

# Étude expérimentale

## Visualisation d'informations en réalité virtuelle

Alexandre Amrani

Responsable : Florent Levillain



<b>Introduction</b>	<b>1</b>
<b>État de l’art</b>	<b>1</b>
Perception	1
Réalité Virtuelle	3
Cognition incarnée	3
<b>Expérience</b>	<b>4</b>
Protocole	4
Déroulement	5
<b>Discussions</b>	<b>10</b>
<b>Conclusion</b>	<b>12</b>
<b>Référence</b>	<b>13</b>
<b>Annexes</b>	<b>15</b>
Annexe 1	15



## Introduction

À ce jour, l'essentiel des données est représenté aux utilisateurs dans un espace à deux dimensions : les smartphones, ordinateurs et tablettes ne sont que des plans sur lesquels sont disposées des images. La grande majorité des technologies, applications et logiciels utilisés fonctionnent sur cet espace plat. Pourtant, de plus en plus de solutions rendent le monde tri-dimensionnel transposable au digital, grâce notamment aux casques de réalité virtuelle.

Ainsi, il est intéressant d'observer comment les utilisateurs évoluent au sein de ces environnements immersifs : alors qu'ils ont l'habitude d'évoluer numériquement sur un plan, comment réagissent-ils à la tri-dimensionnalité numérique ? Retrouvent-ils leurs réflexes de la réalité ? Dans cette étude, le regard sera porté sur la visualisation de données à travers une question : comment l'espace tri-dimensionnel affecte la perception des utilisateurs dans la visualisation d'images, à travers des tâches de recherche d'images ?

La visualisation d'images se fait aujourd'hui essentiellement sur un plan et les logiciels essaient de reproduire une perception de profondeur. En effet, les outils étant multi-fonctions, différentes informations doivent pouvoir être affichées sur les écrans : de même qu'il est possible de disposer des éléments autour de soi dans la réalité, le digital permet aux logiciels de se superposer, aux images de défiler de haut en bas ou de droite à gauche. Malgré cela, toute l'information est contenue sur un plan aux dimensions finies, sur un écran. L'espace à trois dimensions possède bien plus de possibilités d'arrangement. L'humain a acquis des connaissances et compétences cognitives engendrées par les possibilités de l'espace tri-dimensionnel, mais les espaces numériques actuels sont limités. La tri-dimensionnalité étant de plus en plus accessible, il semble donc légitime de recourir à d'autres technologies afin d'exploiter la troisième dimension, créatrice de la *réalité virtuelle*.

## État de l'art

### Perception

La visualisation de données est rendue possible par la perception. Afin d'observer les comportements dans l'espace immersif, il est nécessaire de le comprendre dans la réalité. La perception peut être définie, en suivant les théories sensorimotrices, comme une fonction qui permet à un être vivant de construire une représentation de son environnement sur la base de l'information physique captée par ses organes sensoriels. Sur le plan biologique, le mécanisme de traduction de l'information a lieu au niveau des champs récepteurs, où une énergie physique (lumineuse pour la vision, mécanique, sonore, ...) est traduite en énergie nerveuse et transmise vers les neurones. Observer la localisation des neurones impactés permet d'identifier la nature de l'information. En effet, les neurones sont regroupés par aires sensorielles spécialisées : aire somatosensorielle, aire auditive, aire visuelle, entre autres. Il a été identifié (Goodale & Milner, 1992) que la perception visuelle se compose de deux systèmes de traitement, séparés et fonctionnant en parallèle : la voie ventrale et la voie dorsale. La voie ventrale est le système orienté perception, pour l'identification et le jugement (couleurs, formes, caractéristiques, ...), tandis que la voie dorsale correspond au système orienté action, pour guider et contrôler visuellement les actions (information sur la position de l'objet à saisir, sa distance, ...).

Plusieurs théories expliquent le fonctionnement de la perception humaine. Les plus importantes sont décrites ici, en prenant essentiellement l'idée de perception visuelle.

Tout d'abord, la théorie indirecte considère la perception comme représentation du monde par des inférences. Chaque individu vit alors dans sa propre représentation du monde. Cette perception comporte trois étapes : le traitement sensoriel (construction d'une scène en deux dimensions à partir des stimuli) ; le

traitement de structuration (transformation de cette scène en trois dimensions) ; l'identification (catégorisation de l'élément perçu. Les étapes de traitement sensoriel et de structuration sont dites ascendantes et utilisent les informations de l'environnement pour inférer la structure de l'objet. L'identification (et une partie de la structuration) sont dites descendantes, et infèrent la catégorie de l'objet en utilisant les connaissances de l'individu ainsi que les informations stockées en mémoire. Cette théorie se base ainsi sur l'idée de cerveau prédictif (approche bayésienne) selon laquelle le cerveau utilise des principes statistiques pour déterminer la cause distale. La représentation du monde pour un individu est donc le reflet de la représentation la plus plausible.

Ensuite, la théorie gibsonienne de la perception apporte le concept d'affordance défini comme les possibilités d'actions fournies par une structure de l'environnement à un individu. Pour Gibson, la perception d'affordance est première (on voit en premier lieu les affordances) et immédiate (aucun raisonnement requis). La perception nous permet ainsi de voir ce qu'il est possible de faire avec une structure, les actions possibles.

Enfin, cette notion d'actions est mise en évidence par la théorie sensorimotrice de la perception. Bergson donne dans cette théorie un rôle constitutif à l'action : la perception est modifiée sous l'effet de nos actions. La notion de lois de contingence (régulant la variation entre sensations et actions) permet d'explorer le monde, et ainsi de le percevoir. Cette théorie donne ainsi un sens à la cécité attentionnelle et la cécité au changement. Dans cette théorie, la perception est rendue possible par la disponibilité motrice : la capacité de se mouvoir permet de percevoir notre environnement.

Par ailleurs, la perception visuelle (sujet principal de cette étude) peut être influencée. Le guidage visuel est rendu possible grâce à cinq facteurs (Wolfe & Horowitz, 2017).

Tout d'abord, la saillance visuelle : certains éléments se distinguent par eux-mêmes des autres éléments autour, de par leur couleur, forme, position. L'attention visuelle est alors attirée par ces éléments saillants.

Ensuite, le guidage descendant. Ce facteur est associé au sens et à la structure de la scène. Lorsqu'un individu cherche un élément, il utilise les propriétés de la cible et de l'environnement. En fonction de la scène, le regard se portera aux endroits où il est le plus probable de trouver la cible (s'il faut trouver un humain, il est plus probable de le trouver au sol que dans les airs, le regard se portera donc davantage vers le sol).

De plus, les antécédents de l'individu peuvent influencer le regard. Si l'individu a eu connaissance de la scène en amont, même quelques millisecondes, son regard se portera d'abord sur les nouveaux éléments.

Enfin, la valeur de l'élément recherché joue également un rôle dans le guidage de la perception. En ajoutant de la valeur (par exemple une récompense) à une propriété visuelle (une couleur ou forme), l'attention sera attirée par toutes les instances de cette propriété.

Ainsi, dans le monde réel, la perception visuelle repose sur des informations issues de l'environnement, des connaissances et des possibilités d'actions. Grâce à la technologie, il est possible de construire une réalité alternative, virtuelle, dans laquelle la vision joue un rôle prépondérant à ce jour, puisque les sens du toucher, de l'ouïe et de l'odorat n'ont pas encore leur équivalent digital. Il est donc indispensable de s'interroger sur les caractéristiques de la perception visuelle : sont-elles conservées dans la réalité virtuelle ?

Pour répondre à cette question, il est nécessaire de définir les systèmes de réalité virtuelle et d'observer le comportement des utilisateurs de ces systèmes.

## Réalité Virtuelle

Une définition commune décrit la réalité virtuelle comme un environnement virtuel immersif accessible par des dispositifs d'entrée et de sortie d'informations (Bowman et al., 2001). La sortie s'effectue généralement par un écran placé sur la tête via un casque, présentant l'état du monde virtuel. Il est également possible d'avoir une sortie audio permettant d'entendre les sons de ce monde. De plus, un retour tactile est parfois utilisé pour aider les utilisateurs à sentir leur environnement. En outre, les dispositifs d'immersion en environnement virtuel offrent fréquemment la possibilité de se déplacer dans le monde virtuel.

Les dispositifs d'entrée sont les outils avec lesquels interagissent les utilisateurs afin de modifier le monde virtuel. Ces outils diffèrent selon les besoins et les technologies utilisées, mais des manettes sont couramment utilisées. Ces outils tangibles permettent de donner une information physique à l'information digitale (Ullmer & Ishii, 2000). L'interface utilisateur tangible (TUI, l'ensemble des outils tangibles) est à différencier de l'interface utilisateur graphique (GUI) qui constitue la majorité des interfaces actuelles (utilisées pour les plans visuels en deux dimensions comme un écran d'ordinateur). Le modèle GUI distingue les éléments physiques de leur représentation numérique : le clavier n'a pas son équivalent dans le monde numérique, l'état du clavier physique n'indique rien de l'état du monde virtuel. En revanche, les TUIs possèdent un lien entre l'outil physique et l'état virtuel : l'outil tangible est associé par sa fonction à la représentation digitale de l'information qu'il véhicule (par exemple, la représentation virtuelle d'un cube physique pourrait suivre en temps réel la position d'un cube physique, et il y aurait une adéquation entre l'état du cube physique - sa position dans le monde physique - et l'état du cube virtuel - sa position dans le monde virtuel.). Ainsi, le modèle TUIs considère qu'il n'y a pas de frontière entre les dispositifs d'entrée et de sortie d'informations, comme si la représentation digitale était la représentation physique.

Il est important de noter que le modèle TUIs n'est pas réservé à la réalité virtuelle. Il s'applique à tous les systèmes dont les dispositifs d'entrée sont associés par leur fonction à une représentation numérique, comme un tableau interactif.

L'intérêt des TUIs et ses bénéfices diffèrent selon les dispositifs et les tâches. Ces interfaces sont bénéfiques si les dispositifs d'entrée-sortie sont adaptés aux missions demandées. Par exemple, pour la résolution d'un problème chez les enfants, les TUIs peuvent avoir des résultats bénéfiques sur l'exploration, la collaboration et la représentation mentale (Antle 2012, sur problèmes de puzzle). Pour la visualisation de données sous forme de diagramme, les interfaces tangibles de réalité virtuelle décrites précédemment (casque, manettes, mouvements) semblent permettre une meilleure compréhension globale des données tout en demandant une charge de travail perçue équivalente à une lecture sur écran (Millais & al., 2018). Ces résultats concordent avec la théorie de la cognition incarnée.

## Cognition incarnée

Les fonctions cognitives sont en partie dessinées par la cognition incarnée (embodied cognition). Cette théorie définit les processus cognitifs comme étant engendrés par les interactions du corps avec le monde. Ainsi, le corps joue un rôle prépondérant dans la formation de l'esprit, et de fait, la compréhension des comportements passe par la compréhension du corps et de ses interactions avec le monde. Les caractéristiques de la cognition incarnée peuvent être définies par les notions d'environnement, de contexte, d'action et d'interaction (Wilson, 2002). Une seule de ces notions ne peut à elle seule définir la cognition : c'est l'interaction de ces notions qu'il faut considérer. Pour comprendre le comportement d'un individu, il devient nécessaire de prendre en compte et d'étudier l'environnement et son contexte, ainsi que les actions possibles et les mécanismes de leurs interactions.

Tout d'abord, le contexte influence la cognition en imposant des limites à la réflexion : si une tâche doit être réalisée dans un temps restreint, le processus cognitif de sa résolution ne sera pas identique à

celui d'une tâche sans limite de temps (Kirsh & Maglio, 1994). Ensuite, et de façon similaire, l'environnement influence la réflexion puisque pour une même tâche, changer d'environnement entraîne un changement des objets à disposition pour réaliser la tâche. Inversement, un même environnement sera perçu différemment en fonction de la tâche à réaliser : certains outils ne seront pas considérés voire perçus différemment (Wilson, 2002, Section 4).

Ainsi, l'environnement et la cognition étant liés, le modèle du monde virtuel devient l'une des causes des comportements que pourront avoir les utilisateurs dans ce monde. Par conséquent, il devient nécessaire de développer un modèle TUIs adapté aux tâches.

Pour une tâche d'exploration visuelle dans un environnement virtuel immersif, le système de dispositifs casque et manettes donne des résultats satisfaisants et semble ainsi former un modèle de TUI adapté (Millais & al., 2018).

Les avancées technologiques sur la réalité virtuelle offrent de nombreuses possibilités pour créer de nouveaux mondes virtuels plus complets et plus réalistes. Cependant, peu d'études s'intéressent à leur propriétés internes, comme la représentation des informations ou les modèles d'interactions. Ces questions sont essentielles puisqu'elles sont liées aux interactions corps-environnement et définissent donc les comportements des utilisateurs.

Dans un environnement non immersif, l'espace d'interactions avec les informations est limité à la surface de l'écran et le comportement corporel de l'utilisateur n'a pas de rôle. Ces frontières peuvent être franchies par des dispositifs immersifs. Selon les études passées en revue, il semblerait qu'il soit bénéfique d'utiliser la réalité virtuelle pour visualiser des informations, à condition que le modèle d'interactions soit adapté.

Ainsi, cette étude répond au manque de connaissances sur les caractéristiques des mondes virtuels pour la visualisation et la recherche d'informations visuelles et répond aux questions suivantes : Quelles sont les meilleures manières de représenter des informations en réalité virtuelle ? La réalité virtuelle est-elle plus efficace que la représentation non immersive ? Le modèle d'interface utilisateur tangible composé d'un casque et de manettes est-il pertinent pour la visualisation et la recherche d'informations ?

## **Expérience**

### **Protocole**

Pour tester ces hypothèses, un logiciel a été développé. Les participants ont été testés sur ce logiciel et ont répondu à un questionnaire subjectif.

L'objectif est d'utiliser la réalité virtuelle pour effectuer de la recherche visuelle. Pour cela, la condition en réalité virtuelle (trois dimensions) sera comparée à la condition en deux dimensions, dite "classique" car elle correspond à la recherche classique d'images sur un ordinateur.

Le logiciel se compose donc de deux conditions.

La première condition est en trois dimensions. L'environnement est constitué d'un sol et de 2000 images flottantes (sans support). Pour voir cet environnement, le participant porte un casque avec écran attaché sur la tête. Il peut se déplacer en marchant et regarder autour de lui avec le casque. Il dispose d'une manette afin de zoomer sur les images (avec une gâchette derrière la manette activée par l'index) et les valider (avec un pavé tactile cliquable). Ces images sont configurées soit en cylindre soit en sphère. À



chaque utilisation du logiciel, les images sont positionnées aléatoirement. En configuration cylindrique, les images sont statiques et disposées à environ 2 mètres du centre de la scène et tout autour du centre (360 degrés horizontalement). Les images forment donc un mur vertical qui entoure le centre de la scène. En configuration sphérique, les images forment un nuage (sphérique) et s'écartent de la tête du participant (du casque) et de la manette lorsque le participant approche ces éléments des images, puis reviennent à leur position initiale lorsque le participant s'écarte. Dans les deux configurations, l'orientation des images est dynamique de manière à ce que le participant puisse voir le contenu de l'image, peu importe sa position dans l'espace virtuel (les images se tournent en continu vers le participant). Les interactions sont donc le zoom et la validation. La manette possède sa représentation dans le monde virtuel et est le seul élément tangible représenté.

La seconde condition est en deux dimensions, affichée sur un écran d'ordinateur fixe. L'environnement est un fond blanc sur lequel sont disposées des images en grille. Les images sont organisées par page avec 6 images par ligne. Les dispositifs d'entrée sont une souris filaire d'ordinateur et un clavier. Il est possible de monter et descendre de la page en utilisant la molette ; de zoomer sur une image en utilisant le clic gauche ; de valider l'image en utilisant le clic droit. Il est possible de naviguer entre les pages en utilisant la flèche droite (respectivement gauche) du clavier pour aller à la page suivante (respectivement précédente).

Les images sont organisées par corpus et correspondent à plusieurs catégories (véhicules, animaux, paysages, etc.). Un corpus est composé de 2000 images. L'objectif est de chercher et valider des images correspondantes à une catégorie cible (i.e. motos pour véhicules) dans un temps imparti. Pour chaque corpus, les images à trouver (parmi la catégorie cible i.e. motos) constitue 8% du total ; 16% est constitué d'images similaires (parmi la super-catégorie i.e. véhicules). Les autres images proviennent d'autres super-catégories.

## **Déroulement**

Chaque participant a passé 3 conditions : format non immersif ; réalité virtuelle cylindrique ; réalité virtuelle sphérique. A chaque condition a été associée un corpus différent et une catégorie cible différente. Les participants n'ont pas passé les conditions dans le même ordre et n'ont pas été testés sur les mêmes corpus (Annexe 1).

Après une brève description du travail de recherche, chaque participant a répondu à un premier questionnaire sur leur expertise avec les outils technologiques. Ensuite, une démonstration du logiciel a été donnée ainsi qu'une explication des mécanismes d'interactions, suivie d'une session d'essai. Si leur première configuration de test était dans la réalité virtuelle, la démonstration se déroulait en configuration sphérique, sinon dans la configuration en deux dimensions. La session d'essai s'est effectuée avec un corpus test. La session test n'a pas été limitée dans le temps (aucun participant n'a utilisé plus de 5 minutes). Chaque participant a ensuite été averti de sa tâche de recherche : valider le maximum d'images cibles en 5 minutes tout en gardant en mémoire les catégories des autres images qu'ils ont pu voir pour les restituer à la fin. À la fin du temps imparti, un questionnaire a été donné, portant sur l'intuitivité de la technologie et l'utilité de la configuration (sphérique, cylindrique, ou classique) pour la tâche. Après le dernier questionnaire, un entretien est réalisé avec le participant pour obtenir son ressenti sur l'expérience et sa performance.

La réalité virtuelle étant un environnement immersif, et d'après les données des articles cités, il est attendu que les participants en réalité virtuelle aient de meilleures performances que les participants en condition classique. Il est possible que la saillance visuelle soit mieux répartie en condition cylindrique (les images n'étant pas cachées par celles au premier plan). Il est donc attendu que la condition cylindrique soit perçue comme étant plus intuitive et donne lieu à de meilleurs résultats que la condition sphérique.

## Résultats

Neuf participants ont passé l'expérience, six hommes et trois femmes, âgés de 19 à 44 ans. Leur niveau d'étude varie de 1 à plus de 5 ans après le baccalauréat. Sur une échelle de Likert, quatre participants situent leur niveau de familiarité avec la réalité virtuelle à 2, un participant à 3, un participant à 4, un participant à 5, et deux participants à 7. Un participant, de niveau de familiarité 4, a dit avoir déjà eu un mal de tête à cause de la réalité virtuelle (*motion sickness*). Un autre participant a également dit avoir eu ce mal de tête mais explique s'être vite habitué. Les sept autres participants n'ont pas rencontré de problèmes lorsqu'ils ont utilisé par le passé la réalité virtuelle. Les participants évaluent leur niveau d'aisance avec la manipulation de dispositifs numériques entre 5 et 7 sur une échelle de Likert (deux 5 ; 2 six ; 5 sept). Enfin, quatre participants ne portent pas de lunettes et un est passé en portant ses lunettes sous le casque. Les quatre autres participants, avec des corrections visuelles, n'ont pas porté leurs lunettes (leurs troubles de la vision sont donc reproduits dans la réalité virtuelle).

Les réponses des participants après la passation sont représentées dans le graphique suivant (Figure 1). Les réponses se font sur une échelle de Likert de un à sept points, et leur représentation graphique est de -3 à 3 ("pas du tout" à "tout à fait"). Ainsi, plus la valeur est basse dans le graphique, plus la réponse est négative et tend vers "pas du tout". A l'inverse, plus la valeur est haute, plus la réponse tend vers "tout à fait". Ces résultats montrent que les participants n'ont pas trouvé complexe la manipulation de la technologie en condition classique (2D) et en configuration cylindrique (3D cylindre). Cependant, les participants ont trouvé la complexité moyenne (valeur 0) en configuration sphérique (3D sphère). Les participants affirment, en moyenne, que les trois configurations étaient adaptées à la recherche d'images. Toutefois, les valeurs sont différentes : les participants donnent le plus de "points" à la condition classique, et le moins à la condition sphérique.

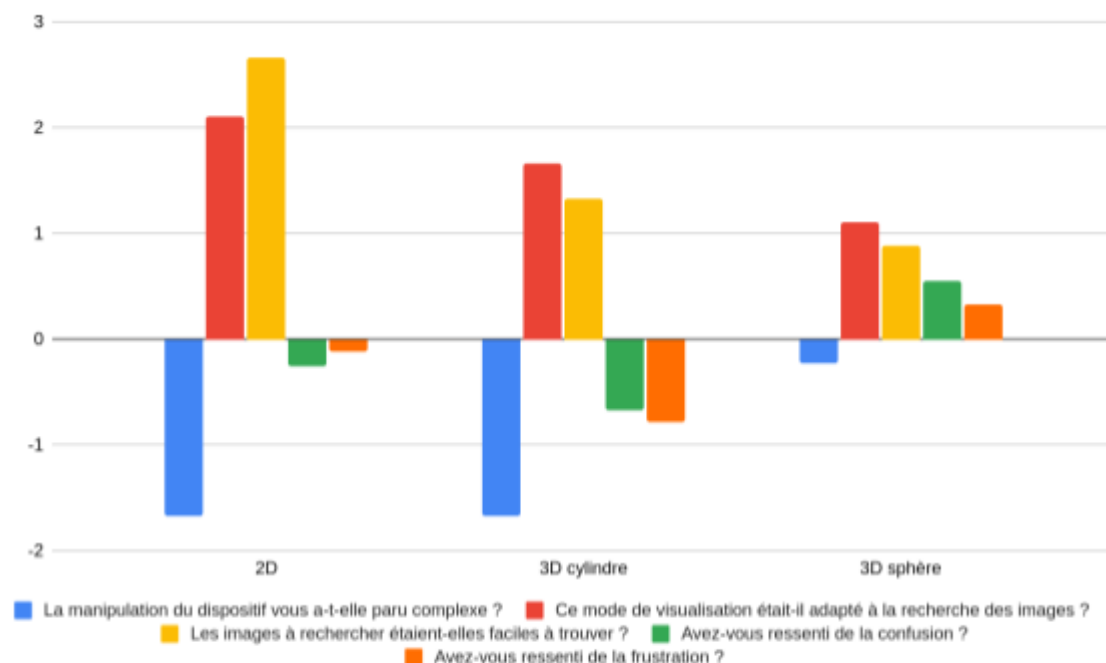


Figure 1 : Moyenne des réponses aux questionnaires en fonction des configurations

Les données étudiées ci-dessous ont été enregistrées automatiquement par le logiciel. L'ensemble des variables comme la position du casque, l'orientation, l'accélération ou encore les images validées ont été enregistrées toutes les 30 millisecondes. L'enregistrement des images a échoué à deux reprises pour la condition en deux dimensions, les restes des données (i.e. la position de la souris) ont pu être exploitées.

Sur les sept conditions en deux dimensions, six se sont révélées avoir de meilleurs résultats que les autres configurations : un nombre d'images validées correctes plus élevé. Pour l'environnement immersif, le nombre d'images validées est plus élevé dans la configuration cylindrique dans six des neuf conditions.

Quatre images validées appartiennent ni à la catégorie-cible ni à la super-catégorie, ce sont donc des images incorrectes. Deux de ces erreurs ont eu lieu dans la condition classique en deux dimensions, une en condition sphérique et une en condition cylindrique. Ces erreurs ont chacune eu lieu dans un corpus différent.

Le nombre d'images hors catégorie-cible mais dans la super-catégorie est au maximum de 9 par session avec une moyenne de 2,48 et une médiane à 2 (toutes conditions confondues).

En configuration sphérique, les participants ont, en moyenne, attendu cinq secondes entre chaque zoom (entre le moment où le participant lâche la gâchette et arrête de zoomer, et le moment où il appuie à nouveau sur la gâchette pour zoomer) et ont en moyenne réalisé 62 actions de zoom. En configuration cylindrique, les participants ont attendu en moyenne 3,88 secondes entre chaque zoom, pour une moyenne de 77 zooms par session. En deux dimensions, les participants n'ont utilisé que 13 fois en moyenne la fonction de zoom, et ont attendu 20 secondes en moyenne entre les zooms, auquel s'ajoute un participant qui n'a jamais utilisé la fonction de zoom (le seul, toutes configurations confondues).

Entre chaque validation s'est passé en moyenne un délai de 9 secondes pour la configuration sphérique ; 7 secondes pour la configuration cylindrique ; 5 secondes pour la configuration en deux dimensions.

Les graphiques du nombre d'images validées en fonction du temps (Figure 2) montrent que les conditions sphériques possèdent davantage de plateaux (que confirment les délais précédents) : périodes pendant lesquelles un participant ne valide pas d'images et est en phase de recherche. La configuration en deux dimensions ne possède que quelques plateaux, et le nombre d'images validées semble linéaire en fonction du temps : les participants valident régulièrement des images. La configuration cylindrique se situe entre les deux autres configurations : certains plateaux apparaissent, mais relativement courts, et hors de ces plateaux le nombre d'images validées augmente rapidement.

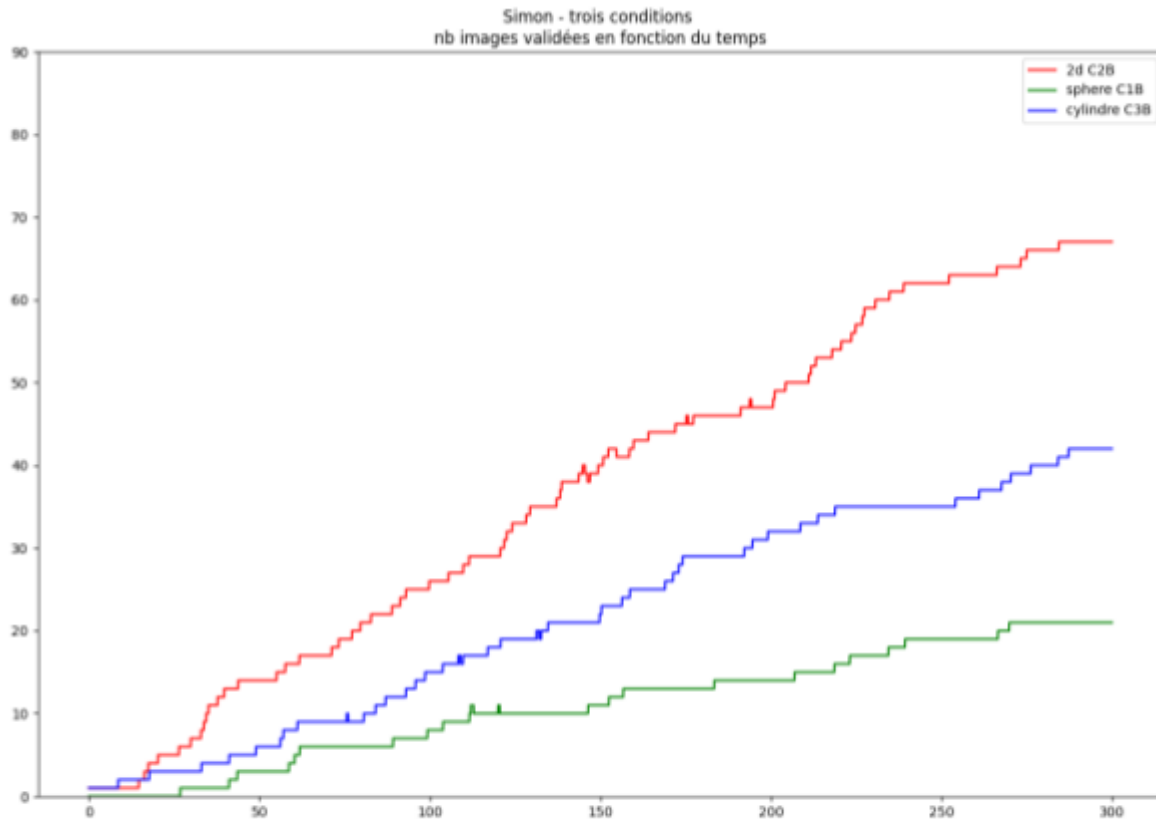


Figure 2 : Nombre d'images validées en fonction du temps (secondes) pour le participant Vincenzo

Les données enregistrées donnent également la variation du nombre d'images dans le champ de vision (Figure 3 et Figure 4), compris au sens large : une image est comptabilisée dans le champ de vision lorsque sa position est dans l'angle de vue du participant, même si le participant ne voit pas cette image (en configuration sphérique, les images peuvent en cacher d'autres ; les images visibles et celles cachées seront considérées comme étant dans le champ de vision).

De plus, la position du casque (i.e. du participant) est représentée sur deux graphiques. D'abord, sa position dans la pièce, vue de haut (Figure 5). Puis, la hauteur du casque (i.e. de la tête du participant) est représentée au cours du temps (Figure 6). Une baisse significative de cette hauteur signifie que le participant s'est accroupi, ou fortement penché. La hauteur de la tête peut, pour certains participants, évoluer entre 2 et 2,5 mètres. Dans ces rares cas, il s'agit d'une erreur de calibrage mais n'a pas d'effet sur la passation de l'expérience.

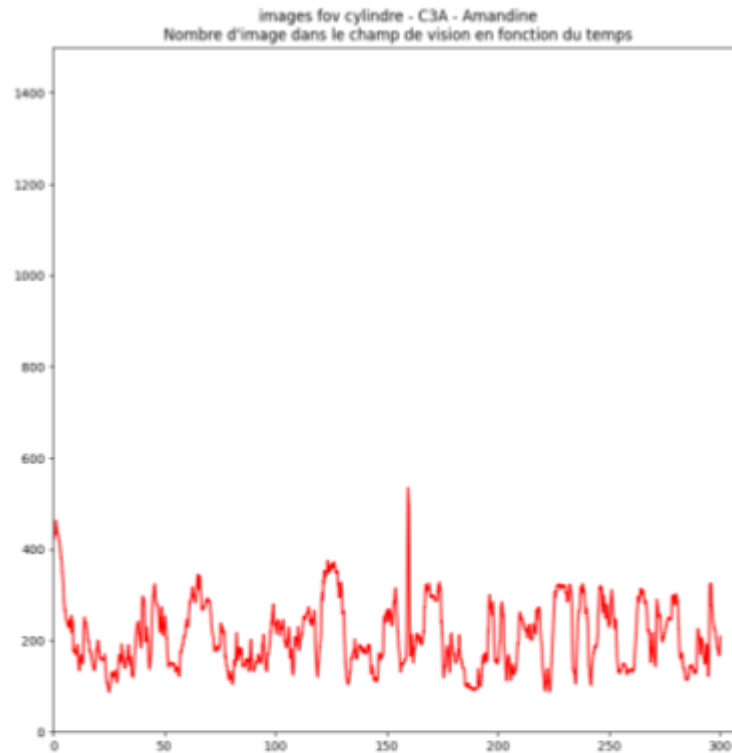


Figure 3 : Nombre d'images dans le champ de vision en fonction du temps pour une configuration cylindrique

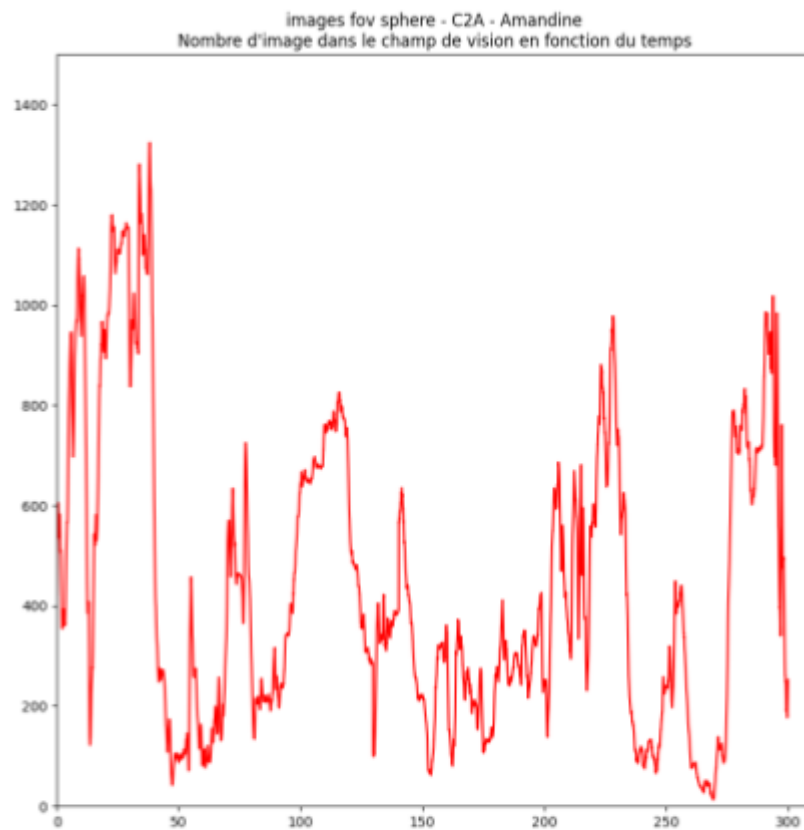


Figure 4 : Nombre d'images dans le champ de vision pour une configuration sphérique

## Discussions

Dans cette partie seront données des interprétations des résultats. L'analyse non littérale des résultats (c'est-à-dire supposée) sera explicitement précisée.

Tout d'abord, pour expliquer les résultats aux questionnaires, il est intéressant d'observer les stratégies de recherche des participants. Les graphiques du nombre d'images dans le champ de vision au cours du temps et les suivis de position peuvent nous renseigner sur les techniques utilisées pour la recherche d'images. En condition cylindrique, le nombre d'images dans le champ de vision se situe en moyenne entre 300 et 400 images. Les variations sont cycliques : pendant quelques dizaines de secondes le nombre d'images augmente, puis pendant quelques dizaines de secondes il descend. En fonction des participants, l'amplitude de variation est plus ou moins grande, mais le schéma reste le même. L'observation des positions des participants nous renseigne sur leur stratégie de recherche en configuration cylindrique. Leur position suit le tracé du cylindre : les participants regardent le mur du cylindre pas à pas, la position au cours du temps forme un cercle. De par la cyclicité des variations des images dans le champ de vision couplée à la hauteur du casque (qui suit le rythme des variations) et les réponses aux entretiens, il est possible de conclure que les participants ont tous opté pour la même stratégie : un parcours vertical, colonne par colonne, du cylindre. Certains participants reviennent par moment à une colonne précédente, peut être un effet *insight*, mais la majorité des participants est méthodique et poursuit linéairement. Un seul participant a détaillé le processus de sa recherche : chercher premièrement les images aux couleurs similaires à son image mentale de la cible, puis une fois le premier tour de cylindre effectué, essayer d'identifier les autres propriétés (forme, environnement possible, etc.) de la cible sur les images.

De même, la configuration classique, en deux dimensions, fait l'objet d'une stratégie de parcours linéaire : les participants parcourent les images en ligne, puis une fois arrivés en bas, passent à la page suivante.

Dans ces configurations, la majorité des participants n'ont pas vu la totalité des images : en configuration classique, ils n'ont pas atteint les dernières pages ; en configuration cylindrique, ils n'ont pas fait le tour du cylindre. Ici, toutes les images sont peut-être rentrées dans leur champ de vision, mais il est supposé que l'attention des participants était, dans les moments où ils regardaient autour d'eux en configuration cylindrique et dans les pages suivantes pour la configuration classique, moins basée sur la nature des images que sur le volume d'images qu'ils leur restaient à parcourir. Ainsi, les participants ont peut-être regardé la totalité des images, mais ils ne les ont pas vues. Si la limite de temps était augmentée, il est certain qu'ils auraient observé toutes les images.

L'étude de la position en configuration sphérique ne traduit pas de stratégie de recherche. La visualisation du nombre d'images dans le champ de vision montre que les participants avaient par moment un important volume d'images devant leurs yeux, et à d'autres moments, seulement une dizaine d'images. Ces variations sont dues à leur position : lorsqu'ils sont à l'extrémité de la sphère et regardent vers l'extérieur, seulement quelques images sont dans leur champ de vision, et lorsqu'ils se retournent, la majorité des images se trouvent devant leurs yeux. Certains participants ont expliqué que les extrémités de la sphère faisaient partie de leur stratégie : ils avaient moins d'images autour d'eux et pouvaient mieux se concentrer et savoir où regarder. D'autres participants n'y sont allés qu'après avoir fini, selon eux, de chercher la partie centrale de la sphère. Les entretiens renseignent sur l'impact de la confusion sur la stratégie : les participants se sont sentis confus par le nuage d'images et pour beaucoup ne savaient pas par où commencer, ils ont donc parcouru l'espace en regardant autour d'eux.

Ensuite, le questionnaire permet de savoir quelles configurations ont été jugées appropriées pour la recherche d'images.

Les réponses des participants montrent que la condition sphérique a été perçue plus difficile à prendre en main que les deux autres conditions (valeurs similaires entre la condition classique et la condition sphérique). De même, la configuration sphérique a été perçue comme la moins adaptée à la recherche d'image et la plus frustrante. Ensuite, cette configuration génère la disposition d'images où les images sont les moins faciles à trouver. Cependant, cette donnée peut être mal interprétée : la disposition des images en configuration sphérique nécessite forcément une recherche plus active que les autres conditions. En effet, les stratégies de balayage visuel des participants pour les configurations cylindriques et classiques transforment leur tâche de recherche en tâche d'identification. Les participants parcourent toutes les images une par une sans être vraiment sélectif comme le montre la cyclicité des variations d'images dans le champ de vision (bien que les cycles soient parfois plus courts ou plus longs, ils restent globalement à la même fréquence ou vitesse), et les entretiens. La tâche de recherche est donc nécessairement plus compliquée dans la condition sphérique, puisque c'est dans cette configuration que la *recherche* au sens de *chercher* est nécessaire. Enfin, la confusion a été plus grande en condition sphérique. Ce résultat peut s'expliquer par l'essence même de la configuration classique : le grand nombre d'images qui entourent le participant est peu habituel dans la réalité. La nécessité de regarder en bas comme en haut n'a pas été jugée très intuitive par les participants qui trouvent ces positions peu ergonomiques. Pareillement, les valeurs plus élevées de frustrations pour la configuration sphérique ne doivent pas accuser cette configuration : la question ne s'adressait que sur la frustration ressentie, sans mentionner de cause. Il est possible que cette frustration soit due au logiciel et non à l'idée d'être entourée par des images. En effet, en interrogeant les participants, beaucoup se disent frustrés de n'avoir pu sélectionner facilement les images qui ne sont pas au premier plan. La frustration pourrait donc relever un manque d'intuitivité de la technique, sans remettre en cause la théorie sphérique.

La configuration classique, en deux dimensions, produit de meilleurs résultats, le nombre d'images validées et correctes étant plus grand. L'affichage par pages incite les participants à finir une page avant de passer à la suivante, et donc de regarder toutes les images. Il devient moins probable de laisser passer des images que dans les autres configurations. Cependant, cette configuration n'est pas la préférée de la majorité des participants. L'installation demande aux participants d'être assis devant un écran et de faire défiler des images. Cette installation a ennuyé beaucoup de participants qui ont préféré la réalité virtuelle et la possibilité de se déplacer et d'interagir avec les images. De plus, certains ont apprécié, dans la configuration cylindrique, pouvoir juger de leur avancée dans la tâche en regardant combien d'images restaient au parcours. Cette fonctionnalité aurait ainsi dû être ajoutée à la configuration classique, en précisant le nombre total de pages et la page actuelle comme c'est le cas sur les GUIs actuelles. L'aspect ennuyeux n'aurait toutefois pas été supprimé par cet ajout. À noter que certains participants ayant des troubles de la vision ont tout de même préféré la configuration classique car ils pouvaient mieux voir les images. Ce problème aurait pu être contourné en utilisant un casque pouvant adapter l'affichage aux troubles de la vision.

Ainsi, dans le cadre de la recherche d'images par catégorie-cible sur une période de cinq minutes, la configuration classique en deux dimensions fournit de meilleurs résultats. De plus, les participants disent pouvoir mieux voir les détails des images et ont jugé qu'ils s'agissaient de la configuration où les images étaient le plus facile à trouver. Cependant, l'ennui de la tâche relevé par les participants permet de supposer que sur le long terme, cette configuration pourrait ne pas être la plus performante : les participants seraient moins attentifs et pourraient perdre le fil de leur recherche et rater des images.

Toutefois, la majorité des participants de cette étude n'est pas familière avec la réalité virtuelle. Des résultats plus performants peuvent être imaginés si davantage de participants accoutumés utilisaient le dispositif. Là où l'espace bi-dimensionnel est limité par la taille de l'écran et les interactions unitaires avec la technique, il est envisageable que l'espace virtuel place la limite de performances plus haute.

En effet, la réalité virtuelle offre de plus grandes possibilités d'affichage de données. En développant d'autres logiciels, il serait possible d'annihiler les frustrations ressenties dans cette étude. Cependant, il serait difficile de contourner l'ennui ressenti par les participants pour la représentation non immersive, les postes d'ordinateur étant fixes. Le dispositif technologique immersif utilisé dans cette étude pour la visualisation d'informations semble pertinent au vu des commentaires des participants. Leurs résultats dans les conditions immersives permettent de réfléchir à d'autres manières de représenter des données, ainsi qu'à d'autres applications.

Pour chercher des images, la configuration cylindrique serait plus adaptée que la configuration sphérique, en associant visibilité et action (et donc longévité de l'attention). Pour permettre une bonne visibilité de toutes les images, il pourrait toutefois être nécessaire de courber les extrémités hautes et basses du cylindre, les images à ces positions étant plus loin de l'utilisateur ; ou de réduire la hauteur du mur en la centrant à hauteur des yeux. Ainsi, ce mur virtuel deviendrait une nouvelle galerie de données, qu'elles soient artistiques, informationnelles, ou autres.

La configuration sphérique pourrait être utilisée à des fins de démonstrations prototypiques : la difficulté de voir l'ensemble des images restreint son utilisation à des objectifs ne nécessitant pas d'être précis sur l'ensemble du corpus.

De plus, il aurait été intéressant d'avoir d'autres interactions avec l'environnement, comme la possibilité de déplacer les images. Cette action naturelle dans la réalité permettrait d'utiliser de nouvelles stratégies pour d'autres tâches que la recherche, comme le tri d'informations, les utilisateurs pouvant faire des "paquets" de données. L'espace virtuel étant infini, les utilisateurs pourraient agencer les images de plusieurs façons, certaines impossibles dans la réalité (impossible par exemple par la vitesse de l'action, la position flottante, toujours visible, etc.). En ajoutant une deuxième manette, le tri, la visualisation, et la recherche pourraient être plus rapides en sélectionnant plusieurs images en même temps pour comparer ou libérer de l'espace avec une main pendant que l'autre maintient plusieurs images. L'ensemble des possibilités ne tient qu'à l'imagination. Bien que l'ensemble des possibilités "réellement" utilisables soit soumises aux tâches, et comme cette étude le montre, à la manière de visualiser les données.

## Conclusion

Cette étude montre que l'utilisation des techniques actuelles pourrait être dépassée par de nouvelles techniques permettant un plus grand nombre d'interactions. L'espace bi-dimensionnel à l'aide d'une souris appliquée dans le cadre d'une recherche d'images et effectuée dans un temps court permet de meilleurs résultats en montrant l'ensemble des images avec une taille assez grande pour voir les détails. Cependant, les résultats de cette étude sous-tendent l'inefficacité des technologies actuelles pour la recherche d'information visuelle sur une longue période de temps. Alors, les technologies immersives semblent combler ce vide technique en proposant de nouvelles dispositions, de nouvelles façons de voir. Ainsi, la visualisation sous format cylindrique est plus intuitive, mieux comprise et plus utile dans une tâche de recherche que le format sphérique. En outre, le système technique de visualisation composé d'un casque et de manettes est approprié pour la visualisation de données : le casque permet une perception similaire au monde réel et les manettes permettent à minima de zoomer sur les informations (comme agrandir la taille ou afficher des détails).

A partir de ces résultats, il est nécessaire de poursuivre l'exploration des possibilités proposées par la réalité virtuelle, en améliorant la solution et la visualisation comme décrit dans cette étude, et fournir de nouvelles manières d'interactions. Les environnements immersifs pourraient alors générer de meilleurs résultats que leurs homologues non immersifs et étendre les utilisations à de nouveaux objectifs.



## Référence

- M. Wilson, 2020, *Six views of embodied cognition*, Psychonomic Bulletin & Review, 9 (4), 625-636
- Kirsh and Maglio, 1994, *On Distinguishing Epistemic from Pragmatic Action*, Cognitive Science, 18, 513-549
- Ullmer & Ishii, 2000, *Emerging frameworks for tangible user interfaces*, IBM Systems Journal, 39, 915-931
- Wolfe & Horowitz, 2017, *Five factors that guide attention in visual search*, Nature Human Behavior
- Bowman et al, 2001, *An introduction to 3-D User Interface Design*, Teleoperators and Virtual Environments, 2001, 10, 96-108
- A. Antle, 2012 *Exploring how children use their hands to think: an embodied interactional analysis*, Behaviour & Information Technology, 32:9, 938-9
- Millais & al., 2018, *Exploring Data in Virtual Reality: Comparisons with 2D Data Visualizations*, Computer and Human Interactions, 18



## Annexes

### Annexe 1

L'alternance des corpus et configurations est donné par le tableau suivant :

Participants	Configuration 1	Configuration 2	Configuration 3
P1	2D (C1a) - canetons	3D sphère (C2a) - barques	3D cylindre (C3a) - rétro
P2	2D (C1b) - colverts	3D cylindre (C2b) - à voile	3D sphère (C3b) - sportives
P3	3D cylindre (C1a) - canetons	2D (C2a) - barques	3D sphère (C3a) - rétro
P4	3D cylindre (C1b) - colverts	3D sphère (C2b) - à voile	2D (C3b) - sportives
P5	3D sphère (C1a) - canetons	3D cylindre (C2a) - barques	2D (C3a) - rétro
P6	3D sphère (C1b) - colverts	2D (C2b) - à voile	3D cylindre (C3b) - sportives
P7	2D (C1a) - canetons	3D sphère (C2a) - barques	3D cylindre (C3a) - rétro
P8	2D (C1b) - colverts	3D cylindre (C2b) - à voile	3D sphère (C3b) - sportives
P9	3D cylindre (C1a) - canetons	2D (C2a) - barques	3D sphère (C3a) - rétro