



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 12328

To link to this article : DOI :10.3166/tsi.32.605-638
URL : <http://dx.doi.org/10.3166/tsi.32.605-638>

To cite this version : Schmitt, Bénédicte and Raynal, Mathieu and Bach, Cédric and Croenne, David and Dubois, Emmanuel *[Vers une unification de l'évaluation de la tâche de pointage en environnement virtuel 3D](#)*. (2013) *Technique et Science Informatiques*, vol. 32 (n° 5). pp. 605-638. ISSN 0752-4072

Any correspondance concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Vers une unification de l'évaluation de la tâche de pointage en environnement virtuel 3D

Bénédicte Schmitt^{1,2}, Mathieu Raynal², Cédric Bach²,
David Croenne¹, Emmanuel Dubois²

1. Global Vision Systems

Immeuble E-evolution, 425 rue Jean Rostand, F-31670 Labège
{prenom.nom}@global-vision-systems.com

2. Université de Toulouse, IRIT - Elipse

118 Route de Narbonne, F-31062 Toulouse cedex 9
{prenom.nom}@irit.fr

RÉSUMÉ. De nouveaux systèmes de visualisation pour de larges ensembles de données complexes se développent. Parmi eux, ceux basés sur des environnements virtuels 3D (EV 3D) s'avèrent être une solution pertinente car ils offrent une vue permettant d'agrèger les multiples données. Pour étudier l'utilisabilité de ces environnements, et notamment la partie liée à la présence d'un EV 3D, des tâches d'interaction en EV 3D ont été répertoriées et sont souvent prises comme référence. Cependant, il manque un protocole unifié pour évaluer les tâches élémentaires dans des EV 3D et favoriser la comparaison des solutions. De plus, il n'existe pas aujourd'hui de base de connaissances répertoriant des résultats d'évaluation de techniques d'interaction en EV 3D. Une norme est disponible pour les tâches de pointage en 2D, mais il n'y a pas d'équivalence en 3D. Dans cet article, nous proposons une adaptation de cette norme pour une tâche de pointage dans un EV 3D. Nous détaillons notre protocole et son instrumentation. Ce protocole a pour but de mesurer la performance, le confort des techniques et la satisfaction utilisateur. Nous présentons enfin une mise en œuvre de ce protocole et les résultats de l'expérimentation utilisateur ainsi conduite.

ABSTRACT. New visualization systems for large and complex datasets are emerging. Among these systems, systems based on 3D Virtual Environments (3D VE) turn out to be a relevant solution, because they provide a view, which brings together many data. To study the usability of these environments, and especially the part regarding 3D data, interaction tasks in these 3D VE have been identified and are often used as reference. Nevertheless, there is a lack of unified protocol to assess these elementary tasks in 3D VE and to encourage comparison of solutions. Moreover there is no knowledge base, which lists evaluation results of interaction techniques in 3D VE. A standard is available for 2D pointing task, but there is no equivalence in 3D. In this paper, we propose an adaptation of this standard to a pointing task in a 3D VE. We detail our protocol, which aims at assessing performance, comfort of techniques and satisfaction of users, and an instrumentation. We also present results of a user experimentation conducted according to this standard's adaptation.

MOTS-CLÉS : tâche de pointage, norme ISO 9241-9, environnement virtuel 3D, étude d'utilisabilité, tests utilisateur.

KEYWORDS: pointing task, ISO 9241-9, 3D virtual environment, usability study, user tests.

1. Introduction

Les environnements virtuels 3D (EV 3D) ont émergé dans différents domaines pouvant aller des jeux (Duke Nukem ou encore Prince of Persia) aux applications industrielles, telles que des systèmes d'informations géographique (SIG, comme ArcGIS de Esri), la conception assistée par ordinateur (CAO, comme Catia de Dassault Systèmes), etc. Un EV 3D peut être défini comme

un espace où un ou plusieurs utilisateurs interagissent en temps réel à travers des moyens physiques et des dispositifs, avec des données 3D générées par un ordinateur (Bach et Scapin, 2003).

Ces environnements peuvent être soit immersifs, tels que des systèmes basés sur la vision stéréoscopique ou encore les systèmes CAVE utilisés en réalité virtuelle, soit non immersifs et ainsi afficher les informations sur des écrans de bureau. Notre étude s'intéresse plus particulièrement aux EV 3D non immersifs : ils ne requièrent pas un espace de travail dédié, ils peuvent être utilisés en situation de travail de bureau, s'adressent à un public plus large et ne nécessitent pas l'utilisation de technologies dédiées ou coûteuses. Récemment de nouvelles technologies, telles que la Wiimote et la Kinect, sont apparues auprès du grand public : ces technologies ont permis de proposer des formes d'interaction avancée et ont rencontré un grand succès auprès des utilisateurs. Majoritairement destinées à des applications de jeux, l'étude de (Gallo *et al.*, 2008) a montré qu'elles s'avèrent également adaptées à une utilisation en EV 3D. Avant tout vouées à des situations ludiques et grand public, ces formes d'interaction avancée doivent prioritairement s'avérer satisfaisantes et confortables.

En parallèle, des dispositifs dédiés, tels que la Rockin'Mouse (Balakrishnan *et al.*, 1997), la Cubic Mouse (Froehlich et Plate, 2000), ont été proposés pour offrir des formes d'interaction avancée avec des environnements 3D dans des contextes plus professionnels. Mais ils sont encore peu répandus et ne font pas partie des dispositifs d'interaction bien acceptés pour ces activités. Ce sont des dispositifs classiques, tels que la souris, le joystick ou la souris 3D qui restent encore les plus utilisés, malgré une adaptation aux tâches et une adéquation limitées. Dans ces contextes professionnels, les techniques d'interaction, classique ou avancée, doivent répondre à des besoins spécifiques tels que la précision, la rapidité d'exécution, la robustesse et la facilité d'apprentissage. La satisfaction et le degré de confort ont aussi un rôle à jouer pour contribuer à l'acceptation des techniques du fait de leur utilisation quotidienne.

Afin d'accompagner et de développer la mise en œuvre de forme d'interaction avancée pour interagir avec des environnements 3D, dans des domaines grands publics ou professionnels, il semble nécessaire de s'appuyer sur des moyens d'évaluation permettant de concilier performance et satisfaction : ces moyens sont nécessaires pour permettre de mesurer et comparer selon différents points de vue complémentaires les techniques d'interaction avancée proposées en EV 3D. Toutefois, les expérimentations menées dans le contexte d'EV 3D visent essentiellement l'étude de la performance des dispositifs utilisés. De plus, ces expérimentations ont pour objectif de comprendre l'adéquation des techniques d'interaction proposées à une ou plusieurs des tâches d'interaction 3D bien établies : sélection, manipulation ou navigation (Bowman, 1999). Enfin, pour réaliser ces tâches, les solutions d'interaction envisageables se multiplient, notamment du fait de l'émergence de nouvelles technologies telles que les interactions tangibles et les systèmes d'interaction mixte (SIM). Les SIM prennent avantage des objets physiques pour interagir avec des entités numériques, qui sont utilisées pour communiquer avec un système et qui constituent une véritable extension de l'utilisateur. Cette multiplication des technologies couplée à un manque de cadre pour mener les évaluations conduisent à réaliser des expérimentations peu reproductibles et à recueillir des résultats complexes à exploiter et comparer. En effet, de nombreux paramètres sont pris en compte et combinés au sein de chaque expérimentation (degrés de liberté, tâches, techniques d'interaction), obérant ainsi la compréhension, la capitalisation et la réutilisation des données collectées et les résultats déduits.

Néanmoins, en se concentrant sur des tâches d'interaction élémentaires (*i.e.* constitutives des tâches de haut niveau telles que la navigation, la sélection et la manipulation), les études expérimentales porteraient sur des situations moins complexes et seraient donc plus facilement reproductibles. Les recommandations construites sur la base d'une tâche élémentaire pourraient ensuite être transposées et prises en compte dans la conception de techniques d'interaction pour des tâches de plus gros grains. Or, le pointage ou le déplacement peuvent être considérés comme des tâches élémentaires d'un système interactif, et ceci même en 3D. La tâche de pointage est d'ailleurs une des tâches les plus répandues dans les applications : elle consiste à déplacer un curseur jusqu'à un endroit spécifique, en modifiant ses coordonnées en translation (tx , ty , tz). La tâche de pointage est inhérente et sous-jacente aux tâches de sélection, qui précèdent elles-mêmes quasi-systématiquement les tâches de manipulation. Certains parallèles ont d'ailleurs été dressés entre les tâches de sélection et de manipulation : les deux concernent la position d'un objet, la première la spécifie tandis que la seconde la modifie (Pan, 2008). Des recommandations extraites des évaluations de la tâche de pointage pourraient donc être transposées aux tâches de sélection et de manipulation. Par ailleurs, l'intérêt de se concentrer prioritairement sur des expérimentations de techniques d'interaction en EV 3D pour une tâche élémentaire est conforté par une étude de (Masliah, 2001). Celle-ci a révélé l'intérêt de décomposer l'évaluation de la tâche de manipulation par l'évaluation séparée des deux sous-tâches définies par (Bowman, 1999) et qui sont une tâche de translation (spécification d'une position) et une tâche de rotation (spécification d'une orientation). En effet, les utilisateurs les considèrent comme deux tâches différentes ; cette étude renforce donc l'intérêt que représente une tâche

élémentaire claire et sans biais pour les utilisateurs comme base pour l'évaluation de l'interaction en EV 3D.

Notre contribution pour la réconciliation des aspects performance et satisfaction dans l'évaluation des techniques d'interaction pour les EV 3D a pour but, d'une part, de dépasser les approches expérimentales *ad hoc* et mal maîtrisées et, d'autre part, de construire une approche ciblée, structurée, guidée par une méthodologie établie. A ces fins, nous proposons un cadre d'évaluation des techniques d'interaction pour la tâche élémentaire de pointage en EV 3D. Ce cadre constitue une approche composite, reproductible, combinant des aspects quantitatifs (*i.e.* la performance, la robustesse et la précision) et des aspects qualitatifs (*i.e.* la satisfaction, la facilité d'apprentissage et le confort). Pour y parvenir, nous proposons une adaptation du protocole de la norme ISO 9241-9. Initialement dédiée au pointage en 1D ou 2D, cette norme combine des mesures quantitatives et qualitatives. Notre adaptation a pour objectif d'étendre son applicabilité aux environnements 3D et à toutes techniques d'interaction en mesurant simultanément la performance et la satisfaction utilisateur.

Après avoir présenté les travaux existants sur l'évaluation de techniques d'interaction en environnements 3D, nous décrivons et justifions les adaptations requises dans les différentes parties du protocole. Nous présentons également un outil logiciel utilisé pour conduire les évaluations de la tâche de pointage en 3D. Enfin, pour montrer l'intérêt de notre cadre, nous présentons une instance de notre adaptation de la norme en détaillant une expérimentation, conduite conformément à cette adaptation. Cette expérimentation, pour laquelle nous proposons une analyse des résultats obtenus au moyen de notre outil, a pour objectif de comparer et d'évaluer des techniques d'interaction pour le pointage en EV 3D.

2. Évaluations menées en environnements 3D

Deux types d'approches ont été relevés dans la littérature concernant les évaluations en EV 3D : des approches structurées se basant sur des normes ou des cadres précis d'une part, et des approches *ad hoc* menées de façon empirique, d'autre part. La norme ISO 9241-9 est couramment utilisée dans les approches structurées pour évaluer, entre autres, la tâche de pointage ; de plus, cette norme commence à être adaptée à la 3D dans certains travaux. Pour les approches *ad hoc*, il s'agit avant tout de tester et de comparer des techniques d'interaction ayant au moins 3 degrés de liberté (DL). Nous allons présenter dans un premier temps une revue des approches structurées puis synthétiser les approches *ad hoc* présentes dans la littérature.

2.1. Évaluation de la tâche de pointage

2.1.1. Loi de Fitts et ses extensions

La loi de Fitts (1954) a été initialement conçue pour prédire le temps mis par un utilisateur pour atteindre une cible en 1D. La formulation couramment utilisée est la suivante :

$$MT = a + b * ID \text{ avec } ID = \log_2(D/W + 1) \quad (1)$$

où a et b sont des constantes dépendantes du contexte d'interaction et sont déterminées de façon empirique. W correspond à la taille de la cible et D est la distance à parcourir pour atteindre la cible souhaitée. Le terme logarithmique est l'indice de difficulté (ID) mesuré en bits.

Suite à cette première loi, des propositions de modèle ont été faites pour étendre cette formule à 2 ou 3 dimensions. Ainsi, MacKenzie et Buxton (1992) ont proposé deux équations différentes pour le calcul de l'ID en 2 dimensions, et notamment l'équation suivante pour prendre en compte les différences potentielles entre hauteur et largeur d'une même cible :

$$ID = \log_2(D/\min(W, H) + 1) \quad (2)$$

Mais, Accot et Zhai (2003) ont montré par la suite certaines limitations à ces modèles, telles que le manque de prise en compte de la contrainte de direction compte tenu de la différence possible entre les valeurs de la hauteur et de la largeur des cibles, et ont eux-mêmes proposé un modèle pour la 2D.

Certaines adaptations de la loi de Fitts ont également été proposées pour la 3D, comme le modèle de Ware et Lowther (1997) qui ont étendu le modèle 2D défini par MacKenzie et Buxton (dans ce modèle W est la taille de la cible, H la hauteur et D la profondeur).

$$ID = \log_2(D/\min(W, H, D) + 1) \quad (3)$$

Cependant, ce modèle souffre des mêmes problèmes que ceux afférents au modèle de référence et d'après nos connaissances, il n'a pas été validé dans une expérience qui manipule explicitement les 3 dimensions de la cible. Par ailleurs, dans le cas d'utilisation d'une sphère en tant que cible, ce modèle reste similaire à la loi de Fitts classique. Un autre modèle proposé est celui de Grossman et Balakrishnan (2004) dont l'étude porte sur les mouvements physiques de l'utilisateur dans le plan XZ pour un affichage volumétrique 3D. Ce modèle a ainsi été validé pour un affichage volumétrique et non un environnement virtuel 3D. De plus, les mouvements étaient contraints sur l'axe (X, Z) avec (Y = 0), ce qui revient à faire de la 2D car tous les axes 3D n'ont pas été évalués.

Jusqu'à présent, les modèles alternatifs à celui de Fitts n'ont pas permis de proposer un modèle validé ou ne sont pas assez robustes pour prédire le pointage en EV 3D. Le modèle classique est ainsi utilisé quelles que soient les dimensions.

2.1.2. Évaluation du débit d'une technique

Le débit (nommé *throughput* ou TP dans la littérature anglo-saxonne) a été introduit par la norme ISO 9241-9 (ISO 9241-9, 2000) pour comparer la performance de différents dispositifs de pointage dans un contexte donné. Il s'agit d'une mesure globale basée sur le modèle de Fitts, et plus précisément sur les indices de difficultés

effectifs mesurés lors de la conduite d'une expérimentation conforme à la loi de Fitts. Le débit est mesuré en bits par seconde et la formule pour le calculer est :

$$TP = IDe/MT \text{ avec } IDe = \log_2(De/We + 1) \quad (4)$$

où MT est le temps de mouvement effectué par les utilisateurs pour réaliser une tâche d'indice de difficulté effectif (IDe). Ces termes représentent la performance effectivement réalisée par l'utilisateur plutôt que ce qu'il est censé faire en théorie. Ainsi De est la longueur effective entre deux cibles parcourues par les utilisateurs durant l'évaluation et We est la taille effective calculée en fonction de SD qui est la distribution des points de validation de l'utilisateur (cf. figure 1).

$$We = 4.1333.SD \quad (5)$$

$$SD = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 + (y_i - \bar{y})^2}{n-1}} \quad (6)$$

Le TP est proche de l'index de performance (IP) (MacKenzie, 1992), qui correspond à l'inverse de la pente de la droite de régression calculée dans une étude conforme à la loi de Fitts, soit $IP = ID/MT$. La différence majeure est toutefois que le TP n'est pas influencé par le traditionnel rapport entre vitesse et précision, contrairement à l'IP. En effet, le TP prend en compte un ajustement de la taille des cibles en fonction de la distribution des points de validation de l'utilisateur, lorsque le taux d'erreur est supérieur à 4 % : ce sont alors les distances effectives (De dans la relation (4)) et les largeurs effectives (We) qui sont prises en compte dans le calcul du TP. Le TP intègre donc automatiquement des corrections qui par effet de bord lissent les problèmes liés à la vitesse et à la précision. Toutefois, aucune mesure de débit n'est actuellement disponible pour des interactions 3D en EV 3D.

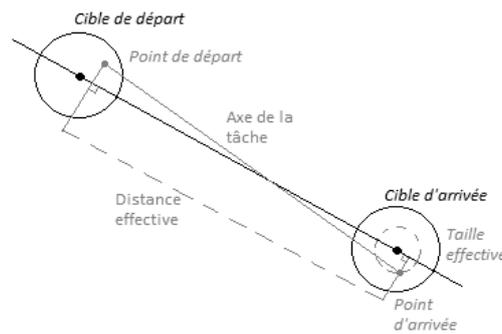


Figure 1. Illustration de la distance effective mesurée d'après le point de départ et le point d'arrivée de la tâche de pointage et de la taille effective (Teather et al., 2009)

2.1.3. Norme ISO 9241-9

La norme ISO 9241-9 (ISO 9241-9, 2000) intitulée « Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV), Partie 9 : Exigences relatives aux dispositifs d'entrée autres que les claviers », fournit des exigences et des recommandations pour la conception et l'évaluation de dispositifs tels que les souris, manettes, et joysticks.

Les 8 articles et les 4 annexes de la norme présentent les différents concepts liés à l'évaluation des dispositifs. Les 2 premiers articles présentent le contexte de mise en œuvre de la norme avec le domaine d'application et les références normatives. Un des articles fournit les définitions (Article 3) de notions telles que le pointage, le feedback, le matériel, la mesure du débit, la posture ou encore les indicateurs d'utilisabilité. Parmi ces définitions, nous soulignons les deux suivantes, particulièrement importantes pour la compréhension du cadre que nous proposons dans la suite :

- le pointage (Article 3. Partie 1.4) est « une opération effectuée dans une interface graphique utilisateur au cours de laquelle le dispositif d'entrée est utilisé pour déplacer une petite représentation graphique (un pointeur, par exemple) vers un point spécifique de l'écran » ;

- la mesure du débit – aussi appelé *throughput* ou TP – (Article 3. Partie 4.11) correspond à la « mesure de la vitesse de transfert des informations lorsqu'un utilisateur commande un pointeur à l'écran à l'aide d'un dispositif d'entrée »).

La norme introduit ensuite :

- des principes directeurs ergonomiques (Article 4), tels que l'opérabilité (Article 4. Partie 2), qui décrit les conditions d'exécution rendant la technique d'interaction utilisable (comme la cohérence ou encore l'efficacité), ou la maniabilité (Article 4. Partie 3), qui définit les aptitudes de la technique à répondre aux attentes des utilisateurs (comme la réactivité ou l'accès aux commandes) ;

- des critères de performance (Article 5), qui permettent de caractériser l'utilisabilité de la technique selon l'objectif et la tâche auxquels la technique doit répondre ;

- des recommandations de conception (Article 6) générales (Article 6. Partie 1) – comme la conception des boutons, la facilité de préhension ou le feedback – puis celles liées aux caractéristiques des dispositifs d'entrée (Article 6. Partie 2) – souris, manettes etc. ;

- des conditions et des conventions de mesurage (Article 7) indiquant la méthode à utiliser pour obtenir les mesures nécessaires à l'évaluation des exigences présentées dans l'article précédent, ainsi que la conformité et le respect de ces exigences (Article 8).

Les éléments déterminants pour la mise en place d'un protocole expérimental se situent dans les annexes. Dans son annexe A, la norme présente des critères de sélection d'un dispositif d'entrée ainsi que les méthodes de mesure et d'analyse de l'utilisabilité (Annexe A). Des éléments tels que l'environnement (Annexe A.

Partie 3.2), l'équipement requis (Annexe A. Partie 3.7) ou encore les procédures d'essai (Annexe A. Partie 3.8) sont ainsi présentés dans cette annexe afin d'aider à monter le protocole expérimental.

L'annexe B introduit les éléments de calculs nécessaires pour mesurer l'efficacité et l'efficacité d'un dispositif d'entrée et notamment : largeur de cible effective (W_e), indice de difficulté (ID), indice effectif de difficulté (IDe) et précision des tâches (Annexe B. Partie 5). Ces différents calculs sont nécessaires afin de déterminer le débit (TP) d'une technique (cf. 2.1.2). Ils sont déterminés en fonction de la direction du mouvement qui détermine l'axe de la tâche : par exemple, selon l'axe x en 1D. L'axe de la tâche correspond au mouvement optimal allant d'une cible de départ jusqu'à une cible d'arrivée. Cette annexe définit également un ensemble d'essais (Annexe B.6), pour lesquels des opérations et un mode opératoire à suivre sont présentés afin d'évaluer chaque tâche : par exemple, l'essai de frappe multidirectionnel (Annexe B. Partie 6.2.2) correspond à une tâche de pointage, en s'approchant des consignes et procédures d'une évaluation Fitts, et concerne notre adaptation.

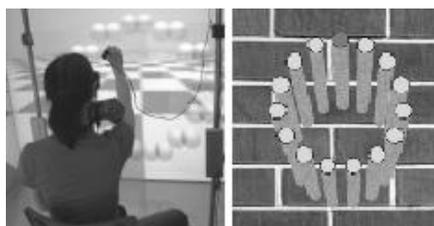
L'annexe C présente les principes de l'étude qualitative. Elle est basée sur un questionnaire pour évaluer le confort de la technique et sur une échelle pour graduer l'effort requis lors de la tâche. Le questionnaire sur le confort peut être utilisé de façon individuelle ou comparative entre des techniques. Ce questionnaire comprend deux parties évaluées sur une échelle de Likert en 5 ou 7 points, selon le questionnaire sélectionné : des indices généraux pour mesurer la perception qu'ont les utilisateurs de certaines caractéristiques des techniques (vitesse, précision par exemple), puis des indices de fatigue pour mesurer la fatigue ressentie par les utilisateurs en fonction des parties de leur corps. Toutefois, le questionnaire de confort ne contient pas d'items concernant la satisfaction de l'utilisateur, ni l'apprentissage. L'échelle permettant de graduer l'effort s'appuie sur l'échelle de Borg (Annexe C.4), mais celle-ci est rarement proposée aux utilisateurs lors des évaluations présentées dans la littérature.

Cette norme a déjà été appliquée à différents cas d'études, dont quelques-uns un peu « exotiques » comme l'évaluation des manettes de jeux (Natapov *et al.*, 2009 ; MacArthur *et al.*, 2009) ou encore d'un eye-tracker (Zhang et MacKenzie, 2007). Cependant, ces exigences sont davantage destinées à des dispositifs à 2 degrés de liberté et les tâches élémentaires ne concernent que des espaces 1D et 2D. La norme ne nous apporte donc pas de recommandations pour évaluer et comparer des techniques d'interaction à au moins 3 degrés de liberté ou situées dans un environnement 3D. Toutefois, certaines tentatives d'adaptation de la norme ISO à un environnement 3D ont été proposées.

2.1.4. Proposition d'adaptation de la norme ISO 9241-9 à un environnement 3D

Des études récentes (Teather *et al.*, 2009 ; Teather et Stuerzlinger, 2011) traitent de l'évaluation de la tâche de pointage basée sur la norme ISO 9241-9. Dans une première étude (Teather *et al.*, 2009), la représentation utilisée est une scène 3D où les cibles sont des cercles 2D placés sur des cylindres 3D. Les utilisateurs ont comme

instruction de pointer le sommet du cylindre pour ainsi sélectionner la cible circulaire (cf. figure 2b). Dans une seconde étude (Teather et Stuerzlinger, 2011), les cibles sont des sphères 3D disposées sur un cercle support 2D, positionné dans un plan vertical (cf. figure 2a). Dans les deux cas, la représentation des cibles ne place pas vraiment l'utilisateur dans un contexte assimilable à de la 3D : sélectionner une cible résulte donc en une combinaison de pointage en 2D sur un plan en utilisant une technique de lancer de rayon pour atteindre la profondeur appropriée. Dans la majorité des cas, les interactions avec le dispositif sont ainsi limitées au plan (X, Y). La profondeur n'est pas réalisée avec un degré de liberté spécifique mais est le résultat d'un calcul logiciel (assez proche de la 2.5D, puisque l'interaction avec la scène 3D se réalise principalement avec 2 degrés de liberté). Enfin, pour ces deux études, les utilisateurs étaient placés dans des environnements immersifs (CAVE – environnement immersif de réalité virtuelle où une scène 3D est projetée sur les murs d'une pièce cubique - ou *fish tank* – image stéréoscopique d'une scène 3D visualisée sur un écran utilisant une projection perspective couplée à la position de la tête de l'utilisateur) et n'interagissaient donc pas avec des environnements virtuels de « bureau ».



(a)

(b)

Figure 2. Environnements de test des études de Teather :
(a) interaction sur un plan vertical (Teather et Stuerzlinger, 2011) ;
et (b) sur un plan horizontal (Teather et al., 2009)

Une étude récente (Ha et Woo, 2010) évalue la manipulation d'objet 3D en utilisant les techniques de main virtuelle. La tâche de manipulation se divise en une sélection suivie d'une translation ou d'une rotation de l'objet 3D. Pour sélectionner l'objet 3D, l'utilisateur doit pointer l'objet puis le positionner sur une position cible idéale déterminée par calcul. Pour calculer cette position, les auteurs se sont basés sur la norme ISO 9241-9 en adaptant la mesure de la largeur effective (We). Cependant, il s'agit d'une évaluation située dans un environnement de réalité augmentée tangible, où l'interaction a lieu dans le monde physique, au sein duquel sont ajoutées des informations numériques. Il ne s'agit donc pas d'une interaction avec un EV 3D à proprement parlé. Ce travail illustre donc l'intérêt de baser une évaluation sur la norme ISO mais les résultats ne peuvent pas être comparés avec des évaluations conduites dans un EV immersif comme ceux de Teather.

Au-delà de ces propositions, nous n'avons pas connaissance d'autres études basées sur le standard et réalisées dans un EV 3D non immersif qui impliquent des

techniques d'interaction avec 3 degrés de liberté (DL) ou plus. Ainsi les résultats fournis par les approches structurées ne présentent pas encore de références réutilisables. Des résultats plus concrets ont par contre été mesurés à travers des études empiriques. La section suivante résume des évaluations menées au sein d'environnements virtuels 3D de manière *ad hoc*, *i.e.* qui n'ont pas été définies d'après la norme.

2.2. Évaluation de techniques d'interaction en environnements virtuels 3D

Parmi les études existantes relatives à l'évaluation des techniques d'interaction à au moins 3 degrés de liberté, nous avons porté notre attention sur trois exemples représentatifs : (1) la technique Airmouse (Ortega et Nigay, 2009), (2) une étude menée par (Bérard *et al.*, 2009), et (3) une étude réalisée par (Zhai, 1993).

Ces études diffèrent en termes de tâches et de paramètres (cf. figure 3). En effet, les études (1) et (2) proposent la manipulation de seulement 3 axes pour les translations alors que l'étude (3) utilise les 6 axes afin de réaliser les translations et les rotations. L'étude (1) (figure 3a) évalue une tâche de pointage : le déplacement d'une petite sphère dans un espace sphérique plus grand. L'étude (2) (figure 3b) évalue une tâche de placement : le déplacement d'un cube aussi proche que possible d'une cible représentée par un cube avec une taille plus petite. L'étude (3) (figure 3c) évalue une tâche de positionnement : manipuler un tétraèdre de façon à le positionner et l'orienter dans un autre tétraèdre cible de même taille. Pour chaque étude, les utilisateurs ont pour instruction de déplacer un curseur ou un objet aussi rapidement que possible à un endroit précis.

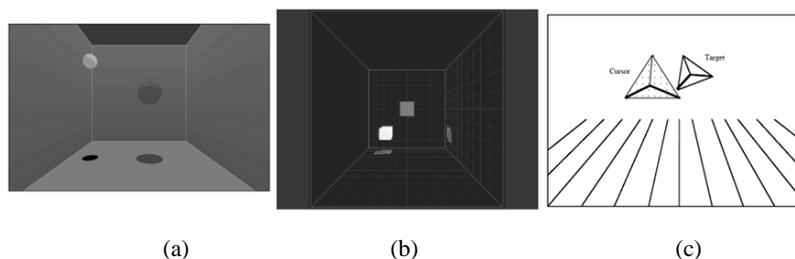


Figure 3. Environnements d'évaluation des études

Même si toutes ces expériences se concentrent sur des résultats quantitatifs, en particulier les temps de performance, chacune s'appuie sur différentes variables. Les études (2) et (3) définissent un taux d'erreur. Les études (1) et (2) utilisent la notion d'indices de difficultés, telle que définie dans la loi de Fitts pour modifier les conditions de la tâche. Seule l'étude (1) fait correspondre indices de difficultés et temps de réalisation de la tâche ; cependant, les paramètres plus classiques *a* et *b* de l'analyse de la loi de Fitts ne sont pas étudiés. En outre, les études (1) et (3) envisagent des mesures qualitatives telles que la fatigue, mais l'étude (2) ne définit

pas de mesures qualitatives. Toutefois, l'évaluation de la fatigue n'est pas similaire dans les deux études. Dans aucune de ces études, il n'est fait cas du débit ; seule l'étude (1) se base sur une des tâches présentées dans la norme 9241-9 pour réaliser les évaluations, pourtant le débit n'est pas calculé pour les techniques testées. Ces études montrent un manque de formalisation dans l'évaluation des techniques d'interaction 3D et soulignent les limites qui ont été mises en avant (le nombre de degrés de liberté considérés, la tâche évaluée, et la méthode utilisée). Par ailleurs, les mesures analysées ne sont pas identiques et il n'existe aucune mesure globale pour comparer facilement ces techniques d'interaction.

Pour résumer, le premier ensemble d'approches, basé sur la norme, donne des recommandations pour évaluer la tâche de pointage de façon structurée, tandis que le deuxième ensemble d'approches, basé sur des mesures ponctuelles et selon un contexte considéré, donne des résultats et des mesures concrets, qui sont complexes à comparer. Notre objectif est donc de prendre en compte les bénéfices de l'approche structurée afin de faciliter et de standardiser la conception de telles expériences. Nous profitons également des variables et mesures proposées dans les approches *ad hoc* ayant des aspects pertinents pour l'expérience utilisateur dans un EV 3D. Notre travail se focalise ainsi sur une extension de la norme ISO.

3. Adaptation de la norme ISO 9241-9 à la tâche de pointage en 3D

Parmi les types de tests couverts par la norme (cf. 2.1.3), nous avons porté notre intérêt sur la tâche de frappe multidirectionnelle étant donné que nous avons l'intention d'évaluer la tâche de pointage. Le mouvement est multidirectionnel vu que nous considérons les mouvements de pointage dans toutes les directions (x, y, z). Nous détaillons dans cette section, l'adaptation de 5 dimensions importantes extraites du standard initial : 3 dimensions quantitatives, qui sont 1) tâche, 2) cibles, 3) variables, puis 2 dimensions qualitatives, qui sont 4) représentation et 5) confort et satisfaction. Pour chaque dimension, nous présentons et justifions l'extension que nous avons apportée afin d'adapter le standard à un contexte de pointage en 3D. Ensuite, nous présentons les éléments de l'instrumentation que nous proposons comme support à la mise en œuvre de notre adaptation de la norme.

3.1. Adaptation du standard d'un point de vue performance

Trois aspects majeurs de la norme structurent l'évaluation quantitative : la tâche, les cibles et les variables.

3.1.1. Tâche

Dans la norme ISO, la tâche consiste à pointer une cible fixe le plus rapidement et le plus précisément possible (ISO 9241-9 - Annexe B.3). La tâche de pointage prend fin lorsque l'utilisateur confirme l'acquisition de la cible en pressant sur un bouton ou une touche. Si le curseur est à l'intérieur de la cible de destination quand l'utilisateur confirme l'acquisition, la tâche de pointage est réussie. Si le curseur est

à l'extérieur de la cible, la validation lève une erreur. Dans tous les cas, appuyer sur une touche ou un bouton marque la fin de la tâche courante et le début de la tâche de pointage suivante. La difficulté de la tâche varie selon les différents indices de difficultés (cf. équation (1)). La taille des cibles et la distance entre les cibles sont modifiées lors de l'expérimentation.

Dans notre adaptation, nous avons conservé l'instruction donnée aux utilisateurs sur la tâche, l'existence d'une cible source et de destination, ainsi que le besoin de confirmation de la sélection. L'adaptation concerne tout d'abord la validation. Par analogie à la tâche 2D, nous avons conservé l'idée que la cible peut être validée dès qu'un pixel du pointeur est en contact avec la cible. Par conséquent, le curseur étant représenté par une sphère (cf. 3.2.1), son volume n'a pas besoin d'être complètement à l'intérieur de la cible pour considérer que la validation est réussie : un contact entre la cible et le curseur est suffisant (cf. figure 4). En absence de contact, la validation lève une erreur. Nous ne forçons donc pas l'inclusion totale du curseur dans la cible pour valider la sélection de la cible. Cette solution n'apporte pas de biais particulier à la tâche proposée car la taille du curseur est prise en compte dans la difficulté de la tâche en plus de la taille de la cible. De plus, elle contourne la contrainte que constituerait le besoin d'inclure le pointeur dans la cible pour valider la sélection.

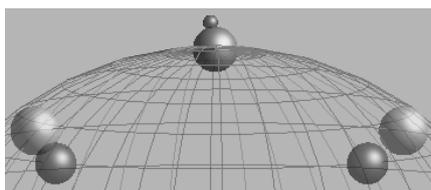


Figure 4. Exemple d'un contact entre le curseur et la cible

3.1.2. Cibles

Même si le standard ne recommande pas explicitement une forme de cibles, les cibles sous forme de cercle sont les plus utilisées dans les évaluations de la littérature, étant donné qu'un cercle n'a qu'une seule et même taille quel que soit le point d'entrée du curseur dans la cible. Dans notre adaptation en 3D, les cibles sont devenues des sphères pour conserver une seule mesure pour la taille quel que soit le mouvement effectué en 3D : la taille des cibles correspondra donc au diamètre des sphères.

La norme ISO donne en revanche des recommandations sur la disposition des cibles et leur position : en 2D, les cibles sont équitablement placées sur la circonférence d'un cercle support (ISO 9241-9 - B.6.2.2). De plus, la place des cibles sur le cercle doit être définie de façon à ce que chaque cible ait une cible symétrique par rapport au centre du cercle support (cf. figure 5). Cela permet d'assurer une distance toujours identique entre deux cibles et que cette distance soit égale au diamètre du cercle support (ISO 9241-9 - B.6.2.2).

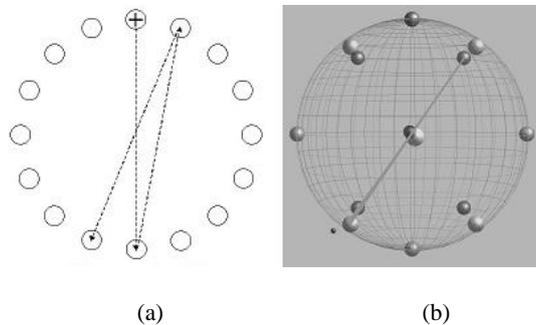


Figure 5. Disposition des cibles pour une tâche de pointage réalisée en 2D (a) et en 3D (b)

Dans notre adaptation, la recommandation sur la disposition des cibles a été revue de manière à prendre en compte la 3D : le cercle support est remplacé par une sphère support. Les cibles sont toujours équitablement disposées sur la circonférence de la sphère support. Mais leur position est définie de façon à ce que chaque cible ait un symétrique par rapport au centre de la sphère support. La distance entre deux cibles symétriques est toujours égale au diamètre de la sphère support (cf. figure 5).

3.1.3. Variables

La norme ISO recommande différentes variables indépendantes et dépendantes (ISO 9241-9 - B.5). Dans le standard, les variables indépendantes inhérentes à ce type d'expérimentation sont les techniques en entrée, la taille des cibles, la distance entre les cibles et par conséquent les indices de difficultés (ID) qui en résultent. Les variables dépendantes citées dans la norme ISO incluent le débit de la technique et le temps moyen de pointage.

Ces différentes variables sont conservées dans notre adaptation, mais celle-ci a tout d'abord un effet sur le calcul de l'axe pris en compte lors de la tâche de pointage et qui considère le mouvement de l'utilisateur durant cette tâche. Cet axe de la tâche est notamment utilisé pour mesurer le débit. Il est représenté par la droite passant par le point de départ et le point d'arrivée de la tâche de pointage. Cet axe a ainsi été adapté de façon à prendre en compte les 3 dimensions (x, y, z) du mouvement à réaliser. Il sert en particulier à calculer la largeur de la cible effective et à déterminer la distance effective entre la cible de départ et la cible d'arrivée.

Notre adaptation nécessite également une révision du mode de calcul des ID car le curseur a désormais un volume. Ce calcul, qui inclut la taille du curseur, est illustré par la formule suivante :

$$ID = \log_2(D/W_{cible} + W_{curseur} + 1)$$

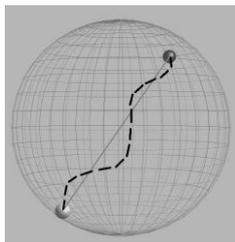
Le calcul des ID effectifs (IDe) est également modifié et dépend des coordonnées du point effectivement atteint par le centre du curseur lors de la validation de la cible par l'utilisateur. La taille effective est calculée d'après la distribution des différents points 3D de validation de l'utilisateur.

$$SD = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 + (y_i - \bar{y})^2 + (z_i - \bar{z})^2}{n - 1}}$$

La distance effective correspond à la distance euclidienne 3D mesurée le long de l'axe de la tâche entre le point de validation de l'utilisateur et le point de destination de validation de l'utilisateur.

$$De = \frac{\sum_{i=1}^n \sqrt{(x_{Endi} - x_{Starti})^2 + (y_{Endi} - y_{Starti})^2 + (z_{Endi} - z_{Starti})^2}}{n}$$

Les tailles et distances effectives, utilisées pour calculer le TP, sont mesurées par une moyenne des essais pour une condition donnée. A ces variables proposées par le standard, notre adaptation ajoute deux variables dépendantes supplémentaires : le taux d'erreur et le ratio d'inefficience. La première variable ajoutée, le taux d'erreur, est nécessaire car il s'agit d'une variable dépendante très souvent utilisée dans les approches *ad hoc* et elle peut permettre d'expliquer et de compléter la mesure du débit. La seconde variable, le ratio d'inefficience, a été défini par (Zhai et Milgram, 1998) et indique la coordination des utilisateurs lors de l'exécution de la tâche en 3D en comparant la longueur parcourue par l'utilisateur et la distance optimale : plus le résultat est proche de 0, meilleure est la coordination (cf. figure 6). Ce type de mesure est assez habituel pour évaluer les techniques à 6 degrés de liberté et le ratio peut montrer des différences allant jusqu'à 2 ou 3 fois la distance parcourue par les utilisateurs (Zhai et Milgram, 1998). Selon (Masliah, 1999), ce ratio complète le taux d'erreur, un des aspects de la performance utilisateur et fournit une analyse de la qualité de la performance en 3D dans le but de mieux la comprendre.



$$\text{Ratio d'inefficience} = \frac{(\text{longueur de l'utilisateur} - \text{longueur optimale})}{\text{longueur optimale}}$$

Figure 6. Exemple de coordination au cours d'une tâche de pointage

3.2. Adaptation de la norme pour les aspects qualitatifs

Deux éléments de la norme structurent l'évaluation qualitative : la représentation et les indices 3D d'une part, le confort et la satisfaction d'autre part.

3.2.1. Représentation et indices 3D

La norme ISO ne recommande aucun élément spécifique en ce qui concerne la représentation. Malgré cela, le pointeur est communément représenté par une croix. Dans notre adaptation, la forme du pointeur a été modifiée car une croix ou un point ne sont pas assez visibles dans un EV 3D, une représentation ayant du volume est donc recommandée. Le pointeur est ainsi représenté par une sphère de couleur rouge, ce qui reste cohérent avec la représentation des cibles. La forme sphérique du curseur répond à une autre contrainte importante : en effet, cette forme permet de faire en sorte que le point de collision entre le curseur et la cible capturée soit toujours le plus proche de la cible à pointer quel que soit le mouvement réalisé en 3D. Le centre du curseur et le centre de la cible sont à même distance l'un de l'autre en cas de collision, ce qui ne serait possible avec un cube étant donné que tous les points du cube ne se situent pas à la même distance du centre. Le diamètre choisi pour le curseur est d'environ 10px dans notre cas et représente ainsi environ 70 % de la plus petite cible à atteindre, ceci afin que le curseur reste visible quelle que soit sa position dans l'environnement 3D.

Notre adaptation à un EV 3D fait face au problème traditionnel de la perception de la profondeur en 3D. Du fait de l'environnement non immersif et au manque de solutions logicielles et matérielles efficaces, les utilisateurs ne sont pas capables de ressentir correctement la profondeur des objets numériques, rendant alors la modélisation de la tâche de pointage en 3D relativement difficile. Pour ces raisons, certains indices visuels 3D ont été définis et sont recommandés pour pallier la difficulté de perception de la profondeur. Tout d'abord, l'effet de perspective fait apparaître une variation des tailles des sphères : les sphères positionnées à l'arrière semblent être plus petites mais cet effet n'est que relatif au rendu et s'applique également au pointeur. En plus de cet effet de perspective, nous nous appuyons sur la transparence des sphères cibles, comme (Zhai *et al.*, 1994) avaient pu le tester dans des environnements 3D. La transparence appliquée aux sphères cibles renforce l'effet de profondeur en réduisant l'occlusion avec les autres sphères ; elle permet en outre aux utilisateurs de distinguer le curseur derrière les cibles. Combinée avec cette transparence, notre adaptation de la norme recommande l'utilisation d'un dégradé de couleurs pour souligner l'effet de profondeur : les couleurs tirant vers le gris sont situées dans le fond et les couleurs plus lumineuses sont à l'avant car elles tendent à sembler plus proches (Kaiser, 2011). A ces indices, notre adaptation préconise pour le rendu de la sphère support, l'utilisation d'un fil de fer transparent. Ces recommandations sont issues de phases de conception itératives réalisées avec la participation d'experts 3D. Par ailleurs, les ombres sont communément utilisées dans les scènes 3D, pourtant notre adaptation n'encourage pas l'utilisation d'ombres pour éviter de surcharger l'environnement. En effet, Hubona *et al.* (1999) ont montré que

les performances des utilisateurs pour une tâche de positionnement diminuent lorsque le nombre des ombres ainsi que la complexité de la scène sont importants.

Ces 4 indices (perspective, dégradé de couleurs, transparence et structure en fil de fer) prônés par notre adaptation permettent de constituer une scène de référence pour l'évaluation de techniques d'interaction pour les tâches de pointage en 3D. Utiliser cette scène de référence évitera l'insertion de biais d'un point de vue de la perception de la profondeur, biais qui pourraient compromettre les résultats des évaluations en EV 3D.

3.2.2. Confort et satisfaction

La norme ISO recommande deux échelles d'évaluation pour évaluer le confort des techniques d'interaction (ISO 9241-9 – Annexe C) : un questionnaire sur le confort pour mesurer le confort et l'utilisabilité des techniques ainsi qu'une échelle d'effort pour estimer l'effort perçu par les utilisateurs lors de la tâche.

Dans notre adaptation, nous avons renforcé l'analyse du confort et de l'utilisabilité du questionnaire de confort en ajoutant aux deux échelles déjà présentes le questionnaire SUS (*System Usability Scale*) (Brooke, 1996) qui couvre en plus la satisfaction utilisateur et l'apprentissage. En effet, le questionnaire initial sur le confort ne traite pas des aspects de l'utilisabilité, tels que la satisfaction utilisateur, l'apprentissage et la préférence utilisateur. Le questionnaire SUS permet d'évaluer l'utilisabilité d'un système, c'est-à-dire qu'il permet de mesurer l'efficacité (savoir comment le but a été atteint), l'efficacités (savoir si le résultat prévu a été atteint) du système et la satisfaction de l'utilisateur. Epruvé depuis plusieurs années, ce questionnaire permet de recueillir rapidement l'avis des participants puisqu'il ne comprend que 10 items. De plus, il mène au calcul d'un score global d'utilisabilité pour chaque technique compris entre 0 et 100, utile pour comparer facilement différentes techniques évaluées dans un même contexte. Nous avons également enrichi l'analyse de l'aspect satisfaction en insérant l'utilisation d'items du SUMI. Le questionnaire SUMI (*Software Usability Measurement Inventory*) permet d'évaluer la qualité d'un système du point de vue de l'utilisateur. Il comprend un ensemble de 50 items permettant de mesurer l'efficacité, l'influence, la capacité d'aide, la contrôlabilité ainsi que l'apprentissage d'un système. Nous avons sélectionné 6 items parmi les 50 de façon à compléter les questionnaires précédents et à qualifier de façon plus précise les dimensions évaluées. En effet, des items sur la frustration ou la stimulation ont été ajoutés afin de recueillir le ressenti des utilisateurs et de caractériser plus finement la technique. Enfin, nous avons complété cet aspect qualitatif en demandant aux utilisateurs les trois points les plus positifs et les trois points les plus négatifs pour chaque technique d'interaction. Recueillir les avis des utilisateurs sur la technique a pour objectif de déterminer les caractéristiques de la technique qui ont le plus marqué les utilisateurs et qui pourront expliquer et compléter leurs réponses aux questionnaires.

Cette adaptation constitue donc un support pour guider la définition d'un protocole d'évaluation. Pour aider les expérimentateurs à mettre en œuvre cette

adaptation, nous avons développé un outil dont nous présentons les principales caractéristiques dans la section suivante.

3.3. Instrumentation

Pour accompagner notre adaptation de cette norme, nous proposons un outil logiciel support au déploiement du protocole, à la réalisation des tâches de pointage décrites dans l'adaptation et à l'enregistrement des traces d'interaction au cours de cette tâche, nécessaires pour les calculs des différentes mesures recommandées. Il constitue un environnement logiciel de référence pour mener les expérimentations basées sur notre adaptation de la norme en outillant de façon concrète le protocole, *i.e.* en appliquant et respectant les recommandations de la norme et les adaptations apportées pour la 3D. Cet outil logiciel rend possible le paramétrage de la taille des cibles, la distance entre ces cibles, le nombre de cibles et le control-gain. L'utilisation de cet outil logiciel garantit donc une mise en œuvre identique de la norme et de ses adaptations, d'un utilisateur à un autre et surtout d'une étude portant sur un ensemble de techniques d'interaction à une autre : les mesures, la scène 3D, les questionnaires et les résultats calculés seront les mêmes, et toujours obtenus de la même façon. L'outil logiciel est composé d'un environnement 3D (cf. figure 7) et d'un panneau de configuration. Le logiciel graphique 3D est écrit en C++ en utilisant la toolkit OSG. Le panneau de configuration est écrit en C#, et afin de proposer un outil logiciel évolutif, les composants de l'outil logiciel communiquent *via* un bus logiciel appelé Ivy.

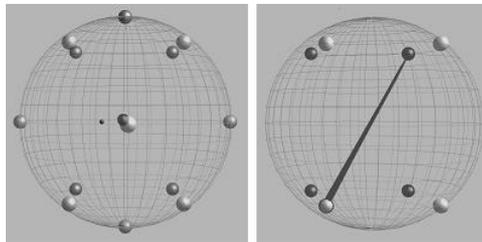


Figure 7. Illustration de l'instrumentation : disposition des cibles sur le support et illustration de la profondeur entre les cibles dans la scène 3D

Ce panneau de configuration est nécessaire à la configuration et au pilotage de l'exécution d'une expérimentation. En effet, la tâche de pointage est composée par un certain nombre de cibles que l'expérimentateur peut gérer : les cibles peuvent être placées à une longitude et une latitude spécifiques. De plus, un bloc est composé par une série de tâches de pointage, qui ont le même ID : l'expérimentateur doit toutefois configurer la taille de la cible (largeur) et la taille du support (distance) pour définir l'ID du bloc. Au lancement de l'expérimentation, ce panneau permet de définir l'ordre de passage des techniques ainsi que l'emplacement des fichiers de traces.

Au cours de l'expérimentation avec ce logiciel, exécuter la tâche sur un bloc revient à un processus en trois étapes : 1) la tâche de pointage commence au moment où le participant choisit la sphère de départ, qui est colorée en bleu dans notre instrumentation, sans commettre d'erreur. 2) Cette action met ensuite en évidence la sphère de destination qui est symétrique et colorée en jaune. 3) La tâche se termine lorsque le participant choisit la sphère de destination, même si une erreur est commise. La sphère cible devient orange lorsque le curseur et la cible sont en contact. Durant la tâche, l'outil logiciel enregistre des données temporelles, telles que le début et la fin de chaque tâche de pointage, la validation de la sélection et le temps d'exécution. Il enregistre également la position 3D du curseur et les erreurs lorsque la validation a été réalisée sans contact entre le curseur et la cible. Tous les événements de l'expérimentation sont datés et enregistrés dans un format XML.

3.4. Synthèse

Notre adaptation est générique et propose un protocole structuré pour conduire des évaluations de la tâche de pointage 3D. Nous proposons ainsi une évaluation composite mêlant des aspects de performance et de satisfaction utilisateur. Dans le tableau 1 sont synthétisés les ajouts et modifications que nous avons apportés à la norme de façon à proposer cette évaluation composite. Par ailleurs, nos différentes mesures couvrent l'ensemble des 6 critères définis par Zhai (1998) pour caractériser l'utilisabilité d'une technique d'entrée 3D : vitesse, précision, fatigue, persistance du dispositif et son acquisition, coordination et enfin facilité d'apprentissage. Ainsi, dans notre adaptation, les critères de vitesse et précision sont mesurés par le débit ; la fatigue mais aussi la persistance et l'acquisition sont évaluées par le questionnaire sur le confort de la norme ISO ainsi que par les entretiens semi-directifs. La coordination est mesurée avec le ratio d'inefficience que nous avons ajouté dans notre adaptation et qui utilise les distances effectives. Pour mesurer la facilité d'apprentissage, nous avons ajouté le score global d'utilisabilité et la satisfaction des utilisateurs mesurés grâce au questionnaire SUS.

Ces différents points d'adaptation et d'ajout constituent une réponse aux limites que nous avons relevées dans l'état de l'art (cf. 2.2) concernant les études *ad hoc* : le nombre de degrés de liberté considérés est clairement établi et contrôlé, nous proposons de n'utiliser que 3 degrés de liberté puisque ce nombre est suffisant pour réaliser la tâche de pointage ; la tâche est définie et stable, les mesures identifiées se complètent et répondent aussi bien à des exigences de performance que de satisfaction utilisateur. La méthode d'évaluation ainsi que son protocole de mise en place sont cernés pour partie par les amendements faits dans la norme et pour partie grâce au logiciel de conduite de l'expérimentation que nous proposons et qui rend possible l'uniformisation du protocole (préoccupations liées au rendu notamment incluses dans une scène 3D pouvant servir de scène 3D de référence). Notre adaptation a également pour but de proposer un lien entre l'interaction physique et les données numériques en présentant des interactions avec 3 degrés de liberté et une scène 3D à réaliser avec des techniques d'interaction avancée. Ainsi, l'adaptation de

la norme 9241-9 a pour objectif d'apporter une unification de l'évaluation de techniques d'interaction en EV 3D.

Pour montrer concrètement l'utilisation de ces recommandations, nous présentons le cas d'étude sur lequel nous nous sommes basés puis le protocole d'évaluation de deux techniques d'interaction, étudiées pour ce cas d'étude. Enfin, nous analysons et discutons les résultats de cette évaluation.

Tableau 1. Synthèse de l'adaptation à la norme 9241-9 pour le pointage en EV 3D

Norme			Etat
Sections 1 à 8			Original
Annexe A			Original
Annexe B			Modifié
Annexe B.5.2			
Critères	Mesures	Variables dépendantes	
Effizienz	Rapidité	Moyenne des temps	Original
Efficacité	Précision	Taux d'erreur	Ajouté
	Coordination de l'utilisateur	Ratio d'inefficacité	Ajouté
Performance	Mesure globale	Throughput/débit	Modifié
Annexe B.6.2.2			
Rendu			
Forme des cibles			Modifié
Forme du curseur			Modifié
Arrangement des cibles			Original
Indices 3D	Transparence		Ajouté
	Dégradé de couleur		Ajouté
	Support sous forme de fil de fer		Ajouté
Annexe C			Modifié
Critères	Mesures	Variables dépendantes	
Confort	Echelle de confort	Questionnaire de confort de la norme ISO 9241-9	Original
Satisfaction utilisateur	Mesure globale	SUS	Ajouté
	Echelle de satisfaction utilisateur	Items du SUMI	Ajouté

4. Utilisation concrète de l'adaptation à la norme et résultats de l'évaluation

Le cas d'étude considéré est issu d'un contexte industriel, porté par la société GVS – Toulouse, centrée sur le développement d'applications expertes pour le monitoring d'activités et de systèmes de données complexes. Après l'avoir introduit,

nous présentons le protocole d'expérimentation détaillée et nous reportons les résultats quantitatifs et qualitatifs de cette évaluation, menée selon l'adaptation de la norme et avec le support de l'instrumentation présentée précédemment.

4.1. Cas d'étude : conception de tests statiques pour avion

Notre étude de cas est basée sur une application concrète qui permet de concevoir et gérer des campagnes d'essais statiques pour l'aéronautique (cf. figure 8). Les concepteurs ont à manipuler et positionner des capteurs sur une maquette numérique 3D de l'avion. Cette maquette numérique 3D représente un avion réel, y compris en ce qui concerne les dimensions et le volume. Les interactions avec ces modèles permettent aux utilisateurs d'effectuer des tâches en 3D comme s'ils manipulaient l'appareil réel. Les utilisateurs peuvent appliquer de la transparence sur des parties de la maquette numérique pour interagir plus facilement avec des éléments de la structure 3D normalement cachés : des techniques comme le ray-casting ne sont donc pas envisageables étant donné la complexité de la maquette numérique, et sa structure multicouche. La figure 8 illustre par exemple la mesure de la distance entre un hublot (A) de la carlingue et un élément placé au milieu sous le cockpit (B). La transparence ne suffit pas pour accéder à ces deux points car au niveau du point A une grande quantité de points candidats existent : la carlingue extérieure, etc.

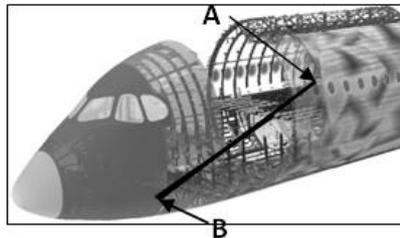


Figure 8. Illustration de la mesure de la distance entre deux points d'une maquette numérique dans l'application EasySensorManager

Les éléments d'adaptation de la norme sont présents dans ce cas d'étude sous les formes suivantes :

– Tâche : la tâche de pointage réalisée en 3D est régulièrement effectuée par les utilisateurs pour placer ou manipuler des capteurs dans un avion 3D. Les concepteurs doivent fréquemment sélectionner de façon successive deux éléments dans la maquette numérique pour mesurer une distance par exemple : cela correspond à la tâche de sélection décrite dans la norme entre deux points de l'avion qui sont positionnés à des profondeurs différentes dans la maquette (cf. figure 8 - points A et B).

– Cibles : les cibles sont typiquement des éléments de la structure de l'avion (par exemple, les extrémités des ailes) ou des capteurs. Les concepteurs peuvent sélectionner des capteurs positionnés à l'intérieur ou à l'extérieur de la surface de la carlingue d'un avion. Le volume de ces éléments et la distance entre eux sont variables et ainsi ces variations correspondent à différentes difficultés de la tâche (ID) de la norme.

– Variables : les concepteurs doivent positionner de façon précise un grand nombre de capteurs pour un même test ; un faible temps de mouvement ou un faible taux d'erreur sont donc importants pour réduire le temps de la tâche.

– Représentation et indices 3D : pour aider à percevoir la 3D dans la maquette qui s'avère complexe et pour éviter de biaiser sa compréhension, aucun élément de décoration, tel que des ombres, n'est ajouté à la maquette numérique. Les objets sont alors représentés par des couleurs et des transparences différentes pour mettre en avant les différentes profondeurs présentes dans la maquette numérique. Ceci est lié à ce qui est proposé comme recommandation dans notre adaptation.

– Confort et satisfaction : le positionnement des capteurs est une activité journalière pour les concepteurs et ils peuvent passer de longues heures à réaliser cette activité : il est en effet important qu'ils se sentent confortables et satisfaits de la technique utilisée. Le score SUS et l'échelle de confort aideront à répondre à ce cas d'utilisation.

4.2. Hypothèses

L'objectif principal de l'évaluation de la tâche de pointage en EV 3D est de comparer deux techniques d'interaction en vue d'une utilisation dans notre cas d'étude. Dans un premier temps, nous allons étudier l'applicabilité de la loi de Fitts dans un tel contexte afin d'évaluer l'efficacité de chaque technique : notre première hypothèse (H1) est que la loi de Fitts s'applique toujours dans un environnement 3D et avec des formes d'interaction avancée. Dans un second temps, nous allons déterminer la meilleure technique pour réaliser la tâche en termes de performance et de satisfaction utilisateur. Pour évaluer la rapidité et la précision des techniques, notre deuxième hypothèse (H2) est que la technique isotonique est plus performante que la technique isométrique et a une meilleure mesure de débit. Pour formuler cette hypothèse, nous nous basons sur des études menées d'après la norme ISO 9241-9 sur des techniques d'interaction, où les techniques isotoniques ont eu de meilleures performances (Soukoreff et MacKenzie, 2004). Nous avons aussi formulé comme hypothèse que les utilisateurs préfèrent interagir dans un EV 3D avec une technique isotonique (H3) étant donné que l'interaction semble plus usuelle qu'avec une technique isométrique. Pour mener à bien cette expérimentation, nous nous sommes appuyés sur l'outil logiciel proposé (cf. 3.3) et rendant possible la conduite d'évaluations selon le protocole d'expérimentation qui met en œuvre notre adaptation de la norme.

4.3. Participants

Quatorze utilisateurs non payés (3 femmes et 11 hommes) ont été recrutés pour participer à cette étude afin de tester les deux techniques. Mais nous n'avons analysé les résultats pour seulement 13 utilisateurs, car les logs pour un utilisateur étaient aberrants et inutilisables. Les participants avaient entre 22 et 39 ans (moyenne : 27,5, SD : 5,26). Utilisateurs réguliers d'ordinateur, ils n'étaient pas familiers avec nos techniques d'interaction en 3D et ils ont utilisé leur main dominante (12 droitiers et 1 gaucher).

4.4. Contexte d'évaluation

L'expérience a été menée en utilisant un ordinateur portable HP EliteBook 6930p portable avec Windows 7 et un écran LCD avec une résolution de 1920 * 1080. Le processeur de l'ordinateur était un 2,4 GHz Intel Core Duo CPU avec 4Go de RAM et la carte graphique une ATI Mobility Radeon HD 3450 256Mo.

Deux techniques ont été conçues et mises en œuvre pour soutenir cette activité et déterminer celle qui s'avère la mieux adaptée (cf. figure 9). Concrètement, la première technique est une technique isotonique basée sur une manipulation d'objet dans un espace physique 3D afin de déplacer le pointeur numérique dans la scène 3D. La position de l'objet dans l'espace physique est captée par la technique et correspond à la position du pointeur dans l'environnement 3D. Cette technique est basée ainsi sur un mode de contrôle de position et utilise la technologie du *Polhemus Patriot*. Les utilisateurs manipulent la technique en positionnant leur bras dans un espace physique qui a été calibré. La seconde technique est une technique plus classique basée sur la force de l'utilisateur captée pour déplacer le pointeur numérique. La force d'entrée est transformée en vitesse pour modifier la position du pointeur dans l'environnement 3D. Cette technique est basée ainsi sur un mode de contrôle de vitesse et correspond à un contrôleur isométrique utilisant un *SpaceNavigator*. Les utilisateurs manipulent le joystick du contrôleur, qui se trouve posé sur une table.

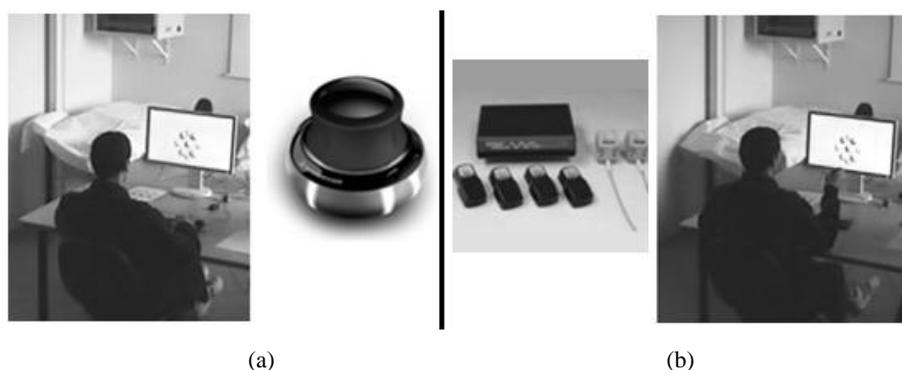


Figure 9. Posture des utilisateurs manipulant les deux techniques d'interaction qui sont le SpaceNavigator, dispositif isométrique (a) vs. le Polhemus Patriot, dispositif isotonique (b)

Une pré-expérience a été menée afin de déterminer les paramètres optimaux pour les deux modes de contrôle, et notamment un control-gain de 0.45 qui correspond au coefficient permettant de déplacer le curseur à l'écran en fonction du déplacement réalisé avec la technique d'interaction dans le monde physique. Les utilisateurs ont visualisé les données sur un écran LCD indépendant et non avec celui de l'ordinateur portable, servant à configurer l'expérimentation.

4.5. Procédure

Au début de l'expérimentation, chaque participant remplissait un formulaire de consentement éclairé et un questionnaire démographique. Les participants s'asseyaient ensuite devant le bureau où étaient disposés l'écran LCD et les deux techniques d'interaction. Les instructions, conformes à celles spécifiées dans l'adaptation à la norme (cf. 3.1.1), étaient données aux utilisateurs et ils pouvaient s'exercer pendant 15 min avant le début de l'évaluation. Comme mentionné dans la partie « tâche » de l'adaptation (cf. 3.1.1), la tâche de pointage débutait lorsque les participants sélectionnaient la sphère de départ. La tâche se terminait lorsque les participants sélectionnaient la sphère de destination. Les participants pouvaient se reposer à tout moment entre deux tâches de pointage. Lorsque toutes les tâches de pointage étaient réalisées avec la première technique, ils remplissaient une échelle pour évaluer le confort de la technique ainsi qu'un questionnaire de satisfaction pour évaluer l'utilisabilité de la technique, afin de se conformer à la partie « confort et satisfaction » de notre adaptation de la norme (cf. 3.2.2). Les participants suivaient la même procédure pour la seconde technique. A la fin de l'expérimentation, ils remplissaient un questionnaire de comparaison des deux techniques. Des entretiens complémentaires, comme recommandés dans notre adaptation, étaient menés afin de recueillir les préférences des utilisateurs et d'identifier les trois points les plus positifs et les trois points les plus négatifs de chaque technique.

4.6. Conception de l'évaluation

Comme recommandé dans la section 3.1.3 de notre adaptation de la norme, différents ID sont spécifiés. Dans notre évaluation, les ID ont été définis comme une combinaison de la largeur de la cible (taille du curseur + taille de la cible = 25px, 40px, 55px) et de la distance à parcourir pour atteindre une cible (distance = 150x, 450px, 750px). Les 9 ID résultants étaient compris entre 1.898 et 4.954. La distance utilisée pour réaliser le calcul des ID correspond à une distance euclidienne mesurée en 3D. Dans cette instance de la norme adaptée, le nombre de cibles à pointer a été fixé à 14 cibles : 6 cibles ont été disposées sur les extrémités des axes (x, y, z) et 8 en longitude et latitude $\pi/4$ (cf. figure 7). Les participants ont été divisés en deux groupes et chaque groupe a testé les techniques dans un ordre contrebalancé. Un test pour une technique était composé de 9 blocs, à savoir des combinaisons de toutes les tailles et distances ; et chaque bloc consistait en 14 tâches de pointage. Toutes les tâches d'un bloc avaient le même indice de difficulté : la taille de la cible et la distance à parcourir sont restées similaires pour toutes les tâches d'un bloc. Chaque

bloc correspondait donc à un des indices de difficultés défini. L'ordre des blocs était contrebalancé par un carré latin.

Au final, nous avons recueilli 3 276 temps de pointage étant donné que nous avons 13 participants * 2 techniques * 9 blocs * 14 sélections de cible. Par ailleurs, les événements comme le temps, l'erreur et la position courante du pointeur ont été enregistrés pendant la réalisation de la tâche par les utilisateurs.

4.7. Mesures quantitatives

Basés sur les données enregistrées lors de l'évaluation, nous avons analysé les résultats en utilisant des statistiques descriptives et inférentielles (*i.e.* ANOVA).

4.7.1. Temps de mouvement et paramètres de Fitts

La norme ISO inclut le temps moyen de mouvement en tant que variable dépendante se basant sur la loi de Fitts. Nous avons donc tout d'abord analysé le temps moyen requis pour sélectionner la cible : la mesure du temps moyen pour la technique isométrique (*SpaceNavigator*) est de 10 681 ms et pour la technique isotonique (*Patriot*) de 6 265 ms. Ces temps sont assez proches de ceux trouvés dans l'étude de (Froehlich *et al.*, 2006). L'analyse montre une différence significative entre les deux techniques ($F_{(1,12)} = 179.05, p < 0.0001$). Le *Patriot* permet ainsi aux utilisateurs de réaliser la tâche de pointage significativement plus rapidement que le *SpaceNavigator*. Nous pensons que cette efficacité plus élevée de la technique isotonique en comparaison de la technique isométrique est due à son utilisation plus usuelle et aux mouvements du bras moins contraints.

En plus du temps moyen global, nous avons analysé les mouvements des utilisateurs afin de mieux comprendre l'impact de la 3D sur les résultats de cette expérience. Nous avons étudié l'impact des mouvements réalisés dans une dimension le long d'un seul axe (mouvements simples) et ceux réalisés en 3 dimensions (mouvements complexes) afin de prendre en compte des cas qui puissent être plus contraignants et correspondre à plusieurs types de trajectoire. Une première analyse montre une différence significative entre les mouvements en termes de temps ($F_{(1,12)} = 13.67, p < 0.001$) et de distance ($F_{(1,12)} = 16.91, p < 0.0001$) quelle que soit la technique. Les utilisateurs effectuent la tâche plus rapidement et avec une plus petite longueur lors de la réalisation des mouvements simples que lors de la réalisation des mouvements complexes. Cela signifie que les utilisateurs ont moins de difficulté à manipuler le pointeur lors des mouvements simples que lors des mouvements complexes. Une deuxième analyse, centrée sur les mouvements simples, montre une différence significative entre les deux techniques en termes de temps ($F_{(1,12)} = 56.74, p < 0.0001$) et de distance ($F_{(1,12)} = 4.81, p < 0,05$). Concernant les mouvements complexes, une différence significative existe également entre les deux techniques pour le temps ($F_{(1,12)} = 123.86, p < 0.0001$) et pour la distance ($F_{(1,12)} = 47.68, p < 0,0001$). Dans les deux cas, il ressort également que le *Patriot* permet aux utilisateurs d'effectuer la tâche plus rapidement (6 001 ms contre 9 505 ms pour les mouvements simples et 6 464 ms contre 11 563 ms pour les mouvements complexes respectivement

pour le *Patriot* et pour le *SpaceNavigator*) et avec une plus petite longueur (944px contre 1043px pour les mouvements simples et 983px contre 1250px pour les mouvements complexes respectivement pour le *Patriot* et pour le *SpaceNavigator*) pour les deux types de mouvements. Ces résultats confirment les meilleures performances de la technique isotonique.

Une deuxième partie de l'analyse consiste à calculer le temps moyen requis pour réaliser la tâche pour un certain ID. Dans la figure 10, l'axe X et l'axe Y représentent respectivement les indices de difficultés (ID) et la moyenne des temps mesurés (en millisecondes) pour chaque ID, tout utilisateur confondu. Conformément au protocole d'analyse menée dans des expérimentations Fitts, ces valeurs sont données pour les 9 ID de l'expérimentation. Les droites solides et en pointillés sont les droites de régression obtenues respectivement pour la technique isométrique ($MT = 5267.8 \times ID - 7543.4$ avec $R^2 = 0.8632$) et pour la technique isotonique ($MT = 2973.1 \times ID - 4019.7$ avec $R^2 = 0.8163$). Nous remarquons que le temps moyen s'accroît plus rapidement pour la technique isométrique quand l'ID augmente : une interprétation nous conduit à conclure que cette technique n'est pas adaptée aux situations les plus difficiles (petites cibles, longues distances). Cette analyse met également en évidence un problème : la corrélation entre l'ID et le temps moyen est assez faible pour les deux techniques ($R^2 < 0.9$). Nous observons aussi que la tâche de pointage pour les ID comprenant une distance élevée (3.87, 4.3 et 4.954) est plus difficile à réaliser pour les deux techniques, qui ne semblent donc pas adaptées à ce type de situation d'interaction.

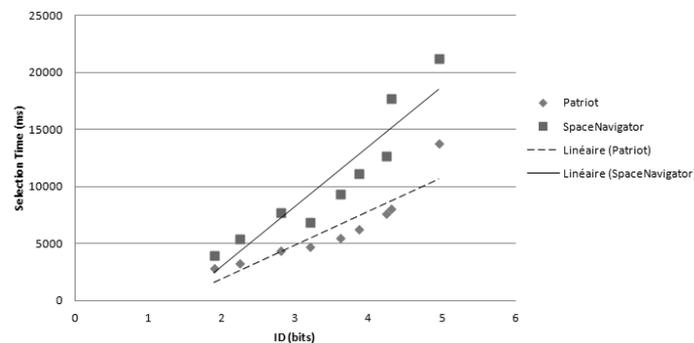


Figure 10. Régression linéaire MT vs ID

Le graphique révèle enfin que le temps moyen de mouvement est au-dessus de la droite de régression pour 3 ID (3.87, 4.3 et 4.954), ID pour lesquels la distance entre deux cibles est égale à 750px (la plus grande distance). En utilisant la méthode réalisée dans (Raynal *et al.*, 2010), nous avons calculé une seconde équation de régression sans ces 3 ID (cf. figure 11). Ce second calcul contribue à augmenter la corrélation. Les nouvelles équations de régression sont ($MT = 3392 \times ID - 2500.7$ avec $R^2 = 0.9244$) pour la technique isométrique et ($MT = 1917.9 \times ID - 1044$ avec

$R^2 = 0.9559$) pour la technique isotonique. Nous remarquons que la corrélation est plus élevée pour les deux techniques sans la distance la plus élevée.

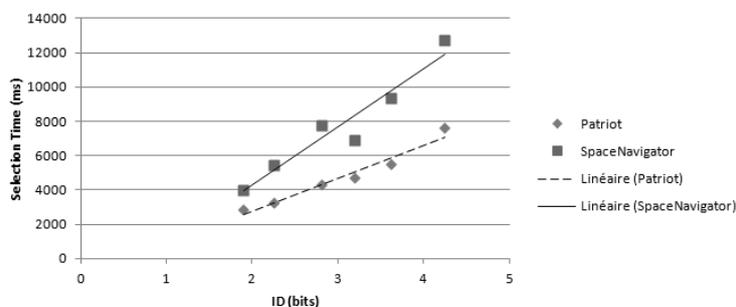


Figure 11. Régression linéaire MT vs ID sans les cibles situées à une distance de 750 px. Il reste donc 6 ID obtenus (1.89, 2.24, 2.8, 3.19, 3.61, 4.24) par la combinaison des tailles (25,40 et 55) et des distances (150 et 450)

Dans les deux cas (avec ou sans la distance de 750px), le profil des droites de régression est similaire : la droite de la technique isométrique est au-dessus de celle de la technique isotonique et s'accroît plus rapidement. Cette tendance confirme l'efficacité de la technique isotonique au regard de la technique isométrique en fonction des ID testés.

Notre première hypothèse (H1) est ainsi confirmée par cette analyse et notamment par les coefficients de corrélation proches de 0.95 : la loi de Fitts peut s'appliquer aux environnements 3D et aux techniques d'interaction avancée, même si le paramètre a est négatif. Ceci peut être expliqué par une activité musculaire incontrôlable pour l'utilisateur au début ou à la fin du mouvement de la tâche selon (MacKenzie, 1992). De plus, le coefficient assez élevé de la pente (le paramètre b) compense ou annule le paramètre a négatif. Le paramètre b de la technique isotonique est plus faible que celui de la technique isométrique ; ainsi le *Patriot* semble être moins sensible à la difficulté de la tâche à réaliser que le *SpaceNavigator*. Nous pensons que la manipulation séparée de chaque degré de liberté est plus simple avec la technique isotonique (*Patriot*) et que cela explique cette différence.

Nous pouvons également noter que les paramètres a et b obtenus pour nos équations varient fortement d'une technique à une autre et fournissent ainsi un moyen clair et uniforme pour les différencier et les comparer.

4.7.2. Taux d'erreur

Notre adaptation de la norme propose également la mesure du taux d'erreur en tant que variable dépendante. Nous avons donc analysé le taux d'erreur pour les deux techniques afin de mesurer la réussite de la tâche de pointage et l'efficacité de ces techniques. Le taux d'erreur mesuré pour la technique isométrique est plus faible

que celui de la technique isotonique (4,4 % vs 7,6 %). L'analyse montre une différence significative entre ces techniques ($F_{(1,12)} = 20.77$, $p < 0.005$). La technique isométrique permet aux utilisateurs de réaliser de petits mouvements pour sélectionner les cibles avec précision et permet également de maintenir une position neutre contrairement à la technique isotonique, pour laquelle les utilisateurs ont des difficultés à garder le pointeur en contact avec les cibles ; ces raisons peuvent expliquer la plus grande précision de cette technique isométrique. Ainsi, la technique isométrique est moins sujette aux erreurs que la technique isotonique.

4.7.3. Ratio d'inefficience

Pour affiner l'étude de la performance, notre adaptation de la norme recommande de considérer le ratio d'inefficience. Le calcul du ratio d'inefficience pour les deux techniques révèle que le ratio d'inefficience de la technique isotonique (1,06) est inférieur à celui de la technique isométrique (1,56). Cela signifie que les trajectoires réalisées avec la technique isotonique sont 106 % plus longues que la distance optimale et celles obtenues avec la technique isométrique sont 156 % plus longues que la distance optimale. L'analyse montre une différence significative entre ces techniques ($F_{(1,12)} = 218,75$, $p < 0.0001$). Nous en déduisons que la technique isotonique supporte une meilleure coordination que la technique isométrique.

4.7.4. Débit

Parmi les variables dépendantes identifiées dans la norme (cf. 3.1.3), le débit est une mesure de performance globale qui permet de comparer facilement deux techniques. Nous avons ainsi calculé le débit pour chaque technique d'interaction en utilisant les tailles des cibles et distances effectives, qui correspondent respectivement aux tailles ajustées des cibles en fonction des sélections des utilisateurs (largeur effective – We) et aux distances réalisées au cours de la tâche par les utilisateurs (distance effective – De). Ce calcul utilise une base par participant et par bloc afin de proposer le « débit général ». Les débits pour la technique isotonique et pour la technique isométrique sont respectivement de 2.1bps et 1.69bps. L'effet majeur de la technique de pointage sur le débit est statistiquement significatif ($F_{(1,12)} = 5.76$, $p < 0.05$). Étant donné la signification du débit, c'est-à-dire le taux de transfert d'informations lors de la tâche, nous en déduisons que la technique isotonique est plus performante que la technique isométrique. Cela renforce nos résultats précédemment détaillés (cf. 4.7.1) et confirme notre deuxième hypothèse (H2) sur la plus grande efficacité de la technique isotonique.

4.8. Mesures qualitatives

Notre approche composite inclut également des aspects qualitatifs pour l'évaluation de ces techniques d'interaction. Pour l'analyse des résultats qualitatifs, nous avons utilisé des statistiques non paramétriques (*i.e.* test de Mann and Whitney, test du signe), plus adaptées pour ce type de données.

4.8.1. Évaluation du confort

Comme recommandé dans la partie confort et satisfaction du standard (cf. 3.2.2), les participants ont complété un questionnaire sur le confort, qui inclut deux types d'indices : des indices généraux et des indices sur la fatigue. Selon la procédure décrite dans la section 4.5, les participants ont complété, dans une première phase, un questionnaire après avoir testé la première technique afin d'évaluer le confort de cette technique. Ce questionnaire est basé sur une échelle de Likert à 5 niveaux allant de « très négatif » (1) à « très positif » (5). Les résultats de ce premier questionnaire n'ont pas montré de différence significative pour tous les indices selon un test de Mann and Whitney, hormis pour la fatigue du bras. L'indice de fatigue du bras révèle une différence significative marginale ($p=0.091$) : la technique isotonique semble plus fatigante pour le bras que la technique isométrique. Le score moyen pour tous les indices de la première phase est 3,41 pour la technique isométrique et 3,54 pour la technique isotonique : ce résultat montre que les utilisateurs ont tendance à avoir un meilleur ressenti sur le confort pour la technique isotonique (*Patriot*). Après avoir testé la seconde technique dans une seconde phase, les participants ont rempli une échelle pour comparer les deux techniques et ainsi déterminer si la dernière technique testée est pire, identique ou meilleure que la première d'après les mêmes indices. Nous avons noté qu'aucun effet d'apprentissage n'affecte l'analyse de ces questionnaires car les techniques ont été testées dans un ordre contrebalancé. Un test du signe, utilisé pour comparer des échantillons appariés, ne montre aucune différence significative ($|Z|=-0.28$, $p>0.5$) entre les deux techniques. Mais les résultats montrent que la technique isotonique tend à entraîner plus de fatigue pour le bras et l'épaule que la technique isométrique.

4.8.2. Satisfaction des utilisateurs

L'adaptation de la norme recommande d'analyser l'utilisabilité des techniques avec un questionnaire SUS en plus du questionnaire sur le confort. Nous avons calculé un score SUS pour mesurer la facilité d'utilisation de chaque technique. Le score de la technique isométrique (52,5) est marginal à un faible niveau alors que le score de la technique isotonique (76,3) est acceptable, d'après l'échelle définie dans (Bangor *et al.*, 2008). Ceci montre que l'utilisabilité des techniques peut être améliorée, même si l'utilisabilité de la technique isotonique (*Patriot*) est bonne. Les deux scores SUS permettent d'établir que l'utilisabilité actuelle des techniques n'ont pas compromis leur utilisation lors des tâches de pointage. Comme suggéré par l'adaptation, des éléments du questionnaire SUMI ont également été soumis aux utilisateurs. Leurs réponses aux items du questionnaire SUMI soulignent que la technique isotonique est perçue comme plus stimulante et moins frustrante que la technique isométrique.

4.8.3. Préférences des utilisateurs

Comme proposé dans l'adaptation du standard, nous avons utilisé des entretiens pour affiner l'analyse de la préférence des utilisateurs : les réponses des utilisateurs sont d'ordinaire des indicateurs fiables en ce qui concerne les développements possibles permettant d'atteindre un meilleur niveau d'utilisabilité. En ce qui

concerne la technique isométrique (basée sur le *SpaceNavigator*), les 3 points les plus positifs les plus fréquemment cités sont : un bon esthétisme du dispositif, la possibilité de réaliser des mouvements en 3D et la facilité d'utilisation. Les 3 points les plus négatifs les plus fréquemment cités sont : une association inappropriée entre les interactions en entrée et en sortie, en particulier pour les interactions en profondeur ; les difficultés à n'utiliser qu'un seul axe durant la tâche et enfin, son fonctionnement général. Concernant la technique isotonique (basée sur le *Patriot*), les 3 points les plus positifs les plus fréquemment cités sont : la facilité d'utilisation, la possibilité de réaliser des mouvements en 3D et la facilité de trouver des repères dans l'espace. Les 3 points les plus négatifs les plus fréquemment cités sont : la fatigue entraînée par la technique durant la tâche, le faible design du dispositif et le fait de garder le bras levé durant la tâche. Ces résultats mettent en évidence deux différences majeures à propos de ces points : (1) la technique isométrique est principalement perçue comme « attrayante » et la technique isotonique comme « intuitive » ; (2) le point le plus négatif pour la technique isométrique est que les utilisateurs n'ont pas bien perçu la profondeur alors qu'ils ont pu trouver des repères dans l'espace avec la technique isotonique pour se déplacer facilement en 3D. Même si le *Patriot* semble être plus « intuitif », les utilisateurs ont ressenti plus de fatigue du bras avec cette technique qu'avec le *SpaceNavigator*, il serait donc pertinent d'envisager de porter un nouvel intérêt à son design et à son utilisation. Par ailleurs, les utilisateurs n'ont fait aucun commentaire négatif sur la représentation et les indices 3D suggérés dans notre adaptation de la norme (cf. 3.2.1).

Concernant les préférences utilisateur, 85 % des utilisateurs ont aimé interagir avec la technique isotonique et 77 % d'entre eux ont estimé que cette technique était plus utilisable que la technique isométrique. Cette évaluation confirme donc notre 3^e hypothèse (H3). En effet, ils sont 69 % à estimer que la technique isométrique est plus contraignante en termes de liberté de mouvement.

Pour conclure sur cette expérimentation comparative de nos deux techniques d'interaction pour une tâche de pointage en 3D, la mise en place de cette évaluation composite a permis de prendre en considération des aspects de performance et de satisfaction. L'objectif recherché est bien entendu de proposer une technique qui réponde au mieux aux différents besoins des utilisateurs : rapidité, précision, facilité d'apprentissage, satisfaction et confort. L'intérêt de recueillir différentes mesures couvrant ces besoins est donc de déterminer la technique la plus adaptée à la tâche évaluée aussi bien en termes de performance que de satisfaction utilisateur. La prise en compte de ces besoins est nécessaire pour un cas d'étude comme celui que nous avons présenté car les tâches sont répétitives et quotidiennes. D'après nos résultats, la technique isotonique pourrait ainsi être recommandée car elle est plus rapide, permet une meilleure coordination et a été préférée par les utilisateurs. Il resterait alors à proposer une solution pour limiter la fatigue ressentie par les utilisateurs.

5. Conclusion

Dans cet article, nous avons présenté une adaptation de la norme ISO 9241-9, qui propose un protocole structuré pour mener l'évaluation de techniques d'interaction avancée pour une tâche de pointage en EV 3D.

L'adaptation de la norme comprend trois aspects principaux d'enrichissement qui tentent de répondre à ce besoin d'évaluer des caractéristiques particulières. Le premier aspect est l'ensemble des variables de l'expérience, qui est enrichi par une mesure de coordination afin de mieux comprendre les mouvements des utilisateurs en 3D et l'analyse explicite de la précision. Le deuxième aspect est l'ajustement de la tâche et des cibles pour s'adapter à des contextes de pointage en EV 3D : les formes, les arrangements, les représentations et les indices 3D sont proposés afin d'aider les utilisateurs à mieux percevoir la profondeur. Le rendu possède bien une perspective mettant en avant la profondeur des objets 3D et proposant ainsi un vrai rendu 3D. Le troisième aspect est l'ajout du questionnaire SUS au questionnaire sur le confort afin de permettre une comparaison facile entre les techniques en termes d'utilisabilité ; il fournit également une mesure globale, similaire au débit, mais liée à la satisfaction et non à la performance. En marge de ces modifications majeures, des adaptations relatives au calcul du débit (*throughput*), de l'axe de la tâche et des distances ont été explicitées pour répondre aux contraintes apportées par le contexte 3D. Outre l'extension de cette norme et donc de la description du protocole d'évaluation, un outil logiciel a été proposé : il vise à instrumenter et à supporter l'exécution de l'expérimentation conformément à l'adaptation de la norme. Il consiste en un environnement virtuel 3D, dans lequel la tâche de pointage est effectuée en 3D, et un panneau de configuration conçu de manière à gérer les différents paramètres (par exemple, la largeur ou la distance des cibles), et à configurer l'enregistrement des traces. Cet environnement de mise en œuvre de l'adaptation de la norme permet de tendre vers une uniformisation de la réalisation du protocole décrit dans l'adaptation de la norme.

Pour illustrer l'utilisation et l'applicabilité de cette adaptation, nous nous sommes inspirés d'une tâche industrielle : la manipulation de capteurs sur une maquette 3D pour une campagne d'essais. D'après les résultats de l'expérience, la technique isotonique (*Patriot*) est plus efficace, plus utilisable et permet aux utilisateurs d'effectuer une tâche de façon plus coordonnée que la technique isométrique (*SpaceNavigator*) : en particulier, la moyenne des temps, les scores de débit et du SUS ainsi que le ratio d'inefficience sont meilleurs pour la technique isotonique. Par ailleurs, les entretiens montrent que les utilisateurs préfèrent employer la technique isotonique pour déplacer le pointeur car la manipulation semble plus naturelle. Cependant, cette technique peut être moins confortable à utiliser pour le bras ou l'épaule puisqu'elle a tendance à provoquer plus de fatigue que la technique isométrique.

L'évaluation des techniques pour le pointage en EV 3D respecte le cadre proposé : par conséquent, la performance, la précision, la satisfaction et le confort ont été considérés. L'approche que nous proposons de l'évaluation d'une tâche de pointage en EV 3D permet ainsi de combiner des aspects quantitatifs et qualitatifs dans un protocole d'évaluation basé sur la norme ISO 9241-9. En adaptant certains

éléments de la norme, nous proposons un cadre structuré qui constitue une première étape vers une formalisation de ce type d'évaluation et la création d'une base de connaissances cohérentes sur l'utilisabilité de techniques à au moins 3 degrés de liberté pour le pointage dans un EV 3D. Certes, le protocole concerne uniquement une tâche de pointage en EV 3D ; toutefois, comme souligné précédemment, le pointage est sous-jacent à de nombreuses interactions en 3D. Ce type d'évaluation basée sur la loi de Fitts permet l'étude d'un grand nombre de cas possibles en faisant varier l'intervalle des indices de difficultés (ID). De plus, cela permet de comparer des résultats globaux car ils ne sont pas dépendants d'un contexte de tâche particulier, et de se focaliser sur une caractéristique de la technique, qui une fois optimisée, contribuera à l'adéquation globale de la technique au domaine visé. Cette étude fine permettra, si besoin, de cerner le contexte d'utilisation qui s'avère le plus approprié pour une technique d'interaction plutôt que pour une autre : par exemple, pour un score SUS et un débit (TP) équivalents entre deux techniques, l'analyse atomique du pointage selon les ID pourra révéler une meilleure adéquation à l'application cible d'une des techniques d'interaction et celle-ci pourra alors être privilégiée sur ce motif. Il en va de même avec la combinaison des indices de performance (temps, ratio d'inefficience et TP) et de satisfaction (SUS, confort) : selon le type de contexte, les uns pourront s'avérer prédominants, mais les autres ne seront pas pour autant omis du protocole d'expérimentation.

Bibliographie

- Accot J., Zhai S. (2003). Refining Fitts' law models for bivariate pointing. *Proceedings of CHI'03*, p. 193-200.
- Bach C., Scapin D. (2003). Ergonomic criteria adapted to human virtual environment interaction. *Proceedings of IHM'03*, p. 24-31.
- Balakrishnan R., Baudel T., Kurtenbach G., Fitzmaurice G. (1997). The Rockin' Mouse: integral 3D manipulation on a plane. *Proceedings of CHI'97*, p. 311-318.
- Bangor A., Kortum P.-T., Miller J.-T. (2008). An empirical evaluation of the System Usability Scale. *Journal of human-computer interaction*, p. 574-594.
- Bérard F., Ip J., Benovoy M., El-Shimy D., Blum J.R., Cooperstock J. R. (2009). Did "Minority Report" Get it Wrong? Superiority of the Mouse over 3D Input in a 3D Placement Task. *Proceedings of INTERACT'09*.
- Bowman D. A., Kruijff E., LaViola J. J., Poupyrev I. (2004). *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc.
- Bowman D. A. (1999). *Interaction Techniques for Common Tasks in Immersive Virtual Environments: Design, Evaluation, and Application*. Ph.D Dissertation.
- Brooke J. (1996). SUS: a "quick and dirty" usability scale. P. W. Jordan, B. Thomas, B. A. Weerdmeester & A. L. McClelland. *Usability Evaluation in Industry*. London: Taylor and Francis.

- Fitts P. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, p. 381-391.
- Froehlich B., Hochstrate J., Skuk V., Huckauf A. (2006). The GlobeFish and the GlobeMouse: two new six degree of freedom input devices for graphics applications. *Proceedings of CHI'06*, p. 191-199.
- Froehlich B., Plate J. (2000). The cubic mouse: a new device for three-dimensional input. *Proceedings of CHI'00*, p. 526-531.
- Gallo L., De Pietro G., Marra I. (2008). 3D interaction with volumetric medical data: experiencing the Wiimote. *Proceedings of Ambi-Sys'08*.
- Grossman T., Balakrishnan R. (2004). Pointing at trivariate targets in 3D environments. *Proceedings of CHI'04*, p. 447-454.
- Ha T., Woo W. (2010). An empirical evaluation of virtual hand techniques for 3D object manipulation in a tangible augmented reality environment. *Proceedings of 3DUI'10*, IEEE Computer Society, p. 91-98.
- Hubona G. S., Wheeler P. N., Shirah G. W., Brandt M. (1999). The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Trans. Comput.-Hum. Interact.*, p. 214-242.
- ISO 9241-9. (2000). *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices*. ISO.
- Kaiser P. (2011). *The Joy of Visual Perception: A Web Book*. York University.
- MacKenzie I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, vol. 7, p. 91-139.
- MacKenzie I.S., Buxton W. (1992). Extending Fitts' law to two-dimensional tasks. *Proceedings of CHI'92*, p. 219-226.
- Masliah M. (2001). *Measuring the allocation of control in 6 degree of freedom human-computer interaction tasks*. Thèse de doctorat, Université de Toronto.
- Masliah M. (1999). <http://etclab.mie.utoronto.ca/people/moman/timeandspace/timeandspace.html>.
- McArthur V., Castellucci S.J., MacKenzie I.S. (2009). An empirical comparison of "wiimote" gun attachments for pointing tasks. *Proceedings of EICS'09*, p. 203-208.
- Natapov D., Castellucci S.J., MacKenzie I.S. (2009). I.S. ISO 9241-9 evaluation of video game controllers. *Proceedings of GI'09*, p. 223-230.
- Ortega M., Nigay L., (2009). AirMouse: Finger Gesture for 2D and 3D Interaction. *Proceedings of INTERACT'09*, p. 214-227.
- Pan Q. (2008). *Techniques d'interactions mixtes isotonique et élastique pour la sélection 2D et la navigation/manipulation 3D*. Thèse de doctorat, Université de Lille.
- Raynal M., Gauffre G., Bach C., Schmitt B., Dubois E. (2010). Tactile Camera vs. Tangible Camera: taking advantage of small physical artefacts to navigate into large data collection. *Proceedings of NordiCHI 2010*, p. 373-382.

- Soukoreff R.W., MacKenzie I.S. (2004). Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research. *HCI Int. J. Hum.-Comput. Studies*, vol. 61, n° 6, p. 751-789.
- SUMI. <http://sumi.ucc.ie/>
- Teather R.J., Stuerzlinger W. (2011). Pointing at 3D targets in a stereo head-tracked virtual environment. *Proceedings of 3DUI'11*, IEEE Computer Society.
- Teather R.J., Pavlovych A., Stuerzlinger W., MacKenzie I.S. (2009). Effects of tracking technology, latency, and spatial jitter on object movement. *Proceedings of IEEE 3D User Interfaces 2009*, p. 43-50.
- Ware C., Rose J. (1999). Rotating virtual objects with real handles. *ACM Transactions on CHI*, vol. 6, n° 2, p. 162-180.
- Ware C., Lowther K. (1997). Selection using a one-eyed cursor in a fish tank VR environment. *ACM Trans. Comput.-Hum. Interact.*, p. 309-322.
- Zhai S., Buxton W., Milgram P. (1994). The "Silk Cursor": Investigating transparency for 3D target acquisition. *Proceedings of CHI'94*, p. 459-464.
- Zhai S. (1993) Investigation of feel for 6 DOF inputs: isometric and elastic rate control for manipulation in 3D environments. *Proceedings of The Human Factors and Ergonomics Society 37th Annual Meeting*.
- Zhai S. (1998). User performance in relation to 3D input device design. *SIGGRAPH Comput. Graph*, vol. 32, n° 4, p. 50-54.
- Zhai S., Milgram P. (1998). Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. *Proceedings of CHI'98*, p. 320-327.
- Zhang X., MacKenzie I.S. (2007). Evaluating eye tracking with ISO 9241 - part 9. *Proceedings of HCI'07*, p. 779-788.

Article reçu le 23 décembre 2011

Accepté après révisions le 27 septembre 2012

Bénédicte Schmitt est doctorante en thèse CIFRE pour l'entreprise Global Vision Systems et a intégré l'équipe ELIPSE de l'IRIT. L'objectif de sa thèse est de proposer des techniques d'interaction adaptées aux tâches que les utilisateurs réalisent dans les applications développées par Global Vision Systems. Ces techniques d'interaction doivent s'avérer performantes et satisfaisantes afin de favoriser leur acceptation par les utilisateurs et faciliter l'accès aux données.

Mathieu Raynal est enseignant-chercheur en Informatique à l'Université de Toulouse, et membre de l'équipe ELIPSE de l'IRIT. Ses travaux de recherche portent sur la modélisation, la conception et l'évaluation de nouveaux systèmes de saisie de texte et dispositifs de pointage pour des contextes d'interaction « dégradée » dans lesquels la saisie et le pointage deviennent des tâches plus complexes : par exemple en situation de mobilité, ou dans le cas d'une déficience motrice, perceptive ou cognitive.

Cédric Bach est assistant de recherche à l'Université de Toulouse au sein de l'IRIT (Institut de Recherche en Informatique de Toulouse). Ses travaux de recherche portent sur les méthodes pour l'utilisabilité des systèmes interactifs et leur expérience d'usage (User eXperience). Il intervient régulièrement, en tant qu'ergonome expert, dans différents secteurs industriels et culturels, pour analyser, modéliser, concevoir et évaluer des interactions homme-machine.

David Croenne est le directeur technique de la jeune entreprise innovante Global Vision Systems. Ingénieur Centralien, il complète ses études par un DEA de physique en 2003. C'est dans le centre de commande spatial allemand de Munich qu'il met en place avec Baptiste Gendron les premiers outils de supervision industrielle en 3D temps réel. Ils fondent Global Vision Systems en 2008 avec pour objectif de développer et commercialiser des IHM innovantes pour les industriels. Global Vision Systems est aujourd'hui une équipe forte de vingt ingénieurs, répartis entre Toulouse et Montréal, et a déjà convaincu de nombreux industriels de l'apport de sa technologie.

Emmanuel Dubois est Professeur en informatique à l'Université de Toulouse. Il est responsable de l'équipe Elipse de l'Institut de Recherche en Informatique de Toulouse (IRIT, UMR CNRS 5505). Les travaux de l'équipe sont centrés sur les interactions homme-machine avancées (réalité augmentée, tangible, suppléance, sonore 3D) : ils visent à mieux les comprendre pour les concevoir, et à les utiliser pour comprendre la cognition humaine. Dans ce contexte, Emmanuel Dubois s'intéresse tout particulièrement aux approches de conception basées modèles pour les systèmes interactifs mixtes, en particulier dans le domaine des environnements 3D et des situations grands publics.