour donner les instructions avec différents indices spatiaux

Les résultats ont mis en évidence deux populations : lorsque les indices distaux étaient présents, sept participants ne les ont pas du tout utilisés.

En l'absence des indices distaux, les participants ont utilisé l'avatar comme référence, c'est-à-dire qu'ils ont centré leurs énoncés sur le destinataire, dans 98% des cas.

En présence des indices distaux, les cadres de référence utilisés sont influencés par la position de la cible par rapport au participant et à l'avatar. Ces deux effets sont principalement dus à la position des panneaux et au fait que l'avatar ne fait jamais face aux panneaux. Les cibles situées en face de panneaux de couleur correspondent aux cibles situées en diagonale devant et derrière l'avatar.

Les participants ont davantage utilisé les panneaux lorsque la cible était directement en face de l'un d'eux que lorsqu'elle était située entre deux panneaux. Dans ce dernier cas, les participants ont alors composé entre cadres de référence centré sur le destinataire, multiple, exocentré et égocentré. D'ailleurs, la seule différence entre les cibles situées entre deux panneaux porte sur l'usage du cadre de référence égocentré exclusivement pour la cible directement en face du participant. Le second résultat intéressant est qu'il n'y a pas de différence de répartition des cadres de référence utilisés pour décrire les cibles situées en face des panneaux.

Deux informations sont à retenir sur l'effet de la position de la cible par rapport à l'avatar. La première est qu'il n'y a pas de différence dans les cadres de référence utilisés lorsque la cible est en diagonale devant ou derrière l'avatar. Autrement dit, il n'y a pas de différence lorsque la cible est située en face de l'un des panneaux de couleur. La seconde information importante, est que lorsque la cible se situe entre deux panneaux, le plan dans lequel la cible se situe par rapport à l'avatar est un facteur important. En effet, les participants ont plus souvent utilisé les caractéristiques intrinsèques de l'avatar lorsque la cible était alignée avec le plan sagittal que lorsqu'elle était alignée avec le plan frontal. Dans cette dernière configuration, les participants ont plus fait référence aux panneaux de couleur.

2.2.2 Exigence mentale

L'ANOVA à trois facteurs (4 types d'indices * 4 orientations de l'avatar * 8 positions de la cible par rapport au participant) montre des interactions entre le type d'indices spatiaux et l'orientation de l'avatar, d'une part ($F(9, 189) = 2.8$; $p<0.01$), et entre le type d'indices spatiaux et la position de la cible, d'autre part ($F(21, 441) = 4.3$; $p<0.001$). L'ANOVA ne montre pas d'effet global des indices spatiaux sur l'exigence mentale (un effet global des indices spatiaux ($F(3, 63) =2.3$; $p=.08$, Figure 57).

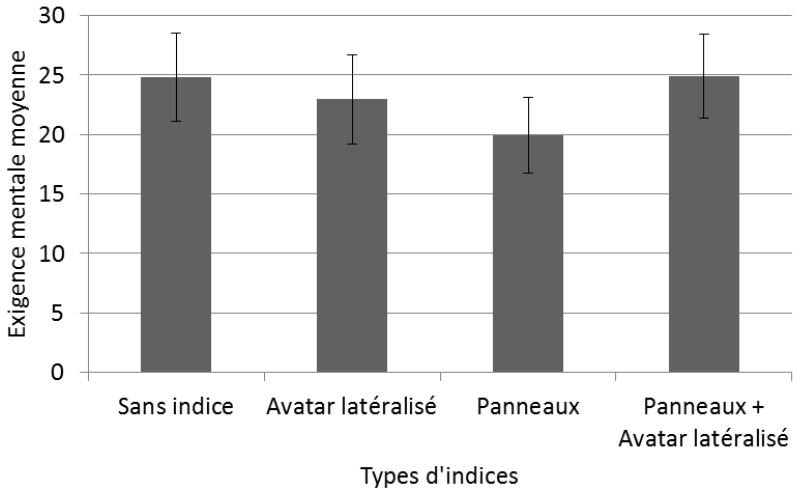


Figure 57 : Exigence mentale moyenne pour chacun condition d'indiçage. * p<.05

Effets de l'interaction entre le type d'indices spatiaux et la position de la cible par rapport au participant sur l'exigence mentale

Comme pour l'étude précédente, nous avons regroupé les données selon l'axe de symétrie du plan sagittal du participant. Nous avons réalisé une nouvelle ANOVA (4 types d'indices spatiaux * 5 positions de la cible par rapport au participant) qui confirme l'interaction entre ces deux variables ($F(12, 252) = 5.1$, $p<.001$, Figure 58). Cette interaction est principalement due au fait que pour les cibles induisant l'exigence mentale la plus faible (A, C et E), les panneaux et l'avatar latéralisé augmentent l'exigence mentale. Inversement, pour les cibles induisant l'exigence mentale la plus forte (B et D), les panneaux seuls diminuent l'exigence mentale. L'ensemble de analyses post-hoc sont présentées dans l'Annexe 14.

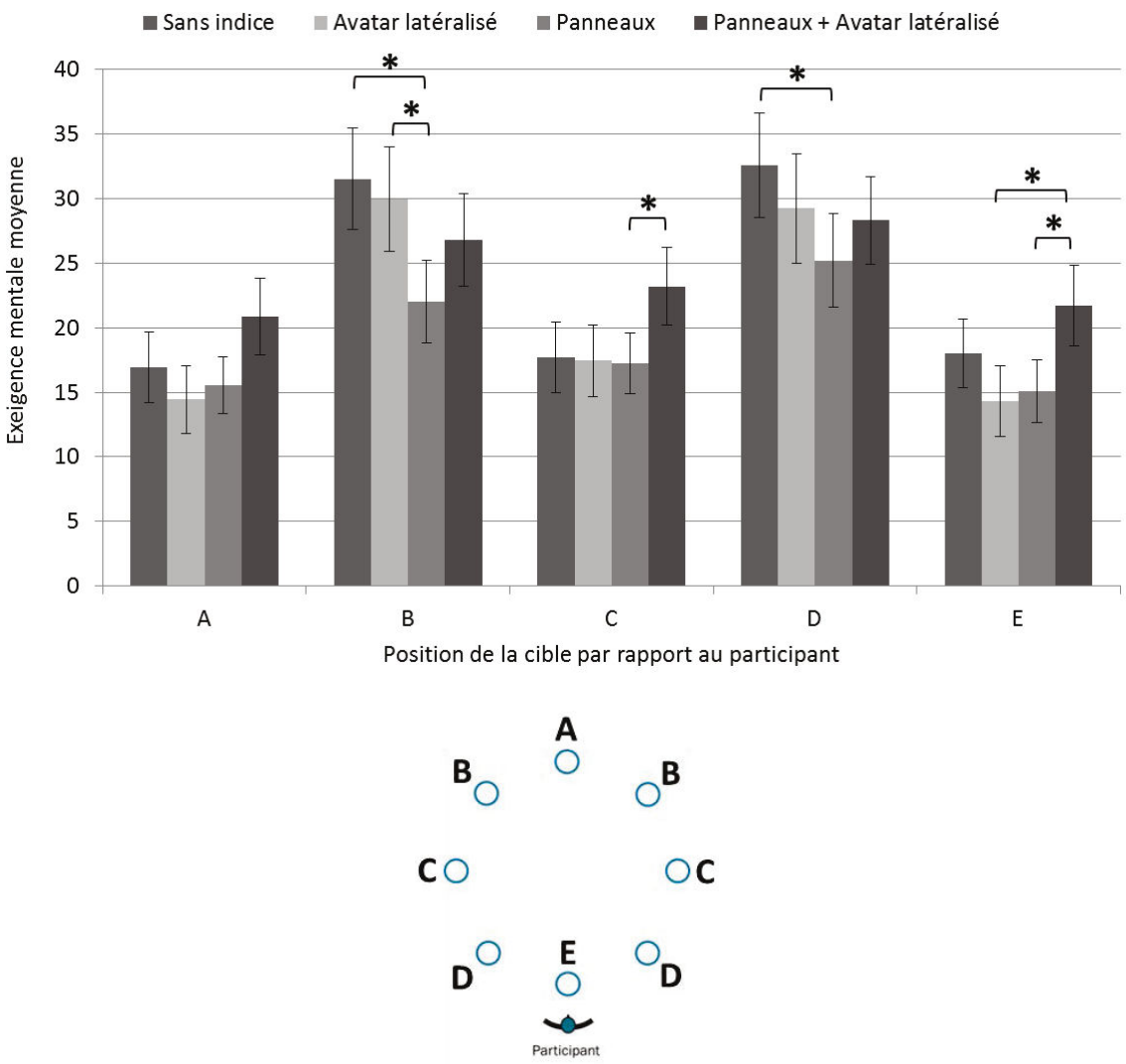


Figure 58 : Exigence mentale moyenne en fonction de la position de la cible par rapport au participant et aux conditions d'indiçage. *p<.05, **p<.01

Effets de l'interaction entre le type d'indices spatiaux et l'orientation de l'avatar sur l'exigence mentale

Les analyses post-hoc de l'interaction entre le type d'indices spatiaux et l'orientation de l'avatar ont montré des différences dans le pattern présenté Figure 57 selon l'orientation de l'avatar. Ainsi, lorsque le participant et l'avatar sont dans la même direction (0°), l'exigence mentale est jugée plus faible lorsque les panneaux sont présents que lorsqu'il n'y a pas d'indice spatial ou lorsque les indices incarnés et distaux sont présents (Figure 59). L'exigence mentale est également plus faible lorsque seul l'avatar latéralisé est présent par rapport aux situations sans indice, ou lorsque les deux types d'indices sont présents. Lorsque l'avatar est tourné, que ce soit à 90° ou 180°, la situation la plus simple par rapport à toutes les autres est lorsque les indices distaux sont présents seuls (*Panneaux*). De plus, lorsque l'avatar est tourné à 90°, la combinaison des deux types d'indices est jugée plus difficile que l'avatar latéralisé seul, et lorsque l'avatar est tourné à 180°, la présence de l'avatar latéralisé diminue l'exigence mentale par rapport aux situations sans indice. Les analyses post-hoc sont présentées dans l'Annexe 15.

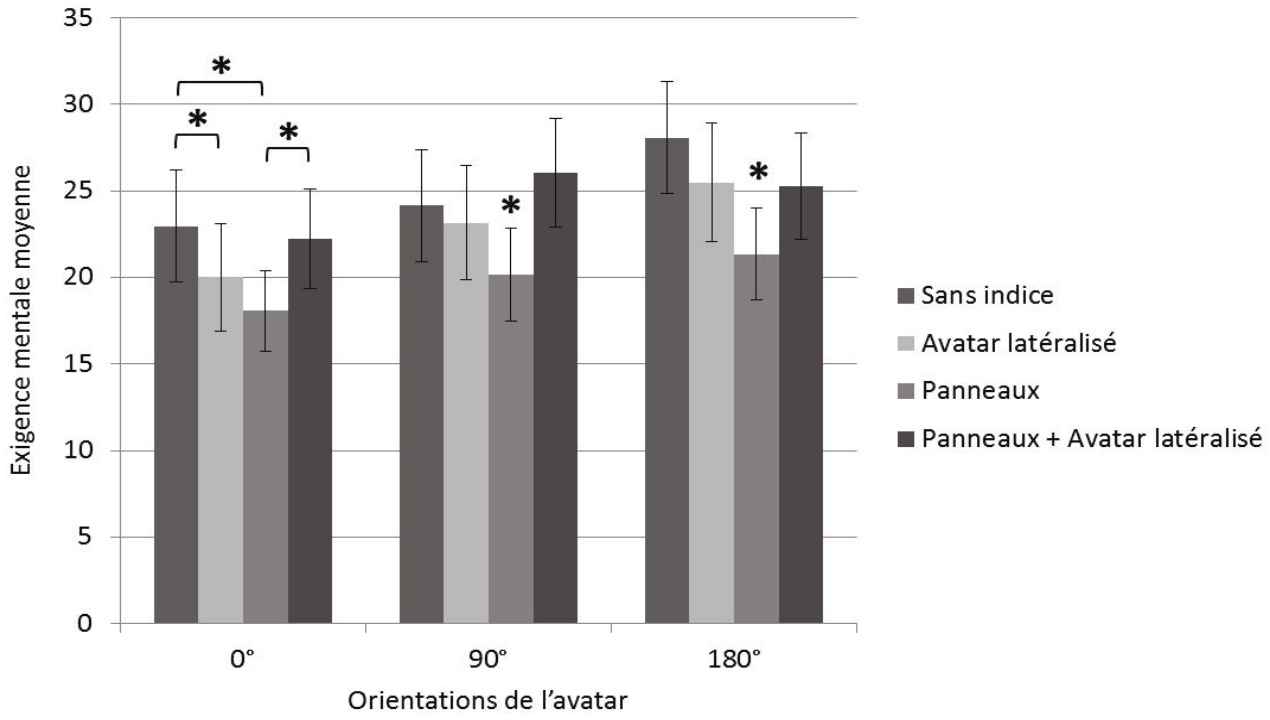


Figure 59 : Exigence mentale moyenne en fonction de l'orientation de l'avatar pour les différentes conditions d'indiçage. * p<.05

Effets de l'interaction entre le type d'indices spatiaux et la position de la cible par rapport à l'avatar sur l'exigence mentale

L'analyse de l'exigence mentale montre également une interaction entre le type d'indices et la position relative de la cible par rapport à l'avatar ($F(9, 126) = 3{,}4$; $p<.001$). L'ensemble des analyses post-hoc est présenté dans l'Annexe 16. Cette interaction est principalement due aux indices distaux (Figure 60). En effet, en leur absence, c'est-à-dire dans les situations *Sans indice* et *Avatar latéralisé*, l'alignement de la cible avec le plan sagittal induit une exigence mentale plus faible que dans les autres configurations. De plus, lorsque la cible est en diagonale derrière l'avatar, l'exigence mentale est plus forte que lorsque la cible est alignée avec le plan frontal. Ces deux effets de la position de la cible par rapport à l'avatar ne se retrouvent pas dans les deux autres conditions (*Panneaux* et *Panneaux + Avatar latéralisé*). Par ailleurs, dans les conditions *Panneaux*, la seule différence porte sur l'alignement de la cible avec le plan sagittal de l'avatar qui induit une exigence mentale plus faible que lorsque la cible est située en diagonale derrière lui. Lorsque les deux types d'indices sont présents (*Panneaux + Avatar latéralisé*), l'alignement de la cible avec le plan sagittal induit une exigence mentale plus faible que lorsqu'elle est alignée avec le plan frontal et lorsqu'elle se situe en diagonale derrière l'avatar.

D'autre part, si l'on regarde les effets des indices selon la position de la cible par rapport à l'avatar, on constate que les indices spatiaux n'ont pas d'effet lorsque la cible est dans le plan sagittal de l'avatar. Lorsque la cible est en diagonale devant, l'exigence mentale est plus faible dans les conditions *Panneaux* que dans les conditions *Sans indice*. Lorsque la cible est sur le plan frontal de l'avatar, l'exigence mentale est plus forte dans les conditions *Panneaux + Avatar latéralisé* par rapport aux conditions *Avatar latéralisé* seul et *Panneaux* seuls. Enfin, lorsque la cible est en diagonale derrière l'avatar, l'exigence mentale est plus faible en présence des indices distaux seuls (*Panneaux*) que dans les conditions *Sans indice* et *Avatar latéralisé*.

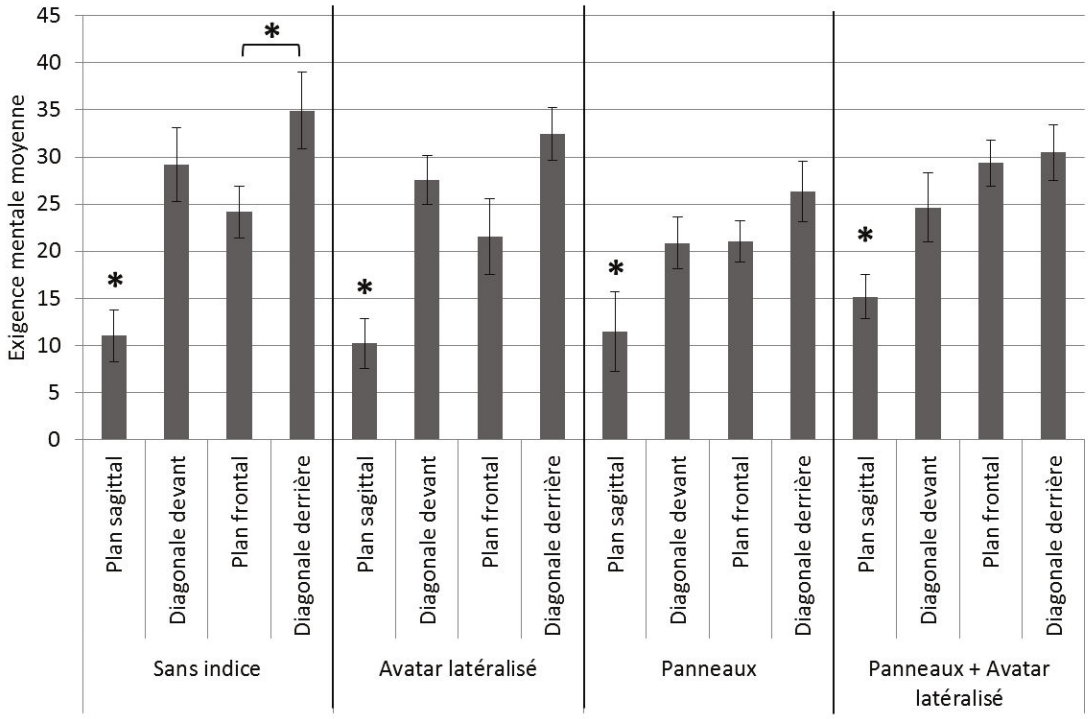


Figure 60 : Exigence mentale moyenne pour les participants ayant utilisé les panneaux, en fonction des différentes conditions d'indiçage et de la position relative de la cible par rapport à l'avatar. * p<.05

Synthèse des résultats sur l'exigence mentale requise pour donner les instructions avec différents indices spatiaux

L'exigence mentale est modulée différemment par la position de la cible par rapport au participant, l'orientation de l'avatar et la position de la cible par rapport à l'avatar, selon les indices présents dans l'EV.

L'interaction des indices avec la position de la cible par rapport au participant montre que les indices proximaux seuls (*Avatar latéralisé*) induisent le même pattern que la condition *Sans indice*, à savoir que les cibles alignées avec les plans intrinsèques de l'avatar induisent une exigence mentale plus faible que les cibles situées en diagonale de l'avatar. Inversement, en présence des indices distaux la position de la cible par rapport au participant n'a qu'un faible effet sur l'exigence mentale, à savoir que dans la condition *Panneaux* la cible la plus proche du participant induit une exigence mentale plus faible que les cibles situées juste à côté de la cible la plus proche du participant.

L'interaction des indices avec l'orientation de l'avatar montre que par rapport aux situations *Sans indice*, les indices proximaux seuls (*Avatar latéralisé*) diminuent l'exigence mentale lorsque l'avatar est dans la même orientation que le participant alors que les indices distaux diminuent l'exigence mentale pour toutes les orientations de l'avatar.

Enfin, l'interaction des indices spatiaux avec la position de la cible par rapport à l'avatar montre que lorsque la cible est alignée avec le plan sagittal de l'avatar, les indices spatiaux n'ont pas d'effet. Pour les autres positions de la cible, les indices distaux seuls (*Panneaux*) diminuent l'exigence mentale lorsque la cible est située en diagonale par rapport à l'avatar.

2.2.3 Temps de préparation des énoncés

Les analyses ne montrent pas d'effet des indices sur le temps de préparation des énoncés.

2.2.4 Temps de production des énoncés

L'ANOVA à trois facteurs (4 types d'indices * 4 orientations de l'avatar * 8 positions de la cible par rapport au participant) montre une interaction entre le type d'indices spatiaux et la position de la cible par rapport au participant sur le temps de production des énoncés ($F(21, 441) = 2.28$; $p<.001$).

Effets de l'interaction entre le type d'indices spatiaux et la position de la cible par rapport au participant sur le temps de production des énoncés

Comme pour l'étude précédente, nous avons regroupé les données selon l'axe de symétrie du plan sagittal du participant. Nous avons réalisé une nouvelle ANOVA (4 types d'indices spatiaux * 5 positions de la cible par rapport au participant) qui confirme l'interaction entre ces deux variables (F(12, 252) = 2.6, p<.01, Figure 61). L'ensemble des analyses post-hoc sont récapitulées dans l'Annexe 17. Tout comme pour l'exigence mentale, cette interaction est principalement due aux indices distaux. En effet, en leur absence (*Sans indice* et *Avatar latéralisé*), les cibles A, C et E induisent un temps de production plus court que les cibles B et D. Cet effet n'est pas du tout retrouvé en présence des indices distaux seuls (*Panneaux*).

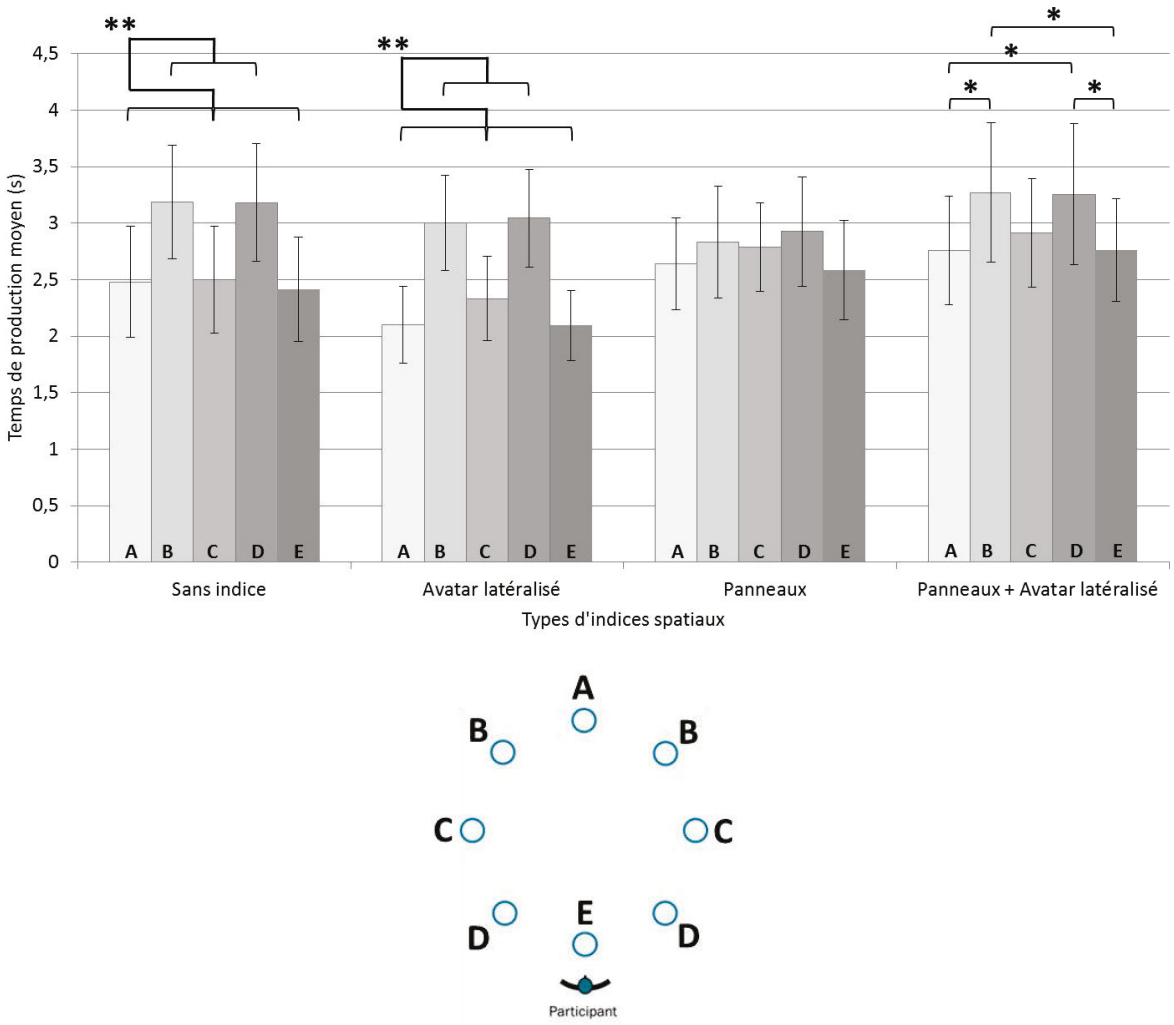


Figure 61 : Temps de production moyen en fonction de la position de la cible par rapport au participant et des indices spatiaux. *p<.05, **p<.01

Effets de l'interaction entre le type d'indices spatiaux et la position de la cible par rapport à l'avatar sur le temps de production des énoncés

L'ANOVA à trois facteurs (4 types d'indices * 4 positions de la cible par rapport au destinataire) montre une interaction entre les indices spatiaux et la position de la cible par rapport au destinataire ($F(9, 126) = 3.37$; $p<.001$). L'ensemble des analyses post-hoc sont présentées dans l'annexe 17. L'interaction est principalement due à des temps de production moyens plus courts en présence de l'avatar latéralisé seul qu'en présence des indices distaux (*Panneaux* et *Panneaux + Avatar latéralisé*), uniquement lorsque la cible est alignée avec un des plans intrinsèques de l'avatar (plan sagittal et plan frontal), comme illustré sur la Figure 62.

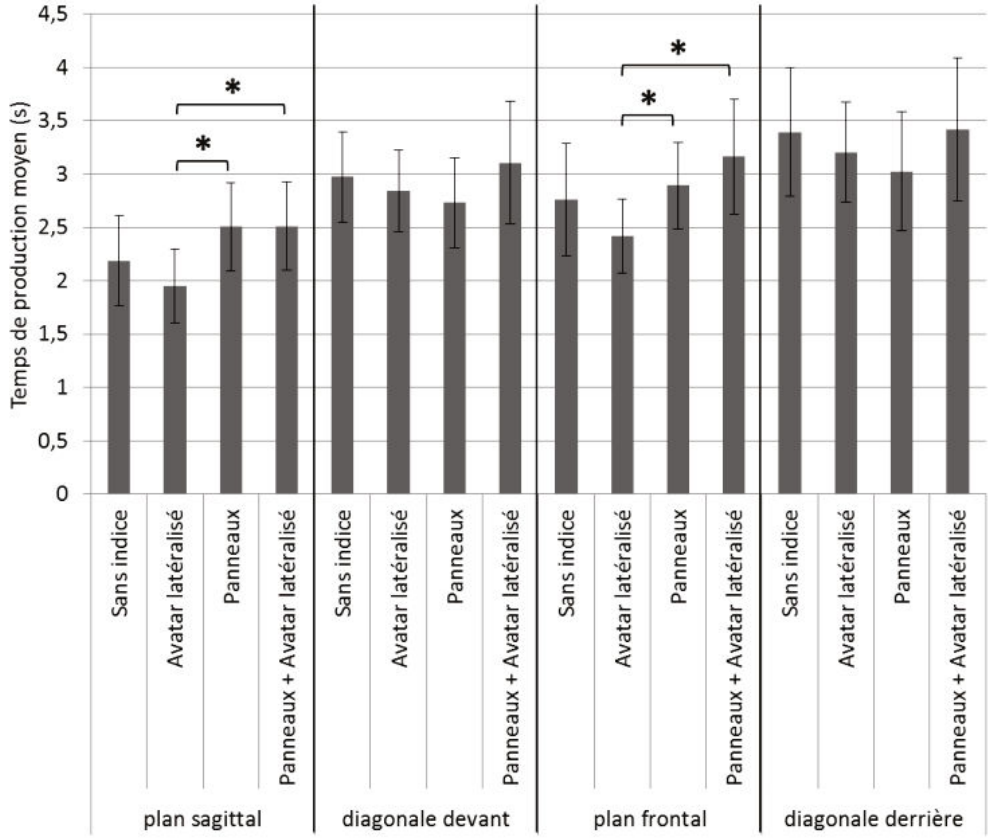


Figure 62 : temps de production moyen pour les participants ayant utilisé les panneaux, en fonction des indices spatiaux et de la position relative de la cible par rapport à l'avatar.* p<.05

Synthèse des résultats sur le temps de production des instructions avec différents indices spatiaux

Sur l'ensemble des essais, le temps de production des énoncés n'est globalement pas affecté par les indices spatiaux mis à disposition des participants. Cependant, les indices spatiaux modulent les temps de production selon la position de la cible à la fois par rapport au participant et à l'avatar. Deux résultats intéressants sont à retenir ici.

Le premier est que le temps de production ne varie pas en fonction de la position de la cible par rapport au participant en présence des indices distaux seuls, contrairement aux autres conditions d'indiçage.

Le second résultat intéressant est que les temps de production des énoncés sont plus courts en présence des indices proximaux seuls qu'en présence des indices distaux seuls uniquement lorsque la cible est dans un des plans intrinsèques de l'avatar (sagittal et frontal).

2.2.5 Corrélation

La comparaison entre les moyennes des variables pour chaque condition a mis en évidence une corrélation entre le temps de production des énoncés et le nombre de mots ($r=.92$, $p<.001$). Pour ne pas surcharger la partie résultats, nous ne présentons pas les données relatives au nombre de mots puisque les résultats sont similaires à ceux présentés pour le temps de production des énoncés.

2.2.6 Questionnaire et autres comparaisons

Les participants ont estimé se sentir présent à 4.8/10 (ET=.5), la présence de leur collaborateur à 5.1/10 (ET=.7) et présent dans la salle réelle à 5.3/10 (ET =.6). Il n'y a pas de corrélation entre les sentiments de présence et de coprésence. Mais une corrélation négative entre le sentiment de présence dans l'environnement virtuel et le sentiment de présence dans la salle réelle ($r=-.54$, $p<.01$).

L'analyse des données du questionnaire a mis en évidence une corrélation négative entre l'âge des participants et leur résultat au test des rotations mentales (MRT-A ®). On constate ainsi que plus l'âge augmente, moins les performances à ce test sont bonnes ($r=-.54$, $p<.05$).

2.3 Synthèse et discussion – Production d'énoncés avec indices spatiaux

Cette étude a complété l'étude précédente en fournissant un grand nombre de résultats intéressants que nous résumons et discutons ici.

Le premier constat est que la présence des indices distaux n'induit pas forcément leur utilisation. En effet, 32% des participants n'ont pas du tout utilisé les indices distaux. Cependant, plusieurs interactions ont pu être constatées entre les indices, la position de la cible (par rapport au participant et à l'avatar), et l'orientation de l'avatar.

Interaction entre le type d'indices spatiaux et la position de la cible par rapport au participant

Les cadres de référence ont mis en évidence pour les conditions *Panneaux* et *Panneaux + Avatar latéralisé*, l'usage du cadre de référence égocentré uniquement lorsque la cible est directement en face du participant. Pour les autres cibles, la différence majeure dans les cadres de référence est la position de la cible par rapport aux indices distaux : lorsque la cible est face à un panneau, le cadre de référence exocentré est utilisé en grande majorité, alors que lorsque la cible est située entre deux panneaux, le cadre de référence centré sur le destinataire est préféré, mais le cadre de référence exocentré est également utilisé.

Concernant l'exigence mentale, les conditions *Sans indice* et *Avatar latéralisé* induisent le même pattern, à savoir que les cibles alignées avec les plans intrinsèques de l'avatar induisent une exigence mentale plus faible que les cibles situées en diagonale de l'avatar. Inversement, en présence des indices distaux la position de la cible par rapport au participant n'a qu'un faible effet sur l'exigence mentale, à savoir que dans la condition *Panneaux* la cible la plus proche du participant induit une exigence mentale plus faible que les cibles situées juste à côté de la cible juste en face du participant.

Concernant les temps de production des énoncés, la position de la cible par rapport au participant n'affecte pas les conditions avec les indices distaux seuls, par rapport aux trois autres conditions d'indiçage.

Interaction entre le type d'indices spatiaux et l'orientation de l'avatar

La seule variable affectée par l'orientation de l'avatar est l'exigence mentale. Cet effet montre que les indices distaux diminuent l'exigence mentale pour toutes les orientations, alors que les indices proximaux n'ont d'effet que lorsque l'avatar est aligné avec le participant.

Interaction entre le type d'indices spatiaux et la position de la cible par rapport à l'avatar

L'analyse des cadres de référence montre que les énoncés diffèrent selon que la cible est face à un panneau ou entre deux panneaux. Le cadre de référence exocentré est préféré lorsque la cible est face à un panneau, sans distinction qu'elle soit en diagonale devant ou derrière l'avatar. Lorsque la cible est entre deux panneaux, c'est-à-dire alignée avec l'un des plans intrinsèques de l'avatar, les participants utilisent plus le cadre de référence centré sur le destinataire, même s'il leur arrive d'utiliser également le cadre de référence exocentré. D'ailleurs, la proportion de cadres de référence centrés sur le destinataire est plus forte lorsque la cible est alignée avec le plan sagittal que lorsqu'elle est alignée avec le plan frontal de l'avatar.

En ce qui concerne l'exigence mentale l'interaction des indices spatiaux avec la position de la cible par rapport à l'avatar montre que lorsque la cible est alignée avec le plan sagittal de l'avatar, les indices spatiaux n'ont pas d'effet. Pour les autres positions de la cible, les indices distaux seuls (*Panneaux*) diminuent l'exigence mentale lorsque la cible est située en diagonale par rapport à l'avatar.

Concernant les temps de production de l'énoncé, il est intéressant de constater que la description des cibles en présence des indices proximaux seuls est plus courte qu'en présence des indices distaux uniquement lorsque la cible est alignée avec un des plans intrinsèques de l'avatar.

2.3.1 Indices spatiaux : aide ou contrainte ?

La première conclusion est que donner des indices spatiaux est bénéfique dans la mesure où il n'y en a pas trop. En effet, les résultats soulignent à plusieurs reprises que la présence simultanée des indices distaux et proximaux annule les effets bénéfiques de l'un ou l'autre des indices présents de manière isolée. Ce phénomène est certainement lié à la surcharge visuelle imposée par les différents indices. Il est donc important de choisir un nombre limité d'indices, selon les besoins de la tâche.

La seconde conclusion est que les indices environnementaux distaux sont, dans plusieurs situations, plus efficaces pour diminuer l'exigence mentale que les indices incarnés. Nous allons discuter les processus sur lesquels s'appuient ces indices pour diminuer l'exigence mentale dans la partie suivante.

2.3.2 Aide apportée par les indices spatiaux : sur quels processus s'appuie-t-elle ?

Les résultats confortent l'hypothèse selon laquelle les indices incarnés et distaux s'appuient sur des étapes différentes du processus de production d'énoncés spatialisés pour faciliter la tâche. Ainsi, les indices incarnés faciliteraient la prise de perspective, alors que les indices environnementaux distaux favoriseraient les codages non centrés sur les interlocuteurs.

La présence d'indices incarnés n'a pas d'effet sur les cadres de référence choisis pour donner les instructions. En effet, les locuteurs ont utilisé uniquement des énoncés centrés sur le destinataire, quelle que soit la position de la cible. Il semble donc que l'aide apportée par ce type d'indices s'inscrive dans les processus cognitifs permettant la prise de perspective. De plus, les résultats montrent que l'avatar latéralisé diminue l'exigence mentale lorsque les interlocuteurs sont alignés (0°). L'avantage donné par l'avatar latéralisé n'est donc que de faible importance.

Contrairement aux indices incarnés, les indices environnementaux distaux affectent les cadres de références utilisés pour localiser la cible. En effet, les panneaux offrent de nouvelles références pour localiser la cible, que certains participants ont utilisées pour donner leurs énoncés. Par ailleurs, les résultats confirment l'hypothèse selon laquelle les indices distaux diminueraient l'exigence mentale pour les cibles situées en diagonale derrière l'avatar. Nous attribuons ce résultat au fait que les participants n'ont pas besoin d'opérer de transformation mentale pour utiliser les panneaux comme référence, contrairement au positionnement par rapport à l'avatar. Cela est notamment confirmé par l'absence d'effet de l'orientation de l'avatar sur l'exigence mentale lorsque les panneaux sont présents. Ainsi, l'usage de cadres de référence exocentrés, nommés neutres par Schober (1995), serait effectivement plus simple et éviterait les transformations mentales coûteuses.

Cependant, il serait faux de dire que les participants n'ont utilisé qu'une représentation allocentrée de la situation en présence des panneaux. En effet, bien qu'il soit très faible, un effet de la position de la cible par rapport à l'avatar persiste pour la situation *Panneaux*. Cet effet peut être mis en relation avec le fait que, malgré la présence des panneaux, les participants ont tout de même, dans certaines situations, utilisé le destinataire comme référence. Cette utilisation d'énoncés centrés sur le destinataire semble s'appuyer fortement sur la prédominance des plans intrinsèques, et particulièrement la prédominance du plan sagittal. Ainsi, les résultats de cette étude montrent que les panneaux offrent bien de nouvelles références que les locuteurs peuvent utiliser pour décrire la position de la cible, mais ces résultats montrent également que les panneaux n'enlèvent pas le fait que l'avatar est également une référence possible. Cette référence possède des directions prédominantes, préférées parfois aux indices distaux. L'étude ne répond d'ailleurs pas à la question de savoir si les plans intrinsèques de l'avatar

prédominent sur les indices distaux puisqu'ils n'ont pas réellement été en compétition. Cependant, cette étude permet de montrer que les indices distaux facilitent la tâche lorsque la position de la cible est difficile à décrire par rapport à l'avatar, et même lorsqu'elle est dans le plan frontal de l'avatar.

Par rapport aux processus cognitifs, ces résultats montrent que plusieurs codages peuvent coexister, privilégiant les directions organisatrices prédominantes. Par ailleurs, le fait que la scène soit décrite dans des cadres de référence différents selon les conditions d'indiçage n'affecte par le temps de préparation des énoncés. Ce résultat suggère que le codage de la scène reste le même : la cible serait codée par rapport au locuteur, et à l'avatar, même en présence des indices distaux.

Ainsi, le principal résultat de cette étude est que la présence d'indices spatiaux n'affecte pas les situations déjà identifiées comme étant les plus simples : lorsque la cible est alignée avec le plan sagittal du destinataire. Ainsi, l'avatar est une référence dont l'axe avant/arrière organise l'espace au niveau perceptif et langagier (Kessler & Rutherford, 2010). Pour les autres situations, les panneaux évitent les mécanismes coûteux de prise de perspective, en s'appuyant sur une représentation allocentrée de l'environnement.

Lors de ces deux études sur la production d'énoncés spatialisés, nous avons pu constater que les participants usaient de cadres de référence variés pour partager l'information spatiale avec leur collaborateur. Dans l'étude qui suit, l'objectif est d'étudier l'effet de ces cadres de référence sur la charge mentale liée à la compréhension de ces énoncés.

3. Etude 3 : Compréhension d'énoncés spatialisés

Comme nous l'avons vu dans les deux études précédentes, la production d'instructions spatialisées dans un environnement virtuel peut être modulée par la position de la cible par rapport à l'interlocuteur, mais également par des indices présents dans l'environnement virtuel. Ainsi, nous avons pu mettre en évidence que la situation demandant l'effort cognitif le plus important pour le locuteur est lorsque la cible se situe en diagonale derrière l'interlocuteur. Cette difficulté peut être amoindrie en présence d'indices distaux, tels que des panneaux de couleurs. Nous avons également constaté que les locuteurs utilisaient différents cadres de référence pour produire leurs énoncés.

Dans cette troisième étude nous étudions la communication spatialisée du point de vue de celui à qui est destinée l'instruction et qui doit l'utiliser pour trouver la cible. Les objectifs de cette étude sont d'évaluer la charge mentale de différents types d'énoncés et d'identifier les configurations les plus difficiles pour la compréhension des énoncés par le destinataire.

3.1 Matériels et méthodes

3.1.1 Participants

Seize personnes ont participé à cette étude : cinq femmes et onze hommes. Ils étaient âgés de 18 à 59 ans, dont la moyenne d'âge était de 24,2 ans.

3.1.2 Matériels

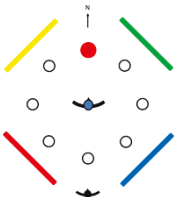
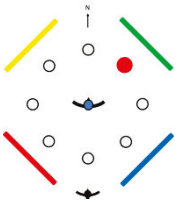
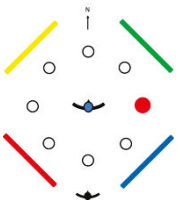
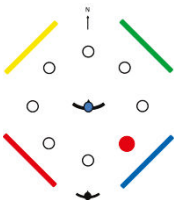
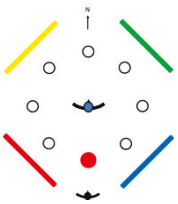
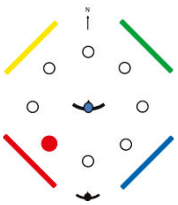
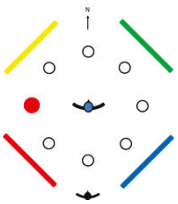
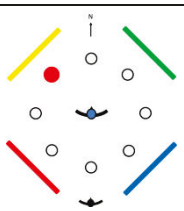
L'environnement est le même que pour la condition *Panneaux* de l'étude précédente. Le participant incarne l'avatar qui est situé au milieu de la pièce.

A la fin des tests, les participants ont répondu à un questionnaire, dont les questions spécifiques sont présentées dans l'annexe 20. Ces questions portent sur la difficulté ressentie par rapport aux types de consignes données et si le participant aurait souhaité que son collaborateur lui donne l'information autrement. Tout comme pour les études précédentes, les participants ont passé des tests d'aptitudes spatiales : Santa Barbara Sense of Direction Scale et le test des rotations mentales MRT-A ©. Ces tests sont présentés dans les annexes 6 et 7. Pour compléter les données, nous avons ajouté un test d'orientation spatiale : short object perspective test (De Beni et al 2014). Ce test est basé sur les images présentées dans la partie bibliographique (p. 42, Figure 16). La version courte comprend une page d'entrainement et douze pages de test. Le score total est calculé à partir de la moyenne des scores obtenus pour chaque test. Un score étant le degré d'écart entre la bonne réponse et la réponse donnée.

3.1.3 Procédure

Nous avons utilisé le même scénario que pour les deux études précédentes mais le participant est cette fois-ci le destinataire des énoncés. Les consignes données au participant sont présentées dans l'annexe 21. Le participant doit sélectionner la cible d'après l'énoncé qui lui est donné. Nous avons utilisé des énoncés utilisant quatre cadres de références différents, résumés Table 7. Deux cadres de référence sont centrés sur le destinataire, l'un utilise uniquement les caractéristiques intrinsèques du destinataire (**Intrinsèque**), l'autre utilise une combinaison de caractéristiques intrinsèques et extrinsèques. Nous avons choisi d'utiliser les degrés (0°, 45°, 135°, 180°) associés aux directions droite et gauche (**Degrés**). Nous avons également utilisé deux cadres de référence exocentrés, l'un utilisant les panneaux de couleur (*Panneaux*), l'autre utilisant les directions cardinales réelles (*Points cardinaux*). Pour que les participants puissent comprendre les énoncés *Points cardinaux*, nous avons précisé que le nord se situait en face de leur collaborateur, entre les panneaux jaune et vert.

Table 7 : Enoncés utilisés dans chaque cadre de référence pour décrire les huit cibles

Intrinsèque	Degrés	Points cardinaux	Panneaux	Exemple, si le participant fait face au nord
La bouteille piégée est sur la table…				
directement devant toi	*à zéro degré*	*au nord*	*entre les panneaux jaune et vert*	
devant toi sur la droite	*à 45 degrés à droite*	*au nord-est*	*à côté du panneau vert*	
directement sur ta droite	*à 90 degrés à droite*	*à l'est*	*entre les panneaux vert et bleu*	
derrière toi sur la droite	*à 135 degrés à droite*	*au sud-est*	*à côté du panneau bleu*	
directement derrière toi	*à 180 degrés*	*au sud*	*entre les panneaux bleu et rouge*	
derrière toi sur la gauche	*à 135 degrés à gauche*	*au sud-ouest*	*à côté du panneau rouge*	
directement sur ta gauche	*à 90 degrés à gauche*	*à l'ouest*	*entre les panneaux rouge et jaune*	
devant toi sur la gauche	*à 45 degrés à gauche*	*au nord-ouest*	*à côté du panneau jaune*	

Le plan expérimental de cette étude comprend trois facteurs : le cadre de référence des énoncés {quatre modalités} * l'orientation du participant {cinq modalités Figure 63} * la position de la cible {huit modalités}, ce qui donne 160 conditions expérimentales, chacune répétée deux fois. A la deuxième répétition, le participant devait évaluer l'exigence mentale de la situation (de la même manière que dans les études précédentes).

Les essais ont été répartis de la manière suivante : toutes les conditions expérimentales pour un même cadre de référence et une même orientation étaient présentées successivement et de manière aléatoire. Le contrebalancement des orientations et des cadres de référence a été contrôlé entre les participants. Nous avons choisi de procéder de cette manière pour ne pas imposer aux participants des changements d'orientation trop rapides, pour éviter les problèmes du mal des simulateurs. Pour permettre aux participants de se familiariser avec l'environnement, les participants ont pu naviguer librement à chaque changement de cadre de référence. Avant les cinq essais d'entrainement présents au début de chaque cadre de référence, l'expérimentateur demandait au participant d'indiquer avec la main la position du collaborateur, du nord et de chacun des panneaux. Ce questionnement n'était pas répété à chaque changement d'orientation, mais les participants ont pris d'eux-mêmes quelques instants pour se repérer dans l'espace. Pendant l'expérience, si le participant se trompait dans les touches, l'essai était annulé et reporté à la fin de la série.

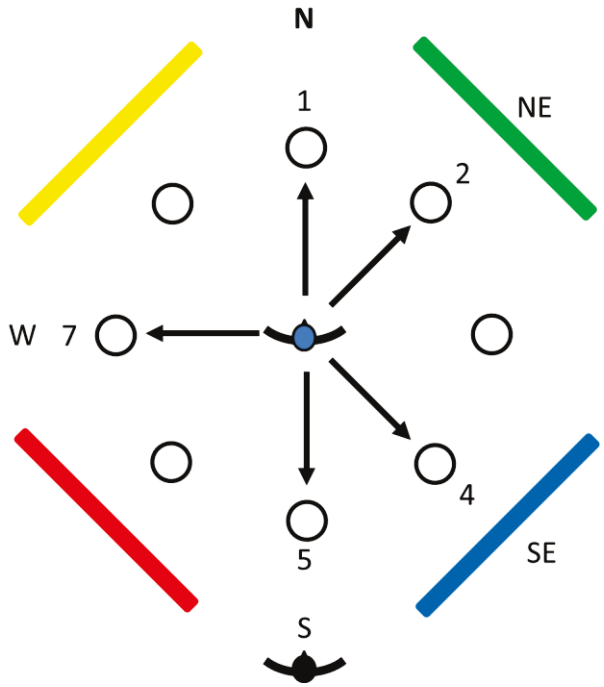


Figure 63 : Cinq orientations prises par le participant durant l'étude 3 - Compréhension d'énoncés spatialisés.

Afin de limiter l'effet de la durée de l'énoncé, et de comparer uniquement la sélection de la cible, nous avons procédé en deux étapes. Les participants devaient dans un premier temps presser la touche 5 du pavé numérique une fois qu'ils avaient compris de quelle cible il s'agissait. Ensuite, ils devaient sélectionner la cible en pressant un chiffre du pavé numérique, comme illustré sur la Figure 64. Comme le 5 se situe au milieu du pavé numérique, la distance est à peu près la même pour chacune des huit autres touches. A chaque fois que la cible 1 a été sélectionnée (touche 8 du pavé numérique), la même commande motrice a été générée. Si des différences apparaissent, elles ne seront pas liées au

mouvement pour aller appuyer sur la touche, mais uniquement aux processus liés à la compréhension de l'énoncé.



Figure 64 : Utilisation du pavé numérique pour sélectionner la cible. La touche 8 correspond toujours à la cible située en face au participant, codée 1 par la suite.

3.1.4 Traitements statistiques

Pour les temps de compréhension et de sélection de la cible, les valeurs des deux essais ont été moyennées. Cette répétition n'entre donc pas dans le plan d'analyse. Pour ces deux variables et l'exigence mentale, les données ont été comparées par des analyses de variance à mesures répétées (ANOVA), dont les facteurs seront précisés lors de la présentation des résultats. Lorsque l'analyse de variance se révélait significative, des analyses post-hoc (Tukey HSD) ont été réalisées. Le nombre d'erreurs par cadre de référence, ont été comparé grâce à des tests du Khi2. Pour tous les tests, le niveau de significativité utilisé était $p<.05$.

3.2 Résultats - Compréhension d'énoncés spatialisés

3.2.1 Exigence mentale

L'analyse de variance à mesures répétées à trois facteurs (ANOVA) « 4 cadres de référence » * « 5 orientations du participant » * « 8 positions de la cible par rapport au participant » montre un effet du cadre de référence ($F(3, 45)=21,28$, $p<.001$), une interaction entre le cadre de référence et l'orientation du participant ($F(12, 180)=4,62$, $p<.001$), une interaction entre le cadre de référence et la position de la cible ($F(21, 315)=4,3$, $p<0,001$) et une interaction entre les trois facteurs : cadre de référence x orientation x cible ($F(84, 1260)=4,25$, $p<.001$).

Effet global des cadres de référence sur l'exigence mentale

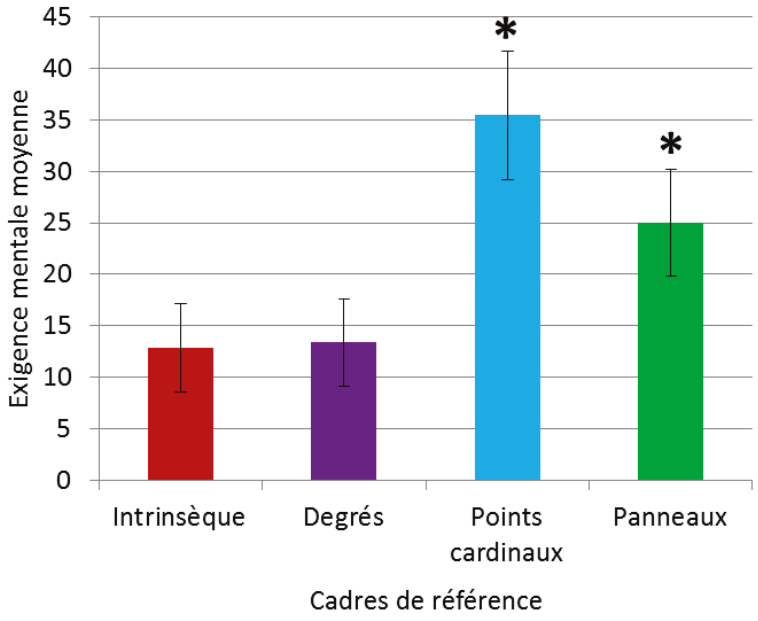


Figure 65 : Exigence mentale moyenne pour les quatre types de cadres de référence. * p<0,05

De manière générale, pour l'ensemble des situations, les énoncés centrés sur le participant (*Intrinsèque* et *Degrés*) engendrent une exigence mentale significativement plus faible que les énoncés exocentrés (*Points cardinaux* et *Panneaux*). Aucune différence n'est observée entre les deux types d'énoncés centrés sur le destinataire. Par contre, les énoncés utilisant les panneaux produisent une exigence mentale plus faible que les énoncés utilisant les points cardinaux (p=0,025). Cet effet global des cadres de référence utilisés pour donner les instructions sur l'exigence mentale est illustré Figure 65.

Effets de l'interaction entre les cadres de référence et l'orientation du participant sur l'exigence mentale

Cependant, ces différences varient selon l'orientation du participant. L'ensemble des analyses post-hoc est présenté dans l'annexe 22. Cette interaction est principalement due à l'effet de l'orientation du participant sur les énoncés exocentrés utilisant les directions cardinales (*Points cardinaux*). En effet, lorsque le participant fait face au nord, il n'y a plus de différence significative entre les énoncés utilisant les points cardinaux et les autres cadres de référence (Figure 66). Seule persiste une différence entre les énoncés utilisant les panneaux et les énoncés centrés sur l'interlocuteur. Ce résultat est présenté Figure 66. C'est la seule orientation pour laquelle survient une différence avec le schéma global présenté Figure 65. Par ailleurs, pour les énoncés utilisant les points cardinaux, l'orientation face au nord est plus simple que toutes les autres orientations ($p<0,001$ pour toutes les comparaisons).

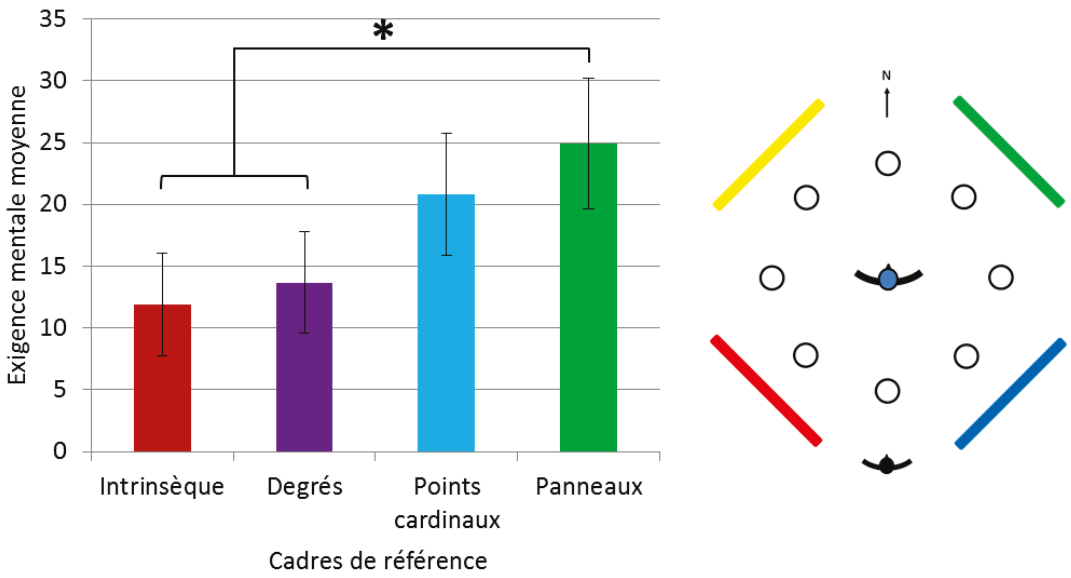


Figure 66 : Exigence mentale moyenne pour les quatre cadres de référence lorsque le participant fait face au nord. * p<.05

Effets de l'interaction entre les cadres de référence et la position de la cible sur l'exigence mentale

L'exigence mentale est également affectée différemment selon le cadre de référence de l'énoncé et la cible sélectionnée (Figure 67). L'ensemble des analyses post-hoc est présenté dans l'annexe 23. Aucune différence n'est mise en évidence en ce qui concerne les comparaisons entre les deux cadres de référence centrés sur le destinataire. Des différences apparaissent entre les énoncés centrés sur le destinataire et les énoncés exocentrés. Les énoncés *Points cardinaux* induisent une exigence mentale plus forte que les énoncés centrés sur le participant (*Intrinsèques* et *Degrés*) pour l'ensemble des huit cibles. En ce qui concerne les énoncés *Panneaux*, l'exigence mentale est plus forte que les énoncés centrés sur le destinataire pour les cibles 3, 4, 5, 6 et 7, c'est-à-dire les cibles situées sur le côté et derrière le participant. Il n'y a donc pas de différence d'exigence mentale entre les énoncés centrés sur le destinataire et *Panneaux* lorsque les cibles sont devant (1) et en diagonale devant (cibles 8 et 2) le participant. Les analyses ont également mis en évidence des différences entre les deux types d'énoncés exocentrés. Les énoncés *Points cardinaux* induisent une exigence mentale plus forte que les énoncés *Panneaux* pour les cibles 7, 8, 1, 2 et 3, c'est-à-dire les cibles situées sur le côté et devant le participant.

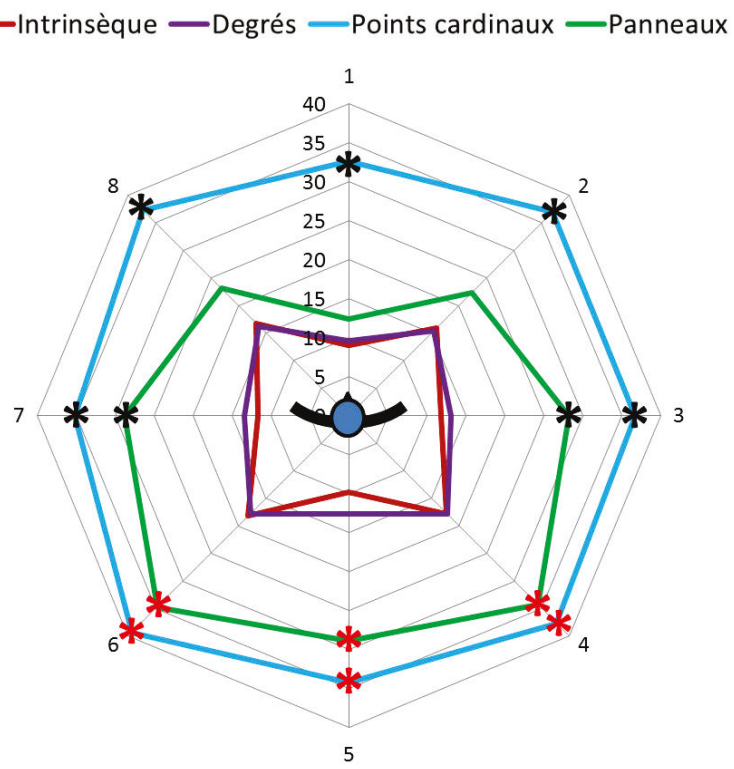


Figure 67 : Exigence mentale moyenne pour chacune des huit réponses attendues pour les quatre cadres de référence. * différence significative avec les trois autres cadres de référence, * différence significative avec les cadres de référence centrés sur l'interlocuteur.

De plus, les analyses post-hoc ont également mis en évidence des différences entre les cibles au sein de chaque catégorie d'énoncés. Pour les énoncés *Intrinsèques*, les cibles 4 et 6 (« derrière toi sur la droite/gauche ») engendrent une exigence mentale significativement supérieure aux cibles 1 (« directement devant toi ») et 5 (« directement derrière toi). Pour les énoncés *Degrés*, les cibles 4 et 6 (« à cent trente-cinq degrés à droite/gauche ») engendrent une exigence mentale significativement supérieure à la cible 1 (« à zéro degrés »). Pour les énoncés *Points cardinaux*, il n'y a pas de différence globale entre les différentes cibles. Enfin, pour les énoncés *Panneaux*, la cible située juste devant le participant (1) induit une exigence mentale inférieure à toutes les autres cibles. De plus, les cibles situées en diagonale devant le participant (2 et 8) engendrent également une exigence mentale plus faible que les cibles situées en diagonale derrière le participant (4 et 6).

Effets de l'interaction entre les cadres de référence, l'orientation du participant et la position de la cible sur l'exigence mentale

Les résultats présentés ci-avant sont à moduler en fonction de l'orientation du participant pour les énoncés exocentrés. En effet, pour les énoncés centrés sur le destinataire (*Intrinsèque* et *Degrés*), l'orientation du participant n'a pas d'effet sur l'exigence mentale induite par chacune des huit cibles. En ce qui concerne les énoncés *Points cardinaux*, les cibles nord et sud induisent une exigence mentale plus faible que d'autres cibles lorsque le participant fait face au Nord Est (NE) et au Sud (S). En ce qui concerne les énoncés *Panneaux*, les effets sont différents selon que le participant fait face à un panneau ou s'il regarde entre deux panneaux. De manière générale, la ou les cibles présentes dans le champ de vision du participant induisent une exigence mentale plus faible que les autres cibles. L'ensemble des analyses post-hoc est présenté dans l'annexe 24, ainsi qu'une figure récapitulative.

Synthèse des résultats sur l'exigence mentale requise pour comprendre les instructions

L'exigence mentale est modulée par plusieurs facteurs qui interagissent.

Globalement, les énoncés centrés sur le destinataire induisent une exigence mentale plus faible que les énoncés exocentrés. De plus, les énoncés *Points cardinaux* induisent l'exigence mentale la plus forte.

Cependant cet effet ne se vérifie pas lorsque le participant est orienté vers le nord. Dans cette orientation, le participant est aligné avec la direction organisatrice principale utilisée dans le cadre de référence *Points cardinaux*. Les résultats montrent que dans cette orientation, l'exigence mentale induite par les énoncés *Points cardinaux* n'est pas différente de l'exigence mentale induite par les autres cadres de référence.

En ce qui concerne les énoncés exocentrés utilisant les indices distaux (*Panneaux)*, l'exigence mentale est modulée par l'accessibilité visuelle des indices utilisés. Lorsque le participant a visuellement accès à la référence utilisée, l'exigence mentale est plus faible que lorsqu'il n'y a pas accès.

Enfin, un dernier résultat intéressant est à souligner ici. L'exigence mentale pour les énoncés centrés sur le destinataire varie selon la position de la cible par rapport au participant. Les cibles situées en diagonale derrière induisent une exigence mentale plus forte que les deux cibles alignées avec le plan sagittal du participant pour les énoncés *Intrinsèques*. Cette différence est vraie uniquement pour la cible située juste devant le participant pour les énoncés *Degrés*.

3.2.2 Temps de compréhension

L'ANOVA à trois facteurs (4 cadres de référence * 5 orientations du participant * 8 positions de la cible par rapport au participant) montre un effet du cadre de référence ($F(3, 45)=25.1$, $p<.001$), une interaction entre le cadre de référence et l'orientation du participant ($F(12, 180)=3.94$, $p<.001$), une interaction entre le cadre de référence et la position de la cible ($F(21, 315)=3$, $p<.001$) et une interaction entre les trois facteurs : cadre de référence x orientation x cible ($F(84, 1260)=1.72$, $p<.001$).

Effet global des cadres de référence sur le temps de compréhension des énoncés

Les analyses post-hoc montrent des différences significatives entre les cadres de référence centrés sur le destinataire (*Intrinsèques* et *Degrés*) et les cadres de référence exocentrés (*Points cardinaux* et *Panneaux*). Ce résultat est présenté Figure 68. Aucune autre différence significative n'apparait.

Effets de l'interaction entre les cadres de référence et l'orientation du participant sur le temps de compréhension des énoncés

L'ensemble des analyses post-hoc est présenté dans l'annexe 25. L'interaction est due aux variations des temps de compréhension des énoncés exocentrés *Points cardinaux* selon l'orientation du participant. En effet, lorsque le participant fait face au nord, le temps de compréhension des énoncés *Points cardinaux* ne diffère pas des temps de compréhension des énoncés centrés sur le destinataire. Ce résultat est représenté Figure 69.

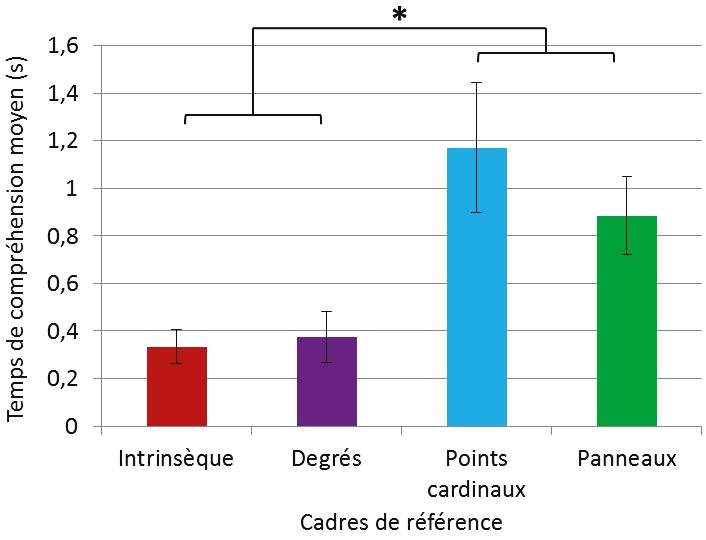


Figure 68 : Temps de compréhension moyens en fonction des cadres de référence utilisés pour localiser la cible. * p<.05

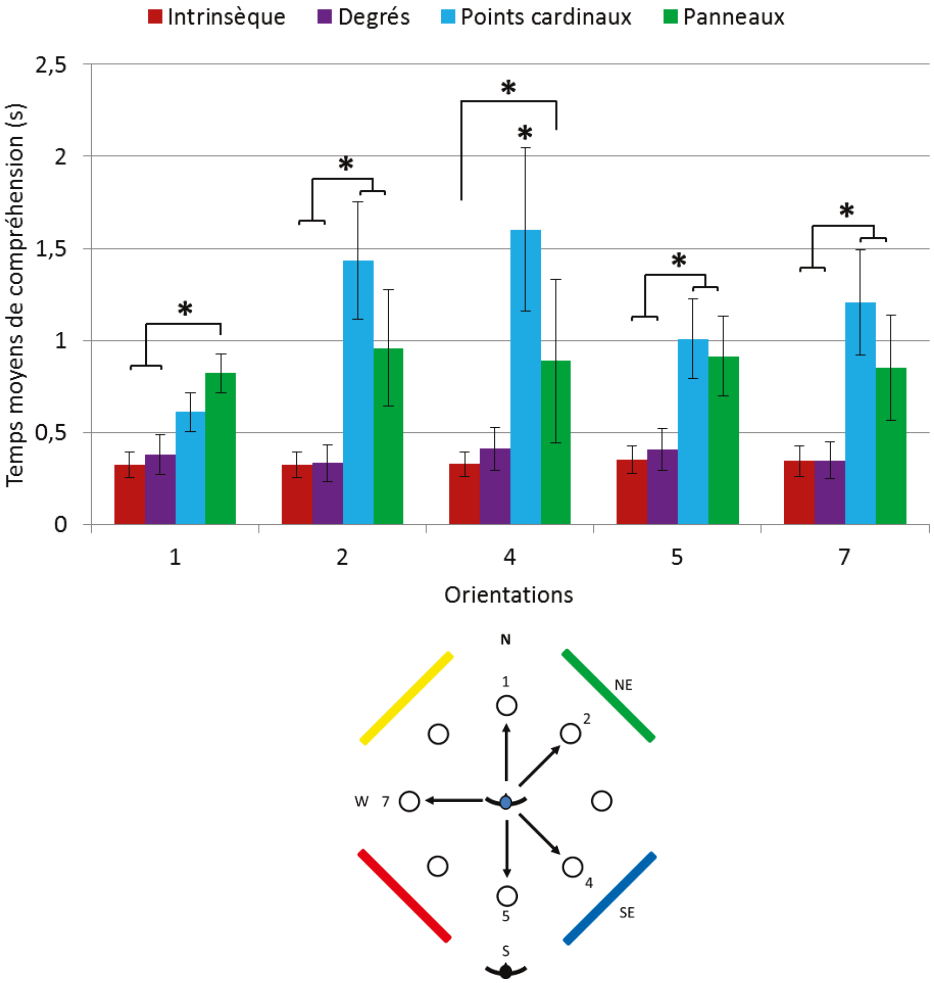


Figure 69 : Temps de compréhension moyens en fonction des cadres de référence utilisés pour localiser la cible et l'orientation du participant. * p<.05 L'interaction est due aux variations de temps de compréhension pour les énoncés *Points cardinaux*, particulièrement court pour l'orientation 1 (N), et particulièrement long pour l'orientation 4 (SE).

Effets de l'interaction entre les cadres de référence et la position de la cible sur le temps de compréhension des énoncés

L'ensemble des analyses post-hoc est présenté dans l'annexe 26. Cette interaction est principalement due aux énoncés *Panneaux* pour lesquels les temps de compréhension sont plus courts lorsque la cible est vers l'avant du participant. Les différences entre les quatre cadres de référence pour les huit positions de cibles sont illustrées Figure 70.

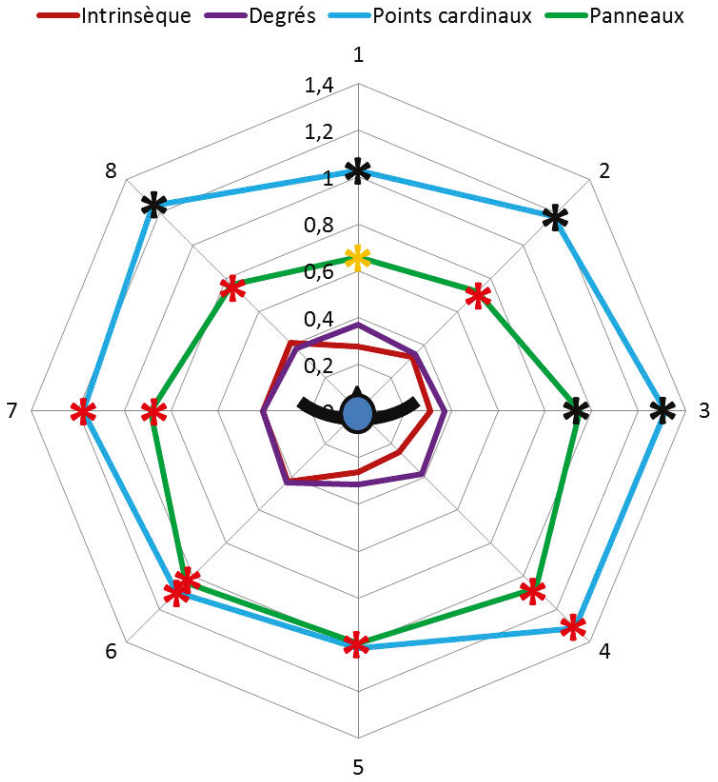


Figure 70 : Temps de compréhension moyen (seconde) pour chacune des huit réponses attendues pour les quatre cadres de référence. * différence significative avec les trois autres cadres de référence, * différence significative avec les cadres de référence centrés sur le destinataire, * différence significative avec le cadre de référence Intrinsèque.

Effets de l'interaction entre les cadres de référence, l'orientation du participant et la position de la cible sur le temps de compréhension des énoncés

Cette interaction est principalement due aux effets de l'orientation du participant et la position de la cible pour les énoncés *Points cardinaux*. Les orientations induisant les temps de compréhension les plus courts sont celles où le participant est aligné avec la direction organisatrice nord/sud. L'effet est d'autant plus marqué lorsque le participant fait face au nord. L'ensemble des analyses post-hoc est présenté dans l'annexe 27, ainsi qu'une figure récapitulative.

Synthèse des résultats sur le temps de compréhension des instructions

Tout comme l'exigence mentale, les temps de compréhension des instructions sont modulés par plusieurs facteurs qui interagissent.

L'effet global des cadres de référence indique que les temps de compréhension sont plus longs pour les deux types d'énoncés exocentrés. Ce résultat n'est plus vrai lorsque le participant est orienté vers le nord. Dans cette situation, seuls les énoncés *Panneaux* induisent un temps de compréhension plus long que les énoncés centrés sur le destinataire. Ce résultat et d'autres mettent en évidence l'importance de la direction organisatrice nord/sud dans l'utilisation des points cardinaux, et plus précisément l'alignement du destinataire avec cette direction.

En ce qui concerne les énoncés exocentrés utilisant les indices distaux (*Panneaux)*, les effets sont tout à fait similaires à ceux mis en évidence pour l'exigence mentale et confirment l'importance de l'accessibilité visuelle des indices utilisés.

3.2.3 Temps de sélection de la cible

L'ANOVA à trois facteurs (4 cadres de référence * 5 orientations du participant * 8 positions de la cible par rapport au participant) montre un effet du cadre de référence ($F(3, 45)=7.24$, $p<.001$), une interaction entre le cadre de référence et l'orientation du participant ($F(12, 180)=2.87$, $p<.01$), une interaction entre le cadre de référence et la position de la cible ($F(21, 315)=2.88$, $p<.001$) et une interaction entre les trois facteurs : cadre de référence x orientation x cible ($F(84, 1260)=1.59$, $p<.01$).

Effet global des cadres de référence sur le temps de sélection de la cible

Les analyses post-hoc montrent que les cadres de référence centrés sur le destinataire (*Intrinsèques* et *Degrés*) induisent des temps de sélection plus courts que les cadres de référence exocentrés (*Points cardinaux* et *Panneaux*). Aucune autre différence significative n'apparait. Ces résultats sont présentés Figure 71.

Effets de l'interaction entre les cadres de référence et l'orientation du participant sur le temps de sélection de la cible

Cette interaction est due aux mêmes facteurs que pour l'exigence mentale et temps de compréhension : l'effet de l'orientation du participant face au nord pour les énoncés *Points cardinaux*. Contrairement aux autres variables, le temps de sélection des cibles dans l'orientation 1 (nord) sont plus longs pour les énoncés *Panneaux* que pour les trois autres types d'énoncés, entre lesquels il n'y a d'ailleurs pas de différence. Ce résultat est illustré Figure 72.

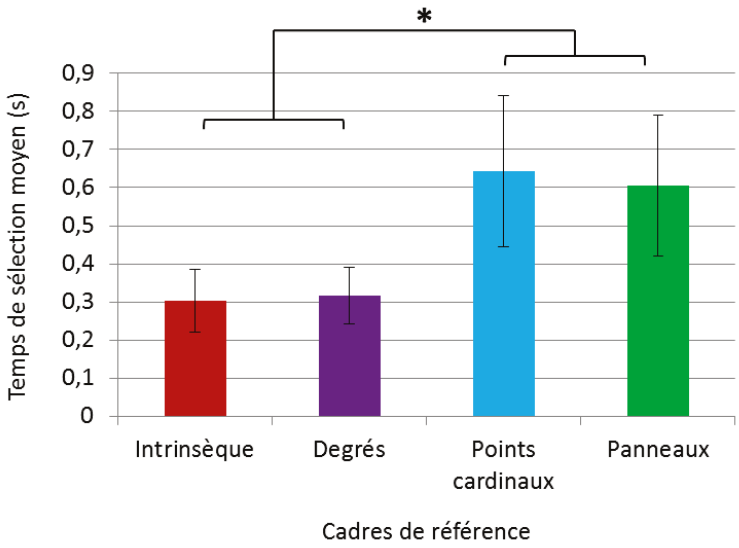


Figure 71 : Temps de sélection moyen de la cible en fonction des cadres de référence. * p<.05

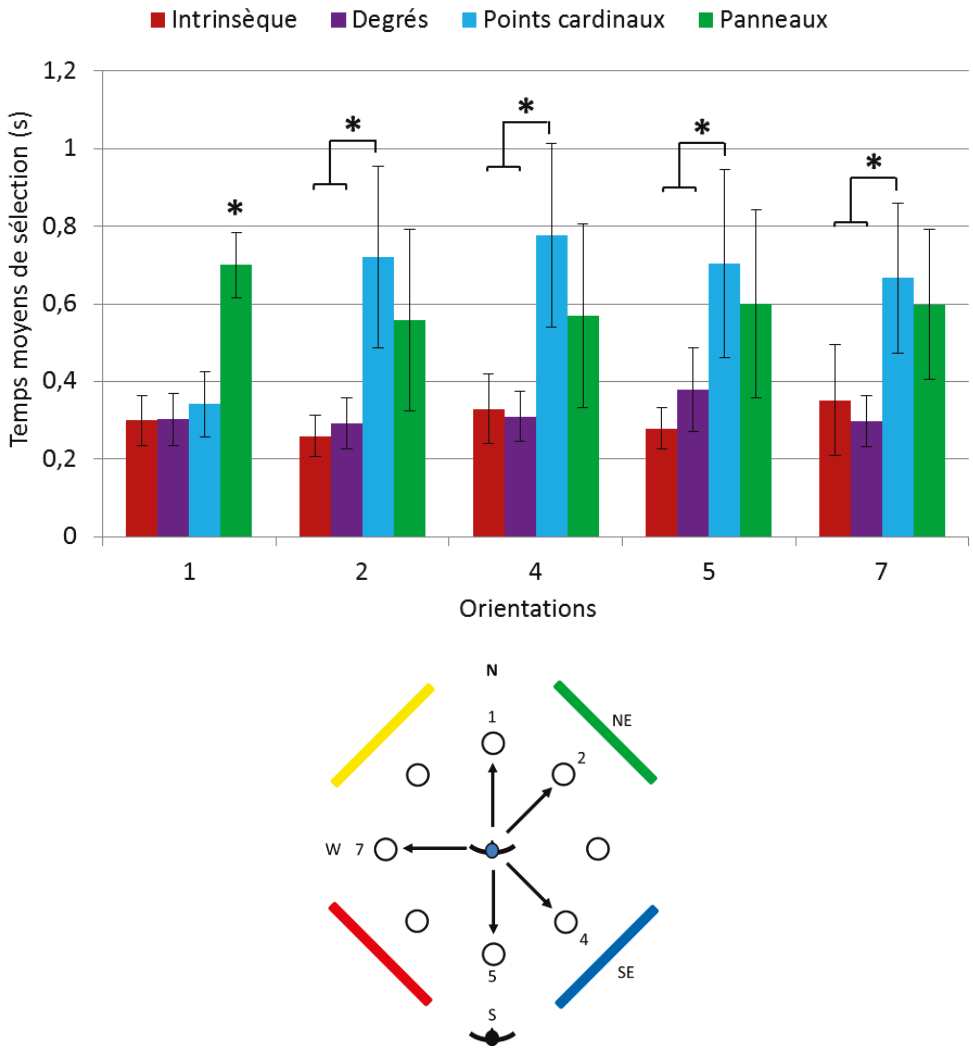


Figure 72 : Temps de sélection moyen de la cible en fonction de l'orientation du participant pour les quatre cadres de référence. * p<.05 L'interaction est due aux variations de temps de sélection de la cible pour les énoncés *Points cardinaux*, particulièrement court pour l'orientation 1 (N).

Effets de l'interaction entre les cadres de référence et la position de la cible sur le temps de sélection de la cible

Cette interaction est principalement due aux cibles présentes dans le champ de vision du participant pour les énoncés *Panneaux* et à la cible qui est en face du participant pour les énoncés *Points cardinaux*. Cet effet est illustré Figure 73 et l'ensemble des analyses post-hoc est présenté dans l'annexe 28.

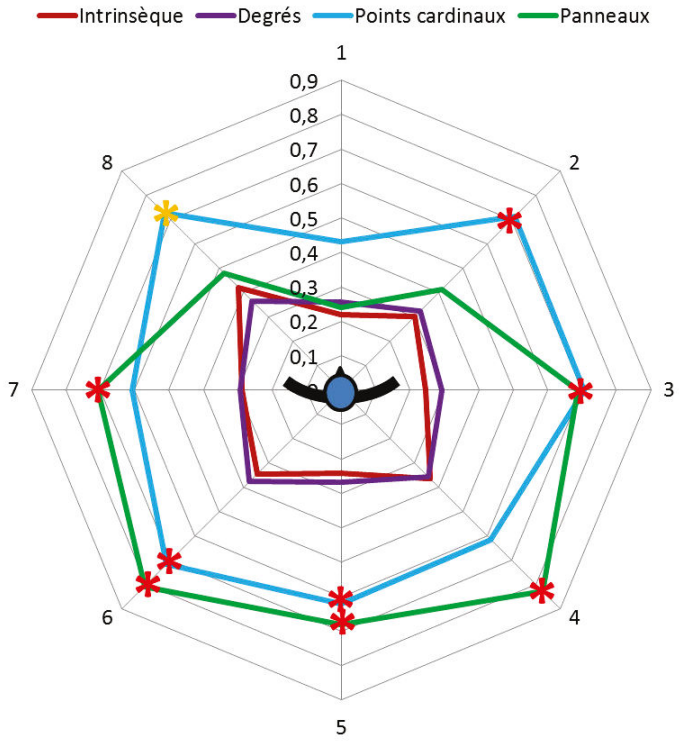


Figure 73 : Temps de sélection de la cible moyen (seconde) pour chacune des huit réponses attendues pour les quatre cadres de référence. * différence significative avec les cadres de référence centrés sur le destinataire, * différence significative avec le cadre de référence Intrinsèque.

Effets de l'interaction entre les cadres de référence, l'orientation du participant et la position de la cible sur le temps de sélection de la cible

Tout comme pour l'exigence mentale et le temps de compréhension, cette interaction est principalement due aux variations pour les cadres de référence exocentrés. L'ensemble des analyses post-hoc est présenté dans l'annexe 29, accompagné d'une figure récapitulatrice. En ce qui concerne le cadre de référence *Points cardinaux*, l'orientation du participant face au nord induit des temps de sélection comparables aux temps de sélection pour les cadres de référence centré sur le destinataire. Par ailleurs, lorsque le participant fait face au sud, les cibles nord et sud sont sélectionnées plus rapidement que d'autres cibles. En ce qui concerne les énoncés *Panneaux*, les cibles en face ou dans le champ de vision sont sélectionnées plus vite que les cibles en dehors du champ de vision du participant.

Synthèse des résultats sur le temps de sélection de la cible

Le temps de sélection d'une cible est influencé par le cadre de référence utilisé pour localisé la cible, et selon ce cadre, la position de la cible désignée et l'orientation du destinataire influencent également ce temps.

L'effet global des cadres de référence indique que les deux types d'énoncés exocentrés induisent des temps de sélection plus long que les énoncés centrés sur le destinataire. Tout comme pour l'exigence mentale et le temps de compréhension, ce résultat n'est plus vrai lorsque le destinataire fait face au nord. Dans ce cas, les temps de sélection pour les énoncés *Points cardinaux* sont similaires à ceux centrés sur le destinataire. Ce résultat confirme l'avantage de l'alignement du destinataire avec la direction organisatrice principale du cadre de référence utilisant les directions cardinales. D'autres résultats viennent confirmer la prédominance de la direction nord/sud, et notamment de ses bornes sur les autres directions. Les résultats sur l'interaction entre les cadres de référence *Points cardinaux* et la position de la cible montrent également que les participants vont plus vite pour sélectionner la cible qui est dans la direction dans laquelle ils sont, même lorsque cette direction n'est pas prédominante.

Enfin, en ce qui concerne le cadre de référence *Panneaux*, la sélection est plus rapide lorsque la cible est dans le champ de vision du participant. Tout comme pour l'exigence mentale et le temps de compréhension, l'accessibilité visuelle de la ou des références utilisées pour localiser la cible est un facteur déterminant dans le processus enclenché.

3.2.4 Nombre d'erreurs

Les participants ont fait entre 3 et 48 erreurs sur 320 essais, soit en moyenne 13,94 erreurs par participant. Cela représente un total de 223 erreurs sur 5120 essais (4.4%). Leur répartition est décrite Figure 74 : 15 erreurs (6,7%) ont été commises avec le cadre de référence *Intrinsèque*, 33 (14,8%) avec le cadre de référence *Degrés*, 118 (52,9%) avec le cadre de référence *Points cardinaux* et 57 (25,6%) avec le cadre de référence *Panneaux*. Toutes les comparaisons sont significatives, dont les résultats au test du Khi2 sont présentés Table 8.

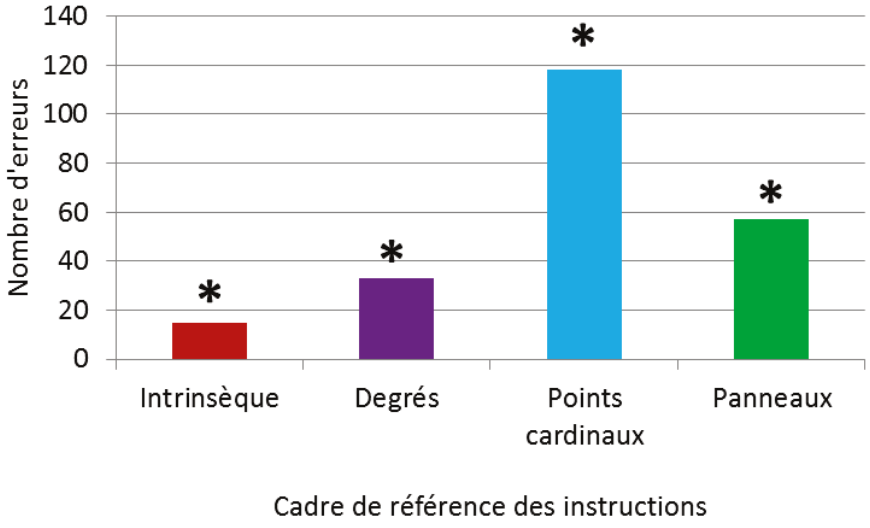


Figure 74 : Nombre d'erreurs commises par les participants en fonction du cadre de référence. * p<.05

Table 8 : Résultats aux tests du Khi2 pour le nombre d'erreurs par cadre de référence * p<.05, ** p<.01.

	Intrinsèque	Degrés	Points cardinaux	Panneaux
Intrinsèque		6,75 *	79,8 **	24,5 **
Degrés			48 **	6,4 *
Points cardinaux				21 **
Panneaux				

3.2.5 Questionnaire et autres comparaisons

L'analyse de variance de l'évaluation des différents cadres de référence dans le questionnaire montre un effet significatif ($F(3, 45)= 107.33$, p<.001). Les analyses post-hoc montrent que les participants ont estimé les énoncés *Points cardinaux* comme plus difficiles que les trois autres types d'énoncés. Les énoncés *Panneaux* sont également considérés comme plus difficiles que les énoncés centrés sur le destinataire. La différence entre les énoncés *intrinsèques* et *Degrés* n'est pas significative (p=.06). Ces résultats sont illustrés Figure 75.

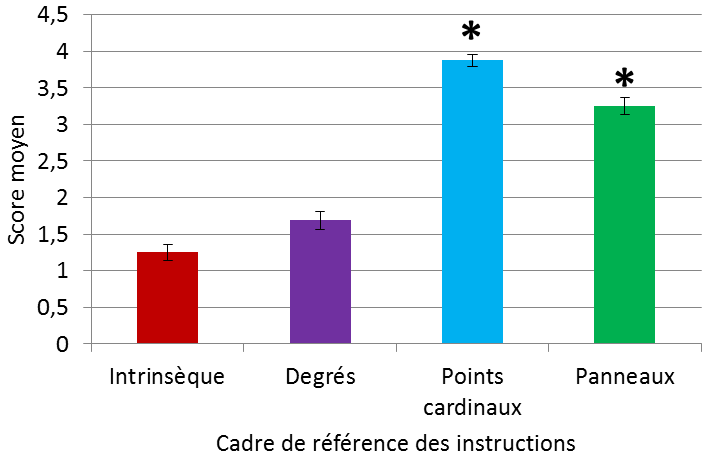


Figure 75 : Notation de la difficulté de chacun des cadres de référence : plus le score est élevé, plus le type d'énoncé a été jugé comme difficile. * p<.05

Les participants ont estimé se sentir présent à 4.6/10 (ET=.7), la présence de leur collaborateur à 7/10 (ET=.7) et présent dans la salle réelle à 6.4/10 (ET =.8). Il n'y a pas de corrélation entre les sentiments de présence et de coprésence. Mais nous observons une corrélation négative entre le sentiment de présence dans l'environnement virtuel et le sentiment de présence dans la salle réelle (r=-.55, p<.05). Comme pour les études précédentes, on note également une corrélation négative entre l'âge des participants et leur résultat au test des rotations mentales (MRT-A ®). On constate ainsi que plus l'âge augmente, moins les performances à ce test sont bonnes (r=-.54, p<.05). Mais on note également une corrélation entre l'âge et le temps de sélection de la cible (r=.91, p<.001), ainsi qu'avec le nombre d'erreurs (r=.69, p<.01), et les résultats au test SCOT (r=.89, p<.001). L'analyse des résultats a également mis en évidence une corrélation entre le temps de compréhension des énoncés et le nombre d'erreurs commises (r=.5, p<.05), et les résultats au test SCOT (r=.53, p<.05). De plus, le temps de sélection de la cible est également corrélé au nombre d'erreurs (r=.74, p<.01), aux résultats au test MRT (r=-.65, p<.01) et à celui SCOT (r=.84, p<.001). Enfin, le nombre d'erreur est corrélé aux résultats au test SCOT (r=.87, p<.001).

3.3 Synthèse et discussion – Compréhension d'énoncés spatialisés

Cette étude a fourni un grand nombre de résultats intéressants sur la compréhension d'énoncés spatialisés, que nous résumons ici selon les facteurs influençant les différentes variables et les interactions observées entre ces facteurs.

Effet global des cadres de référence

Les cadres de référence utilisés pour donner les instructions ont affecté toutes les variables mesurées, mais de manières légèrement différentes. Ainsi, pour les temps de compréhension et de sélection de la cible les temps mesurés sont plus longs pour les énoncés exocentrés que pour les énoncés centrés sur le destinataire, sans différence entre les cadres de référence au sein de ces catégories. Les résultats de l'exigence mentale montrent également des valeurs plus fortes pour les énoncés exocentrés que les énoncés centrés sur le destinataire, mais ils mettent également en évidence une exigence mentale plus forte pour les énoncés *Points cardinaux* que *Panneaux*. Il n'y a pas de différence d'exigence mentale entre les énoncés centrés *Intrinsèques* et *Degrés*. Enfin, le nombre d'erreurs montre que les participants ont fait plus d'erreurs avec les énoncés *Points cardinaux* qu'avec les trois autres cadres de référence. Les participants ont également fait plus d'erreur avec les énoncés *Panneaux* qu'avec les énoncés centrés sur le destinataire. Enfin, les participants ont fait moins d'erreurs avec les énoncés *Intrinsèques* qu'avec les énoncés *Degrés*.

Interaction entre les cadres de référence et l'orientation du participant

L'exigence mentale et les temps de compréhension et de sélection de la cible sont influencés de la même manière par l'orientation du participant. En effet, pour chacune des variables l'interaction est due à des mesures plus faibles pour les énonces *Points cardinaux* lorsque le participant est face au nord. Cet effet est d'autant plus marqué pour le temps de sélection de la cible. En effet, pour l'exigence mentale et le temps de compréhension, lorsque le participant est orienté vers le nord, les énoncés *Points cardinaux* induisent des valeurs non différentes des autres types d'énoncés. Alors que dans cette orientation, le temps de sélection avec les énoncés *Points cardinaux* est similaire aux temps de sélection des énoncés centrés sur le destinataire, et plus court que celui des énoncés *Panneaux*.

Interaction entre les cadres de référence et la position de la cible

L'exigence mentale et les temps de compréhension et de sélection de la cible montrent tous des mesures plus faibles pour les énoncés *Panneaux* pour les cibles situées dans le champ de vision du participant. L'exigence mentale est également affectée pour les énoncés centrés sur le destinataire. En effet, les cibles situées en diagonale derrière le participant induisent une exigence mentale plus forte que les situées dans le plan sagittal du participant. Cet effet est d'ailleurs plus marqué pour les énoncés *Intrinsèques* que pour les énoncés *Degrés*. Enfin, le temps de sélection de la cible montre également des temps plus courts pour sélectionner la cible juste en face du participant pour les énoncés *Points cardinaux*.

Interaction entre les cadres de référence, l'orientation du participant et la position de la cible

En ce qui concerne les énoncés *Points cardinaux*, les résultats de l'exigence mentale et des temps de compréhension et de sélection de la cible montrent tous l'importance des directions nord et sud. En ce qui concerne le cadre de référence *Panneaux*, l'exigence mentale et le temps de sélection de la cible indiquent l'importance des indices distaux présents dans le champ de vision. En effet, lorsque le participant fait face à un panneau, seule la cible située en face de ce panneau induit une exigence mentale plus faible et un temps de sélection plus court. Lorsque le participant est orienté entre deux panneaux de couleur, les cibles situées dans le champ de vision du participant sont sélectionnées plus vite

et induisent une exigence mentale plus faible. Enfin, aucune interaction entre l'orientation du participant et la position de la cible n'est constatée pour les cadres de références centrés sur le destinataire.

L'ensemble de ces résultats confirme que les énoncés centrés sur le destinataire sont globalement plus simples à comprendre que les énoncés exocentrés. Cependant, dans certaines situations, les énoncés *Points cardinaux* et *Panneaux* peuvent induire des performances similaires aux énoncés centrés sur le destinataire. Pour les énoncés *Points cardinaux*, c'est l'alignement du participant avec le nord qui est le principal élément facilitateur. Pour les énoncés *Panneaux*, c'est l'accessibilité visuelle de la ou des références utilisées.

D'après les deux études sur la production d'énoncés spatialisés décrites plus tôt dans ce manuscrit, les locuteurs choisissent les énoncés pour faciliter la tâche du destinataire, ce qui se traduit par l'utilisation de cadres de référence différents selon les situations. Les résultats de cette étude montrent que les locuteurs ont raison d'opérer des choix car tous les cadres de référence ne se valent pas, mais la question que nous allons discuter ici est : sur quels mécanismes la compréhension des énoncés s'appuie-t-elle ? Nous reviendrons dans la discussion générale sur la question de savoir si les locuteurs ont fait les bons choix.

3.3.1 Les énoncés centrés sur le destinataire sont les plus simples à comprendre

Le premier constat est que choisir la perspective du destinataire est un très bon choix pour minimiser l'effort de compréhension. En effet, les énoncés centrés sur le destinataire ont induit des exigences mentales plus faibles et de meilleures performances que les énoncés exocentrés. Ces résultats s'expliquent d'une part au niveau de la représentation occurrente sous-tendant l'action, et d'autre part de par la nature de la tâche. En effet, les énoncés centrés sur le destinataire donnent l'information dans un référentiel égocentré pour le destinataire. Il n'a donc pas à opérer de "traduction" entre son référentiel et celui supporté par le cadre de référence de l'énoncé. De plus, le référentiel de la tâche est également égocentré. En effet, le participant doit sélectionner la cible par rapport à lui. Donc, pour l'exécution de la tâche, là non plus le destinataire n'a pas de traduction à opérer.

Par ailleurs, les analyses globales pour les trois variables mesurées (exigence mentale, temps de compréhension et de sélection) indiquent que les participants n'ont pas fait de différence entre les énoncés utilisant uniquement leurs propriétés intrinsèques et les énoncés utilisant une combinaison de propriétés intrinsèques et extrinsèques (*Degrés*). Cependant, dans l'évaluation finale, bien que la différence entre énoncés *Intrinsèques* et *Degrés* ne soit pas significative, les résultats semblent pointer vers une préférence des énoncés utilisant exclusivement les caractéristiques intrinsèques du participant. Cette tendance se confirme dans les erreurs commises par les participants. En effet, les participants ont fait significativement plus d'erreurs dans le cadre de référence *Degrés* qu'*Intrinsèque*. Ainsi, l'introduction de termes extrinsèques, tels que les degrés, peut compliquer la compréhension des énoncés et induire des erreurs. Cette mécompréhension peut être liée au fait que le destinataire de l'information est plus ou moins familier de ces notions géométriques. Ainsi, il est nécessaire pour le locuteur de s'assurer que son destinataire a bien les mêmes notions que lui : cela fait partie du phénomène de grounding. On peut ainsi supposer que si un locuteur choisit spontanément un cadre de référence combiné utilisant des degrés, s'il se rend compte que son interlocuteur ne le comprend pas, alors il ajustera son choix de cadre de référence pour assurer une meilleure compréhension de la part de son collaborateur.

De plus, tout comme la position de la cible par rapport au destinataire affecte les coûts de production des énoncés, elle affecte également les coûts de compréhension. En effet, toutes les cibles n'induisent pas la même exigence mentale. Ces résultats sont à mettre en relation avec les résultats de l'étude de Carlson et Van Deman (2008) dans laquelle ils ont mis en évidence que certains plans sont

prédominants dans la compréhension d'énoncés spatialisés et que les plans intrinsèques d'un système de référence peuvent être accessibles indépendamment des autres plans. Par ailleurs, on retrouve au niveau de la compréhension les effets de la *Spatial Framework Theory* (Tversky, 2005b). En effet, les cibles situées en diagonale derrière le participant demandent une exigence mentale plus forte que les cibles situées dans le plan sagittal. Ce phénomène est liée au fait qu'une cible située en diagonale derrière n'est ni alignée avec les plans intrinsèques du référentiel égocentré du participant, ni dans son champ de vision. Par ailleurs, il est intéressant de noter que cet effet diffère légèrement entre les cadres de référence *Intrinsèques* et *Degrés* : lorsque la cible est décrite avec les degrés, seule la différence entre les positions "135° droite ou gauche" et "0°" est significative, alors que lorsque les cibles sont décrites avec les caractéristiques intrinsèques, les énoncés "derrière toi sur la droite", ou "sur la gauche" sont différents à la fois des énoncés "devant", mais également des énoncés "derrière". Ce résultat soulève la dimension linguistique : l'expression "180°" fait référence à l'écart par rapport à la direction avant de l'axe avant/arrière, mais elle ne fait pas directement référence à son opposé, comme l'expression "derrière" le fait. Dans un cas les mots organisent l'espace d'après toutes les directions intrinsèques de la personne, à savoir devant, droite, gauche et derrière, alors que dans l'autre les mots sont organisés autour de la direction avant (0°) et décrivent l'écart à cette direction.

Enfin, ces effets spécifiques aux énoncés centrés sur le destinataire ne transparaissent que sur l'exigence mentale et pas sur les mesures de la performance : les temps de compréhension et de sélection sont les mêmes pour toutes les cibles. Ce résultat montre que les participants ont mobilisé plus de ressources pour atteindre un niveau de performance identique.

3.3.2 Les efforts de compréhension des énoncés exocentrés varient selon plusieurs facteurs

Les analyses globales ont mis en évidence que les énoncés les plus difficiles sont ceux utilisant les directions cardinales. Ce résultat global est à nuancer en fonction de l'orientation de la personne à qui est destiné l'énoncé. En effet, dans la situation où le destinataire fait face au nord, les instructions ne sont pas plus difficiles à comprendre que les énoncés utilisant les cadres de référence centrés sur le destinataire. Certains locuteurs avaient d'ailleurs choisi cette stratégie dans les études sur la production d'énoncés en considérant que leur destinataire faisait toujours face au nord. Ainsi, lorsque le destinataire fait face au nord, les énoncés *Points cardinaux* peuvent être considérés comme centrés sur le destinataire utilisant des directions organisatrices extrinsèques. La difficulté pour le destinataire dépendra alors de sa maîtrise des directions cardinales.

Pour les autres orientations du destinataire, l'effort de compréhension est fort puisqu'il requiert un effort mnésique important. En effet, le participant doit se rappeler dans quelle direction il est, où est le nord et déduire où est la position de la cible. Cet effort n'est d'ailleurs supporté par aucun indice visuel, sauf lorsqu'ils faisaient face à l'avatar, mais dans cette situation ils étaient orientés vers le sud. Dans les commentaires des participants, on constate d'ailleurs qu'ils ont essayé de minimiser cet effort en mettant en place des stratégies (repérer les directions par rapport aux touches notamment).

La difficulté pour les énoncés *Points cardinaux* doit être mise en relation avec la nature de la tâche. En effet, comme nous l'avons évoqué précédemment, les participants devaient sélectionner la cible dans un référentiel égocentré. Ils devaient opérer une traduction entre le cadre de référence de l'énoncé et le référentiel égocentré de la tâche. Si le référentiel des réponses avait été centré sur le nord, comme sur une carte, cela aurait certainement été plus simple.

En ce qui concerne les énoncés s'appuyant sur les panneaux, ils ont induit, de manière générale, une exigence mentale et des temps de réaction plus longs. Cependant, ces données sont fortement influencées par la position de la cible à localiser. En effet, l'exigence mentale est plus faible lorsque les

indices utilisés pour donner les instructions sont présents dans le champ de vision que lorsque les références étaient non visibles. Ainsi, il semble que les cibles les plus compliquées soient celles pour lesquelles le participant doit faire appel à sa mémoire pour situer la référence dans l'espace. Cela renvoie à la différence entre l'imagerie mentale et la perception, pour laquelle Denis (1982) souligne que détecter une cible dans son champ de vision est plus simple que se remémorer des éléments perçus mais non présents. Par ailleurs ce résultat est à mettre en relation avec la particularité de la tâche : les participants ne pouvaient se déplacer dans l'environnement au moment de sélectionner la cible. Les résultats auraient peut-être été différents si les participants avaient pu vérifier la couleur des panneaux présents autour d'eux. En effet, dans la réalité il est très simple de se tourner pour voir de quelle couleur sont les rideaux situés derrière soi.

L'ensemble de ces trois études doit maintenant être discuté au regard des choix opérés par les locuteurs et de leurs effets sur les destinataires. Il sera également intéressant de discuter ces résultats au regard de la première étude présentée dans la première partie expérimentale. C'est ce que nous allons faire dans la discussion générale.

Chapitre 5 Discussion générale

Ce travail de recherche s'est intéressé aux situations collaboratives réalisées à distance grâce à des environnements virtuels collaboratifs immersifs. Nous nous sommes particulièrement intéressés à la construction du référentiel commun dans sa dimension spatiale. Ainsi, il nous a paru important d'étudier la manière dont les collaborateurs distants se comprennent mutuellement dans de tels environnements lors de dialogues comportant des informations spatialisées. Pour cela nous avons proposé une série d'études basées sur des scénarios plus ou moins proches des situations réelles de travail. La première partie expérimentale a permis d'identifier le rôle majeur de la prise en compte des contraintes individuelles dans la construction du référentiel spatial commun. Nous avons mis ces résultats en relation avec le principe du moindre effort collaboratif, selon lequel chacun des collaborateurs participe à la minimisation de l'effort global de la tâche et pas uniquement de sa propre charge mentale. Cette étude a également permis de soulever le fait que le lien entre les cadres de référence utilisés pour partager une information de nature spatialisée, les transformations mentales requises pour produire cet énoncé et la charge mentale associée, était mal connu. Dans la seconde partie expérimentale, nous nous sommes attachés à étudier ce lien entre prise de perspective et charge mentale pour les processus de production et de compréhension des énoncés. Les résultats de ces études ont permis de mettre en évidence la difficulté relative des situations en fonction de plusieurs paramètres pour le locuteur et le destinataire du message. Les résultats de ces études ont également permis de compléter les théories sur les processus cognitifs à l'œuvre lors de la production et la compréhension d'énoncés spatialisés. Dans ce chapitre, nous allons discuter les résultats dans leur ensemble en les confrontant aux théories et résultats d'autres études.

Nous reviendrons dans un premier temps sur les processus à l'œuvre lors de la production d'énoncés spatialisés, puis ceux à l'œuvre lors de la compréhension des tels énoncés. Nous remettrons ensuite ces processus dans leur contexte plus large de la collaboration, avec la construction du référentiel spatial commun et la gestion collective de la charge mentale. Enfin nous proposerons des conclusions sur la dimension technique.

1. La production d'énoncés spatialisés : deux niveaux de prise de perspective

La production d'un énoncé spatialisé nécessite d'avoir une représentation de la scène à décrire. Comme toute représentation occurrente, elle n'est pas simplement perception de l'environnement, même si cette perception est déterminante, elle est également une construction (Endsley, 1995; Hoc, 2003).

Les résultats de l'étude 1 Production d'énoncés sans indice ont permis de mettre en évidence l'influence du niveau perceptif sur l'exigence mentale. En effet, l'exigence mentale est modulée par l'excentricité de la cible dans le champ de vision (page 80) : plus la cible est proche du participant, et au centre de son champ de vision, plus elle est détectée rapidement. L'influence de ce facteur sur l'exigence mentale montre que le participant a dû produire un effort pour aller chercher l'information. Cet effort est dû au fait que si l'information n'est pas située juste en face de lui, le participant doit aller la chercher en déplaçant son regard. Etant donné la nature symétrique de la scène, il se peut que le participant commence du côté opposé à celui où la cible est apparue, ce qui lui demandera de balayer l'ensemble de la scène du regard. Selon nous ce résultat montre également le caractère égocentré de la représentation, au moins concernant son premier niveau, celui de la perception.

Pouliquen-Lardy, Lauriane. Collaboration à distance :
étude de la compréhension mutuelle dans les environnements virtuels collaboratifs immersifs : le cas de la communication spatiale - 2016

Etant donné que le cadre de référence égocentré est très peu utilisé, nous pensons que ce référentiel est inhibé au profit d'une autre représentation, exocentrée (*object-to-object*), basée sur les éléments de la scène, dont l'avatar. Ainsi, la scène est codée dans le référentiel centré sur l'avatar dont les caractéristiques intrinsèques organisent l'espace en zones plus ou moins accessibles. Mou & McNamara (2002) suggèrent que les représentations non-égocentrées n'imposent pas de coût cognitif supplémentaire. Les résultats des deux études sur la production d'énoncés présentées dans ce manuscrit montrent que les coûts associés aux représentations exocentrées varient selon plusieurs paramètres : la position de la cible par rapport au destinataire d'une part, et le processus cognitif permettant la prise de perspective d'autre part. Ces deux phénomènes étant liés, nous les traiterons de manière combinée.

Lorsque la cible est alignée avec le plan sagittal du destinataire, elle est détectée très rapidement et sa position peut être décrite dans un cadre de référence centré sur le destinataire sans qu'aucune transformation mentale ne soit nécessaire. En effet, les résultats de l'étude sur la production d'énoncés spatialisés sans indice, ont mis en évidence l'absence d'effet de l'orientation de l'avatar sur la charge mentale lorsque la cible est alignée avec le plan sagittal de l'avatar. Cela signifie que le locuteur n'a pas à se mettre mentalement à la place de son interlocuteur pour déterminer que la cible est devant ou derrière lui. Ce résultat est à mettre en relation avec le modèle de la *Spatial Framework theory* (Tversky, 2005b) et avec le modèle des deux niveaux de prise de perspective (Kessler & Rutherford, 2010). En effet, selon le modèle de la *Spatial Framework theory*, cette prédominance de l'axe avant/arrière est d'abord liée à l'asymétrie de cet axe dans le référentiel égocentré. Cette asymétrie détermine notre rapport au monde : l'espace en avant du corps est le lieu de la perception, de la marche, de la manipulation. Cette prédominance se retrouve dans toute référence disposant d'un axe asymétrique. Du point de vue de la représentation, la position de la cible est localisée par rapport à un système de référence à une direction, et non par rapport à un système de coordonnées complet (Carlson & Van Deman, 2008). Seules la position et l'orientation de l'avatar sont prises en compte, les autres caractéristiques intrinsèques n'étant pas indispensables à la description de la scène. Cette importance s'est inscrite dans notre vocabulaire pour pouvoir décrire ces directions facilement. Une cible située dans le plan sagittal du destinataire induit une représentation allocentrée où le destinataire est un point de repère orienté. L'utilisation du cadre de référence centré sur le destinataire est facilité puisque cette direction est supportée par du vocabulaire spécifique. Dans le modèle des deux niveaux de prise de perspective, les auteurs différentient un premier niveau de prise de perspective qui correspond à la ligne de visibilité (déterminer si une cible est visible ou non par une autre personne), et un second niveau qui consiste à opérer une transformation mentale pour mentalement prendre la perspective désirée. Alors que dans les études sur ce sujet les tâches associées pour étudier ces deux niveaux sont bien séparées, dans les études présentées ici, la tâche est la même. Ainsi, nos résultats suggèrent que lorsque la cible est alignée avec le plan sagittal du destinataire, les locuteurs n'utilisent que le niveau 1 de la prise de perspective, non pas pour déterminer la ligne de visibilité, mais pour désigner l'emplacement d'une cible avec les caractéristiques intrinsèques asymétriques du destinataire.

Pour les autres cibles, situées sur les côtés de l'avatar, deux phénomènes interviennent, le positionnement de la cible dans le référentiel du destinataire et la transformation mentale permettant la prise de perspective (niveau 2 dans le modèle de prise de perspective). Ces deux phénomènes sont liés : le positionnement de la cible d'un côté ou l'autre du destinataire nécessite de faire coïncider le référentiel égocentré avec le référentiel du destinataire. Ce processus est coûteux et l'effort augmente avec le degré de rotation à opérer (Michelon & Zacks, 2006; Roberts & Aman, 1993; Tamborello et al., 2012). Cette démarche est nécessaire car l'axe droite/gauche est symétrique. L'attribution des mots "droite" et "gauche" est arbitraire et relative à notre corps : rien mis à part le fait que nous avons appris que la main

droite est à droite et que la main gauche est à gauche, ne permet de les différencier. L'attribution des directions droite et gauche n'est donc possible que si on latéralise la référence d'après notre propre latéralité. Nous avons observé, lors de l'étude de la première partie expérimentale, mais également dans les études sur la production d'énoncés spatialisés, cette dimension incarnée de la prise de perspective : certains participants changeaient physiquement d'orientation pour se rapprocher de l'orientation de leur collaborateur. Nous avons observé un participant-guide levant le bras droit ou gauche pour s'assurer de l'exactitude de son énoncé. Ainsi, il est probable que le processus de prise de perspective de niveau 2 (Kessler & Rutherford, 2010) soit lié aux aptitudes spatiales. Cependant, nous n'avons pas trouvé de relation entre les résultats aux tests d'aptitude spatiale, les cadres de référence utilisés ou les performances. Cette absence d'effet peut s'expliquer par le niveau d'aptitude spatiale des participants qui serait suffisant par rapport à la demande de la tâche. De plus, lors de la seconde étude sur la production d'énoncés spatialisés (avec indices spatiaux), la présence de l'avatar latéralisé a facilité cette prise de perspective, renforçant le caractère incarné de ce processus. Le fait de rendre asymétrique l'axe droite/gauche facilite la latéralisation de la référence. Ce type d'indice peut faciliter la prise de perspective, mais s'appuie sur le processus cognitif le plus coûteux, celui requérant une transformation mentale. Ainsi, l'aide apportée par l'avatar latéralisé n'est que très limitée.

Le phénomène de prise de perspective n'est pas le seul à intervenir dans le processus de représentation et de description des cibles situées sur un des côtés de l'avatar. Après l'alignement avec le plan sagittal, deux facteurs de prédominance interviennent : l'alignement de la cible avec le plan frontal et le degré d'accessibilité. Cette combinaison conduit au fait que les cibles situées en diagonale derrière l'avatar sont celles demandant le plus d'effort et induisant les performances les moins bonnes. Ces cibles ont la particularité de ne pas être dans l'espace *vista*, c'est-à-dire qu'elles ne sont pas directement perceptibles et manipulables (Tenbrink et al., 2011). De plus, étant donné que ces directions ne sont pas primordiales dans notre rapport au monde, nous n'avons pas dans notre vocabulaire de mot pouvant facilement les désigner. Le locuteur doit donc trouver dans son répertoire langagier une solution parmi un large choix de possibilités. Le choix des mots pour donner l'instruction nécessite d'inhiber les autres possibilités, ce qui a un coût (Carlson & Van Deman, 2008). Par ailleurs, certains participants l'ont évoqué dans leur questionnaire, l'absence de retour de la part de leur collaborateur induisait un doute sur la pertinence de leurs choix.

Le choix du cadre de référence centré sur le destinataire s'appuie donc sur une représentation exocentrée plus ou moins complexe, dans laquelle le référentiel du destinataire est composé uniquement de son orientation (celle du plan sagittal) ou de l'ensemble de ses caractéristiques intrinsèques. L'étude sur la production d'énoncés spatialisés en présence d'indices spatiaux a permis de mettre en évidence que dans certaines situations, les locuteurs utilisaient des cadres de référence exocentrés non centrés uniquement sur l'avatar pour donner leurs instructions. Ces cadres de référence s'appuient sur d'autres processus que la prise de perspective.

En effet, la présence des indices distaux a mis en évidence que choisir un cadre de référence exocentré dont l'information est basée sur les repères présents dans la scène, évitait le phénomène de prise de perspective. Les résultats de l'étude sur la production d'énoncés avec indices spatiaux ont mis en évidence un effet minime de l'orientation de l'avatar sur l'exigence mentale (page 107). Ainsi, les cibles sont positionnées dans une représentation allocentrée : la cible est située par rapport aux panneaux. Les résultats montrent également que l'avatar fait également partie des points de repères codés dans cette représentation. L'avatar est ici une référence dont l'unique caractéristique intrinsèque prise en compte est l'orientation de l'axe avant/arrière (McNamara, 2003). Dans cette représentation allocentrée, une

cible alignée avec cet axe du destiantaire va enclencher plus souvent le choix du cadre de référence centré sur le destinataire, étant plus simple à utiliser.

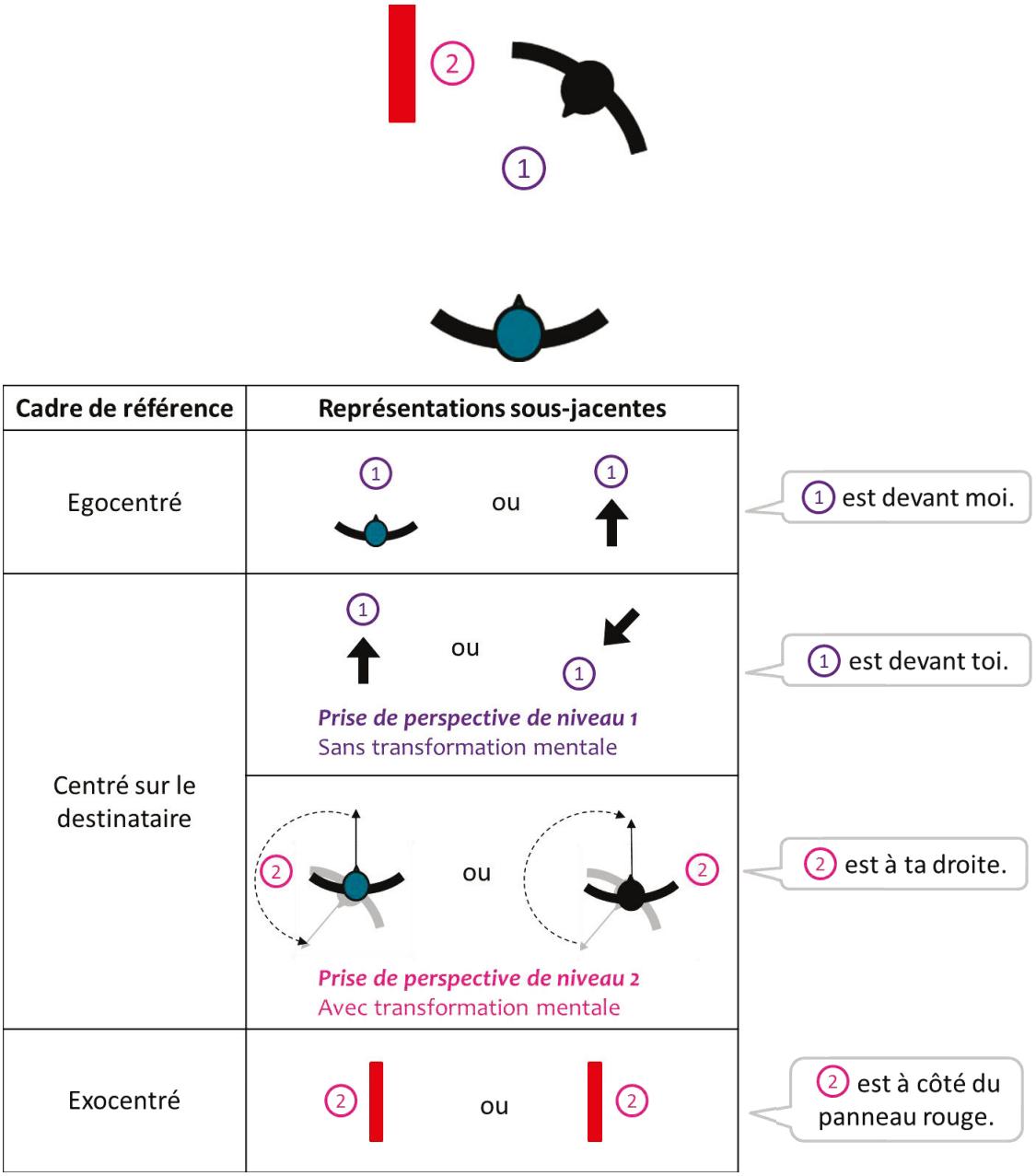


Figure 76 : Quatre énoncés possibles pour décrire la scène proposée en haut, selon le cadre de référence utilisé et la position de la cible à décrire, les représentations et processus sous-jacents diffèrent.

Ainsi, les travaux réalisés au cours de ce travail de recherche ont permis d'apporter des données factuelles sur les liens entre cadre de référence, référentiels spatiaux et charge mentale. La conclusion principale est qu'étudier les cadres de référence dans lesquels sont données des informations spatiales n'est pas suffisant pour conclure sur les coûts cognitifs liés à la production de ces énoncés. Pour résumer, la production d'énoncés centrés sur une autre référence que soi-même est principalement influencée par deux phénomènes : la position de la cible par rapport à cette autre référence, et l'orientation de la référence, comme illustré Figure 76. Pour les cibles situées dans le plan sagittal de la référence, le locuteur n'opère pas de transformation mentale : ce sont les cibles les plus simples à décrire. Les

représentations sous-jacentes conservent uniquement la direction avant/arrière de la référence et le positionnement de la cible par rapport à cet axe. Pour les autres cibles, le locuteur opère une prise de perspective (niveau 2) impliquant la simulation du changement d'orientation de son propre référentiel. Le coût associé à cette transformation mentale augmente avec l'écart entre sa propre orientation et l'orientation de la référence (Michelon & Zacks, 2006; Schober, 2009b; Tamborello et al., 2012).

2. Compréhension et utilisation d'énoncés spatialisés

Une fois produits, les énoncés spatialisés doivent être interprétés pour être utilisés. Cette interprétation demande l'activation d'un référentiel, par inhibition des autres référentiels possibles (Carlson et Van Deman, 2008). Les résultats de la troisième étude sur la compréhension d'énoncés spatialisés montrent que certains cadres de référence sont plus simples à interpréter que d'autres.

Le premier constat est à mettre en relation avec la limite de cette étude : le référentiel de la tâche est le référentiel égocentré. Il aurait pu en être autrement, mais nous avons choisi d'uniformiser les modes de réponses et d'utiliser le même mode de sélection de la cible pour tous les cadres de référence. Ce choix nous a permis de comparer la difficulté de compréhension et d'utilisation de différents cadres de référence pour un usage donné, ici égocentré. Cette étude apporte un premier éclairage sur les processus de traitement de l'information spatialisée. Il serait intéressant de répliquer cette étude en modifiant les modes de sélection de la cible. Pour les résultats discutés ici, il est donc nécessaire de garder en mémoire que la tâche s'effectue dans un référentiel égocentré.

Nous avons observé que, de manière générale, les énoncés utilisant un cadre de référence centré sur le destinataire sont plus simples à comprendre et à appliquer que les énoncés exocentrés. En effet, nous avons observé que les énoncés centrés sur le destinataire sont ceux qui demandent le moins d'effort et induisent les meilleures performances. Cette observation peut s'expliquer par le fait que ces énoncés peuvent être traduits directement dans le référentiel égocentré du destinataire, qui est également le référentiel de la tâche. Il est donc effectivement plus simple pour un destinataire de ne pas avoir à faire de traduction entre le cadre de référence et le référentiel de la tâche.

Concernant les énoncés exocentrés utilisant les points cardinaux, le principal résultat est l'importance de l'alignement du référentiel allocentré sur le référentiel de la tâche. En effet, lorsque le destinataire fait face au nord, ses performances et l'effort demandé pour comprendre l'énoncé ne sont pas différents des performances et de l'effort demandé par les énoncés centrés sur le destinataire. Ainsi, lorsque le référentiel allocentré est aligné avec le référentiel égocentré, l'usage des points cardinaux par le locuteur s'apparente à l'usage d'un cadre de référence centré sur le destinataire extrinsèque. Certains participants de l'étude sur la production d'énoncés spatialisés sans indices ont d'ailleurs utilisé les directions cardinales, en considérant le destinataire comme une boussole, c'est-à-dire faisant toujours face au nord. Le résultat associé sur la compréhension des énoncés spatialisés vient confirmer que l'usage de cadres de référence centrés sur le destinataire extrinsèques, utilisant les directions cardinales, est pertinent et efficace. Nous reviendrons sur cette question dans la partie sur le référentiel commun.

Pour les autres orientations du destinataire, les résultats suggèrent que les participants ont dû opérer des transformations mentales coûteuses pour aligner mentalement le référentiel exocentré et le référentiel égocentré. On constate de nouveau l'influence du référentiel de la tâche sur la performance et l'effort nécessaire pour réaliser la tâche (Verjat, 1994). L'utilisation de cinq orientations parmi les huit possibles ne permet pas de conclure sur la nature de ces transformations mentales. Cependant les résultats montrent que les énoncés spatialisés exocentrés demandent un effort considérable au destinataire, pour des performances médiocres, lorsque la direction organisatrice principale du cadre de

référence, ici le nord, n'est pas alignée avec la direction organisatrice principale du référentiel de la tâche, ici l'avant du participant. Notre hypothèse est que si les participants avaient dû utiliser l'énoncé dans un référentiel centré sur les directions cardinales, les résultats auraient été aussi bons quelle que soit l'orientation du destinataire dans l'environnement. En effet, il n'y aurait pas eu de traduction à opérer entre le cadre de référence et le référentiel de la tâche.

Enfin, concernant les énoncés allocentrés, c'est-à-dire basés sur les repères distaux présents dans l'environnement, les résultats montrent que les participants ne s'appuient plus sur un processus d'alignement de référentiel, mais sur des processus perceptif et mnésique. En effet, lorsque les repères distaux (panneaux) utilisés pour situer la cible sont présents dans le champ de vision du participant, les performances et l'effort pour comprendre l'énoncé ne sont pas différents des énoncés centrés sur le destinataire. Inversement, les performances sont moins bonnes et l'effort plus important lorsque les références sont situées en dehors du champ de vision. Les participants s'appuient sur une représentation occurrente basée, dans un cas sur les indices visuels accessibles, alors que dans l'autre, elle est basée sur la mémoire de l'environnement. On constate ici que la représentation occurrente n'active les éléments pertinents qu'au dernier moment. Cela permet une charge mentale globale modérée puisque l'activation coûteuse des connaissances mémorisées n'est sollicitée que lorsque les références sont absentes du champ de vision. Par ailleurs, ce résultat est à modérer par rapport aux contraintes techniques de la tâche. En effet, les participants n'avaient pas la possibilité de bouger pour sélectionner la cible. Donc dans les situations où les références utilisées pour situer la cible n'étaient pas visibles, le fait de se tourner aurait pu diminuer considérablement l'exigence mentale. Les participants n'auraient pas eu besoin de réactiver leurs connaissances pour avoir une représentation fonctionnelle de la situation. Cependant, l'exigence mentale aurait été affectée par la dimension technique supportant le mouvement dans l'environnement. Dans un environnement réel, il est facile d'aller chercher l'information derrière soi, dans un environnement virtuel, cela dépend des outils à disposition. Par rapport aux choix des locuteurs dans l'étude sur la production d'énoncés avec indices spatiaux, nous n'avons pas constaté de différence selon l'accessibilité visuelle des indices pour le destinataire. Ainsi, l'utilisation des indices distaux pour décrire les cibles situées derrière le destinataire, ne semble pas être pertinente pour minimiser l'effort du destinataire. Il est donc possible que ce soit bien l'effort du locuteur qui soit privilégié dans ce choix. Cependant, et nous y reviendrons dans la partie sur le référentiel commun, les locuteurs n'avaient pas connaissance des contraintes de la tâche de leur collaborateur.

Ainsi, les résultats de l'étude sur la compréhension des énoncés a permis d'apporter, pour un contexte technique et de tâche très spécifiques, des données factuelles sur le lien entre cadre de référence et coûts cognitifs associés à leur compréhension. Pour une tâche de sélection de cible dans un référentiel égocentré, dans un environnement virtuel sans possibilité de déplacement, les énoncés centrés sur le destinataire sont les plus simples à comprendre. Mais l'usage de cadres de référence exocentrés, dans certaines circonstances, l'est tout autant.

3. Le référentiel commun permet la gestion collective de la charge mentale

3.1 Le principe du moindre effort collaboratif se manifeste sous plusieurs formes

Lors d'une activité collaborative, la performance est sous-tendue par des activités invisibles dont les coûts cognitifs sont plus ou moins élevés. Nous avons vu notamment que lors d'un dialogue contenant des informations spatialisées, les interlocuteurs doivent assumer des coûts liés à la production et à la compréhension des énoncés. Nous avons vu également que les interlocuteurs ne minimisent pas uniquement les coûts individuels. Ainsi la charge mentale est gérée collectivement.

Le premier phénomène a été observé dans la première partie expérimentale (page 55). Dans cette étude nous avons mis en évidence un déséquilibre dans l'effort individuel demandé à chacun pour participer à la réalisation du but commun. Les manipulateurs ont estimé une charge mentale globale plus importante que les guides. Par ailleurs, nous avons constaté que les guides assumaient plus de rotations mentales dans le dialogue. Il semble ainsi que celui qui a la plus faible charge mentale globale essaie de minimiser, là où il le peut, la charge mentale de celui qui doit assumer une charge mentale plus importante. Et il semble que cette stratégie soit performante puisqu'elle est renforcée au cours du temps. En effet, nous avons observé que l'écart entre énoncés égocentrés et centrés sur le destinataire était plus conséquent à la fin qu'au début des sessions collaboratives (Figure 30, (Roger et al., 2013; Schober, 1995). De plus, étant donné que le référentiel de l'action des manipulateurs est centré sur eux, l'étude sur la compréhension des énoncés spatialisés confirme que, les énoncés centrés sur le destinataire sont plus simples à comprendre pour eux. Les manipulateurs n'ont donc pas de traduction à opérer pour appliquer ce qui leur est dit. Ainsi le choix de privilégier la perspective du manipulateur, leur permettrait de minimiser les ressources allouées à la compréhension pour les diriger vers les tâches demandant le plus d'effort. Ce constat rejoint le fait que les contraintes techniques influencent la collaboration (Spante et al., 2004).

En ce qui concerne les cadres de référence, l'étude sur la production d'énoncés sans indice a permis de mettre en évidence qu'un même cadre de référence pouvait prendre des formes différentes selon les caractéristiques utilisées pour donner l'information. En effet, pour l'utilisation du cadre de référence centré sur le destinataire, certains participants-locuteurs ont préféré utiliser uniquement les caractéristiques intrinsèques du destinataire, alors que d'autres ont choisi des caractéristiques extrinsèques, telles que les heures, les degrés ou les directions cardinales. Les résultats sur la compréhension montrent que l'utilisation des caractéristiques intrinsèques offre de meilleures performances que l'utilisation de caractéristiques extrinsèques telles que les degrés. Selon nous, cette différence de performance révèle un manque de familiarité de certains participants sur ces notions de trigonométrie. Un participant-locuteur ayant utilisé cette stratégie fait part du choix qu'il a pu faire, supposant que son collaborateur connaissait mieux les degrés que les heures, mais cela a soulevé un doute sur l'efficacité de ce choix. On constate ici que bien qu'il n'y ait pas d'interaction avec le destinataire, il fait partie des préoccupations du locuteur. Nous reviendrons sur ce phénomène dans la partie sur le référentiel commun. Du point de vue du principe du moindre effort collaboratif, il nous semble que l'utilisation de caractéristiques extrinsèques dans un cadre de référence centré sur le destinataire est pertinente et efficace si les collaborateurs sont tous les deux à l'aise avec ces caractéristiques. Notre hypothèse est que les locuteurs ont choisi ces caractéristiques car elles sont plus simples à utiliser pour eux, mais qu'ils auraient adaptés leurs énoncés s'ils avaient constatés des défauts de compréhension de la part de leurs collaborateurs.

Par ailleurs, le fait de minimiser l'effort collectif n'empêche pas de minimiser les efforts individuels. En effet, les guides ont mis en place des stratégies pour minimiser leurs efforts de production. Le fait de se déplacer et de s'aligner avec la perspective choisie facilite la prise de perspective. D'autre part, comme l'ont mis en évidence les études sur la production et la compréhension d'énoncés spatialisés, l'utilisation de repères visuellement partagés facilite la tâche des locuteurs et des destinataires (Dumazeau & Karsenty, 2008). Dans le travail présenté ici, les participants-destinataires semblent également avoir minimisé leurs propres efforts. En effet, lorsque les indices distaux sont utilisés, les informations stockées en mémoire ne sont sollicitées que lorsque cela est nécessaire, c'est-à-dire lorsque la référence n'est pas accessible dans le champ de vision. Ainsi, la représentation occurrente ne garde que les informations jugées pertinentes, et l'effort n'est produit que s'il est nécessaire. C'est un effet similaire

que Tamborello et collaborateurs (2012) ont mis en évidence. Les participants devaient orienter un canon pour l'aligner avec la cible. Lorsque les cibles étaient plus nombreuses pour un cadre de référence, il était privilégié au détriment de l'autre. Ainsi, la gestion individuelle de la charge mentale suppose d'orienter les ressources, notamment attentionnelles, vers ce qui est le plus pertinent pour la tâche. Cette stratégie permet à moindre coût d'assurer de bonnes performances dans la plupart des cas.

Ainsi, comme évoqué dans l'état de l'art (page 28) le principe du moindre effort collaboratif consiste à trouver le meilleur compromis entre les différentes sources de coûts cognitifs (Raufaste et al., 2004; Roger et al., 2011). Avec les résultats mis en avant dans ce travail, nous avons pointé le fait que le principe du moindre effort collaboratif s'appuie sur une gestion collective des ressources. Les participants-guides ont minimisé les efforts de leur collaborateur en vue de lui permettre d'allouer des ressources, notamment attentionnelles, sur les aspects de la tâche demandant un effort particulier : la manipulation de l'objet.

3.2 Construction du référentiel commun

Pour ajuster au mieux les efforts investis dans la tâche commune, les opérateurs ont besoin d'intégrer des informations concernant leur collaborateur dans la représentation occurrente de la situation. Ces informations constituent une partie du référentiel commun. Ainsi les opérateurs, pour ajuster leurs actions et notamment les choix de cadre de référence, prennent en compte une évaluation de la charge mentale déployée par leur collaborateur. Selon la situation cette évaluation se base soit sur ce que renvoie le collaborateur, soit sur une inférence. Dans les études sur la production des énoncés spatialisés, puisqu'il n'y a pas de retour de la part du collaborateur, les stratégies, si elles prennent en compte les coûts associés à la compréhension, ne s'appuient que sur une inférence. Etant donné que la majorité des énoncés sont produits dans le cadre de référence du destinataire (page 78), il est probable que les participants-locuteurs aient inféré que le coût de compréhension est plus faible si les énoncés sont centrés sur le destinataire. Les résultats de l'étude sur la compréhension des énoncés spatialisés valide cette inférence. Cependant, dans l'étude sur la production d'énoncés spatialisés en présence d'indices spatiaux, les participants-locuteurs ont également utilisé des références hors du champ de vision du destinataire (page 103). Les résultats sur la compréhension des énoncés spatialisés montrent que ce choix ne semble pas optimal pour minimiser le coût de compréhension. L'évaluation faite par les participants-locuteurs de la charge déployée par les destinataires pour comprendre leurs énoncés est faussée par le fait qu'ils n'ont pas eux-mêmes réalisé la tâche et qu'ils n'ont pas d'indication sur la manière dont le destinataire doit réaliser la sélection de la cible. Ainsi, ils n'ont pas de connaissance des contraintes de leurs destinataires. En effet, s'ils avaient eux-mêmes eu à sélectionner la cible, ils auraient constaté la difficulté de sélectionner une cible d'après des références non visibles, et auraient probablement changé de cadre de référence (Spante et al., 2004). C'est sans doute pour cette raison que certains participants ont choisi d'utiliser malgré tout le destinataire comme référence, même en présence des indices distaux.

Contrairement aux études de la deuxième partie expérimentale où la collaboration est simulée, les participants ayant participé à la première partie expérimentale ont été réellement mis en situation collaborative. Ils ont ainsi pu évaluer la charge mentale déployée par leur collaborateur au cours de l'activité et adapter leurs stratégies. En effet, si le collaborateur montre des signes d'incompréhension, alors il faut ajuster les métainformations sur le collaborateur et choisir des stratégies plus opérationnelles. C'est ce que Schober (2009) a montré dans son étude où les capacités spatiales des interlocuteurs ont influencé les prises de perspectives verbales. Une fois les stratégies mises en place, si leur effet bénéfique est confirmé, alors il semble qu'elles soient renforcées. C'est ce que nous avons observé pour les énoncés

utilisant les perspectives de l'un ou l'autre des participants. Même si la perspective des manipulateurs est privilégiée chez les participants dès le début de la session collaborative, elle est renforcée au cours de l'activité. Ce résultat confirme que le référentiel commun s'appuie sur des métainformations concernant la charge de chacun.

Alors que le choix des cadres de référence s'est fait naturellement et implicitement dans la première partie expérimentale, l'absence de dialogue dans les études de la seconde partie expérimentale a induit différents comportements. D'une part les participants-locuteurs ont souvent explicité le cadre de référence dans lequel ils donnaient leurs instructions au début, lors des essais d'entrainement. Cela leur permettait de rendre commun explicitement ce cadre générale. Cette explicitation leur permettait ensuite d'automatiser leurs énoncés. La tâche des locuteurs étant très répétitive, pour garder leur concentration et être les plus efficaces possibles ils ne donnaient que les informations pertinentes par rapport au contexte partagé (Karsenty & Falzon, 1993; Raufaste et al., 2004). Ainsi, l'étude comparée des résultats de la seconde partie expérimentale suggère que si les collaborateurs avaient pu échanger réellement, certaines stratégies auraient été privilégiées et d'autres abandonnées. C'est le cas notamment des énoncés centrés sur le destinataire utilisant des directions combinées (*Degrés*). Le fait que les participants-destinataires aient fait plus d'erreurs dans ces conditions suggère que cette stratégie aurait au moins été discutée, voir abandonnée pour une stratégie partagée et plus efficace.

4. Limites des études

Il est important de discuter ici des limites des études présentées dans ce travail.

La première limite concerne le choix d'étudier indépendamment les phases de production et de compréhension des énoncés spatialisés. Ce choix est discutable car il décontextualise la production et la compréhension des énoncés et empêche tout phénomène d'ajustement d'apparaître (Bezuidenhout, 2013). Cependant, sans ce choix, nous n'aurions pas pu contrôler les paramètres tels que nous l'avons fait, notamment le choix des cadres de référence. En effet, les participants se seraient influencés mutuellement, et c'est bien de cela qu'il sera question par la suite, puisque dans tout dialogue les interlocuteurs sont actifs (Giboin, 2004; Hoc, 1998; Karsenty & Falzon, 1993). Mais nous n'aurions pas pu conclure sur la relation entre cadres de référence, transformations mentales et coûts cognitifs. En effet, avec les résultats des études présentées dans ce manuscrit, il n'est plus possible de conclure que l'utilisation d'un cadre de référence centré sur le destinataire entraîne une augmentation de la charge mentale. Selon la situation, et surtout selon la position de la cible dans le référentiel du destinataire, cela est plus ou moins vrai.

Par ailleurs, l'étude sur la compréhension des énoncés spatialisés s'est appuyée sur les énoncés produits dans les études sur la production d'énoncés spatialisés. Nous n'avons ainsi pas testé d'énoncés centrés sur le locuteur. Les conclusions sur la compréhension ne sont donc valables que pour les comparaisons opérées, et pas pour l'ensemble des énoncés possibles.

Enfin, comme déjà évoqué ci-avant, les participants-destinataires dans l'étude sur la compréhension d'énoncés spatialisés, ont dû réaliser la tâche de sélection de la cible dans un référentiel égocentré. Les résultats auraient sans doute varié significativement avec une tâche réalisée dans un autre référentiel.

5. Implications techniques et perspectives

Bien que la distance induise de nouvelles contraintes sur le travail collaboratif, l'usage des environnements virtuels collaboratifs immersifs offre de nombreuses possibilités de développement. Les aides apportées par les outils numériques sont plus ou moins faciles à mettre en place, d'où l'importance de bien cibler les besoins des utilisateurs (Burkhardt, 2003b; Giboin, 2004). Les études présentées dans ce manuscrit ont soulevé plusieurs éléments permettant d'orienter les futurs développements de ces environnements virtuels collaboratifs immersifs.

Les informations fournies par l'EV peuvent servir de référentiels communs externes, et supporter directement les échanges (Dumazeau & Karsenty, 2008; Hoc, 2003). C'est ce que nous avons constaté avec la mise en place des indices distaux. Leur présence a permis de soulager les locuteurs en offrant une alternative à la prise de perspective du destinataire. Mais l'usage de ces indices a induit des coûts supplémentaires pour les destinataires. Ces coûts sont principalement liés aux contraintes techniques imposées par la tâche. En effet, ces indices visuels auraient été réellement efficaces pour les deux interlocuteurs (locuteur et destinataire) si le destinataire avait pu simplement se tourner dans l'environnement. Le besoin ici n'était pas de pouvoir se déplacer librement comme dans la première partie expérimentale (Annexe 2). En effet, avec cet outil il n'est pas simple de se retourner facilement, cela demande un apprentissage. Par contre, une simple commande de rotation, ne modifiant pas la position du participant dans l'environnement aurait suffi. L'utilisation des fonctionnalités est donc à mettre en relation avec les coûts nécessaires pour les utiliser : s'il est trop compliqué (coûteux en temps et en énergie) de réaliser telle ou telle action, alors l'opérateur va compenser en produisant des énoncés plus complexes ou en demandant plus de gage de compréhension de la part de son collaborateur (Schober, 1993). L'illustration extrême de cette balance gain/coût est l'observation décrite dans le contexte industriel (page 9 - Contexte industriel et approche ergonomique) : les efforts demandés aux opérateurs pour accéder aux fonctionnalités sont tels qu'ils préfèrent déléguer cette dimension technique aux responsables de la réalité virtuel habitués et performants. Cependant cette stratégie implique plus d'échanges entre les opérateurs et les spécialistes techniques.

Du point de vue des dialogues spatialisés, il semble que les fonctionnalités nécessaires pour faciliter les activités individuelles diffèrent selon les rôles attribués aux différents collaborateurs. Convertino et ses collaborateurs (2005) ont d'ailleurs mis en évidence qu'une combinaison des vues spécifiques à chacun des rôles, complétées par des vues communes, semble être un bon compromis pour gérer les informations métier et les informations constituant le référentiel commun. En effet, chaque métier a besoin d'avoir accès à des informations spécifiques, qui n'ont pas besoin d'être toutes partagées car cela créerait une charge cognitive trop importante. Cependant les échanges doivent porter sur des informations mutualisées. Ainsi, il serait possible d'imaginer des menus équipés de fonctionnalités dédiées aux métiers des différents collaborateurs. Les personnes qui seront amenées à guider un collaborateur pourraient avoir accès à des fonctionnalités telles que le point de vue déporté, l'affichage de points de repères. Alors que les personnes amenées à réaliser des tâches spécifiques pourraient bénéficier d'autres modes de visualisation (plan, rétroviseurs, par exemple).

Le projet PIVIPP s'est conclu sur une dernière phase de tests pour valider l'usage de différentes solutions techniques par des opérateurs d'Airbus. Pour réaliser ces tests, nous avons mis en place cinq scénarios, très proches de situations réelles de travail. Les opérateurs, en binômes distants, devaient prendre des décisions de manière collective. Les observations ont permis de différentier deux types d'outils : ceux supportant les échanges au niveau macroscopique (localisation dans la maquette) et ceux supportant les échanges au niveau local (pièce d'intérêt). Ces deux niveaux d'échanges ne requièrent pas

les mêmes besoins. En effet, il est apparu que les opérateurs Airbus n'avaient absolument aucune difficulté pour se repérer dans la maquette numérique de l'avion : ils ont utilisé les mêmes références que celles utilisées dans leur environnement de travail réel (numéro des cadres structurant l'avion). Inversement, les opérateurs ont beaucoup utilisé les outils permettant des échanges précis : le laser, la surbrillance ou le fait d'isoler une pièce. Ces outils ont permis aux opérateurs de faire entrer dans le référentiel commun les objets et lieux d'intérêt de manière précise et efficace.

Enfin, cette dernière phase de test a permis de soulever un point particulièrement important pour l'utilisation de la réalité virtuelle pour la collaboration distante. Les situations requièrent parfois des échanges sur des pièces situées dans la dimension verticale. Pour accéder visuellement à des zones situées en hauteur ou sous plancher, l'utilisation du mur immersif contraint énormément les opérateurs. En effet, là où dans un CAVE l'opérateur n'a qu'à lever ou pencher la tête pour voir ce qu'il y a au-dessus ou sous lui, devant un mur immersif l'accès à ces informations demande aux opérateurs d'adapter leurs stratégies de positionnement. Bien que nous ne nous soyons pas intéressé à la dimension verticale dans les études réalisées dans ce travail de recherche, les auteurs semblent s'accorder sur la prédominance de l'axe tête/pied sur tous les autres (Tversky, 2005a). Par ailleurs, cet axe est très structurant, notamment parce qu'il est renforcé par le référentiel gravitaire. Ainsi, il serait intéressant de poursuivre les travaux présentés ici afin de déterminer les coûts cognitifs pour localiser des cibles dans cette dimension, et quelles métaphores ou outils virtuels permettraient de supporter cette dimension des dialogues spatialisés sans perturber la verticalité des opérateurs.

REFERENCES BIBLIOGRAPHIQUES

Albaret, J. M., & Aubert, E. (1996). Etalonnage 15-19 ans du test de rotation mentale de Vandenberg. *Evolutions psychomotrices*, 206–209.

Ameisen, J. C. (2012). *Sur les épaules de Darwin-: Les battements du temps* (Vol. 1). Éditions Les Liens qui libèrent.

Amorim, M.-A., Isableu, B., & Jarraya, M. (2006). Embodied spatial transformations:« Body analogy » for the mental rotation of objects. *Journal of Experimental Psychology: General*, *135*(3), 327.

Avraamides, M. N., & Kelly, J. W. (2008). Multiple systems of spatial memory and action. *Cognitive processing*, *9*(2), 93–106.

Axelsson, A.-S., Abelin, A., Heldal, I., Schroeder, R., & Wideström, J. (2001). Cubes in the cube: A comparison of a puzzle-solving task in a virtual and a real environment. *CyberPsychology & Behavior*, *4*(2), 279–286.

Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments*, *14*(4), 379–393.

Basdogan, C., Ho, C.-H., Srinivasan, M. A., & Slater, M. (2000). An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *7*(4), 443–460.

Benali, K., Bourguin, G., David, B., Derycke, A., & Ferraris, C. (2002). Collaboration/Coopération. *Information-Interaction-Intelligence, Actes des deuxièmes Assises nationales du GDR I3, J. Le Maître (Ed.)*, 247–261.

Benford, S., Bowers, J., Fahlén, L. E., & Greenhalgh, C. (1994). Managing mutual awareness in collaborative virtual environments. In *Proceedings of the conference on Virtual reality software and technology* (p. 223–236).

Benford, S., Brown, C., Reynard, G., & Greenhalgh, C. (1996). Shared spaces: transportation, artificiality, and spatiality. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work* (p. 77–86). ACM.

Benford, S., Greenhalgh, C., Rodden, T., & Pycock, J. (2001). Collaborative Virtual Environments. *Communications of the ACM*, *44*(7), 79‑85.

Berthoz, A. (2005). Espace perçu, espace vécu, espace conçu. In *Les Espaces de l'homme* (p. 127). Paris: Odile Jacob.

Berthoz, A. (2013). La manipulation mentale des points de vue, un des fondements de la tolérance. In C. Ossola & B. Stock (éd.), *La pluralité interprétative : Fondements historiques et cognitifs de la notion de point de vue*. Paris: Collège de France.

Bezuidenhout, A. (2013). Perspective taking in conversation: A defense of speaker non-egocentricity. *Journal of Pragmatics*, *48*(1), 4–16.

Bourgeois, J., Farnè, A., & Coello, Y. (2014). Costs and benefits of tool-use on the perception of reachable space. *Acta psychologica*, *148*, 91–95.

Brown-Schmidt, S. (2012). Beyond common and privileged: Gradient representations of common ground in real-time language use. *Language and Cognitive Processes*, *27*(1), 62–89.

Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. *Dialog and Discourse*, *2*, 11–33.

Bryant, D. J., & Tversky, B. (1999). Mental representations of perspective and spatial relations from diagrams and models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 137.

Bryant, D. J., Tversky, B., & Lanca, M. (2001). Retrieving spatial relations from observation and memory. *Conceptual structure and its interfaces with other modules of representation*, 116–139.

Bulu, S. T. (2012). Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education*, *58*(1), 154–161.

Burigo, M., & Sacchi, S. (2013). Object Orientation Affects Spatial Language Comprehension. *Cognitive science*.

Burkhardt, J.-M. (2003a). Immersion, réalisme et présence dans la conception et l'évaluation des Environnements Virtuels. *Psychologie française*, *48*(2), 35‑42.

Burkhardt, J.-M. (2003b). Réalité virtuelle et ergonomie : quelques apports réciproques. *Le travail humain*, *66*(1), 65. http://doi.org/10.3917/th.661.0065

Burkhardt, J.-M. (2007). Immersion, représentation et coopération : discussion et perspectives de recherches empiriques pour l'ergonomie cognitive de la Réalité Virtuelle., *1*(45), 59‑87.

Burkhardt, J.-M., Détienne, F., & Wiedenbeck, S. (1997). Mental representations constructed by experts and novices in object-oriented program comprehension. In *Human-Computer Interaction INTERACT'97* (p. 339–346). Springer.

Burkhardt, J.-M., & Fuchs, P. (2006). 1 Introduction à la réalité virtuelle. In *Le traité de la réalité virtuelle*. Paris: Ecole des Mines de Paris.

Carletta, J., Hill, R. L., Nicol, C., Taylor, T., De Ruiter, J. P., & Bard, E. G. (2010). Eyetracking for two-person tasks with manipulation of a virtual world. *Behavior Research Methods*, *42*(1), 254–265.

Carlson, L. A., & Van Deman, S. R. (2008). Inhibition within a reference frame during the interpretation of spatial language. *Cognition*, *106*(1), 384–407.

Carlson-Radvansky, L. A., & Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, *46*(3), 223–244.

Carlson-Radvansky, L. A., & Radvansky, G. A. (1996). The influence of functional relations on spatial term selection. *Psychological Science*, 56–60.

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, *57*(8), 1241–1261.

Cazamian, P., Hubault, F., & Noulin, M. (1996). *Traité d'ergonomie* (Octares). Toulouse: JSTOR.

Chellali, A. (2009). *Etude des interactions homme-homme pour l'élaboration du référentiel commun dans les environnements virtuels collaboratifs*. Université de Nantes, Nantes.

Chellali, A., Milleville-Pennel, I., & Dumas, C. (2013). Influence of contextual objects on spatial interactions and viewpoints sharing in virtual environments. *Virtual Reality*, *17*(1), 1‑15. http://doi.org/10.1007/s10055-012-0214-5

Churchill, E. F., Snowdon, D., & Munro, A. J. (2001). *Collaborative virtual environments*. Springer.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on socially shared cognition*, *13*(1991), 127–149.

Convertino, G., Ganoe, C. H., Schafer, W. A., Yost, B., & Carroll, J. M. (2005). A multiple view approach to support common ground in distributed and synchronous geo-collaboration. In *Coordinated and*

Multiple Views in Exploratory Visualization, 2005.(CMV 2005). Proceedings. Third International Conference on (p. 121–132).

Cooper, G. E., & Harper Jr, R. P. (1969). *The use of pilot rating in the evaluation of aircraft handling qualities*. DTIC Document.

Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., & Thompson, W. B. (2005). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. *Perception*, *34*(2), 191–204.

Daniel, M.-P., & Denis, M. (1998). Spatial descriptions as navigational aids: A cognitive analysis of route directions. *Kognitionswissenschaft*, *7*(1), 45–52.

Darses, F., & Hoc, J.-M. (2004). *Psychologie ergonomique: tendances actuelles*. Presses Universitaires de France « Le Travail humain ».

Denis, M. (1982). Représentation imagée et résolution de problèmes. *Revue française de pédagogie*, *60*(1), 19‑29. http://doi.org/10.3406/rfp.1982.1748

Denis, M. (1994). *Image et cognition* (2e édition). Presses Universitaires de France.

Denis, M. (2012). *La psychologie cognitive*. [Paris]: Éditions de la Maison des sciences de l'homme.

De Terssac, G., & Chabaud, C. (1990). Référentiel opératif commun et fiabilité. *Les facteurs humains de la fiabilité dans les systèmes complexes*, 111–139.

Detienne, F., Barcellini, F., & Burkhardt, J.-M. (2007). Une approche multidimensionnelle de la participation dans les communautés de développement de logiciels libres. In *Colloque de l'Association pour la Recherche Cognitive-ARCo'07: Cognition–Complexité–Collectif* (p. 279–290).

Diessel, H. (2014). Demonstratives, Frames of Reference, and Semantic Universals of Space. *Language and Linguistics Compass*, *8*(3), 116–132.

Dumazeau, C., & Karsenty, L. (2008). Communications distantes en situation de travail : favoriser l'établissement d'un contexte mutuellement partagé. *Le travail humain*, *71*(3), 225‑252.

Duran, N. D., Dale, R., & Kreuz, R. J. (2011). Listeners invest in an assumed other's perspective despite cognitive cost. *Cognition*, *121*(1), 22–40.

Durlach, N., & Slater, M. (2000). Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators & Virtual Environments*, *9*(2), 214–217.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64.

Farnè, A., Serino, A., & Làdavas, E. (2007). Dynamic size-change of peri-hand space following tool-use: determinants and spatial characteristics revealed through cross-modal extinction. *Cortex*, *43*(3), 436–443.

Fuchs, P., Moreau, G., & Papin, J.-P. (2006). *Le traité de la réalité virtuelle*. Presses de l'Ecole des Mines.

Galati, A., & Avraamides, M. N. (2014). Social and Representational Cues Jointly Influence Spatial Perspective-Taking. *Cognitive science*.

Galati, G., Pelle, G., Berthoz, A., & Committeri, G. (2010). Multiple reference frames used by the human brain for spatial perception and memory. *Experimental brain research*, *206*(2), 109–120.

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., & Sasse, M. A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (p. 529–536).

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*(2), 292–304.

Gergle, D., Kraut, R. E., & Fussell, S. R. (2004). Action as language in a shared visual space. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (p. 487–496).

Gerhard, M., Moore, D., & Hobbs, D. (2001). Continuous presence in collaborative virtual environments: Towards a hybrid avatar-agent model for user representation. In *Intelligent Virtual Agents* (p. 137–155).

Giboin, A. (2004). Chapitre 6. La construction de référentiels communs dans le travail coopératif. *Le Travail humain*, 119–139.

Goebbels, G., Lalioti, V., & Göbel, M. (2003). Design and evaluation of team work in distributed collaborative virtual environments. In *Proceedings of the ACM symposium on Virtual reality software and technology* (p. 231–238).

Goschler, J., Andonova, E., & Ross, R. J. (2008). Perspective use and perspective shift in spatial dialogue. In *Spatial Cognition VI. Learning, Reasoning, and Talking about Space* (p. 250–265). Springer.

Gronier, G. (2006). *Psychologie ergonomique du travail collectif assisté par ordinateur: l'utilisation du collecticiel dans les projets de conception de produits*. Université de Franche-Comté.

Gronier, G., & Sagot, J.-C. (2007). Le rôle des communications dans les projets de conception de produits: Comparaison des situations de travail collectif en présence et à distance. In *Colloque de l'Association pour la Recherche Cognitive-ARCo'07: Cognition–Complexité–Collectif* (p. 291–302).

Gutwin, C., & Greenberg, S. (2004). The importance of awareness for team cognition in distributed collaboration. *Team cognition: Understanding the factors that drive process and performance*, *201*, 1–33.

Hagège, C. (2005). Espace et cognition à la lumière des choix faits par les langues humaines. In *Les espaces de l'homme: Symposium annuel* (p. 241). Odile Jacob.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 50, p. 904–908). Sage Publications.

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition*, *55*(1), 39–84.

Head, H., & Holmes, G. (1911). Sensory disturbances from cerebral lesions. *Brain*, *34*(2-3), 102–254.

Hoc, J.-M. (1998). L'ergonomie cognitive: un compromis nécessaire entre des approaches centres sur la machine et des approches centrées sur l'homme. In *Actes du colloque «Recherche Ergonomie*.

Hoc, J.-M. (2001). Towards a cognitive approach to human–machine cooperation in dynamic situations. *International Journal of Human-Computer Studies*, *54*(4), 509–540.

Hoc, J.-M. (2003). Coopération humaine et systèmes coopératifs. *Ingénerie Cognitive. France: Lavoisier*.

Holmes, D. L., Cohen, K. M., Haith, M. M., & Morrison, F. J. (1977). Peripheral visual processing. *Perception & Psychophysics*, *22*(6), 571–577.

Honoré, J., Richard, C., & Mars, F. (2002). Perception de l'espace du corps et action. In *Percevoir, s' orienter et agir dans l'espace: approche pluridisciplinaire des relations perception-action* (p. 135–146).

Hubault, F., Noulin, M., & Rabit, M. (1996). L'analyse du travail en ergonomie. In *Traité d'ergonomie* (Octares, p. 289‑309). Toulouse.

Iachini, T., Coello, Y., Frassinetti, F., & Ruggiero, G. (2014). Body Space in Social Interactions: A Comparison of Reaching and Comfort Distance in Immersive Virtual Reality. *PLoS ONE*, *9*(11). http://doi.org/10.1371/journal.pone.0111511

Kaltner, S., Jansen, P., & Riecke, B. E. (2014). Embodied mental rotation: a special link between egocentric transformation and the bodily self. *Cognition*, *5*, 505.

Karsenty, L., & Falzon, P. (1993). L'analyse des dialogues orientés-tâche: Introduction à des modèles de la communication. *Les aspects collectifs du travail. Toulouse: Octarès*.

Kessler, K., & Rutherford, H. (2010). The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. *Frontiers in psychology*, *1*, 108. http://doi.org/10.3389/fpsyg.2010.00213

Keysar, B., Barr, D. J., Balin, J. A., & Paek, T. S. (1998). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language*, *39*(1), 1–20.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological bulletin*, *129*(5), 723.

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, *29*(5), 745–756.

Krauss, R. M., & Fussell, S. R. (1991). Perspective-taking in communication: Representations of others' knowledge in reference. *Social Cognition*, *9*(1), 2–24.

Kuan, S.-T., Chang, W.-J., Wang, M., Tsai, W.-L., Hsu, W., Yen, J.-R., & Ho, H. (2007). Investigation on correlation between mental workload and situation awareness of pilots by simulation experiment.

Kvan, T. (2000). Collaborative design: what is it? *Automation in construction*, *9*(4), 409–415.

Lackner, J. R., & Dizio, P. (1998). Spatial Orientation as a Component of Presence: Insights Gained from Nonterrestrial Environments. *Presence: Teleoperators & Virtual Environments*, *7*(2), 108‑115.

Latham, K., & Whitaker, D. (1996). A Comparison of Word Recognition and Reading Performance in Foveal and Peripheral Vision *. *Vision Research*, *36*(17), 2665‑2674. http://doi.org/10.1016/0042-6989(96)00022-3

Leplat, J. (1991). Activités collectives et nouvelles technologies. *Revue internationale de Psychologie sociale*, *4*(3/4), 335–356.

Leplat, J., & Sperandio, J.-C. (1967). La mesure de la charge de travail par la technique de la tâche ajoutée. *L'Année psychologique*, *67*(1), 255–277.

Levinson, S. C. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. *Language and space*, 109–169.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity* (Vol. 5). Cambridge University Press.

Linebarger, J. M., & Kessler, G. D. (2002). The effect of avatar connectedness on task performance. *Lehigh Univ TR*.

Loomis, J. M., Klatzky, R. L., Avraamides, M., Lippa, Y., & Golledge, R. G. (2007). Functional equivalence of spatial images produced by perception and spatial language. In *Spatial processing in navigation, imagery and perception* (p. 29–48). Springer.

Lund, K., Rossetti, C., & Metz, S. (2006). Les facteurs internes à la coopération, influencent-ils l'activité médiatisée à distance. In *Premières journées communication et apprentissages instrumentés en réseau* (p. 310–329).

Luyat, M. (2009). *La perception*. Paris: Dunod.

MacMillan, J., Entin, E. E., & Serfaty, D. (2004). Communication overhead: The hidden cost of team cognition. *Team cognition: Process and performance at the interand intra-individual level. American Psychological Association, Washington, DC.*

Majid, A., Bowerman, M., Kita, S., Haun, D., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in cognitive sciences*, *8*(3), 108–114.

Mars, F., Honoré, J., Richard, C., & Coquery, J.-M. (1998). Effects of an illusory orientation of the head on straight-ahead pointing movements. *Current Psychology of Cognition*, *17*(4-5), 749–762.

McNamara, T. P. (2003). How are the locations of objects in the environment represented in memory? In *Spatial cognition III* (p. 174–191). Springer.

Metzler, J., & Shepard, R. N. (1974). Transformational studies of the internal representation of three-dimensional objects.

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception & psychophysics*, *68*(2), 327–337.

Mishra, R. C., Dasen, P. R., & Niraula, S. (2003). Ecology, language, and performance on spatial cognitive tasks. *International Journal of Psychology*, *38*(6), 366–383.

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 162.

Müller, R., Helmert, J. R., Pannasch, S., & Velichkovsky, B. M. (2013). Gaze transfer in remote cooperation: Is it always helpful to see what your partner is attending to? *The Quarterly Journal of Experimental Psychology*, *66*(7), 1302–1316.

Navarro, C. (2001). Partage de l'information en situation de coopération à distance et nouvelles technologies de la communication: bilan de recherches récentes. *Le travail humain*, *64*(4), 297–319.

Newlands, A., Anderson, A., Thomson, A., Ion, B., & Dickson, N. (2002). Using Speech Related Gestures to Aid Referential Communication in Face-to-face and Computer-Supported Collaborative Work. In *Proceedings of the First congress of the International Society for Gesture Studies tool place* (p. 5–8).

Nova, N. (2004). *Socio-cognitive functions of space in collaborative settings: a literature review about Space, Cognition and Collaboration*.

Nowak, K. L. (2001). Defining and differentiating copresence, social presence and presence as transportation. In *Presence 2001 Conference, Philadelphia, PA* (p. 686‑690).

Osiurak, F., Morgado, N., & Palluel-Germain, R. (2012). Tool use and perceived distance: when unreachable becomes spontaneously reachable. *Experimental Brain Research*, *218*(2), 331‑339. http://doi.org/10.1007/s00221-012-3036-5

Ott, D., & Dillenbourg, P. (2001). Using Proximity and View Awareness to Reduce Referential Ambiguity in a Shared 3D Virtual Environment. In *Proceedings of CSCL 2001*.

Paillard, J. (1991). Motor and representational framing of space. *Brain and space*, 163–182.

Pallascio, R., Talbot, L., Allaire, R., & Mongeau, P. (1990). L'incidence de l'environnement sur la perception et la représentation d'objets géométriques. *Revue des sciences de l'éducation*, *16*(1), 77–90.

Pavard, B., & Karsenty, L. (1997). Différents niveaux d'analyse du contexte dans l'étude ergonomique du travail collectif. *Réseaux*, *15*(85), 73–99.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test-different versions and factors that affect performance. *Brain and cognition*, *28*(1), 39–58.

Pylyshyn, Z. W. (1979). The rate of « mental rotation » of images: A test of a holistic analogue hypothesis. *Memory & Cognition*, *7*(1), 19–28.

Rabardel, P., & Beguin, P. (2005). Instrument mediated activity: from subject development to anthropocentric design. *Theoretical Issues in Ergonomics Science*, *6*(5), 429–461.

Raufaste, E., Daurat, A., Mélan, C., & Ribert-van de Weerdt, C. (2004). Chapitre 8. Aspects intensifs de la cognition en situation de travail. *Le Travail humain*, 175–199.

Retz-Schmidt, G. (1988). Various views on spatial prepositions. *AI magazine*, *9*(2), 95.

Richard, J.-F. (1990). Les activités mentales. *Comprendre, raisonner, trouver des solutions. Paris: Armand Colin*.

Riva, G., Davide, F., & IJsselsteijn, W. A. (2003). Being There: The experience of presence in mediated environments. *Being there: Concepts, effects and measurement of user presence in synthetic environments*, 5.

Roberts, D., Wolff, R., Otto, O., & Steed, A. (2003). Constructing a Gazebo: supporting teamwork in a tightly coupled, distributed task in virtual reality. *Presence: Teleoperators & Virtual Environments*, *12*(6), 644–657.

Roberts, R. J., & Aman, C. J. (1993). Developmental differences in giving directions: spatial frames of reference and mental rotation. *Child development*, *64*(4), 1258–1270.

Rodriguez, N., Jessel, J.-P., & Torguet, P. (2002). A virtual reality tool for teleoperation research. *Virtual Reality*, *6*(2), 57–62.

Roger, M., Bonnardel, N., & Le Bigot, L. (2011). Landmarks' use in speech map navigation tasks. *Journal of Environmental Psychology*, *31*(2), 192–199.

Roger, M., Knutsen, D., Bonnardel, N., & Le Bigot, L. (2013). Landmark Frames of Reference in Interactive Route Description Tasks. *Applied Cognitive Psychology*, *27*(4), 497‑504. http://doi.org/10.1002/acp.2927

Sallnäs, E.-L., Rassmus-Gröhn, K., & Sjöström, C. (2000). Supporting presence in collaborative environments by haptic force feedback. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *7*(4), 461–476.

Schafer, W. A., & Bowman, D. A. (2004). Evaluating the effects of frame of reference on spatial collaboration using desktop collaborative virtual environments. *Virtual Reality*, *7*(3), 164–174.

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, *47*(1), 1–24.

Schober, M. F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations? *Discourse Processes*, *20*(2), 219–247.

Schober, M. F. (1996). Addressee-and object-centered frames of reference in spatial descriptions. In *American Association for Artificial Intelligence, Working Notes of the 1996 AAAI Spring Symposium on Cognitive and Computational Models of Spatial Representation* (Vol. 47, p. 92–100).

Schober, M. F. (2009). Spatial dialogue between partners with mismatched abilities. *Spatial language and dialogue*, *1*, 23–39.

Schroeder, R., Steed, A., Axelsson, A.-S., Heldal, I., Abelin, A., Wideström, J., … Slater, M. (2001). Collaborating in networked immersive spaces: as good as being there together? *Computers & Graphics*, *25*(5), 781–788.

Schuemie, M. J., Van Der Straaten, P., Krijn, M., & Van Der Mast, C. A. P. G. (2001). Research on presence in virtual reality: A survey. *CyberPsychology & Behavior*, *4*(2), 183–201.

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, *43*(4), 274–310.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701‑703.

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence*, *8*(5), 560–565.

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.

Slater, M., & Usoh, M. (1993). Presence in immersive virtual environments. In *, 1993 IEEE Virtual Reality Annual International Symposium, 1993* (p. 90‑96). http://doi.org/10.1109/VRAIS.1993.380793

Slater, M., Usoh, M., & Chrysanthou, Y. (1995). *The influence of dynamic shadows on presence in immersive virtual environments*. Springer.

Spante, M., Schroeder, R., Axelsson, A.-S., & Christie, M. (2004). How Putting Yourself into the Other Persons Virtual Shoes Enhances Collaboration. In *Proceeding of the 7th International Workshop on Presence, Valencia, Spain* (p. 190–196).

Sperandio, J.-C. (1996). L'ergonomie face aux changements technologiques et organisationnels du travail humain. *Octares, Toulouse, France*.

Steed, A., Slater, M., Sadagic, A., Bullock, A., & Tromp, J. G. (1999). Leadership and collaboration in shared virtual environments. In *Virtual Reality, 1999. Proceedings., IEEE* (p. 112–115).

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, *42*(4), 73–93.

Tamborello, F. P., Sun, Y., & Wang, H. (2012). Spatial reasoning with multiple intrinsic frames of reference. *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)*, *59*(1), 3–10.

Tenbrink, T., Coventry, K. R., & Andonova, E. (2011). Spatial strategies in the description of complex configurations. *Discourse Processes*, *48*(4), 237–266.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, *28*(5), 675–690.

Tversky, B. (2005a). Functional significance of visuospatial representations. *Handbook of higher-level visuospatial thinking*, 1–34.

Tversky, B. (2005b). La cognition spatiale: incarnée et désincarnée. In *Les Espaces de l'homme* (p. 161–184). Paris: Odile Jacob.

Tversky, B. (2008). Spatial Cognition Embodied and Situates.

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*(1), 124–129.

Vacherand-Revel, J. (2007). Enjeux de la médiatisation du travail coopératif distribué dans les équipes de projets de conception. *Pistes*, *9*(2).

Verjat, I. (1994). Confrontation de deux approches de la localisation spatiale. *L'année psychologique*, *94*(3), 403‑423. http://doi.org/10.3406/psy.1994.28774

Wang, X., & Tsai, J. J.-H. (Éd.). (2011). *Collaborative Design in Virtual Environments* (Vol. 48). Springer Science+ Business Media.

Watson, M. E., Pickering, M. J., & Branigan, H. P. (2004). Alignment of reference frames in dialogue. In *Proceedings of the 26th annual conference of the Cognitive Science Society* (p. 2353–2358). Lawrence Erlbaum Mahwah, NJ.

Wickens, C. D. (2008). Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 397–403.

Witt, J. K. (2011). Action's effect on perception. *Current Directions in Psychological Science*, *20*(3), 201–206.

Zacklad, M. (2003). Approche cognitive et sociale de la coopération pour guider la conception de systèmes de CSCW. In *Actes de la Conférence Epique* (p. 149–162).

Zacks, J. M., & Michelon, P. (2005). Transformations of visuospatial images. *Behavioral and Cognitive Neuroscience Reviews*, *4*(2), 96–118.

Zahorik, P., & Jenison, R. L. (1998). Presence as being-in-the-world. *Presence*, *7*(1), 78–89.

Zhao, S. (2003). Toward a taxonomy of copresence. *Presence: Teleoperators & Virtual Environments*, *12*(5), 445–455.

Liste des annexes

Annexe 1 : Version du questionnaire NASA-TLX utilisée pour la première partie expérimentale..... 159
Annexe 2 : Fiche technique pour la première partie expérimentale 161
Annexe 3 : Explication du rôle d'expert conformité (manipulateur) 163
Annexe 4 : Explication du rôle de responsable d'atelier (guide)..... 164
Annexe 5: Questionnaire pour la première partie expérimentale..... 165
Annexe 6 : Version du Santa Barbara Sense of Scale utilisé dans la première partie expérimentale 170
Annexe 7 : Exemple de ligne de test du MRT-A 171
Annexe 8 : Questionnaire pour l'étude 1 de la seconde partie expérimentale 172
Annexe 9 : Consignes données aux participants de l'étude 1 de la seconde partie expérimentale... 173
Annexe 10 : Table des résultats des analyses post-hoc (Tukey HSD) réalisées sur l'exigence mentale en fonction de l'orientation de l'avatar et la position de la cible par rapport à l'avatar. 175
Annexe 11 : Table des résultats des analyses post-hoc (Tukey HSD) réalisées sur les temps de préparation des énoncés en fonction de l'orientation de l'avatar et la position de la cible par rapport à l'avatar. 175
Annexe 12 : Table des résultats des analyses post-hoc (Tukey HSD) réalisées sur les temps de production des énoncés en fonction de l'orientation de l'avatar et la position de la cible par rapport à l'avatar. 176
Annexe 13 : Question ajoutée au questionnaire de l'étude 1 de la seconde partie expérimentale (Annexe 8) 177
Annexe 14 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des conditions d'indiçage et la position de la cible par rapport au participant 178
Annexe 15 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des conditions d'indiçage et l'orientation de l'avatar. 178
Annexe 16 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des conditions d'indiçage et la position de la cible par rapport à l'avatar 179
Annexe 17 : Analyses post-hoc (Tukey HSD) pour les valeurs de temps de production des énoncés en fonction des conditions d'indiçage et la position de la cible par rapport au participant. 180
Annexe 18 : Analyses post-hoc (Tukey HSD) pour les valeurs de temps de production des énoncés en fonction des conditions d'indiçage et la position de la cible par rapport à l'avatar - Etude 2 Production d'énoncés avec indices. Cases rouges : p<.05 181
Annexe 19 : Analyses post-hoc (Tukey HSD) pour les valeurs de nombre de mots en fonction des conditions d'indiçage et la position de la cible par rapport au participant. 182
Annexe 20 : Analyses post-hoc (Tukey HSD) pour les valeurs de nombre de mots en fonction des conditions d'indiçage et la position de la cible par rapport à l'avatar. Etude 2 - Production d'énoncés avec indices. Cases rouges : p<.05 183
Annexe 21 : Questions spécifiques de l'étude 3 de la seconde partie expérimentale (Annexe 8)..... 184
Annexe 22 : Consignes données aux participants de l'étude 1 de la seconde partie expérimentale. 185
Annexe 23 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des cadres de référence et l'orientation du participant. 187
Annexe 24 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des cadres de référence et la position de la cible..... 187

Annexe 25 : Exigence mentale moyenne pour chaque orientation du participant, chaque cadre de référence et les huit cibles.. 189
Annexe 26 : Analyses post-hoc (Tukey HSD) pour les valeurs du temps de compréhension en fonction des cadres de référence et l'orientation du participant. 190
Annexe 27 : Analyses post-hoc (Tukey HSD) pour les valeurs du temps de compréhension en fonction des cadres de référence et de la position de la cible. 190
Annexe 28 Temps moyen de compréhension pour chaque orientation du participant, chaque cadre de référence et les huit cibles... 191
Annexe 29 : Analyses post-hoc (Tukey HSD) pour les temps de sélection de la cible en fonction des cadres de référence et la position de la cible..... 192
Annexe 30 : Temps moyen de sélection pour chaque orientation du participant, chaque cadre de référence et les huit cibles.. 193

Annexe 1 : Version du questionnaire NASA-TLX utilisée pour la première partie expérimentale

Code participant :

Pour chaque item tracez un trait à l'emplacement qui vous semble le plus approprié.

1:
EXIGENCE MENTALE : Dans quelle mesure des opérations mentales et perceptives ont-elles été requises (par ex.: penser, décider, calculer, se rappeler, regarder, chercher, etc.) ? La tâche était-elle plutôt facile ou difficile, simple ou complexe, abordable ou exigeante ?

Faible (0) Forte (100)

2:
EXIGENCE PHYSIQUE : Dans quelle mesure des opérations physiques ont-elles été requises (par ex.: pousser, tirer, tourner, superviser, activer, etc.) ? Avez-vous trouvé la tâche plutôt facile ou difficile, lente ou rapide, lâche ou vigoureuse, reposante ou ardue ?

Faible (0) Forte (100)

3:
EXIGENCE TEMPORELLE : Quelle quantité de pression liée au temps à cause du rythme ou de l'allure des tâches ou de l'apparition des éléments de la tâche avez-vous ressenti ? L'allure était-elle lente et tranquille ou rapide et frénétique ?

Faible (0) Forte (100)

4:
EFFORT : Quelle a été la difficulté d'accomplir (mentalement et physiquement) la tâche avec un niveau de performance tel que le votre ?

Faible (0) Forte (100)

5:
PERFORMANCE : Quelle réussite vous attribuez-vous en ce qui concerne l'atteinte des buts de la tâche fixés par l'expérimentateur (ou par vous-même) ? Dans quelle mesure êtes-vous satisfait de votre performance dans l'accomplissement de ces buts ?

Bonne (0) Faible (100)

6:
FRUSTRATION : Au cours de la tâche, quel sentiment d'être peu sûr de vous, découragé, irrité, stressé et agacé avez-vous ressenti contrairement au fait d'être sûr de vous, satisfait, content, détendu et complaisant ?

Faible (0) Forte (100)

Code participant :

Entourez l'item qui contribue le plus, selon vous, à la charge mentale que vous avez ressenti durant la session collaborative.

1. Effort ou performance

2. Exigence temporelle ou frustration

3. Exigence temporelle ou effort

4. Exigence physique ou frustration

5. Performance ou frustration

6. Exigence physique ou exigence temporelle

7. Exigence physique ou performance

8. Exigence temporelle ou exigence mentale

9. Frustration ou effort

10. Performance ou exigence mentale

11. Performance ou exigence temporelle

12. Exigence mentale ou effort

13. Exigence mentale ou exigence physique

14. Effort ou exigence physique

15. Frustration ou exigence mentale

Changement de Mode : Toujours pointer la manette vers l'écran, droit devant, bras bien tendu comme sur la Figure 1.

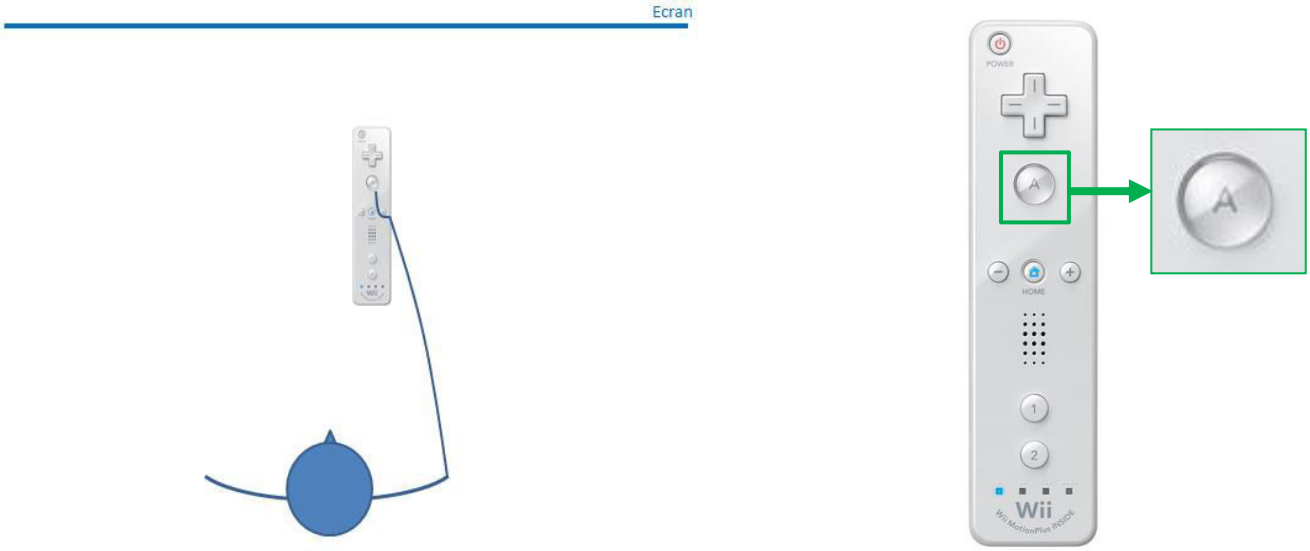


Figure 1 : bras tendu, viser l'écran et appuyer sur Home ou A

Déplacements :

Vous êtes en mode Vol, vous pouvez vous déplacer avec les flèches.
Bien mettre la manette vers l'écran (Figure 2)

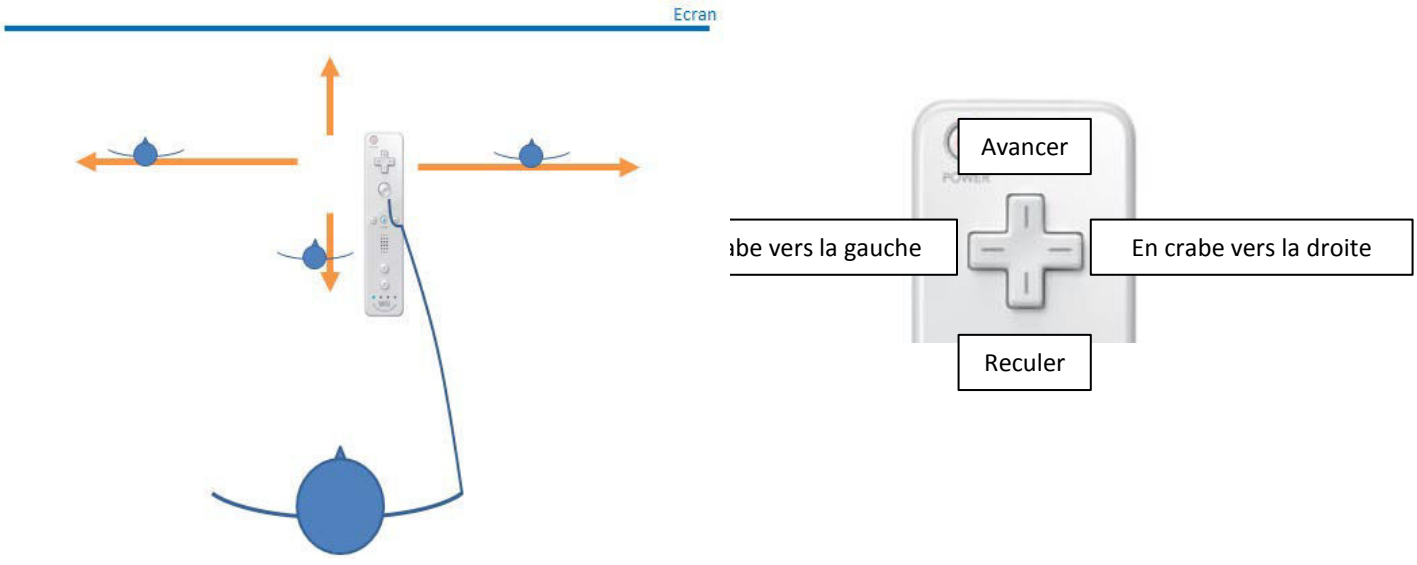
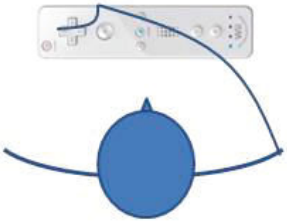


Figure 2 : Déplacement avec les flèches : on reste toujours dans la même direction

Pour tourner : tourner la manette dans la direction où vous souhaitez aller et appuyer sur la flèche tout droit (comme ci-dessous).







Pour aller vers le haut ou le bas = comme pour tourner → mettre la manette vers la direction dans laquelle vous souhaitez aller = vers le haut dans le monte-charge par exemple, puis appuyez sur la flèche tout droit

Sélectionner un objet :

Appuyez sur  , apparaît alors le menu suivant :



Sélectionner le Mode Main :



Le Mode Main permet de saisir les objets et de les déplacer à l'aide de la manette.
Pour cela : une fois en Mode Main :
Viser l'objet que vous souhaitez déplacer → 1er Coup de Gâchette = l'objet est sélectionné [Les bords de l'objet sont alors orange] et un repère apparaît
Pour saisir l'objet : placer le repère dans l'objet sélectionné → 2ème Coup de Gâchette = l'objet est saisi = il suit les mouvements de la manette
En refaisant un 3ème Coup de Gâchette → vous lâchez l'objet

Conseil : pour vous déplacer avec l'objet plus facilement : saisissez l'objet assez bas (vous pourrez porter l'objet plus haut) et avec la manette bien parallèle au sol et perpendiculaire à l'écran.

Annexe 3 : Explication du rôle d'expert conformité (manipulateur)

Ces informations vous sont données uniquement à vous

Expert Conformité

Vous êtes expert conformité chez Airbus : vous intervenez dans la gestion des problèmes de sureté. Vous êtes basé sur Toulouse.

Votre chef vient de vous apprendre que deux pièces montées sur un des avions en cours d'assemblage à Nantes sont suspectées d'être endommagées et **de mettre en danger la sécurité des passagers** : une série de tests doit être mise en place pour vérifier la conformité de ces pièces (résistance à la chaleur, aux vibrations et aux chocs).

Le changement d'une pièce coûte 500 000€. Les tests sont fiables à 100%, il est donc nécessaire de venir tester les pièces pour décider si les pièces doivent être changées ou non.

Pour réaliser les tests, vous allez devoir placer un gros équipement (décrit ci-après) juste en face de sorties de tuyau (photos ci-jointes), situées à deux endroits différents dans l'avion (croix rouges sur le plan).

Cet équipement est très fragile et ne peut être porté (c'est-à-dire qu'il ne peut pas passer par un escalier) : il se déplace sur roulettes (possibilité de monter par un plan légèrement incliné). Celles-ci permettent une rotation autour de l'axe vertical.

La venue sur site pour réaliser ces tests **immobilise tout l'atelier**, il est donc **impératif** de trouver, avant votre venue, la meilleure solution pour exécuter le test **le plus rapidement possible**.

Votre chef vous a donné un plan de l'atelier de Nantes, ci-joint, avec l'emplacement des deux pièces à tester (croix rouges). Vous devez réaliser les mesures dans l'ordre, donc accéder au point de mesure situé au niveau 1 puis au point de mesure situé au niveau 2.

Afin d'anticiper votre venue sur le site de Nantes, vous allez participer à une session de travail distante avec un collaborateur : il s'agit du responsable de l'atelier de Nantes. Vous allez pouvoir préparer cette visite, depuis Toulouse, et mettre en place la procédure qui sera appliquée lors de votre venue réelle sur site (dans une semaine).

Vous n'avez pas besoin de savoir quelles seront les manipulations réelles pour effectuer les tests, vous devez, dans cette session, uniquement transporter l'équipement à l'emplacement où ils auront lieu.

Avant de commencer la session collaborative, vous allez pouvoir vous entraîner : l'objectif de cette session d'entraînement est que vous maitrisiez l'outil de réalité virtuelle et preniez connaissance de l'outillage que vous allez devoir déplacer. Vous serez le seul à pouvoir manipuler les objets au cours de la session.

Le but de cette session en réalité virtuelle, entre Toulouse (là où vous êtes) et Nantes (là où est le responsable atelier)**, est de trouver la procédure la plus rapide et la moins chère à mettre en place avant votre venue, réelle, sur le site de Nantes.**

Voici des informations utiles concernant l'équipement :

Dimensions : Largeur : 0,7 m – Longueur : 1 m – Hauteur : 1,70 m (roulettes comprises)

Poids : 30 kg

Voici la liste des opérations à réaliser dans la phase d'entraînement :

- sélectionner l'équipement
- se déplacer, seul et avec l'équipement
- ouvrir et fermer des barrières

Annexe 4 : Explication du rôle de responsable d'atelier (guide)

Ces informations vous sont données uniquement à vous

Responsable Atelier

Vous travaillez chez Airbus en tant que responsable atelier du site de production d'avions de Nantes. Plusieurs pièces montées sur un avion de votre site sont suspectées d'être endommagées et de mettre en danger la sécurité des passagers : une série de tests doit être mise en place pour vérifier la conformité de ces pièces.

Pour réaliser les tests, un expert conformité de chez Airbus va venir sur votre site. **Les tests à réaliser immobilisent tout l'atelier**, il est donc **impératif** de trouver la meilleure solution pour exécuter ces tests **le plus rapidement possible**. Une heure d'immobilisation coûte environ 20 000€.

Afin d'anticiper cette venue sur site, vous allez participer à une session de travail distante avec l'expert conformité Airbus qui viendra faire les tests dans une semaine.

Le but de cette session en réalité virtuelle, entre Nantes (là où vous êtes) et Toulouse (là où est l'expert conformité)**, est de trouver la procédure la plus rapide et la moins chère à mettre en place avant sa venue. réelle. sur le site de Nantes.**

Voici les informations qui vous seront utiles pendant la session :

Les armoires gris clair sont immobiles et **ne peuvent pas être déplacées**.

Un portique vert (Niveau 1) peut être ouvert selon des caractéristiques propres (expliquées pendant la phase d'entrainement), et son ouverture coûte 1 000 €.

Les armoires gris foncé (Niveau 2) peuvent être bougées MAIS uniquement en cas d'extrême nécessité : le fait de les déplacer nécessite une opération de débranchement extrêmement coûteuse ! (Elles permettent de réaliser une série de tests importants à cette phase de construction de l'avion.) Il faut avoir éliminé toutes les autres options pour valider leur déplacement (votre chef vous demandera des justificatifs en cas de décision de déplacement de ces armoires). Cette opération prend environ 30 minutes avec 4 opérateurs et coûte environ 50 000€ (main d'œuvre incluse).

Vous avez à votre disposition des rampes d'accès pour chacun des accès avec marche (pour monter uniquement une ou deux marches : pas un escalier entier). Leur installation prend 15 minutes et coûte environ 1 000€ (main d'œuvre incluse). Elles ne sont pas représentées dans l'environnement virtuel, il suffit que vous sachiez où les disposer.

Avant de commencer la session collaborative, vous allez pouvoir vous entraîner : l'objectif de cette session d'entraînement est que vous connaissiez bien l'environnement virtuel de la session. Pendant la phase d'entraînement, vous devez vous déplacer dans l'environnement et être attentif à la circulation dans l'espace. Nous vous demandons de vous intéresser uniquement au premier avion de l'atelier.

Juste après l'entrainement, il vous sera demandé de faire un plan de l'environnement par niveaux, en précisant pour les accès : l'encombrement devant les accès, le nombre de marches et le coût pour passer par chaque accès.

Annexe 5: Questionnaire pour la première partie expérimentale

Questionnaire Participants :

Veuillez remplir le questionnaire suivant avec la plus grande assiduité.

Si la question n'a pas de sens pour vous, n'hésitez pas à demander plus de précisions.

Pour les questions du type :

Quel niveau de fatigue ressentez-vous après avoir réalisé cette tâche ?

Veuillez placer un curseur sur la ligne, comme sur l'exemple :

Nul ——————|—————— Très fort

Nom : _____ Prénom : _____

Tel : _____ Adresse mail : _____

Sexe : ☐ Masculin ☐ Féminin **Age : _____ Profession :** _____

Vous êtes : ☐ Gaucher ☐ Droitier ☐ Ambidextre

1. Quel niveau de fatigue ressentez-vous après avoir réalisé cette tâche ?
   Nul —————————————— Très fort
2. Avez-vous des douleurs (maux de têtes) ou sensations désagréables (nausées, vertiges,…) suite à l'exercice ?
   ☐ Oui ☐ Non

Si oui : Quels symptômes : _____

3. Comment qualifieriez-vous l'utilité de la phase d'entraînement pour la réalisation de la tâche ?
   Nulle —————————————— Très forte
4. Vous sentiez-vous suffisamment à l'aise après la phase d'entraînement pour réaliser l'exercice ?
   ☐ Oui totalement ☐ Oui partiellement ☐ Non
5. Auriez-vous eu besoin de plus de temps d'entrainement pour vous sentir prêt pour réaliser l'exercice ?
   ☐ Oui ☐ Non
6. Auriez-vous eu besoin de plus de temps pour vous entraîner pour maîtriser la manette ?
   ☐ Oui ☐ Non
7. Auriez-vous eu besoin de plus de temps pour vous entraîner pour maitriser le changement de fonctionnalités ?
   ☐ Oui ☐ Non
8. Auriez-vous eu besoin de plus de temps pour vous entraîner pour vous repérer dans l'espace ?
   ☐ Oui ☐ Non
10. Durant l'exercice aviez-vous la sensation d'être dans une situation réelle ?
    ☐ Oui ☐ Non
11. Comment était votre sensation « d'être là» dans l'environnement ?
    Nulle —————————————— Très forte
12. Comment était votre perception (conscience) de **l'environnement réel** (bruits, autres personnes présentes, mobilier,…) ?
    Nulle —————————————— Très forte
13. Avez-vous touché l'écran ?

☐ Oui ☐ Non

Si oui, précisez combien de fois et quand (entrainement ou session collaborative) : _____

14. Comment noteriez-vous votre distractibilité par des bruits ou événements survenus dans le **monde réel** ?

Nulle ——— Très forte

15. Comment noteriez-vous le sentiment que vous avez eu que votre partenaire était présent avec vous dans le **monde réel** ?

Nul ——— Très fort

Pourquoi ?_____

Comment noteriez-vous le sentiment que vous avez eu que votre partenaire était présent avec vous dans **l'Environnement Virtuel** ?

Nul ——— Très fort

Pourquoi ?_____

Avez-vous eu la sensation de réaliser la même chose (même type d'actions) que votre partenaire ?

☐ Oui tout le temps ☐ Oui parfois ☐ Rarement ☐ Jamais

16. Comment était votre sentiment de contrôler la situation ?

Nul ——— Très fort

17. Comment était votre sentiment que votre partenaire contrôlait la situation ?

Nul ——— Très fort

18. Qui de vous deux contrôlait le plus la situation ?

☐ Vous ☐ Votre partenaire ☐ De manière équivalente

19. Comment était votre sentiment de collaborer véritablement avec votre partenaire ?

Nul ——— Très fort

20. Comment était votre sentiment que vos actions étaient dépendantes de votre partenaire ?

Nul ——— Très fort

21. Comment était votre sentiment que les actions de votre partenaire étaient dépendantes de vous ?

Nul ——— Très fort

22. Comment était votre sentiment que votre partenaire essayait de vous aider ?

Nul ——— Très fort

23. Comment était votre sentiment d'aider votre partenaire ?

Nul ——— Très fort

24. Concernant les prises de décision, selon vous, elles étaient prises :

☐ Par vous uniquement ☐ Par votre partenaire uniquement

☐ Conjointement parfois ☐ Toujours conjointement

25. Sur quel(s) élément(s) vous êtes-vous basé(e) pour vous orienter dans l'espace ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)

☐ Votre partenaire (personne physique)

☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar

☐ Autre : _____

26. Sur quel(s) élément(s) vous êtes-vous basé(e) pour orienter votre partenaire ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)

☐ Votre partenaire (personne physique)

☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar

☐ Autre : _____

27. Sur quel(s) élément(s) vous êtes-vous basé(e) pour réaliser la tâche ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)
☐ Votre partenaire (personne physique)
☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar
☐ Autre : _____

28. Sur quel(s) élément(s) vous êtes-vous basé(e) pour comprendre les actions de votre partenaire ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)
☐ Votre partenaire (personne physique)
☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar
☐ Autre : _____

29. Vous diriez que **la communication** avec votre partenaire, était

Nulle _____ Très forte

30. Comment noteriez-vous votre **niveau de compréhension** de **ce que vous disait votre partenaire** ?

Nul _____ Très fort

31. Selon vous, quel était le **niveau de compréhension** de votre partenaire concernant **ce que vous lui disiez** ?

Nul _____ Très fort

32. Comment était **le niveau de compréhension** entre vous et votre partenaire **concernant les actions à effectuer** dans l'Environnement Virtuel ?

Nul _____ Très fort

33. Comment était votre compréhension des **actions de votre partenaire** ?

Nulle _____ Très forte

34. Selon vous, quel était le niveau de compréhension de votre partenaire **concernant vos propres actions** ?

Nul _____ Très fort

35. Sur quel(s) élément(s) vous êtes-vous basé(e) pour expliquer quoi faire à votre partenaire ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)
☐ Votre partenaire (personne physique)
☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar
☐ Autre : _____

36. Sur quel(s) élément(s) vous êtes-vous basé(e) pour expliquer ce que vous faisiez à votre partenaire ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)
☐ Votre partenaire (personne physique)
☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar
☐ Autre : _____

37. Sur quel(s) élément(s) vous êtes-vous basé(e) pour communiquer avec votre partenaire **lors des phases de déplacement** ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)
☐ Votre partenaire (personne physique)
☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar
☐ Autre : _____

38. Sur quel(s) élément(s) vous êtes-vous basé(e) pour communiquer avec votre partenaire **lors des phases de manipulation d'objet** ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)
☐ Votre partenaire (personne physique)
☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar
☐ Autre : _____

39. Sur quel(s) élément(s) vous êtes-vous basé(e) pour communiquer avec votre partenaire **lors des phases de planification ou de raisonnement** ? (plusieurs réponses possibles)

☐ L'environnement réel ☐ Vous-même (personne physique)
☐ Votre partenaire (personne physique)
☐ L'environnement virtuel ☐ Votre avatar ☐ Son avatar
☐ Autre : _____

40. Comment jugeriez-vous **le plaisir** que vous avez eu **à effectuer la tâche avec votre partenaire** ?

Nul _____ Très fort

41. Selon vous, quel est le **niveau de réussite** de la tâche ?

Nul _____ Très fort

42. Quel est votre **niveau de satisfaction** concernant **votre participation** à la tâche ?

Nul _____ Très fort

43. Quel est votre **niveau de satisfaction** concernant **la participation de votre partenaire** à la tâche ?

Nul _____ Très fort

44. Comment noteriez-vous la **qualité du système audio** ?

Nulle _____ Très bonne

45. Comment noteriez-vous la **qualité de l'image** ?

Nulle _____ Très bonne

46. Auriez-vous souhaité plus d'éléments ou d'indications dans la préparation de la session collaborative distante ?

☐ Oui ☐ Non

Si oui : Lesquels : _____

47. Auriez-vous souhaité que le dispositif technique vous offre plus de fonctionnalités ou possibilités d'action ?

☐ Oui ☐ Non

Si oui : Lesquelles : _____

48. Avez-vous des suggestions concernant **l'accès aux fonctionnalités** ?

☐ Oui ☐ Non

Si oui : Lesquelles : _____

49. Avez-vous des attentes suite à cette session collaborative distante ?

☐ Oui ☐ Non

Si oui : Lesquelles : _____

50. Souhaitez-vous nous faire remonter d'autres remarques ?

☐ Oui ☐ Non

Si oui : Lesquelles : _____

53 - **Avez-vous un ordinateur ?** ☐ Oui ☐ Non

54 - **Avez-vous une console de jeu ?** ☐ Oui ☐ Non

Si oui laquelle ?_____

55 - **Comment situeriez-vous votre pratique des jeux vidéo :**

☐ Ne pratique pas du tout
☐ Quelques fois par an
☐ Une à deux fois par mois
☐ Une à deux fois par semaine
☐ Tous les jours

Si oui de quels types de jeux s'agit-il ?_____

56 - **Aviez-vous déjà utilisé une wiimote (manette de la wii, utilisée pendant l'exercice) ?**

☐ Oui, souvent ☐ Oui, parfois ☐ Oui, rarement ☐ Non, jamais

57 - **Aviez-vous déjà été immergé dans un Environnement Virtuel Immersif (avec lunettes 3D) ?**

☐ Oui ☐ Non

Si oui, - à quelle fréquence ? _____
- sur quelle période ? _____
- à quelle occasion ? _____
- pour faire quoi ? _____
- dans quel type d'environnement ? _____

58 - **Aviez-vous déjà été dans tout ou partie de l'Environnement Réel représenté par l'Environnement virtuel ?**

☐ Oui ☐ Non

Si oui, où êtes-vous allé ?

☐ Au niveau 0
☐ Autour des avions (passerelles Niveau 1 et/ou 2)
☐ Dans les avions (quel que soit le niveau)
☐ Autre, précisez : _____

59 - **Utilisez-vous souvent des logiciels de modélisation 3D (3DS Max, Maya, Google sketchup, …) ?**

☐ Oui ☐ Non

Si oui, le(s)quel(s) ?_____

60 - **Comment situeriez-vous votre pratique des logiciels de modélisation 3D :**

☐ Ne pratique pas du tout
☐ Quelques fois par an
☐ Une à deux fois par mois
☐ Une à deux fois par semaine
☐ Tous les jours

61 – **En voiture, souffrez-vous du mal des transports ?**

Lorsque vous conduisez :

Jamais ——— A chaque trajet

Lorsque vous êtes passager :

Jamais ——— A chaque trajet

62 – S**ouffrez-vous du mal des transports dans d'autres moyens de transport ?**

Jamais ——— A chaque trajet

Lesquels ?_____

Annexe 6 : Version du Santa Barbara Sense of Scale utilisé dans la première partie expérimentale

De manière générale, vous diriez que :

1. Je suis doué pour indiquer des directions.

Entièrement d'accord ——— Désaccord complet

2. Je ne mémorise pas bien là où je laisse mes affaires.

Entièrement d'accord ——— Désaccord complet

3. Je suis très doué pour juger des distances.

Entièrement d'accord ——— Désaccord complet

4. Mon sens de l'orientation est très bon.

Entièrement d'accord ——— Désaccord complet

5. J'ai tendance à penser mon environnement en termes de directions cardinales (nord, sud, est, ouest).

Entièrement d'accord ——— Désaccord complet

6. Je me perds facilement dans une nouvelle ville.

Entièrement d'accord ——— Désaccord complet

7. J'aime lire des cartes.

Entièrement d'accord ——— Désaccord complet

8. J'ai des difficultés à comprendre les directions.

Entièrement d'accord ——— Désaccord complet

9. Je suis très doué pour lire des cartes (type routières).

Entièrement d'accord ——— Désaccord complet

10. Je ne me souviens pas bien des routes empruntées lorsque je suis passager.

Entièrement d'accord ——— Désaccord complet

11. Je n'aime pas donner des directions.

Entièrement d'accord ——— Désaccord complet

12. Ce n'est pas grave si je ne sais pas où je suis.

Entièrement d'accord ——— Désaccord complet

13. Je laisse les autres préparer les itinéraires pour les longs trajets.

Entièrement d'accord ——— Désaccord complet

14. Je me souviens d'une route, même après l'avoir faite qu'une seule fois.

Entièrement d'accord ——— Désaccord complet

15. Je n'ai pas une « carte mentale » de mon environnement très bonne.

Entièrement d'accord ——— Désaccord complet

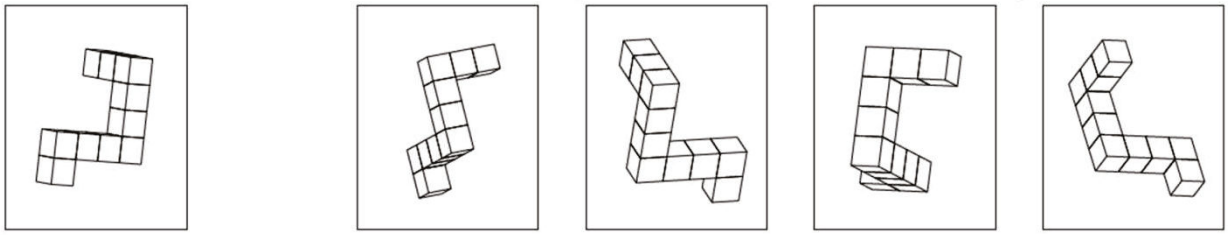
Annexe 7 : Exemple de ligne de test du MRT-A

Nous ne pouvons présenter le test dans son ensemble puisque le concepteur demande explicitement à ce qu'il ne soit pas diffusé. La raison est que si les participants ont facilement accès à ce test, les résultats peuvent être biaisés puisque plus les gens y sont exposés, plus leurs résultats sont bons.

Ce test est accessible sur simple demande via le site

http://spatiallearning.org/index.php/testsainstruments.

Le test consiste, pour chaque ligne telle que celle présentée ci-dessous, à cocher les deux images parmi les quatre propositions qui sont une vue différente de l'objet présenté sur l'image de gauche. Le test comprend 24 lignes, le participant a quatre minutes pour faire 12 lignes, puis quatre minutes pour faire les 12 lignes restantes. Le score final est calculé par la somme de bonnes réponses (une bonne réponse correspond au choix des deux bonnes réponses).



Annexe 8 : Questionnaire pour l'étude 1 de la seconde partie expérimentale

Nom : _____ Prénom : _____

Sexe : ☐ Masculin ☐ Féminin **Age : _____ Profession :** _____

Vous êtes : ☐ Gaucher ☐ Droitier ☐ Ambidextre

Pendant la séance, aviez-vous l'impression d'être dans la pièce présentée dans les scènes virtuelles ?

Oui totalement _____ Non pas du tout

Remarque : _____

Pendant la séance, aviez-vous l'impression que votre collaborateur était présent dans la pièce virtuelle ?

Oui totalement _____ Non pas du tout

Remarque : _____

Pendant la séance, aviez-vous l'impression d'être dans la salle B13bis de l'Ecole Centrale ?

Oui totalement _____ Non pas du tout

Remarque : _____

Pendant la séance, aviez-vous une stratégie pour optimiser votre performance (aller le plus vite possible tout en étant précis) ?

☐ Oui ☐ Non

Si oui, laquelle ? _____

Pendant la séance, aviez-vous une stratégie pour vous faciliter la tâche, diminuer l'exigence mentale ?

☐ Oui ☐ Non

Si oui, laquelle ? _____

Avez-vous un ordinateur chez vous ? ☐ Oui ☐ Non

Si oui, à quelle fréquence utilisez-vous cet ordinateur :

☐ Moins d'une fois par mois ☐ Une à deux fois par semaine

☐ Une à deux fois par mois ☐ Tous les jours

Utilisez-vous un ordinateur dans votre travail ? ☐ Oui ☐ Non

Si oui, combien de temps passez-vous sur cet ordinateur par jour :

☐ Moins d'une heure

☐ Entre une heure et une 3h

☐ Toute la journée

Avez-vous une console de jeu ? ☐ Oui ☐ Non

Si oui laquelle ?_____

Comment situeriez-vous votre pratique des jeux vidéo :

☐ Ne pratique pas du tout ☐ Une à deux fois par semaine

☐ Quelques fois par an ☐ Tous les jours

☐ Une à deux fois par mois

Si oui de quels types de jeux s'agit-il ?_____

Annexe 9 : Consignes données aux participants de l'étude 1 de la seconde partie expérimentale

A lire attentivement

Tout d'abord merci de participer à cette étude sur la collaboration à distance.

Vous allez devoir réaliser une tâche en collaboration avec un autre participant installé dans une autre pièce de l'Ecole Centrale.
Vous allez être immergés tous les deux dans le même environnement virtuel dans lequel vous serez tous les deux représentés par des avatars (personnage virtuel).
Grâce à la touche espace du clavier qui sera devant vous, vous pourrez parler à votre collaborateur (déclenchement du micro). Malheureusement, lui ne pourra pas vous parler.

Le but de la séance est que vous et votre collaborateur supprimiez un maximum de bouteilles piégées dans différentes scènes virtuelles. Vous serez donc immergés successivement dans différentes situations. Une seule bouteille est piégée par scène. Vous serez situé toujours au même endroit, et votre collaborateur le sait.

Les rôles sont répartis de la manière suivante :

- vous seul recevez l'information sur la localisation de la bouteille piégée (vous verrez la table sur laquelle est posée la bouteille piégée changer de couleur) ;
- seul votre collaborateur peut éliminer les bouteilles. Il n'a droit qu'à une seule tentative par scène.

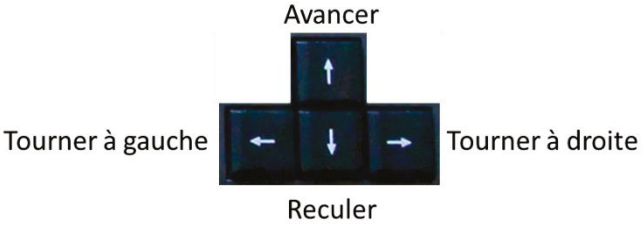
Vous devrez donc donner à votre collaborateur des consignes pour qu'il puisse localiser la bouteille piégée, avec deux contraintes :

être le plus précis et le plus rapide possible.

Vous obtiendrez un point à chaque fois que vous remplirez correctement la mission.
Vous n'aurez votre score qu'une fois l'ensemble des scènes fini.
Le but est que vous obteniez un maximum de points.

La séance totale dure environ 1h15, et est organisée de la manière suivante :
Visite du 1er environnement – 8 scènes test – 5 minutes avec différentes scènes – pause – 10 minutes (scènes + questions*) – idem avec 4 environnements différents – pause – questionnaire de fin sur papier (20 minutes).

Pendant les phases de visite d'environnement, vous devrez vous déplacer dans l'environnement afin de vous familiariser avec l'endroit, à l'aide des flèches du clavier :



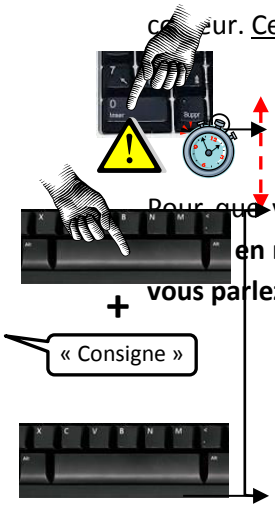
* : explications données au verso

Ensuite vous ferez 8 scènes test pour vous entrainer. Toutes les scènes se déroulent de la même manière :



- Un écran vous invite à appuyer sur le bouton 0 du pavé numérique afin de faire apparaître l'environnement virtuel.
- La scène apparaît : prenez-en connaissance. Votre collaborateur est représenté par l'avatar situé au milieu de la pièce. Notez qu'il ne sera pas toujours dans la même position.

Quand vous êtes prêt : appuyez sur 0 du pavé numérique pour que l'ordinateur vous indique quelle est la bouteille piégée. Pour cela la table sur laquelle est posée la bouteille piégée va changer de couleur. Ce changement de couleur ne sera visible que par vous.



- La table sur laquelle est située la bouteille piégée change de couleur. Vous devez donner la consigne à votre collaborateur **de la manière la plus précise et le plus rapidement possible** afin qu'il sélectionne et élimine la bouteille piégée.

Pour que votre collaborateur puisse vous entendre, **vous devez appuyer sur la barre espace du en même temps que vous donnez votre consigne, et maintenir la touche enfoncée tant que vous parlez**.

- Une fois votre consigne donnée, relâchez la barre espace.
- La scène disparaît pendant que votre collaborateur se charge d'éliminer la bouteille piégée. Il est donc important que vous soyez très précis dans la consigne que vous lui donnez pour obtenir un point.
- Un nouvel écran vous indiquera quand vous pouvez lancer une nouvelle scène. Si vous faites une mauvaise manip (appuyer sur le mauvais bouton par exemple) : indiquez-le-moi avant de lancer une nouvelle scène.

* Pour les phases avec questions : la consigne vous sera donnée juste avant. Il s'agira toujours de la même question après chaque scène. Vous pourrez prendre votre temps pour y répondre.
Pour répondre, vous aurez besoin des flèches droite et gauche du clavier. Pour valider votre réponse, vous devrez appuyer sur le bouton "." du pavé numérique (à droite de la touche 0). Une fois votre réponse validée, l'écran qui vous invite à afficher la scène apparaîtra.



Une fois les dernières scènes passées, je vous demanderai de répondre à un petit questionnaire.

Avant de commencer, avez-vous des questions ?

Annexe 10 : Table des résultats des analyses post-hoc (Tukey HSD) réalisées sur l'exigence mentale en fonction de l'orientation de l'avatar et la position de la cible par rapport à l'avatar.
Etude 1 – Production d'énoncés sans indices. Cases rouges : p<.05

		Plan sagittal					Diagonale devant					plan frontal					diagonale derrière				
		0°	45°	90°	135°	180°	0°	45°	90°	135°	180°	0°	45°	90°	135°	180°	0°	45°	90°	135°	180°
Plan sagittal	0°		1,00	0,13	1,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	45°	1,00		0,84	1,00	1,00	0,00	0,00	0,00	0,00	0,00	0,08	0,00	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	90°	0,13	0,84		0,83	0,30	0,02	0,00	0,00	0,00	0,00	1,00	0,08	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	135°	1,00	1,00	0,83		1,00	0,00	0,00	0,00	0,00	0,00	0,08	0,00	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	180°	1,00	1,00	0,30	1,00		0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Diagonale devant	0°	0,00	0,00	0,02	0,00	0,00		1,00	0,85	0,94	0,00	0,49	1,00	0,52	1,00	0,99	0,00	0,00	0,00	0,00	0,00
	45°	0,00	0,00	0,00	0,00	0,00	1,00		1,00	1,00	0,00	0,11	1,00	0,13	1,00	1,00	0,00	0,00	0,00	0,00	0,00
	90°	0,00	0,00	0,00	0,00	0,00	0,85	1,00		1,00	0,40	0,00	0,50	0,00	0,99	1,00	0,01	0,16	0,00	0,00	0,00
	135°	0,00	0,00	0,00	0,00	0,00	0,94	1,00	1,00		0,25	0,00	0,67	0,00	1,00	1,00	0,01	0,09	0,00	0,00	0,00
	180°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,40	0,25		0,00	0,00	0,00	0,00	0,13	1,00	1,00	0,92	0,89	0,33
Plan frontal	0°	0,00	0,08	1,00	0,08	0,01	0,49	0,11	0,00	0,00	0,00		0,84	1,00	0,14	0,00	0,00	0,00	0,00	0,00	0,00
	45°	0,00	0,00	0,08	0,00	0,00	1,00	1,00	0,50	0,67	0,00	0,84		0,86	1,00	0,86	0,00	0,00	0,00	0,00	0,00
	90°	0,00	0,07	1,00	0,07	0,00	0,52	0,13	0,00	0,00	0,00	1,00	0,86		0,15	0,00	0,00	0,00	0,00	0,00	0,00
	135°	0,00	0,00	0,00	0,00	0,00	1,00	1,00	0,99	1,00	0,00	0,14	1,00	0,15		1,00	0,00	0,00	0,00	0,00	0,00
	180°	0,00	0,00	0,00	0,00	0,00	0,99	1,00	1,00	1,00	0,13	0,00	0,86	0,00	1,00		0,00	0,04	0,00	0,00	0,00
Diagonale derrière	0°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	1,00	0,00	0,00	0,00	0,00	0,00		1,00	1,00	1,00	0,98
	45°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,16	0,09	1,00	0,00	0,00	0,00	0,00	0,04	1,00		0,99	0,99	0,63
	90°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,92	0,00	0,00	0,00	0,00	0,00	1,00	0,99		1,00	1,00
	135°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,89	0,00	0,00	0,00	0,00	0,00	1,00	0,99	1,00		1,00
	180°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,33	0,00	0,00	0,00	0,00	0,00	0,98	0,63	1,00	1,00	

Annexe 11 : Table des résultats des analyses post-hoc (Tukey HSD) réalisées sur les temps de préparation des énoncés en fonction de l'orientation de l'avatar et la position de la cible par rapport à l'avatar.
Etude 1 – Production d'énoncés sans indices. Cases rouges : p<.05

		Plan sagittal					Diagonale devant					plan frontal					diagonale derrière				
		0°	45°	90°	135°	180°	0°	45°	90°	135°	180°	0°	45°	90°	135°	180°	0°	45°	90°	135°	180°
Plan sagittal	0°		0,99	0,26	1,00	1,00	0,02	0,03	0,00	0,00	0,00	0,02	0,07	0,36	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	45°	0,99		1,00	1,00	0,67	0,77	0,83	0,01	0,08	0,00	0,79	0,95	1,00	0,28	0,00	0,00	0,00	0,00	0,00	0,00
	90°	0,26	1,00		0,99	0,02	1,00	1,00	0,54	0,89	0,05	1,00	1,00	1,00	0,99	0,14	0,00	0,00	0,00	0,00	0,00
	135°	1,00	1,00	0,99		0,78	0,66	0,73	0,01	0,05	0,00	0,68	0,90	1,00	0,20	0,00	0,00	0,00	0,00	0,00	0,00
	180°	1,00	0,67	0,02	0,78		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Diagonale devant	0°	0,02	0,77	1,00	0,66	0,00		1,00	0,98	1,00	0,48	1,00	1,00	1,00	1,00	0,72	0,00	0,00	0,00	0,00	0,00
	45°	0,03	0,83	1,00	0,73	0,00	1,00		0,97	1,00	0,41	1,00	1,00	1,00	1,00	0,65	0,00	0,00	0,00	0,00	0,00
	90°	0,00	0,01	0,54	0,01	0,00	0,98	0,97		1,00	1,00	0,98	0,87	0,41	1,00	1,00	0,01	0,45	0,01	0,00	0,00
	135°	0,00	0,08	0,89	0,05	0,00	1,00	1,00	1,00		1,00	1,00	0,99	0,80	1,00	1,00	0,00	0,14	0,00	0,00	0,00
	180°	0,00	0,00	0,05	0,00	0,00	0,48	0,41	1,00	1,00		0,46	0,22	0,03	0,92	1,00	0,25	0,98	0,24	0,00	0,00
Plan frontal	0°	0,02	0,79	1,00	0,68	0,00	1,00	1,00	0,98	1,00	0,46		1,00	1,00	1,00	0,70	0,00	0,00	0,00	0,00	0,00
	45°	0,07	0,95	1,00	0,90	0,00	1,00	1,00	0,87	0,99	0,22	1,00		1,00	1,00	0,42	0,00	0,00	0,00	0,00	0,00
	90°	0,36	1,00	1,00	1,00	0,04	1,00	1,00	0,41	0,80	0,03	1,00	1,00		0,98	0,08	0,00	0,00	0,00	0,00	0,00
	135°	0,00	0,28	0,99	0,20	0,00	1,00	1,00	1,00	1,00	0,92	1,00	1,00	0,98		0,98	0,00	0,03	0,00	0,00	0,00
	180°	0,00	0,00	0,14	0,00	0,00	0,72	0,65	1,00	1,00	1,00	0,70	0,42	0,08	0,98		0,11	0,89	0,11	0,00	0,00
Diagonale derrière	0°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,25	0,00	0,00	0,00	0,00	0,11		1,00	1,00	0,23	0,98
	45°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,45	0,14	0,98	0,00	0,00	0,00	0,03	0,89	1,00		1,00	0,00	0,28
	90°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,24	0,00	0,00	0,00	0,00	0,11	1,00	1,00		0,24	0,99
	135°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,23	0,00	0,24		1,00
	180°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,98	0,28	0,99	1,00	

Annexe 12 : Table des résultats des analyses post-hoc (Tukey HSD) réalisées sur les temps de production des énoncés en fonction de l'orientation de l'avatar et la position de la cible par rapport à l'avatar.

Etude 1 – Production d'énoncés sans indices. Cases rouges : p<.05

		Plan sagittal					Diagonale devant					plan frontal					diagonale derrière				
		0°	45°	90°	135°	180°	0°	45°	90°	135°	180°	0°	45°	90°	135°	180°	0°	45°	90°	135°	180°
Plan sagittal	0°		1,00	1,00	1,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	45°	1,00		1,00	1,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	90°	1,00	1,00		1,00	0,99	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	135°	1,00	1,00	1,00		1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	180°	1,00	1,00	0,99	1,00		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Diagonale devant	0°	0,00	0,00	0,00	0,00	0,00		1,00	1,00	1,00	1,00	0,99	0,98	0,04	0,89	1,00	0,02	1,00	0,00	0,01	0,00
	45°	0,00	0,00	0,00	0,00	0,00	1,00		0,88	1,00	0,91	1,00	1,00	0,69	1,00	1,00	0,00	0,40	0,00	0,00	0,00
	90°	0,00	0,00	0,00	0,00	0,00	1,00	0,88		1,00	1,00	0,72	0,61	0,00	0,35	1,00	0,21	1,00	0,02	0,17	0,02
	135°	0,00	0,00	0,00	0,00	0,00	1,00	1,00	1,00		1,00	1,00	1,00	0,18	1,00	1,00	0,00	0,91	0,00	0,00	0,00
	180°	0,00	0,00	0,00	0,00	0,00	1,00	0,91	1,00	1,00		0,78	0,67	0,00	0,41	1,00	0,17	1,00	0,01	0,13	0,01
Plan frontal	0°	0,00	0,00	0,00	0,00	0,00	0,99	1,00	0,72	1,00	0,78		1,00	0,86	1,00	1,00	0,00	0,24	0,00	0,00	0,00
	45°	0,00	0,00	0,00	0,00	0,00	0,98	1,00	0,61	1,00	0,67	1,00		0,92	1,00	1,00	0,00	0,16	0,00	0,00	0,00
	90°	0,00	0,15	0,19	0,04	0,00	0,04	0,69	0,00	0,18	0,00	0,86	0,92		0,99	0,26	0,00	0,00	0,00	0,00	0,00
	135°	0,00	0,00	0,00	0,00	0,00	0,89	1,00	0,35	1,00	0,41	1,00	1,00	0,99		1,00	0,00	0,06	0,00	0,00	0,00
	180°	0,00	0,00	0,00	0,00	0,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,26	1,00		0,00	0,84	0,00	0,00	0,00
Diagonale derrière	0°	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,21	0,00	0,17	0,00	0,00	0,00	0,00	0,00		0,69	1,00	1,00	1,00
	45°	0,00	0,00	0,00	0,00	0,00	1,00	0,40	1,00	0,91	1,00	0,24	0,16	0,00	0,06	0,84	0,69		0,16	0,61	0,16
	90°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,01	0,00	0,00	0,00	0,00	0,00	1,00	0,16		1,00	1,00
	135°	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,17	0,00	0,13	0,00	0,00	0,00	0,00	0,00	1,00	0,61	1,00		1,00
	180°	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,01	0,00	0,00	0,00	0,00	0,00	1,00	0,16	1,00	1,00	

Annexe 13 : Question ajoutée au questionnaire de l'étude 1 de la seconde partie expérimentale (Annexe 8)

Est-ce que les indices visuels vous ont aidé ?

- **Avatar avec les manches de couleur ?** ☐ Oui ☐ Non
- **Murs de couleur ?** ☐ Oui ☐ Non

Pourquoi ? _____

Annexe 14 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des conditions d'indiçage et la position de la cible par rapport au participant

Etude 2 - Production d'énoncés avec indices. Cases rouges : p<.05

		Sans indice					Avatar latéralisé					Panneaux					Panneaux + Avatar latéralisé				
		A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
Sans indice	A		0,00	1,00	0,00	1,00	0,99	0,00	1,00	0,00	0,99	1,00	0,17	1,00	0,00	1,00	0,64	0,00	0,02	0,00	0,26
	B	0,00		0,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,29	0,00	0,91	0,00
	C	1,00	0,00		0,00	1,00	0,90	0,00	1,00	0,00	0,87	1,00	0,46	1,00	0,00	0,99	0,92	0,00	0,08	0,00	0,61
	D	0,00	1,00	0,00		0,00	0,00	0,99	0,00	0,88	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,00	0,48	0,00
	E	1,00	0,00	1,00	0,00		0,80	0,00	1,00	0,00	0,76	0,99	0,61	1,00	0,00	0,96	0,97	0,00	0,14	0,00	0,75
Avatar latéralisé	A	0,99	0,00	0,90	0,00	0,80		0,00	0,95	0,00	1,00	1,00	0,00	0,97	0,00	1,00	0,01	0,00	0,00	0,00	0,00
	B	0,00	1,00	0,00	0,99	0,00	0,00		0,00	1,00	0,00	0,00	0,00	0,00	0,28	0,00	0,00	0,93	0,01	1,00	0,00
	C	1,00	0,00	1,00	0,00	1,00	0,95	0,00		0,00	0,93	1,00	0,34	1,00	0,00	1,00	0,85	0,00	0,05	0,00	0,47
	D	0,00	1,00	0,00	0,88	0,00	0,00	1,00	0,00		0,00	0,00	0,00	0,00	0,60	0,00	0,00	1,00	0,03	1,00	0,00
	E	0,99	0,00	0,87	0,00	0,76	1,00	0,00	0,93	0,00		1,00	0,00	0,96	0,00	1,00	0,01	0,00	0,00	0,00	0,00
Panneaux	A	1,00	0,00	1,00	0,00	0,99	1,00	0,00	1,00	0,00	1,00		0,01	1,00	0,00	1,00	0,11	0,00	0,00	0,00	0,02
	B	0,17	0,00	0,46	0,00	0,61	0,00	0,00	0,34	0,00	0,00	0,01		0,27	0,92	0,00	1,00	0,27	1,00	0,02	1,00
	C	1,00	0,00	1,00	0,00	1,00	0,97	0,00	1,00	0,00	0,96	1,00	0,27		0,00	1,00	0,78	0,00	0,03	0,00	0,39
	D	0,00	0,01	0,00	0,00	0,00	0,00	0,28	0,00	0,60	0,00	0,00	0,92	0,00		0,00	0,46	1,00	1,00	0,94	0,84
	E	1,00	0,00	0,99	0,00	0,96	1,00	0,00	1,00	0,00	1,00	1,00	0,00	1,00	0,00		0,05	0,00	0,00	0,00	0,01
Panneaux + Avatar latéralisé	A	0,64	0,00	0,92	0,00	0,97	0,01	0,00	0,85	0,00	0,01	0,11	1,00	0,78	0,46	0,05		0,04	1,00	0,00	1,00
	B	0,00	0,29	0,00	0,05	0,00	0,00	0,93	0,00	1,00	0,00	0,00	0,27	0,00	1,00	0,00	0,04		0,80	1,00	0,18
	C	0,02	0,00	0,08	0,00	0,14	0,00	0,01	0,05	0,03	0,00	0,00	1,00	0,03	1,00	0,00	1,00	0,80		0,17	1,00
	D	0,00	0,91	0,00	0,48	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,02	0,00	0,94	0,00	0,00	1,00	0,17		0,01
	E	0,26	0,00	0,61	0,00	0,75	0,00	0,00	0,47	0,00	0,00	0,02	1,00	0,39	0,84	0,01	1,00	0,18	1,00	0,01	

Annexe 15 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des conditions d'indiçage et l'orientation de l'avatar.

Etude 2 - Production d'énoncés avec indices. Cases rouges : p<.05

		Sans indice			Avatar latéralisé			Panneaux			Panneaux + Avatar latéralisé		
		0°	90°	180°	0°	90°	180°	0°	90°	180°	0°	90°	180°
Sans indice	0°		1,0	0,0	0,0	1,0	0,2	0,0	0,1	0,8	1,0	0,0	0,3
	90°	1,0		0,0	0,0	1,0	0,9	0,0	0,0	0,1	0,6	0,6	1,0
	180°	0,0	0,0		0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,5	0,1
Avatar latéralisé	0°	0,0	0,0	0,0		0,0	0,0	0,6	1,0	0,9	0,4	0,0	0,0
	90°	1,0	1,0	0,0	0,0		0,3	0,0	0,0	0,7	1,0	0,1	0,4
	180°	0,2	0,9	0,2	0,0	0,3		0,0	0,0	0,0	0,0	1,0	1,0
Panneaux	0°	0,0	0,0	0,0	0,6	0,0	0,0		0,5	0,0	0,0	0,0	0,0
	90°	0,1	0,0	0,0	1,0	0,0	0,0	0,5		1,0	0,5	0,0	0,0
	180°	0,8	0,1	0,0	0,9	0,7	0,0	0,0	1,0		1,0	0,0	0,0
Panneaux + Avatar latéralisé	0°	1,0	0,6	0,0	0,4	1,0	0,0	0,0	0,5	1,0		0,0	0,0
	90°	0,0	0,6	0,5	0,0	0,1	1,0	0,0	0,0	0,0	0,0		1,0
	180°	0,3	1,0	0,1	0,0	0,4	1,0	0,0	0,0	0,0	0,0	1,0	

Annexe 16 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des conditions d'indiçage et la position de la cible par rapport à l'avatar
Etude 2 Production d'énoncés avec indices. Cases rouges : p<.05

		Sans indice				Avatar latéralisé				Panneaux				Panneaux + Avatar latéralisé			
		Pl. sagittal	Diag. devant	Pl. frontal	Diag. derrière	Pl. sagittal	Diag. devant	Pl. frontal	Diag. derrière	Pl. sagittal	Diag. devant	Pl. frontal	Diag. derrière	Pl. sagittal	Diag. devant	Pl. frontal	Diag. derrière
Sans indice	Pl. sagittal		0,00	0,00	0,00	1,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,74	0,00	0,00	0,00
	Diag. devant	0,00		0,37	0,19	0,00	1,00	0,01	0,95	0,00	0,00	0,00	0,98	0,00	0,56	1,00	1,00
	Pl. frontal	0,00	0,37		0,00	0,00	0,93	0,99	0,00	0,00	0,94	0,97	1,00	0,00	1,00	0,32	0,08
	Diag. derrière	0,00	0,19	0,00		0,00	0,01	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,23	0,62
Avatar latéralisé	Pl. sagittal	1,00	0,00	0,00	0,00		0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,43	0,00	0,00	0,00
	Diag. devant	0,00	1,00	0,93	0,01	0,00		0,13	0,43	0,00	0,04	0,06	1,00	0,00	0,98	1,00	0,98
	Pl. frontal	0,00	0,01	0,99	0,00	0,00	0,13		0,00	0,00	1,00	1,00	0,48	0,06	0,97	0,01	0,00
	Diag. derrière	0,00	0,95	0,00	1,00	0,00	0,43	0,00		0,00	0,00	0,00	0,11	0,00	0,01	0,97	1,00
Panneaux	Pl. sagittal	1,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00		0,00	0,00	0,00	0,88	0,00	0,00	0,00
	Diag. devant	0,00	0,00	0,94	0,00	0,00	0,04	1,00	0,00	0,00		1,00	0,24	0,18	0,84	0,00	0,00
	Pl. frontal	0,00	0,00	0,97	0,00	0,00	0,06	1,00	0,00	0,00	1,00		0,29	0,14	0,89	0,00	0,00
	Diag. derrière	0,00	0,98	1,00	0,00	0,00	1,00	0,48	0,11	0,00	0,24	0,29		0,00	1,00	0,97	0,74
Panneaux + Avatar latéralisé	Pl. sagittal	0,74	0,00	0,00	0,00	0,43	0,00	0,06	0,00	0,88	0,18	0,14	0,00		0,00	0,00	0,00
	Diag. devant	0,00	0,56	1,00	0,00	0,00	0,98	0,97	0,01	0,00	0,84	0,89	1,00	0,00		0,50	0,16
	Pl. frontal	0,00	1,00	0,32	0,23	0,00	1,00	0,01	0,97	0,00	0,00	0,00	0,97	0,00	0,50		1,00
	Diag. derrière	0,00	1,00	0,08	0,62	0,00	0,98	0,00	1,00	0,00	0,00	0,00	0,74	0,00	0,16	1,00	

Annexe 17 : Analyses post-hoc (Tukey HSD) pour les valeurs de temps de production des énoncés en fonction des conditions d'indiçage et la position de la cible par rapport au participant.
Etude 2 Production d'énoncés avec indices. Cases rouges : p<.05

		Sans indice					Avatar latéralisé					Panneaux					Panneaux + Avatar latéralisé				
		A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
Sans indice	A		0,0	1,0	0,0	1,0	0,1	0,0	1,0	0,0	0,1	1,0	0,2	0,4	0,0	1,0	0,6	0,0	0,0	0,0	0,6
	B	0,0		0,0	1,0	0,0	0,0	1,0	0,0	1,0	0,0	0,0	0,2	0,1	0,7	0,0	0,0	1,0	0,7	1,0	0,0
	C	1,0	0,0		0,0	1,0	0,1	0,0	1,0	0,0	0,0	1,0	0,3	0,6	0,0	1,0	0,8	0,0	0,0	0,0	0,7
	D	0,0	1,0	0,0		0,0	0,0	1,0	0,0	1,0	0,0	0,0	0,2	0,1	0,8	0,0	0,0	1,0	0,7	1,0	0,0
	E	1,0	0,0	1,0	0,0		0,4	0,0	1,0	0,0	0,3	0,9	0,0	0,1	0,0	1,0	0,2	0,0	0,0	0,0	0,2
Avatar latéralisé	A	0,1	0,0	0,1	0,0	0,4		0,0	0,9	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	B	0,0	1,0	0,0	1,0	0,0	0,0		0,0	1,0	0,0	0,2	1,0	0,9	1,0	0,0	0,8	0,7	1,0	0,8	0,9
	C	1,0	0,0	1,0	0,0	1,0	0,9	0,0		0,0	0,9	0,4	0,0	0,0	0,0	0,8	0,0	0,0	0,0	0,0	0,0
	D	0,0	1,0	0,0	1,0	0,0	0,0	1,0	0,0		0,0	0,1	0,9	0,8	1,0	0,0	0,6	0,9	1,0	1,0	0,6
	E	0,1	0,0	0,0	0,0	0,3	1,0	0,0	0,9	0,0		0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Panneaux	A	1,0	0,0	1,0	0,0	0,9	0,0	0,2	0,4	0,1	0,0		1,0	1,0	0,6	1,0	1,0	0,0	0,7	0,0	1,0
	B	0,2	0,2	0,3	0,2	0,0	0,0	1,0	0,0	0,9	0,0	1,0		1,0	1,0	0,8	1,0	0,0	1,0	0,0	1,0
	C	0,4	0,1	0,6	0,1	0,1	0,0	0,9	0,0	0,8	0,0	1,0	1,0		1,0	1,0	1,0	0,0	1,0	0,0	1,0
	D	0,0	0,7	0,0	0,8	0,0	0,0	1,0	0,0	1,0	0,0	0,6	1,0	1,0		0,2	1,0	0,2	1,0	0,3	1,0
	E	1,0	0,0	1,0	0,0	1,0	0,0	0,0	0,8	0,0	0,0	1,0	0,8	1,0	0,2		1,0	0,0	0,3	0,0	1,0
Panneaux + Avatar latéralisé	A	0,6	0,0	0,8	0,0	0,2	0,0	0,8	0,0	0,6	0,0	1,0	1,0	1,0	1,0	1,0		0,0	1,0	0,0	1,0
	B	0,0	1,0	0,0	1,0	0,0	0,0	0,7	0,0	0,9	0,0	0,0	0,0	0,0	0,2	0,0	0,0		0,2	1,0	0,0
	C	0,0	0,7	0,0	0,7	0,0	0,0	1,0	0,0	1,0	0,0	0,7	1,0	1,0	1,0	0,3	1,0	0,2		0,2	1,0
	D	0,0	1,0	0,0	1,0	0,0	0,0	0,8	0,0	1,0	0,0	0,0	0,0	0,0	0,3	0,0	0,0	1,0	0,2		0,0
	E	0,6	0,0	0,7	0,0	0,2	0,0	0,9	0,0	0,6	0,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	1,0	0,0	

Annexe 18 : Analyses post-hoc (Tukey HSD) pour les valeurs de temps de production des énoncés en fonction des conditions d'indiçage et la position de la cible par rapport à l'avatar - Etude 2 Production d'énoncés avec indices. Cases rouges : p<.05

		Sans indice				Avatar latéralisé				Panneaux				Panneaux + Avatar latéralisé			
		plan sagittal	diag. devant	plan frontal	diag. derrière	plan sagittal	diag. devant	plan frontal	diag. derrière	plan sagittal	diag. devant	plan frontal	diag. derrière	plan sagittal	diag. devant	plan frontal	diag. derrière
Sans indice	plan sagittal		0,00	0,00	0,00	0,94	0,00	0,95	0,00	0,62	0,01	0,00	0,00	0,57	0,00	0,00	0,00
	diag. devant	0,00		0,97	0,15	0,00	1,00	0,01	0,95	0,05	0,93	1,00	1,00	0,06	1,00	0,99	0,09
	plan frontal	0,00	0,97		0,00	0,00	1,00	0,49	0,09	0,90	1,00	1,00	0,87	0,92	0,46	0,20	0,00
	diag. derrière	0,00	0,15	0,00		0,00	0,01	0,00	0,99	0,00	0,00	0,02	0,33	0,00	0,77	0,95	1,00
Avatar latéralisé	plan sagittal	0,94	0,00	0,00	0,00		0,00	0,06	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	diag. devant	0,00	1,00	1,00	0,01	0,00		0,14	0,36	0,52	1,00	1,00	1,00	0,56	0,86	0,60	0,00
	plan frontal	0,95	0,01	0,49	0,00	0,06	0,14		0,00	1,00	0,64	0,05	0,00	1,00	0,00	0,00	0,00
	diag. derrière	0,00	0,95	0,09	0,99	0,00	0,36	0,00		0,00	0,05	0,64	0,99	0,00	1,00	1,00	0,98
Panneaux	plan sagittal	0,62	0,05	0,90	0,00	0,01	0,52	1,00	0,00		0,96	0,26	0,02	1,00	0,00	0,00	0,00
	diag. devant	0,01	0,93	1,00	0,00	0,00	1,00	0,64	0,05	0,96		1,00	0,76	0,97	0,32	0,12	0,00
	plan frontal	0,00	1,00	1,00	0,02	0,00	1,00	0,05	0,64	0,26	1,00		1,00	0,29	0,98	0,85	0,01
	diag. derrière	0,00	1,00	0,87	0,33	0,00	1,00	0,00	0,99	0,02	0,76	1,00		0,02	1,00	1,00	0,22
Panneaux + Avatar latéralisé	plan sagittal	0,57	0,06	0,92	0,00	0,00	0,56	1,00	0,00	1,00	0,97	0,29	0,02		0,00	0,00	0,00
	diag. devant	0,00	1,00	0,46	0,77	0,00	0,86	0,00	1,00	0,00	0,32	0,98	1,00	0,00		1,00	0,64
	plan frontal	0,00	0,99	0,20	0,95	0,00	0,60	0,00	1,00	0,00	0,12	0,85	1,00	0,00	1,00		0,89
	diag. derrière	0,00	0,09	0,00	1,00	0,00	0,00	0,00	0,98	0,00	0,00	0,01	0,22	0,00	0,64	0,89	

Annexe 19 : Analyses post-hoc (Tukey HSD) pour les valeurs de nombre de mots en fonction des conditions d'indiçage et la position de la cible par rapport au participant.
Etude 2 - Production d'énoncés avec indices. Cases rouges : p<.05

		Sans indice					Avatar latéralisé					Panneaux					Panneaux + Avatar latéralisé				
		A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
Sans indice	A		0,0	1,0	0,0	1,0	0,5	0,0	0,9	0,1	0,4	0,9	0,6	0,6	0,4	1,0	0,7	0,0	0,4	0,0	1,0
	B	0,0		0,0	1,0	0,0	0,0	1,0	0,0	1,0	0,0	0,4	0,8	0,8	0,9	0,0	0,7	1,0	0,9	1,0	0,3
	C	1,0	0,0		0,0	1,0	0,5	0,0	0,9	0,1	0,4	0,9	0,6	0,6	0,4	1,0	0,7	0,0	0,4	0,0	1,0
	D	0,0	1,0	0,0		0,0	0,0	1,0	0,0	1,0	0,0	0,3	0,7	0,8	0,8	0,0	0,6	1,0	0,9	1,0	0,2
	E	1,0	0,0	1,0	0,0		0,9	0,0	1,0	0,0	0,8	0,5	0,2	0,2	0,1	1,0	0,3	0,0	0,1	0,0	0,7
Avatar latéralisé	A	0,5	0,0	0,5	0,0	0,9		0,0	1,0	0,0	1,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0
	B	0,0	1,0	0,0	1,0	0,0	0,0		0,0	1,0	0,0	1,0	1,0	1,0	1,0	0,4	1,0	1,0	1,0	1,0	1,0
	C	0,9	0,0	0,9	0,0	1,0	1,0	0,0		0,0	1,0	0,0	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,0
	D	0,1	1,0	0,1	1,0	0,0	0,0	1,0	0,0		0,0	1,0	1,0	1,0	1,0	0,5	1,0	1,0	1,0	1,0	1,0
	E	0,4	0,0	0,4	0,0	0,8	1,0	0,0	1,0	0,0		0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Panneaux	A	0,9	0,4	0,9	0,3	0,5	0,0	1,0	0,0	1,0	0,0		1,0	1,0	1,0	1,0	1,0	0,4	1,0	0,5	1,0
	B	0,6	0,8	0,6	0,7	0,2	0,0	1,0	0,0	1,0	0,0	1,0		1,0	1,0	1,0	1,0	0,8	1,0	0,9	1,0
	C	0,6	0,8	0,6	0,8	0,2	0,0	1,0	0,0	1,0	0,0	1,0	1,0		1,0	1,0	1,0	0,8	1,0	0,9	1,0
	D	0,4	0,9	0,4	0,8	0,1	0,0	1,0	0,0	1,0	0,0	1,0	1,0	1,0		1,0	1,0	0,9	1,0	1,0	1,0
	E	1,0	0,0	1,0	0,0	1,0	0,1	0,4	0,3	0,5	0,0	1,0	1,0	1,0	1,0		1,0	0,0	0,9	0,0	1,0
Panneaux + Avatar latéralisé	A	0,7	0,7	0,7	0,6	0,3	0,0	1,0	0,0	1,0	0,0	1,0	1,0	1,0	1,0	1,0		0,7	1,0	0,8	1,0
	B	0,0	1,0	0,0	1,0	0,0	0,0	1,0	0,0	1,0	0,0	0,4	0,8	0,8	0,9	0,0	0,7		0,9	1,0	0,3
	C	0,4	0,9	0,4	0,9	0,1	0,0	1,0	0,0	1,0	0,0	1,0	1,0	1,0	1,0	0,9	1,0	0,9		1,0	1,0
	D	0,0	1,0	0,0	1,0	0,0	0,0	1,0	0,0	1,0	0,0	0,5	0,9	0,9	1,0	0,0	0,8	1,0	1,0		0,4
	E	1,0	0,3	1,0	0,2	0,7	0,0	1,0	0,0	1,0	0,0	1,0	1,0	1,0	1,0	1,0	1,0	0,3	1,0	0,4	

Annexe 20 : Analyses post-hoc (Tukey HSD) pour les valeurs de nombre de mots en fonction des conditions d'indiçage et la position de la cible par rapport à l'avatar. Etude 2 - Production d'énoncés avec indices. Cases rouges : p<.05

		Sans indice				Avatar latéralisé				Panneaux				Panneaux + Avatar latéralisé			
		plan sagittal	diag. devant	plan frontal	diag. derrière	plan sagittal	diag. devant	plan frontal	diag. derrière	plan sagittal	diag. devant	plan frontal	diag. derrière	plan sagittal	diag. devant	plan frontal	diag. derrière
Sans indice	plan sagittal		0,00	0,00	0,00	0,94	0,00	0,70	0,00	0,33	0,02	0,00	0,00	0,66	0,00	0,00	0,00
	diag. devant	0,00		1,00	0,52	0,00	1,00	0,23	1,00	0,58	0,99	1,00	1,00	0,26	1,00	0,99	0,68
	plan frontal	0,00	1,00		0,13	0,00	1,00	0,70	0,94	0,95	1,00	1,00	1,00	0,73	1,00	0,81	0,22
	diag. derrière	0,00	0,52	0,13		0,00	0,16	0,00	0,99	0,00	0,01	0,67	0,66	0,00	0,50	1,00	1,00
Avatar latéralisé	plan sagittal	0,94	0,00	0,00	0,00		0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00
	diag. devant	0,00	1,00	1,00	0,16	0,00		0,64	0,96	0,93	1,00	1,00	1,00	0,68	1,00	0,85	0,27
	plan frontal	0,70	0,23	0,70	0,00	0,01	0,64		0,01	1,00	0,98	0,15	0,15	1,00	0,25	0,00	0,00
	diag. derrière	0,00	1,00	0,94	0,99	0,00	0,96	0,01		0,05	0,54	1,00	1,00	0,01	1,00	1,00	1,00
Panneaux	plan sagittal	0,33	0,58	0,95	0,00	0,00	0,93	1,00	0,05		1,00	0,44	0,45	1,00	0,60	0,02	0,00
	diag. devant	0,02	0,99	1,00	0,01	0,00	1,00	0,98	0,54	1,00		0,98	0,98	0,99	0,99	0,31	0,03
	plan frontal	0,00	1,00	1,00	0,67	0,00	1,00	0,15	1,00	0,44	0,98		1,00	0,17	1,00	1,00	0,81
	diag. derrière	0,00	1,00	1,00	0,66	0,00	1,00	0,15	1,00	0,45	0,98	1,00		0,17	1,00	1,00	0,80
Panneaux + Avatar latéralisé	plan sagittal	0,66	0,26	0,73	0,00	0,01	0,68	1,00	0,01	1,00	0,99	0,17	0,17		0,28	0,00	0,00
	diag. devant	0,00	1,00	1,00	0,50	0,00	1,00	0,25	1,00	0,60	0,99	1,00	1,00	0,28		0,99	0,66
	plan frontal	0,00	0,99	0,81	1,00	0,00	0,85	0,00	1,00	0,02	0,31	1,00	1,00	0,00	0,99		1,00
	diag. derrière	0,00	0,68	0,22	1,00	0,00	0,27	0,00	1,00	0,00	0,03	0,81	0,80	0,00	0,66	1,00	

Annexe 21 : Questions spécifiques de l'étude 3 de la seconde partie expérimentale (Annexe 8)

Classez les consignes par difficulté croissante (0 : plus facile – 3 : plus difficile) **:**

___ Degrés « la bouteille piégée est sur la table à 90° sur ta droite »

___ Devant, derrière, cotés « la bouteille piégée est sur la table en face de toi »

___ Points cardinaux (nord, sud…) « la bouteille piégée est sur la table au sud-est »

___ Panneaux « la bouteille piégée est sur la table à côté du panneau rouge »

Expliquez, précisez votre classement : _____

Auriez-vous souhaité que votre collaborateur vous donne l'information autrement ?

☐ Oui ☐ Non

Si oui, comment ?_____

Annexe 22 : Consignes données aux participants de l'étude 1 de la seconde partie expérimentale

A lire attentivement

Tout d'abord merci de participer à cette étude sur la collaboration à distance.

Vous allez devoir réaliser une tâche en collaboration avec un autre participant installé dans une autre pièce de l'Ecole Centrale.
Vous allez être immergés tous les deux dans le même environnement virtuel (pièce fermée) dans lequel vous serez tous les deux représentés par des avatars (personnages virtuels).
Malheureusement, vous ne pourrez pas lui parler. Lui seul pourra vous parler, mais uniquement par des énoncés simples sur lesquels il ne pourra pas revenir (énoncés produits en une fois).

Le but de la séance est que vous et votre collaborateur supprimiez un maximum de bouteilles piégées.

Votre collaborateur sera situé toujours au même endroit dans la pièce virtuelle.
Sachez qu'il fait face au nord.

Les rôles sont répartis de la manière suivante :

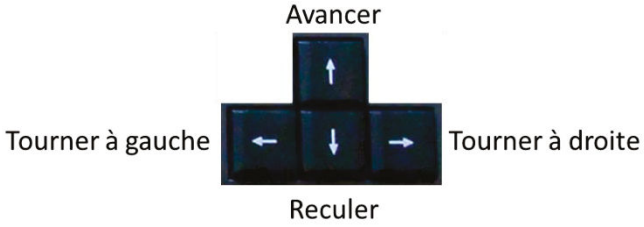
- lui seul recevra l'information sur la localisation de la bouteille piégée ;
- vous seul pourrez éliminer les bouteilles.

Votre collaborateur va vous donner une consigne pour localiser la bouteille à détruire et vous devrez sélectionner la bouteille piégée de manière la plus rapide possible sans vous tromper. Vous n'avez droit qu'à une seule tentative par scène.

Vous obtiendrez un point à chaque fois que vous remplirez correctement la mission.
Vous n'aurez votre score qu'une fois l'ensemble des scènes fini.
Le but est que vous obteniez un maximum de points.

La séance totale dure environ 2h, et est organisée de la manière suivante :
Visite de l'environnement – 5 essais d'entrainement – x situations normales puis y avec questions* : et ce répété 4 fois. Nous ferons des pauses toutes les 10 à 15 minutes, ou selon les besoins. Vous devrez à la fin répondre à un questionnaire papier (20 minutes).

Pendant la phase de visite de l'environnement, vous devrez vous déplacer dans l'environnement afin de vous familiariser avec l'endroit, à l'aide des flèches du clavier :



* : explications données au verso



Ensuite vous ferez 5 essais test pour vous entrainer. Tous les essais se déroulent de la même manière :

- Appuyer sur le bouton 0 du pavé numérique afin de faire apparaître la scène virtuelle.



- La scène apparaît : prenez-en connaissance. Rappelez-vous que votre collaborateur sera toujours à la même place, orienté vers le nord.



- Quand vous êtes prêt : appuyez sur 0 du pavé numérique pour que votre collaborateur vous indique quelle est la bouteille piégée.

- « Consigne »



- Une fois la consigne donnée vous devez sélectionner la bouteille piégée pour qu'elle soit éliminée **le plus rapidement possible**. Cela se passe en deux temps :

  - Appuyer sur la touche 5 du pavé numérique **pour indiquer que vous avez compris de quelle bouteille il s'agit**. **Ne pas appuyer si vous êtes encore en train de réfléchir !!**

  - Sélectionnez ensuite la bouteille à éliminer à l'aide du pavé numérique de la manière suivante :



Vous êtes situé au centre du cercle et faites face à la bouteille située sur la touche 8 du pavé numérique.
Par exemple, si vous pensez que la bouteille piégée est celle située derrière vous : appuyez sur la touche 2 du pavé numérique.

Il est important que vous soyez rapide et précis pour ces deux étapes.

- Une fois que vous avez sélectionné une bouteille la scène disparaît.
- Pour faire apparaître de nouveau la scène et commencer l'essai suivant, appuyez sur 0 et ainsi de suite.

* Pour les phases avec question : il s'agira toujours de la même question qui vous sera expliquée au tout début et posée après chaque scène. Vous pourrez prendre votre temps pour y répondre.
Pour répondre, vous aurez besoin des flèches droite et gauche du clavier. Pour valider votre réponse, vous devrez appuyer sur le point "." du pavé numérique (à droite de la touche 0). Une fois votre réponse validée, la question disparaîtra et vous pourrez afficher la scène de nouveau.



Une fois les dernières scènes passées, je vous demanderai de répondre à un petit questionnaire.
A tout moment, si vous pensez vous être trompé dans les touches, avoir appuyé trop tôt… prévenez-moi : je peux annuler l'essai.
Avant de commencer, avez-vous des questions ?

Annexe 23 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des cadres de référence et l'orientation du participant.

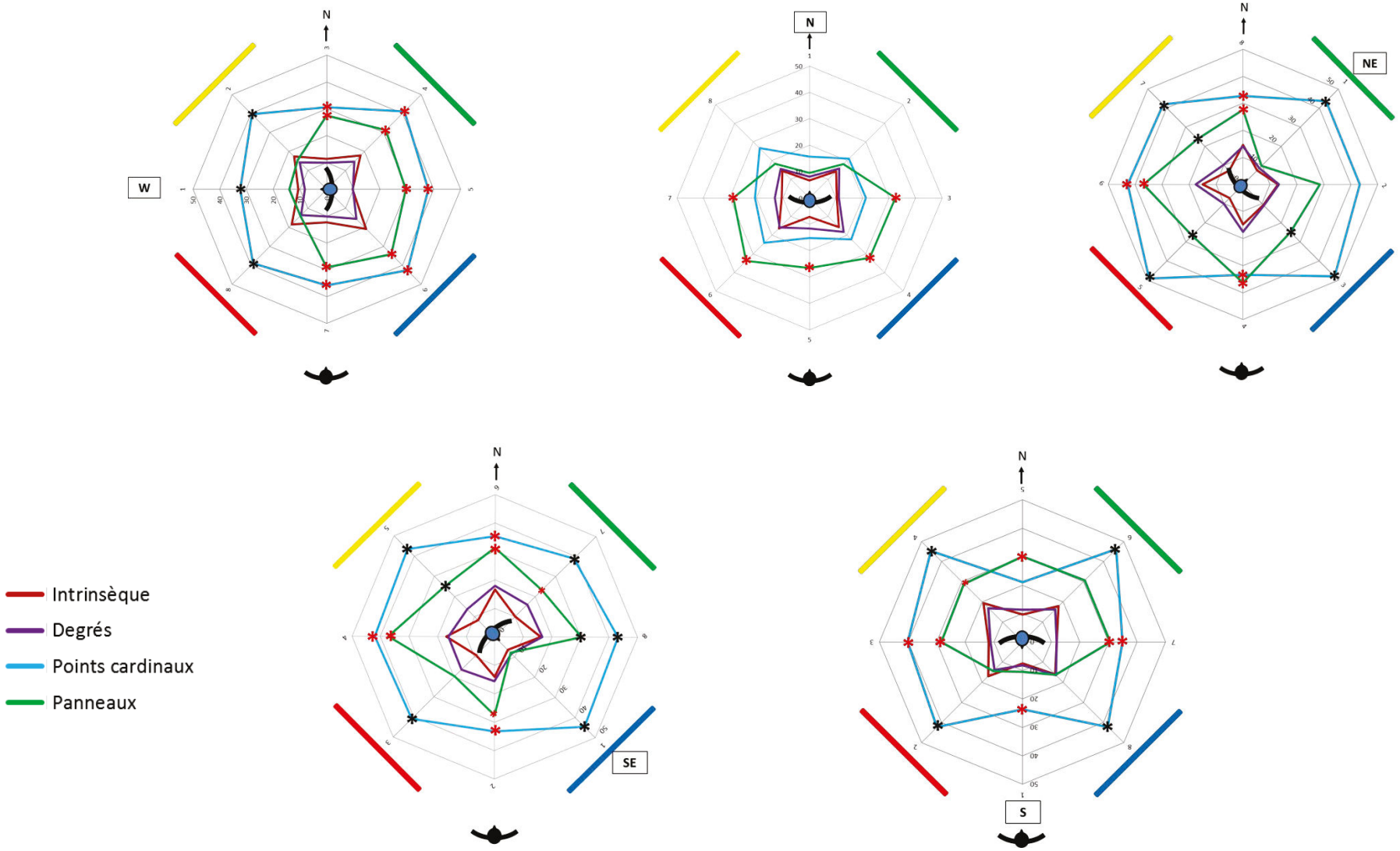
Etude 3 - Compréhension d'énoncés spatialisés. Cases rouges : p<.05

		Intrinsèque					Degrés					Points cardinaux					Panneaux				
		W	N	NE	SE	S	W	N	NE	SE	S	W	N	NE	SE	S	W	N	NE	SE	S
Intrinsèque	W		1,00	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	0,87	0,00	0,00	0,00	0,05	0,04	0,00	0,03	0,06
	N	1,00		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NE	0,99	1,00		1,00	1,00	1,00	1,00	1,00	0,99	1,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SE	1,00	1,00	1,00		1,00	1,00	1,00	1,00	1,00	1,00	0,00	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	S	1,00	1,00	1,00	1,00		1,00	1,00	1,00	1,00	1,00	0,00	0,48	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,01
Degrés	W	1,00	1,00	1,00	1,00	1,00		1,00	1,00	1,00	1,00	0,00	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	N	1,00	1,00	1,00	1,00	1,00	1,00		1,00	1,00	1,00	0,00	0,35	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	NE	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,00	1,00	0,00	0,30	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	SE	1,00	1,00	0,99	1,00	1,00	1,00	1,00	1,00		1,00	0,00	0,88	0,00	0,00	0,00	0,05	0,05	0,00	0,03	0,07
	S	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		0,00	0,34	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Points cardinaux	W	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		0,00	0,99	1,00	1,00	0,00	0,00	0,05	0,00	0,00
	N	0,87	0,09	0,04	0,12	0,48	0,15	0,35	0,30	0,88	0,34	0,00		0,00	0,00	0,00	1,00	0,99	0,73	0,99	1,00
	NE	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,99	0,00		1,00	0,98	0,00	0,00	0,00	0,00	0,00
	SE	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	1,00		1,00	0,00	0,00	0,00	0,00	0,00
	S	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,98	1,00		0,00	0,00	0,08	0,01	0,00
Panneaux	W	0,05	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,05	0,00	0,00	1,00	0,00	0,00	0,00		1,00	1,00	1,00	1,00
	N	0,04	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,05	0,00	0,00	0,99	0,00	0,00	0,00	1,00		1,00	1,00	1,00
	NE	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,73	0,00	0,00	0,08	1,00	1,00		1,00	1,00
	SE	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,99	0,00	0,00	0,01	1,00	1,00	1,00		1,00
	S	0,06	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,07	0,00	0,00	1,00	0,00	0,00	0,00	1,00	1,00	1,00	1,00	

Annexe 24 : Analyses post-hoc (Tukey HSD) pour les valeurs d'exigence mentale en fonction des cadres de référence et la position de la cible.

Etude 3 - Compréhension d'énoncés spatialisés. Cases rouges : p<.05

		Intrinsèque								Degrés								Points cardinaux								Panneaux							
		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Intrinsèque	1		0,2	1,0	0,0	1,0	0,0	1,0	0,1	1,0	0,4	1,0	0,0	1,0	0,0	1,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	2	0,2		1,0	1,0	0,5	1,0	1,0	1,0	0,4	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,3	0,0	0,0	0,0	0,0	0,0	0,1
	3	1,0	1,0		0,5	1,0	0,4	1,0	0,9	1,0	1,0	1,0	0,5	1,0	0,5	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	4	0,0	1,0	0,5		0,0	1,0	0,5	1,0	0,0	1,0	0,9	1,0	0,8	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	1,0	0,0	0,0	0,0	0,0	0,0	0,8
	5	1,0	0,5	1,0	0,0		0,0	1,0	0,2	1,0	0,7	1,0	0,0	1,0	0,0	1,0	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	6	0,0	1,0	0,4	1,0	0,0		0,3	1,0	0,0	1,0	0,9	1,0	0,7	1,0	0,9	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,6	1,0	0,0	0,0	0,0	0,0	0,0	0,9
	7	1,0	1,0	1,0	0,5	1,0	0,3		0,8	1,0	1,0	1,0	0,4	1,0	0,5	1,0	0,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	8	0,1	1,0	0,9	1,0	0,2	1,0	0,8		0,1	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,7	0,0	0,0	0,0	0,0	0,0	0,4
Degrés	1	1,0	0,4	1,0	0,0	1,0	0,0	1,0	0,1		0,6	1,0	0,0	1,0	0,0	1,0	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	2	0,4	1,0	1,0	1,0	0,7	1,0	1,0	1,0	0,6		1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,2	0,0	0,0	0,0	0,0	0,0	0,1
	3	1,0	1,0	1,0	0,9	1,0	0,9	1,0	1,0	1,0	1,0		0,9	1,0	0,9	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	4	0,0	1,0	0,5	1,0	0,0	1,0	0,4	1,0	0,0	1,0	0,9		0,8	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	1,0	0,0	0,0	0,0	0,0	0,0	0,8
	5	1,0	1,0	1,0	0,8	1,0	0,7	1,0	1,0	1,0	1,0	1,0	0,8		0,8	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	6	0,0	1,0	0,5	1,0	0,0	1,0	0,5	1,0	0,0	1,0	0,9	1,0	0,8		1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	1,0	0,0	0,0	0,0	0,0	0,0	0,8
	7	1,0	1,0	1,0	1,0	1,0	0,9	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	8	0,1	1,0	1,0	1,0	0,3	1,0	0,9	1,0	0,3	1,0	1,0	1,0	1,0	1,0	1,0		0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,5	0,0	0,0	0,0	0,0	0,0	0,2
Points cardinaux	1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0		1,0	1,0	0,9	1,0	0,3	1,0	0,9	0,0	0,0	1,0	1,0	1,0	1,0	1,0	0,0
	2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	1,0	0,0	1,0	0,0	0,0
	3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	1,0	0,1	1,0	0,0	0,0
	4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,9	1,0	1,0		1,0	1,0	1,0	1,0	0,0	0,0	0,0	1,0	0,0	1,0	0,0	0,0
	5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	1,0		0,9	1,0	1,0	0,0	0,0	0,5	1,0	0,8	1,0	0,7	0,0
	6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,3	1,0	1,0	1,0	0,9		1,0	1,0	0,0	0,0	0,0	0,9	0,0	1,0	0,0	0,0
	7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	0,0	0,0	0,2	1,0	0,4	1,0	0,4	0,0
	8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,9	1,0	1,0	1,0	1,0	1,0	1,0		0,0	0,0	0,0	1,0	0,0	1,0	0,0	0,0
Panneaux	1	1,0	1,0	1,0	0,7	1,0	0,6	1,0	1,0	1,0	1,0	1,0	0,7	1,0	0,7	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0		0,0	0,0	0,0	0,0	0,0	0,0	0,0
	2	0,0	0,3	0,0	1,0	0,0	1,0	0,0	0,7	0,0	0,2	0,0	1,0	0,0	1,0	0,0	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0		0,6	0,0	0,3	0,0	0,4	1,0
	3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,5	0,0	0,2	0,0	0,0	0,6		0,5	1,0	0,3	1,0	0,8
	4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	1,0	1,0	0,9	1,0	1,0	0,0	0,0	0,5		0,7	1,0	0,7	0,0
	5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,1	0,0	0,8	0,0	0,4	0,0	0,0	0,3	1,0	0,7		0,6	1,0	0,6
	6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,3	1,0	0,6		0,5	0,0
	7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0	0,0	0,7	0,0	0,4	0,0	0,0	0,4	1,0	0,7	1,0	0,5		0,7
	8	0,0	0,1	0,0	0,8	0,0	0,9	0,0	0,4	0,0	0,1	0,0	0,8	0,0	0,8	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,8	0,0	0,6	0,0	0,7	



Annexe 25 : Exigence mentale moyenne pour chaque orientation du participant, chaque cadre de référence et les huit cibles. * différence significative avec les trois autres cadres de référence, * différence significative avec les Cadre de référence centrés sur l'interlocuteur.

Annexe 26 : Analyses post-hoc (Tukey HSD) pour les valeurs du temps de compréhension en fonction des cadres de référence et l'orientation du participant.

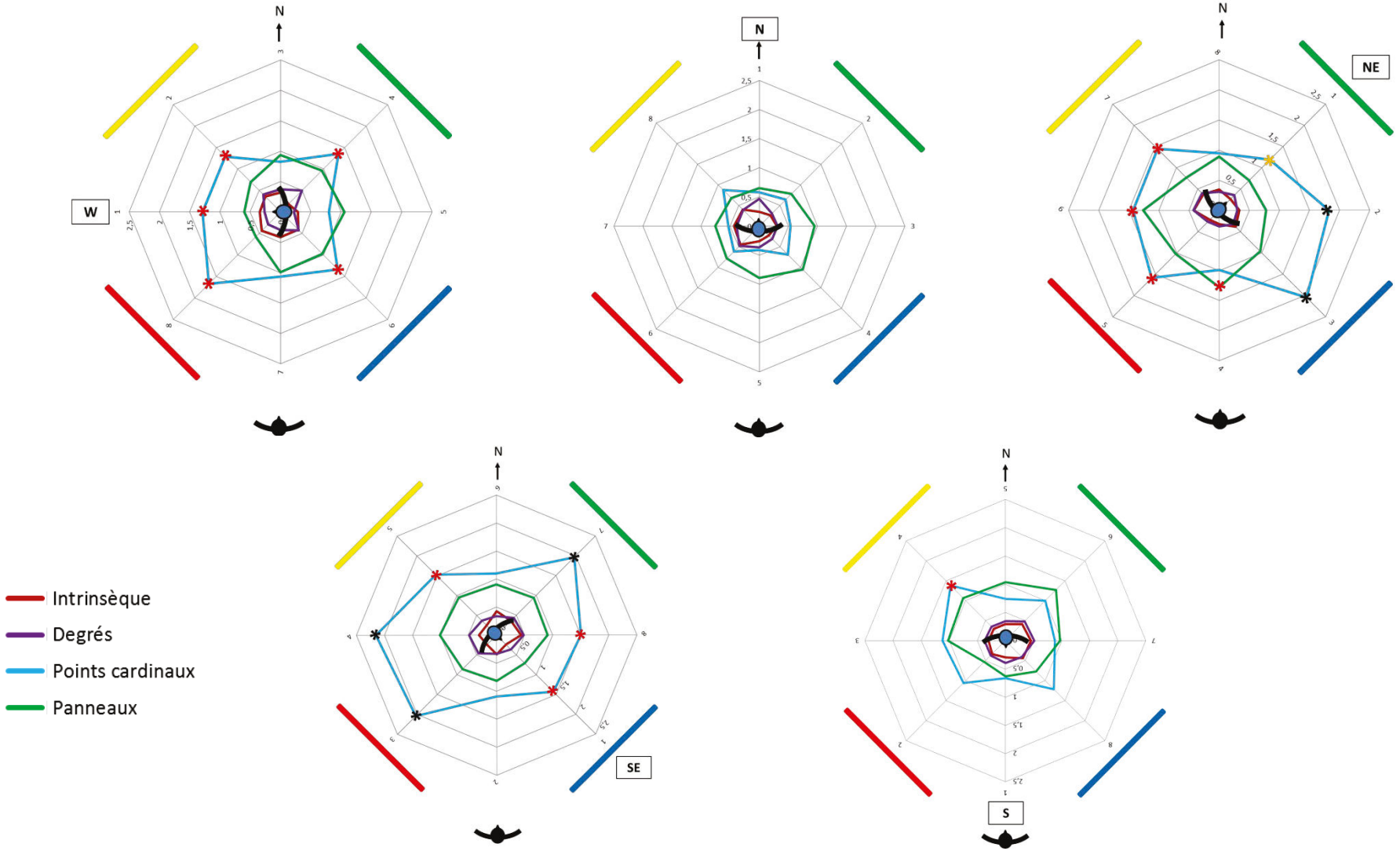
Etude 3 - Compréhension d'énoncés spatialisés. Cases rouges : p<.05

		Intrinsèque					Degrés					Points cardinaux					Panneaux				
		W	N	NE	SE	S	W	N	NE	SE	S	W	N	NE	SE	S	W	N	NE	SE	S
Intrinsèque	W		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	0,93	0,00	0,00	0,00	0,04	0,01	0,00	0,01	0,01
	N	1,00		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	0,88	0,00	0,00	0,00	0,02	0,01	0,00	0,01	0,00
	NE	1,00	1,00		1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	0,87	0,00	0,00	0,00	0,02	0,01	0,00	0,01	0,00
	SE	1,00	1,00	1,00		1,00	1,00	1,00	1,00	1,00	1,00	0,00	0,88	0,00	0,00	0,00	0,02	0,01	0,00	0,01	0,00
	S	1,00	1,00	1,00	1,00		1,00	1,00	1,00	1,00	1,00	0,00	0,95	0,00	0,00	0,00	0,05	0,01	0,00	0,02	0,01
Degrés	W	1,00	1,00	1,00	1,00	1,00		1,00	1,00	1,00	1,00	0,00	0,94	0,00	0,00	0,00	0,04	0,01	0,00	0,01	0,01
	N	1,00	1,00	1,00	1,00	1,00	1,00		1,00	1,00	1,00	0,00	0,98	0,00	0,00	0,00	0,08	0,03	0,00	0,03	0,02
	NE	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,00	1,00	0,00	0,90	0,00	0,00	0,00	0,03	0,01	0,00	0,01	0,00
	SE	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,00	0,00	1,00	0,00	0,00	0,00	0,15	0,06	0,01	0,07	0,04
	S	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		0,00	1,00	0,00	0,00	0,00	0,14	0,05	0,01	0,06	0,04
Points cardinaux	W	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		0,00	0,99	0,32	1,00	0,54	0,79	0,97	0,75	0,86
	N	0,93	0,88	0,87	0,88	0,95	0,94	0,98	0,90	1,00	1,00	0,00		0,00	0,00	0,31	0,98	0,88	0,59	0,91	0,81
	NE	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,99	0,00		1,00	0,20	0,00	0,02	0,07	0,01	0,03
	SE	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,32	0,00	1,00		0,00	0,00	0,00	0,00	0,00	0,00
	S	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,31	0,20	0,00		1,00	1,00	1,00	1,00	1,00
Panneaux	W	0,04	0,02	0,02	0,02	0,05	0,04	0,08	0,03	0,15	0,14	0,54	0,98	0,00	0,00	1,00		1,00	1,00	1,00	1,00
	N	0,01	0,01	0,01	0,01	0,01	0,01	0,03	0,01	0,06	0,05	0,79	0,88	0,02	0,00	1,00	1,00		1,00	1,00	1,00
	NE	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,97	0,59	0,07	0,00	1,00	1,00	1,00		1,00	1,00
	SE	0,01	0,01	0,01	0,01	0,02	0,01	0,03	0,01	0,07	0,06	0,75	0,91	0,01	0,00	1,00	1,00	1,00	1,00		1,00
	S	0,01	0,00	0,00	0,00	0,01	0,01	0,02	0,00	0,04	0,04	0,86	0,81	0,03	0,00	1,00	1,00	1,00	1,00	1,00	

Annexe 27 : Analyses post-hoc (Tukey HSD) pour les valeurs du temps de compréhension en fonction des cadres de référence et de la position de la cible.

Etude 3 Compréhension d'énoncés spatialisés. Cases rouges : p<.05

		Intrinsèque								Degrés								Points cardinaux								Panneaux							
		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Intrinsèque	1		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	2	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	3	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	4	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	5	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	6	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	0,2	0,0	0,0	0,0	0,0	0,0	0,0
	7	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,1	0,0	0,0	0,0	0,0	0,0	0,0
	8	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,1	0,0	0,0	0,0	0,0	0,0	0,0
Degrés	1	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	2	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	3	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	4	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	5	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	6	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	0,2	0,0	0,0	0,0	0,0	0,0	0,0
	7	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,1	0,0	0,0	0,0	0,0	0,0	0,0
	8	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Points cardinaux	1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0		1,0	0,2	0,2	1,0	1,0	1,0	0,8	0,0	0,1	1,0	1,0	1,0	1,0	1,0	0,4
	2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,7	1,0	1,0	1,0	0,2	0,0
	3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	1,0		1,0	0,1	0,8	1,0	1,0	0,0	0,0	0,0	0,6	0,1	1,0	0,0	0,0
	4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	1,0	1,0		0,1	0,8	1,0	1,0	0,0	0,0	0,0	0,6	0,1	1,0	0,0	0,0
	5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,1	0,1		1,0	1,0	0,7	0,0	0,2	1,0	1,0	1,0	1,0	1,0	0,5
	6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,8	0,8	1,0		1,0	1,0	0,0	0,0	1,0	1,0	1,0	1,0	0,8	0,0
	7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	0,0	0,0	0,7	1,0	1,0	1,0	0,2	0,0
	8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,8	1,0	1,0	1,0	0,7	1,0	1,0		0,0	0,0	0,2	1,0	0,5	1,0	0,0	0,0
Panneaux	1	0,0	0,0	0,0	0,0	0,0	0,7	0,5	0,5	0,2	0,1	0,2	0,3	0,0	0,7	0,5	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0		1,0	0,2	0,0	0,0	0,0	0,7	1,0
	2	0,0	0,0	0,0	0,0	0,0	0,2	0,1	0,1	0,0	0,0	0,0	0,0	0,2	0,1	0,0		0,1	0,0	0,0	0,0	0,2	0,0	0,0	0,0	1,0		0,7	0,0	0,3	0,0	1,0	1,0
	3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,7	0,0	0,0	1,0	1,0	0,7	0,2	0,2	0,7		1,0	1,0	0,8	1,0	1,0
	4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,6	0,6	1,0	1,0	1,0	1,0	0,0	0,0	1,0		1,0	1,0	1,0	0,1
	5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,1	0,1	1,0	1,0	1,0	0,5	0,0	0,3	1,0	1,0		1,0	1,0	0,7
	6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,8	1,0	1,0		0,3	0,0
	7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,2	0,0	0,0	1,0	0,8	0,2	0,0	0,7	1,0	1,0	1,0	1,0	0,3		1,0
	8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,4	0,0	0,0	0,0	0,5	0,0	0,0	0,0	1,0	1,0	1,0	0,1	0,7	0,0	1,0	

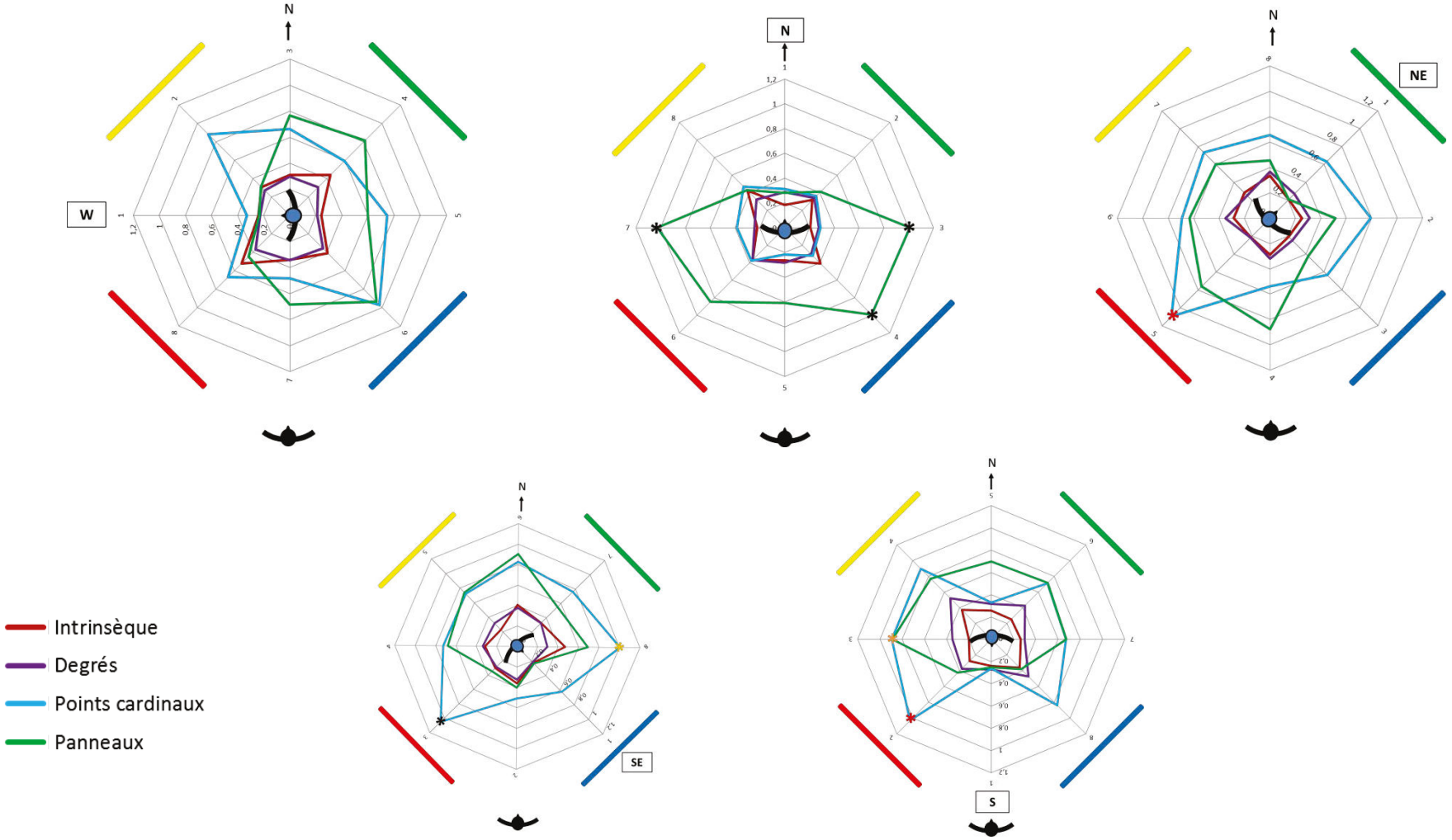


Annexe 28 Temps moyen de compréhension pour chaque orientation du participant, chaque cadre de référence et les huit cibles. * différence significative avec les trois autres cadres de référence, * différence significative avec les Cadre de référence centrés sur l'interlocuteur..

Annexe 29 : Analyses post-hoc (Tukey HSD) pour les temps de sélection de la cible en fonction des cadres de référence et la position de la cible.

Etude 3 Compréhension d'énoncés spatialisés. Cases rouges : p<.05

		Intrinsèque								Degrés								Points cardinaux								Panneaux							
		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Intrinsèque	1		1,0	1,0	1,0	1,0	1,0	1,0	0,9	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,9	0,0	0,0	0,0	0,0	0,0	0,4
	2	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,1	0,1	0,0	0,1	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	1,0
	3	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,6
	4	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,5	0,4	0,0	0,5	0,0	1,0	1,0	0,1	0,0	0,1	0,0	0,0	1,0
	5	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,6
	6	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,3	0,2	0,0	0,4	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	1,0
	7	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,9
	8	0,9	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,1	0,2	0,9	0,9	0,1	1,0	0,1	1,0	1,0	0,3	0,0	0,4	0,0	0,2	1,0
Degrés	1	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,7
	2	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,2	0,1	0,0	0,2	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	1,0
	3	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,1	0,0	0,0	0,1	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,9
	4	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,4	0,3	0,0	0,5	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	1,0
	5	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,8
	6	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	1,0	0,0	0,0	0,6	0,5	0,0	0,7	0,0	1,0	1,0	0,1	0,0	0,1	0,0	0,0	1,0
	7	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	1,0	0,0	0,0	0,1	0,0	0,0	0,1	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,9
	8	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		1,0	0,0	0,0	0,5	0,4	0,0	0,6	0,0	1,0	1,0	0,1	0,0	0,1	0,0	0,0	1,0
Points cardinaux	1	0,8	1,0	1,0	1,0	0,9	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0		0,2	0,3	1,0	0,9	0,2	1,0	0,1	0,9	1,0	0,4	0,0	0,5	0,0	0,2	1,0
	2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2		1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,1	1,0	1,0	1,0	1,0	1,0	0,6
	3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,3	1,0		1,0	1,0	1,0	1,0	1,0	0,0	0,1	1,0	1,0	1,0	1,0	1,0	0,7
	4	0,0	0,1	0,0	0,5	0,0	0,3	0,0	0,9	0,0	0,2	0,1	0,4	0,0	0,6	0,1	0,5	1,0	1,0	1,0		1,0	1,0	1,0	1,0	0,0	0,9	1,0	0,8	1,0	0,9	1,0	1,0
	5	0,0	0,1	0,0	0,4	0,0	0,2	0,0	0,9	0,0	0,1	0,0	0,3	0,0	0,5	0,0	0,4	0,9	1,0	1,0	1,0		1,0	1,0	1,0	0,0	0,8	1,0	0,9	1,0	1,0	1,0	1,0
	6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	1,0	1,0	1,0	1,0		1,0	1,0	0,0	0,1	1,0	1,0	1,0	1,0	1,0	0,6
	7	0,0	0,1	0,0	0,5	0,0	0,4	0,1	1,0	0,0	0,2	0,1	0,5	0,0	0,7	0,1	0,6	1,0	1,0	1,0	1,0	1,0	1,0		1,0	0,0	0,9	1,0	0,8	1,0	0,9	1,0	1,0
	8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	1,0	1,0	1,0	1,0	1,0	1,0		0,0	0,1	1,0	1,0	1,0	1,0	1,0	0,5
Panneaux	1	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0		1,0	0,0	0,0	0,0	0,0	0,0	0,6
	2	0,9	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,1	0,1	0,9	0,8	0,1	0,9	0,1	1,0		0,3	0,0	0,3	0,0	0,1	1,0
	3	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,1	0,4	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,3		1,0	1,0	1,0	1,0	0,9
	4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,8	0,9	1,0	0,8	1,0	0,0	0,0	1,0		1,0	1,0	1,0	0,0
	5	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,4	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,1	0,5	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,3	1,0	1,0		1,0	1,0	0,9
	6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,9	1,0	1,0	0,9	1,0	0,0	0,0	1,0	1,0	1,0		1,0	0,0
	7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	1,0	1,0	1,0	1,0	1,0	1,0	1,0	0,0	0,1	1,0	1,0	1,0	1,0		0,7
	8	0,4	1,0	0,6	1,0	0,6	1,0	0,9	1,0	0,7	1,0	0,9	1,0	0,8	1,0	0,9	1,0	1,0	0,6	0,7	1,0	1,0	0,6	1,0	0,5	0,6	1,0	0,9	0,0	0,9	0,0	0,7	



Annexe 30 : **Temps moyen de sélection pour chaque orientation du participant, chaque cadre de référence et les huit cibles. * différence significative avec les trois autres cadres de référence, * différence significative avec les Cadre de référence centrés sur l'interlocuteur.**

Collaboration à distance : étude de la compréhension mutuelle dans les environnements virtuels collaboratifs immersifs
Le cas de la communication spatiale

Résumé

Les situations de collaboration à distance dans l'industrie induisent de nouvelles contraintes pour les opérateurs. Dans le contexte de l'utilisation d'environnements virtuels collaboratifs immersifs, nous avons mis en place une série d'études portant sur la compréhension mutuelle, et plus particulièrement au partage d'information de nature spatiale. Les résultats de la première étude ont permis de mettre en évidence l'influence du rôle des participants, guides ou manipulateurs, sur les énoncés spatialisés. Les énoncés étaient centrés préférentiellement sur l'action du manipulateur, suggérant la recherche d'un moindre effort collaboratif. Les résultats de deux études sur la production d'énoncés spatialisés ont permis d'identifier que l'exigence mentale pour la production d'énoncés est modulée notamment par la position de la cible à décrire par rapport au destinataire. En effet, selon la position de la cible, le locuteur doit opérer ou non des transformations mentales coûteuses pour prendre la perspective du destinataire. Cet effort peut être amoindri en présence d'indices visuels distaux. La dernière étude, portant sur la compréhension d'énoncés spatialisés, a permis de mettre en évidence que les énoncés centrés sur le destinataire sont les plus simples à comprendre pour une tâche de nature égocentrée. Certains énoncés exocentrés induisent également une moindre exigence mentale, mais uniquement selon certaines conditions. Les résultats sont discutés selon le principe du moindre effort collaboratif et la théorie des cadres de référence. Ce travail a permis de proposer des pistes de développement pour faciliter la collaboration à distance dans les environnements virtuels.

Mots clé : Collaboration, référentiel commun, cadre de référence, exigence mentale.

Remote collaboration: mutual comprehension in immersive collaborative virtual environments
The case of spatial communication

Abstract

Remote collaborative situations in industry involve new constraints for workers. In the context of using immersive virtual environments to collaborate, we set up a series of experiments focusing on the mutual comprehension, and more specifically on the process of sharing spatial information. Results of the first experiment showed the influence of one collaborator's role on spatial statements. Guides and manipulators both used statements preferentially centered on manipulator's action, which supports the least collaborative effort principle. Results of two experiments about spatial statements production allowed to point out that mental workload is modulated by the target position relative to the addressee. According to the target location, the speaker must operate or not cognitively costly mental transformations to take the addressee perspective. However this workload could be lowered by means of visual cues. The last experiment focused on the understanding of spatial statements. It showed that statements centered on the addressee are the easiest to understand when the task is also centered on the addressee. Some exocentered statements could also induce a lower mental workload but only in some conditions. Results are discussed in relation to the least collaborative effort principle and the spatial frames of reference theory. This work opens new leads to facilitate remote collaboration through virtual environments.

Key words: Collaboration, common ground, reference frame, mental workload.