



**HAL**  
open science

# Influence des instructions multimodales sur l'apprentissage par compagnonnage des compétences de manipulation d'outil dans un environnement immersif

Cassandre Simon, Manel Boukli Hacene, Flavien Lebrun, Samir Otmane,  
Amine Chellali

## ► To cite this version:

Cassandre Simon, Manel Boukli Hacene, Flavien Lebrun, Samir Otmane, Amine Chellali. Influence des instructions multimodales sur l'apprentissage par compagnonnage des compétences de manipulation d'outil dans un environnement immersif. 2024. hal-04450132

**HAL Id: hal-04450132**

**<https://hal.science/hal-04450132>**

Preprint submitted on 9 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Influence des instructions multimodales sur l'apprentissage par compagnonnage des compétences de manipulation d'outil dans un environnement immersif

## Influence of multimodal instructions on learning tool manipulation skills through mentoring in an immersive environment

CASSANDRE SIMON\*, Laboratoire IBISC, Université d'Évry Paris-Saclay, France

MANEL BOUKLI-HACENE\*, Laboratoire IBISC, Université d'Évry Paris-Saclay, France

FLAVIEN LEBRUN\*, Laboratoire IBISC, Université d'Évry Paris-Saclay, France

SAMIR OTMANE\*, Laboratoire IBISC, Université d'Évry Paris-Saclay, France

AMINE CHELLALI\*, Laboratoire IBISC, Université d'Évry Paris-Saclay, France



FIG. 1 – Illustration de l'étude expérimentale. Pensant qu'il interagit en temps-réel avec un instructeur distant, l'apprenant reçoit des instructions pré-enregistrées concernant l'amplitude d'un mouvement à reproduire en utilisant quatre combinaisons différentes de modalités : verbale-visuelle, verbale-haptique, visuelle-haptique et verbale-visuelle-haptique. Il tente ensuite de reproduire ce mouvement le plus fidèlement possible.

La formation par compagnonnage permet aux novices d'acquérir des compétences sous la supervision d'experts qui utilisent diverses modalités de communication. Cependant, reproduire ce modèle dans des simulateurs immersifs reste un défi, notamment pour assurer une communication efficace entre experts et novices. Notre étude explore l'impact de la communication multimodale expert-novice pour transmettre des instructions sur l'amplitude des mouvements dans une tâche de manipulation d'outils en environnement immersif. Les résultats révèlent que la combinaison des modalités visuelle-haptique améliore la précision, la vitesse et la qualité des mouvements. De plus, la combinaison verbale-visuelle-haptique renforce le sentiment de présence et de coprésence, et l'expérience d'apprentissage. Ces résultats suggèrent que la combinaison visuelle-haptique est optimale pour améliorer les performances des novices, et que l'intégration de la modalité verbale améliore l'expérience utilisateur. Ces conclusions ouvrent de nouvelles perspectives pour améliorer l'acquisition de gestes techniques par compagnonnage en réalité virtuelle grâce à la communication multimodale.

Mentorship allows novices to acquire skills under the guidance of experts who use various communication modalities. However, reproducing this model in immersive simulators remains challenging, particularly in ensuring effective communication between

\*. All authors contributed equally to this research.

experts and novices. Our study explores the impact of multimodal expert-novice communication to transmit instructions on the amplitude of movements in a tool manipulation task in an immersive environment. The results reveal that combining visual-haptic modalities improves movements' precision, speed, and quality. Additionally, the verbal-visual-haptic combination enhances the sense of presence, co-presence, and learning experience. These results suggest that the visual-haptic combination is optimal for improving novice performance, and integrating the verbal modality enhances the user experience. These findings open new perspectives for improving the acquisition of technical skills through VR mentorship using multimodal communication.

CCS Concepts: • **Human-centered computing** → Virtual reality; User studies; Collaborative interaction.

Mots Clés et Phrases Supplémentaires: Interactions multimodale, Formation par compagnonnage, Apprentissage des gestes

Additional Key Words and Phrases: Multimodal interactions, Mentorship, Skill learning

### Reference:

Cassandra Simon, Manel Boukli-Hacene, Flavien Lebrun, Samir Otmame, and Amine Chellali. 2024. Influence des instructions multimodales sur l'apprentissage par compagnonnage des compétences de manipulation d'outil dans un environnement immersif.

This is the author's draft version of the work. It is posted here for your personal use. Not for distribution. The definitive version of the record is published in IHM '24: Proceedings of the 35th Conference on l'Interaction Humain-Machine. .

## 1 INTRODUCTION

Les nouvelles méthodes de simulation, telles que la réalité virtuelle (RV), ont récemment montré leur efficacité pour l'apprentissage des compétences motrices, particulièrement en médecine [6, 55]. Néanmoins, la plupart des simulateurs existants permettent uniquement un apprentissage en autonomie, où les apprenants pratiquent les compétences de manière individuelle. Cette approche exclut ainsi l'encadrement et les retours des formateurs, éléments jugés essentiels pendant les premières étapes de l'acquisition des compétences motrices [2, 30, 44]. Plusieurs études suggèrent que les instructeurs jouent un rôle crucial pour aider les apprenants à développer et à améliorer durablement ces compétences tout en réduisant la charge mentale [2, 5, 35]. Pour transmettre des instructions ou fournir des rétroactions aux apprenants, les formateurs utilisent diverses modalités de communication verbales ou non verbales [14, 50]. Cependant, la possibilité pour le formateur de communiquer efficacement avec un apprenant dans un environnement virtuel immersif demeure une question ouverte. Cet article explore la conception d'interactions et d'interfaces de RV pour la formation par compagnonnage. Ces outils permettront aux instructeurs de partager leurs compétences motrices et de guider les apprenants dans un cadre immersif afin d'améliorer le transfert de ces compétences dans un environnement sûr et contrôlé. Pour concevoir de tels systèmes, il est essentiel de comprendre comment les experts transmettent ces compétences aux apprenants et l'impact des modalités de communication sur ce processus. Le présent travail est centré sur l'entraînement aux tâches de manipulation d'outils dans un environnement immersif partagé (EIP). Plus précisément, nous étudions comment un instructeur peut aider efficacement les apprenants à connaître la bonne amplitude des mouvements à réaliser, à l'aide d'instructions multimodales (figure 1). Des études antérieures ont exploré l'impact d'utiliser une seule modalité de communication (verbale, visuelle ou haptique) sur les performances et l'expérience d'apprentissage de l'apprenant [46], montrant que chaque modalité a ses propres points forts et points faibles. Ce travail s'appuie sur ces recherches antérieures et vise à étudier l'impact de la combinaison des modalités de communication sur les performances et l'expérience utilisateur. Notre hypothèse centrale est que la communication multimodale améliorerait les performances et l'expérience d'apprentissage de l'apprenant lorsqu'il reçoit des instructions dans un EIP. Les principales contributions de cette étude sont de donner des indications sur les rôles joués par chaque combinaison

de modalités dans le processus de communication instructeur-apprenant et d'extraire des recommandations pour la conception d'environnements d'apprentissage immersifs qui prennent en charge les interactions multimodales.

## 2 ÉTAT DE L'ART

### 2.1 Modèle de compagnonnage et modalités de communication

Dans le domaine médical, le compagnonnage est un modèle reconnu regroupant des cours théoriques, des observations et la pratique sous la supervision d'un pair qui transmet ses connaissances et son savoir-faire par le biais de la démonstration [12, 21, 49]. Ce modèle favorise les interactions entre les formateurs et les apprenants. Pour apprendre des gestes techniques, les apprenants commencent souvent par la phase d'observation, pendant laquelle l'instructeur montre à l'apprenant l'exécution de chaque partie de la procédure. Goffman définit la démonstration comme la "réalisation d'une activité ressemblant à une tâche en dehors de son contexte fonctionnel habituel, afin de permettre à quelqu'un qui n'est pas l'exécutant, d'obtenir une image précise de l'activité en cours" [19]. La démonstration est considérée comme une méthode efficace pour l'apprentissage de nouveaux gestes techniques complexes [13]. Pour démontrer la compétence, le formateur utilise des instructions ou des rétroactions comme moyen d'interaction. Il est important ici de faire la distinction entre (i) les instructions, définies comme des directives fournies par un formateur afin d'aider l'apprenant à intégrer les compétences nécessaires avant d'exécuter la tâche, et (ii) les rétroactions, définies comme les informations fournies aux apprenants pour guider et corriger leurs actions pendant ou après leur exécution de la tâche [5, 23]. La littérature confond souvent ces deux concepts car les deux impliquent des interactions entre l'apprenant et le formateur. Cependant, elles interviennent à des étapes différentes de la réalisation de la tâche, et leur impact sur l'apprentissage peut donc différer. À ce stade, notre étude se focalisera uniquement sur les instructions.

Pour fournir des instructions, le formateur utilise diverses modalités de communication : verbale, visuelle et/ou haptique [17, 36]. La modalité verbale offre des informations explicites sur les exigences de la tâche, les objectifs du mouvement et les stratégies d'exécution, tout en stimulant la motivation, l'auto-efficacité et l'engagement. Cependant, elle peut être insuffisante en raison des compétences motrices impliquant des éléments insaisissables tels que la dynamique des mouvements et les sensations haptiques, qui sont difficiles à articuler verbalement [40, 45]. La modalité visuelle, qui implique l'imitation suivant des démonstrations physiques, est largement utilisée dans le sport et la médecine. Elle engendre une meilleure communication de l'expertise lors de l'exécution d'actions [15, 45]. Néanmoins, les démonstrations visuelles manquent de détails sur les processus internes, tels que les schémas d'activation musculaire, les angles des articulations ou les forces appliquées. Enfin, la communication haptique consiste à guider physiquement les apprenants dans l'exécution du mouvement. Toutefois, son efficacité varie en fonction de la perception, de la sensibilité et de l'adaptation de chacun, pouvant potentiellement influencer l'interprétation et la performance [7, 8].

En résumé, le compagnonnage est un modèle efficace pour enseigner les compétences motrices et améliorer l'expérience d'apprentissage. Ce modèle repose sur une interaction "un à un" et utilise diverses modalités de communication, permettant au formateur de fournir des instructions et des rétroactions personnalisées à l'apprenant. Par conséquent, lors de la conception de systèmes prenant en compte les interactions formateur-apprenant, il est impératif de tenir compte des caractéristiques spécifiques de chaque modalité. Notre étude examine l'interaction entre l'apprenant et le formateur en RV lorsque le formateur utilise le processus de démonstration pour fournir des instructions. Notre objectif est de fournir aux concepteurs des recommandations en analysant comment la combinaison des modalités de communication impacte le modèle de compagnonnage dans un EIP.

## 2.2 Modalités d'interaction dans les Environnements Immersifs Partagés

Les Environnements Immersifs Partagés (EIP) sont des espaces 3D favorisant l'interaction entre plusieurs utilisateurs lors de réunions, sessions d'apprentissage, ou autres activités de groupe, même à distance [10]. Ils permettent des interactions en temps réel ou différé [10], avec des objets virtuels [25] et offrent la possibilité de partager des connaissances [18]. Ces systèmes semblent propices à faciliter les interactions formateur-apprenant au sein de simulations. Cependant, les limitations actuelles, notamment dans la communication entre utilisateurs [11], doivent être considérées lors de la conception d'outils de formation. Par exemple, reproduire fidèlement les expressions faciales, mouvements corporels, ou la direction du regard peut être difficile, et les participants peuvent ne pas partager le même point de vue, limitant la conscience des activités des partenaires [9]. Ces limitations peuvent entraver la communication et la transmission des compétences dans le contexte des interactions pédagogiques. Elles doivent par conséquent être prises en compte lors de la conception de ces systèmes. Récemment, une attention croissante s'est portée sur la conception des modalités d'interactions et de communication, leurs combinaisons, ainsi que leur intégration au sein des EIP. Ainsi, Wu et al. [53] ont montré que l'utilisation combinée de signaux d'orientation verbaux et visuels était plus efficace pour le téléguidage en RV que la condition verbale seule. De même, Gutwin et al. [20] constate que l'ajout d'informations verbales dynamiques renforce la collaboration entre participants lors de tâches de dessin collaboratif. D'autres études [1, 32, 43, 52] ont examiné l'impact de la modalité haptique, combinée aux modalités visuelles et/ou verbales, dans des contextes variés. Ainsi l'ajout d'un retour haptique améliore la performance et le sentiment de présence [43] lors d'une tâche collaborative de co-manipulation d'objets dans un EV. De leur côté, Moll et al. [32] ont étudié différentes combinaisons de modalités (visuelle/haptique, visuelle/audio, visuelle/haptique/audio) dans une tâche collaborative de balayage et de placement. Cette tâche impliquait une collaboration entre les participants afin de repérer 10 cellules vides et de les remplir en disposant des cubes qui tombaient d'en haut. Les résultats soulignent que la combinaison visuelle/haptique/audio améliore significativement les temps de réalisation des tâches, réduit les erreurs, et favorise une collaboration plus fluide. L'étude de Wang et al. [52] a montré que la combinaison des modalités verbale et haptique était plus efficace pour réaliser une tâche collaborative de pointage que la modalité haptique seule. Enfin, AlQassab and Wang [1] ont évalué diverses combinaisons de rétroaction (visuelle, auditive, haptique) dans des tâches de guidage simples, démontrant que le feedback auditif conduit généralement à des temps de réponse plus rapides, une précision accrue, et une réduction de la mentale. En résumé, ces recherches convergent en soulignant l'efficacité d'une combinaison judicieuse de plusieurs modalités pour améliorer la performance, la communication, et la collaboration dans des environnements collaboratifs. Les résultats suggèrent également une réduction de la charge mentale lorsque les instructions sont présentées avec certaines combinaisons de modalités.

## 2.3 Technologies supportant les interactions formateur-apprenant

L'utilisation des EIP pour le compagnonnage et l'apprentissage de nouvelles compétences suscite de plus en plus d'attention, car des études soulignent la nécessité d'interaction entre les instructeurs et les apprenants au sein des EIP pour favoriser l'apprentissage, notamment en médecine [35]. Dans cette perspective, Cecilio-Fernandes et al. [5] ont montré les effets de différentes sources de rétroaction sur l'acquisition et la rétention d'une compétence médicale de manipulation de sonde échographique. Les résultats montrent que la présence de l'expert a contribué à réduire la charge mentale pendant la pratique [5]. Dans une approche visant à explorer l'effet des retours d'information dans des environnements virtuels de formation, Strandbygaard et al. [48] ont examiné l'influence des rétroactions du formateur à l'intérieur d'un simulateur en RV pour une procédure de laparoscopie. Les résultats indiquent que les participants

qui ont reçu les rétroactions (verbales et visuelles) de l'instructeur pendant leur formation ont acquis la maîtrise de la procédure plus rapidement que ceux qui n'en ont pas reçu. De manière similaire, Porte et al. [38] ont comparé les rétroactions verbales provenant d'un expert et les rétroactions générées par un système automatique, pour améliorer l'efficacité des mouvements lors de l'apprentissage d'un geste de suture en chirurgie. Les résultats montrent que les rétroactions verbales d'un expert ont conduit à une amélioration durable des performances [38]. D'autres études [8, 27] ont également exploré l'impact de la modalité haptique, combinée avec les modalités visuelles et verbales sur l'apprentissage. Chellali et al. [8] ont montré que l'intégration d'instructions haptiques a significativement amélioré les performances des apprenants dans un geste de manipulation d'aiguille par rapport aux conditions n'incluant que des instructions visuelles et verbales. Des résultats similaires ont été observés dans d'autres études portant sur l'apprentissage de tâches plus basiques. Ainsi, l'étude de Lu et al. [27] a confirmé que la combinaison de retours haptiques et visuels a amélioré les performances pour une tâche de déplacement d'une sphère dans un tube par rapport aux retours visuels seuls. Plus récemment, une étude a examiné les effets de chaque modalité d'instruction (visuelle, haptique et verbale) sur l'enseignement des compétences basiques de manipulation d'un outil pour le déplacement d'une sphère en 3D [46]. Les résultats montrent que les temps de réalisation de la tâche étaient plus rapides avec les instructions haptiques, mais les apprenants étaient plus précis avec les instructions visuelles. De plus, la modalité verbale a amélioré le sentiment de coprésence avec le formateur [46]. Ainsi, ces précédentes études suggèrent que la modalité visuelle et la modalité haptique sont plus efficaces que la modalité verbale pour améliorer les performances, mais la modalité verbale reste importante pour renforcer le sentiment de coprésence.

En résumé, diverses études soulignent les avantages et inconvénients de chaque mode de communication, mettant en évidence que l'utilisation judicieuse de ces modes favorise l'acquisition de compétences motrices et améliore l'expérience utilisateur dans les EIP. Cependant, plusieurs méta-revues, dont Sigrist et al. [45], Philippe et al. [37], et Martin et al. [29], soulignent la nécessité d'approfondir la compréhension de l'impact de la communication multimodale en RV pour améliorer l'expérience utilisateur et faciliter l'apprentissage des compétences motrices, aussi bien dans des tâches simples que complexes. Ainsi, bien que les EIP montrent un potentiel pour améliorer les interactions instructeur-apprenant et l'apprentissage par compagnonnage, il est essentiel d'évaluer leurs caractéristiques pour éviter les ruptures de communication et les défis d'apprentissage. Les études sur les interactions multimodales indiquent que la combinaison de plusieurs modes de communication peut favoriser la collaboration, l'acquisition des compétences motrices, et la baisse de la charge mentale. Cependant, leur utilisation dépend des compétences à enseigner [8] et de leur disponibilité [42]. Par conséquent, des recherches supplémentaires sont nécessaires pour comprendre de manière plus exhaustive comment les interactions entre ces modalités influent sur l'acquisition des compétences motrices.

Le présent travail vise à contribuer à cette discussion en fournissant des informations sur l'utilisation de combinaisons de modalités (visuelle-verbale, verbale-haptique, visuelle-haptique, verbale-visuelle-haptique) au sein des EIP pour enseigner aux apprenants des compétences motrices de base. Afin de mieux isoler l'impact de la communication multimodale sur l'apprentissage de ces compétences, notre choix s'est porté sur une tâche simple inspirée des tâches de saisie-transfert régulièrement utilisées pour l'apprentissage des gestes moteurs en RV [4, 41].

### 3 ÉTUDE EXPÉRIMENTALE

#### 3.1 Objectifs de l'étude et hypothèses

Notre étude vise à examiner l'impact de la communication multimodale entre un instructeur et un apprenant dans un EIP. Nous évaluons l'impact de différentes combinaisons de modalités pour la transmission d'instructions d'amplitude de

mouvements lors d'une tâche de manipulation d'outil sur les performances des apprenants et leur expérience utilisateur. Nous utilisons les modalités : verbales, visuelles et haptiques, ce qui conduit à trois combinaisons bi-modales différentes et une combinaison tri-modale : verbale-visuelle, verbale-haptique, visuelle-haptique et verbale-visuelle-haptique. La tâche consiste à reproduire fidèlement un mouvement d'outil en fonction des instructions transmises par le formateur selon l'une des combinaisons de modalités. Ces combinaisons sont comparées en fonction de leur impact sur les performances des participants (vitesse et précision), leur charge mentale et leur expérience subjective.

Notre hypothèse principale est que les quatre conditions de communication affecteront différemment les performances des participants, la charge mentale et l'expérience utilisateur. Plus précisément, nous nous attendons à ce que :

- **H1** : La combinaison des modalités visuelles et haptiques améliorera les performances des apprenants par rapport aux conditions où une seule d'entre elles est combinée avec la modalité verbale.
- **H2** : La charge mentale diminue lorsque les instructions sont fournies avec certaines combinaisons de modalités.
- **H3** : Les conditions incluant la modalité verbale améliorent l'expérience utilisateur mesurée à l'aide de questionnaires sur la présence et la coprésence.

### 3.2 Participants

Trente-deux participants ont pris part à cette étude, (6 femmes et 26 hommes), recrutés parmi les étudiants, le personnel universitaire et des participants externes. Aucun critère spécifique d'expertise ou d'expérience préalable avec la RV n'était requis. L'âge moyen des participants était de 25,5 ans (min = 20, max = 47), et tous étaient droitiers. Tous les participants avaient une vision normale ou corrigée, avec 13 d'entre eux portant des lunettes correctrices pendant l'expérience. Vingt-six d'entre eux avaient déjà eu une expérience avec des casques de RV, dont six en étaient des utilisateurs réguliers (une fois par semaine). Seize d'entre eux avaient déjà eu une expérience avec les dispositifs haptiques pendant des démonstrations ou d'autres études expérimentales. Le Comité d'éthique de la recherche de l'Université Paris Saclay a validé le protocole expérimental et tous les participants ont signé un formulaire de consentement éclairé avant d'être inclus dans l'étude.

### 3.3 Tâche expérimentale

Nous avons défini une tâche spécifique de saisie-transfert pour cette étude. La tâche visait à reproduire la manipulation d'un outil après la démonstration des mouvements par l'instructeur. Pour simplifier la tâche, nous nous sommes uniquement concentrés sur la direction et l'amplitude des mouvements en 2D. Cette simplification était nécessaire pour limiter les biais inhérents à une tâche complexe et dont les instructions auraient été complexes à produire à travers les trois modalités. Ainsi, la tâche consistait à déplacer une sphère 3D de sa position initiale à une position finale le long d'un seul axe (l'axe X ou l'axe Y). L'expérience était structurée en deux phases :

- (1) Phase d'instruction : pendant cette phase, l'instructeur devait communiquer au participant la direction et l'amplitude du mouvement. Cette instruction était transmise selon l'une des quatre conditions de communication. Les participants étaient informés que les instructions étaient fournies en temps réel par un instructeur distant se trouvant dans une autre pièce, bien qu'en réalité, celles-ci étaient pré-enregistrées (figure 1). Ceci a permis de contrôler les éventuels biais induits par des instructions variant pour chaque participant [46]. Le choix de dissimuler l'absence de l'instructeur est justifié par des recherches antérieures indiquant que le degré selon lequel une entité virtuelle est perçue comme une représentation d'une personne réelle (agentivité perçue) influence l'expérience utilisateur et la présence sociale dans un environnement virtuel [16, 34].



FIG. 2 – Équipement : il comprenait un casque de RV, une interface haptique (équipé d'un styilet), et un Vive Tracker pour l'étalonnage. Environnement virtuel : il est constitué d'un plan jaune et rouge. La scène du milieu montre l'étape avant de commencer un essai. Les participants peuvent voir les prochaines modalités qui seront utilisées. On observe également la capsule verte qui permet de démarrer l'essai. La scène de droite montre la phase d'instruction où les participants voient une sphère bleue claire et la main virtuelle de l'instructeur seulement lorsque la modalité visuelle est impliquée.

- (2) Phase de manipulation : dans cette phase, les participants devaient reproduire les mouvements décrits d'une manière autonome.

### 3.4 Dispositif expérimental

**3.4.1 Équipement.** L'équipement (voir figure 2) comprenait un casque de RV HTC Vive Pro pour visualiser l'environnement immersif. Ce casque a une résolution de  $2880 \times 1600$  pixels, avec un taux de rafraîchissement de 90 Hz et un champ de vision de 110 degrés. De plus, un dispositif haptique Geomagic Touch, avec un espace de travail de  $160 \text{ mm} \times 120 \text{ mm} \times 70 \text{ mm}$ , était utilisé pour manipuler les objets virtuels et recevoir les instructions haptiques. Il est doté d'un bouton que les participants actionnaient pour saisir et relâcher une sphère virtuelle. Un tracker Vive 3.0 (fixé à la table) était également utilisé pour calibrer le système et garantir la correspondance entre le monde physique et virtuel.

**3.4.2 Environnement virtuel.** L'environnement virtuel a été créé à l'aide de Unity 3D (version 2020.3.15f2) et le plugin SteamVR (version 1.17.3). L'environnement virtuel comprenait deux plans perpendiculaires, mesurant  $44 \text{ cm} \times 15 \text{ cm}$ , colorés en jaune (horizontal) et en rouge (vertical) (voir figure 2), respectivement. Pendant la phase d'instruction, ces deux plans étaient affichés pour toutes les conditions de communication, à l'exception de la condition verbale-haptique (où un écran gris neutre était affiché). L'environnement comprenait également un outil virtuel, une sphère bleue, et une représentation de la main de l'instructeur (voir figure 2 droite). Pendant la phase de manipulation, l'environnement incluait également une sphère bleue à déplacer, un outil contrôlé par le participant via le bras haptique, et une représentation virtuelle de sa main. Pour améliorer la perception de la distance dans l'environnement immersif, la main virtuelle était attachée à l'outil et positionnée pour s'aligner avec la main réelle du participant. Cependant, elle n'était pas animée. De plus, une planche en bois, de la même taille que le plan horizontal jaune virtuel, était placée sur la table. L'outil virtuel correspondait également en couleur, taille et forme au styilet du bras haptique et était positionné pour s'aligner avec lui. Enfin, le dispositif haptique était placé de manière que lorsque le participant touchait le plan virtuel avec l'extrémité de l'outil virtuel, il ressentait la collision entre la planche en bois et le styilet du bras haptique. Par conséquent, l'étape de calibration était nécessaire avant le début de chaque session expérimentale.



**3.4.3 Scènes virtuelles et interactions.** Une série de scènes virtuelles était présentée aux participants en fonction de la phase expérimentale. Dans la scène initiale, un message expliquant comment les instructions allaient être données aux participants était affiché (selon la condition expérimentale). Ensuite, les participants étaient placés dans une scène de transition contenant une capsule 3D et il leur était demandé s'ils étaient prêts à commencer. Ils devaient déplacer le pointeur 3D à l'intérieur de la capsule à l'aide du bras haptique et appuyer sur le bouton du dispositif pour lancer la phase d'instruction (voir figure 2 milieu). Seuls les plans bicolores étaient affichés pendant cette phase (sauf pour la condition verbale-haptique), accompagnés d'un message de synchronisation pour informer les participants que les instructions seraient bientôt fournies. Après chaque instruction, une scène de transition avec deux capsules 3D était affichée, proposant aux participants de répéter l'instruction (une seule fois) ou de passer à la phase de manipulation. Pendant la phase de manipulation, les participants devaient utiliser le bras haptique pour saisir la sphère (en appuyant sur le bouton) et la déplacer conformément à l'instruction reçue jusqu'à la position souhaitée. Ils devaient ensuite appuyer à nouveau sur le bouton du stylet pour relâcher la sphère, indiquant la fin de l'essai. Une nouvelle scène leur permettait alors de passer à l'essai suivant. À la fin de tous les essais dans une condition, une scène de transition apparaissait, les invitant à retirer le casque pour répondre aux questionnaires.

### 3.5 Plan expérimental et conditions

L'expérience a suivi un plan intra-sujet à un seul facteur (la combinaison des modalités de communication) avec quatre conditions : verbale-visuelle, verbale-haptique, visuelle-haptique et verbale-visuelle-haptique. L'ordre de présentation des conditions aux participants a été contrebalancé en utilisant un carré latin pour atténuer les effets d'apprentissage potentiels. Chaque participant a réalisé 14 essais pour chaque condition, avec des mouvements sur les axes X et Y (gauche/droite; haut/bas). Pour garantir que les mouvements restaient dans les limites de l'espace de travail du bras haptique, les amplitudes des mouvements variaient entre 4 et 8 centimètres et étaient sélectionnées de manière aléatoire pour chaque essai tout en étant équilibrées entre les conditions. Cela a abouti à un total de 1792 essais enregistrés (14 essais x 4 conditions x 32 participants). Trois formes d'instructions ont été utilisées :

- **Instructions Verbales** : les instructions décrivant l'amplitude et la direction du mouvement étaient données verbalement aux participants. Par exemple, l'instruction pour un mouvement de 8 centimètres vers la gauche était : "Veuillez déplacer la sphère de 8 cm vers la gauche". Toutes les instructions étaient pré-enregistrées sous forme de clips audio utilisant la voix de l'expérimentateur et étaient diffusées sur les écouteurs du casque de RV.
- **Instructions Visuelles** : les instructions étaient transmises par le biais d'une animation 3D de la main de l'instructeur manipulant l'outil dans l'environnement immersif. Suivant une courbe de Bézier, la main virtuelle se déplaçait pour saisir la sphère et la placer dans une position finale variable pour chaque essai.
- **Instructions Haptiques** : les instructions étaient transmises par le biais du bras haptique saisi par les participants. Le bras se déplaçait de la position de départ à la position finale, suivant les mêmes courbes que la condition visuelle, imitant l'exécution de la tâche par l'instructeur à l'aide du dispositif haptique.

Ces trois formes de communication ont ensuite été combinées pour créer les quatre conditions expérimentales où les instructions étaient jouées simultanément, transmettant la même direction et la même amplitude.

Il convient de noter que pendant les instructions haptiques et visuelles, la sphère 3D et la position du stylet haptique étaient placées à des positions de départ différentes pendant les phases d'instruction et de manipulation pour empêcher la mémorisation des positions finales. De cette manière, les participants étaient encouragés à mémoriser l'amplitude et la direction du mouvement plutôt que la position finale de la sphère ou du stylet.

### 3.6 Procédure expérimentale

La durée moyenne de l'expérience était de 75 minutes par participant. Préalablement à leur entrée dans la salle d'expérimentation, les participants ont rencontré un "faux" instructeur dans une pièce adjacente. Il leur a alors été expliqué que cette personne leur donnerait les instructions à distance pendant l'expérience. Par la suite, les participants accédaient à la salle d'expérimentation, où la première étape consistait à leur présenter l'objectif de l'étude et l'équipement qu'ils allaient utiliser. Puis, il leur a été demandé de lire et de signer un formulaire de consentement pour participer à l'étude. Enfin, ils ont reçu une fiche d'instruction détaillant le fonctionnement du prototype, les actions à effectuer et ce que l'on attendait d'eux. L'étape suivante consistait à remplir un questionnaire démographique, après quoi les participants étaient invités à enfiler le casque immersif pour commencer avec la première condition. Une fois les 14 essais de la condition terminés, les participants ont été invités à retirer le casque pour répondre au questionnaire NASA-TLX et au questionnaire évaluant leur sentiment de présence, de présence sociale, de coprésence et d'expérience d'apprentissage au cours de la condition précédente. Une fois cette étape terminée, ils devaient remettre le casque pour commencer les essais de la condition suivante et répéter le même cycle décrit précédemment. Après avoir terminé les essais et répondu à tous les questionnaires pour toutes les conditions, ils ont été invités à répondre à un questionnaire comparant les quatre conditions sur la base de différents critères.

### 3.7 Mesures et analyses des données

Dans cette expérience, des mesures objectives et subjectives étaient utilisées. Les mesures objectives évaluaient la capacité des participants à reproduire le mouvement demandé et comprenaient trois métriques : le temps moyen de manipulation (de la saisie au positionnement de la sphère), l'erreur moyenne d'estimation de distance (la distance euclidienne moyenne, en centimètres, entre la position finale du centre de la sphère et la position souhaitée), et la qualité de la trajectoire de l'outil basé sur la Déformation Temporelle Dynamique (DTD) [54] évaluant la similitude entre les trajectoires présentées par l'instructeur et celles réalisées par les participants. La DTD est un algorithme conçu pour identifier l'alignement optimal entre deux trajectoires. Il s'avère plus efficace pour évaluer la similitude entre deux trajectoires que les méthodes conventionnelles telles que la mesure de la distance euclidienne point à point. Plus la distance DTD est faible, plus les deux trajectoires sont similaires. Nous avons donc calculé la distance DTD (en utilisant Python et le module `DTAIDistance`<sup>1</sup>) entre la trajectoire du participant et leur trajectoire de référence respective pour chaque participant, chaque essai et pour chaque condition. Nous avons ensuite comparé les distances moyennes de chaque condition. Pour ces trois métriques, des valeurs plus basses indiquent de meilleures performances.

Les mesures subjectives incluaient les réponses des participants à divers questionnaires. Cela comprenait un questionnaire suivant une échelle de Likert à sept points, qui évaluait le sentiment de présence, de présence sociale, de coprésence et l'expérience d'apprentissage des participants (tableau 1). Enfin, l'expérience d'apprentissage a été mesurée à l'aide d'une question extraite du questionnaire de Simon et al. [46].

Le questionnaire NASA TLX [22] a évalué la charge mentale des participants lors de l'exécution de la tâche. Le choix de ce questionnaire se justifie par sa validité, son applicabilité et sa popularité dans des recherches similaires [26, 39, 46, 50]. Enfin, les participants ont répondu à un questionnaire où ils devaient classer les combinaisons de modalités en fonction des mêmes onze critères utilisés dans [46], de la plus appréciée à la moins appréciée.

Le logiciel SPSS a été utilisé pour l'analyse statistique des données. Ces analyses ont été effectuées avec un niveau de confiance de 95%, et dans les cas où des corrections ont été appliquées, nous avons rapporté les valeurs de  $p$  ajustées.

---

1. <https://pypi.org/project/dtaidistance/>

TABLE 1 – Éléments du questionnaire subjectif sur la présence, la présence sociale, la coprésence et l'expérience d'apprentissage. Les échelles de notation vont de 1 à 7.

Q#	Question
<b>Présence [33]</b>	
Q1	Dans quelle mesure l'expérience a-t-elle été immersive ?
Q2	Dans quelle mesure l'expérience a-t-elle été intense ?
Q3	Dans quelle mesure vous êtes vous senti à l'intérieur de l'environnement que vous avez vu/entendu ?
Q4	Dans quelle mesure vous êtes-vous senti immergé dans l'environnement que vous avez vu/entendu ?
Q5	Dans quelle mesure avez-vous eu l'impression d'être entouré par l'environnement que vous avez vu/entendu ?
<b>Présence sociale [33]</b>	
Q6	Dans quelle mesure cela ressemblait-il à une réunion en face à face ?
Q7	Dans quelle mesure aviez-vous l'impression d'être dans la même pièce que l'instructeur ?
<b>Coprésence [33] [3]</b>	
Q8	L'instructeur était-il intensément impliqué dans notre interaction ?
Q9	L'instructeur a-t-il communiqué de la froideur plutôt que de la chaleur ?
Q10	Dans quelle mesure vous êtes-vous senti isolé de l'instructeur dans l'environnement virtuel ?
Q11	Dans quelle mesure aviez-vous le sentiment d'être avec l'instructeur ?
Q12	Dans quelle mesure y a-t-il eu des moments où l'interface semblait disparaître, et vous travailliez directement avec l'instructeur ?
Q13	Dans quelle mesure avez-vous oublié l'instructeur et vous êtes vous concentré uniquement sur la tâche comme si vous étiez la seule personne impliquée ?
Q14	Dans quelle mesure étiez-vous et l'instructeur en harmonie lors de l'exécution de la tâche ?
Q15	Globalement, évaluez dans quelle mesure vous aviez le sentiment qu'il y avait un instructeur qui interagissait avec vous, plutôt qu'une simple machine ?
<b>Expérience d'apprentissage [46]</b>	
Q16	Dans quelle mesure pensez-vous pouvoir acquérir de nouvelles compétences avec cette application ?

## 4 RÉSULTATS

### 4.1 Mesures objectives

Pour chaque participant, nous avons calculé la moyenne de chaque mesure objective par condition, ce qui a donné 128 valeurs par mesure (4 conditions x 32 participants). Nous avons utilisé le test de Shapiro-Wilk pour vérifier la distribution normale des données. Les résultats indiquent que toutes les données d'erreurs d'estimation de distance et des temps d'achèvement étaient normalement distribuées. D'un autre côté, seules les données de distance DTD pour les conditions Visuelle-Haptique et Verbale-Visuelle-Haptique étaient normalement distribuées.

*4.1.1 Erreur d'estimation de distance.* Suite aux résultats des tests de normalité, une analyse de variance (ANOVA) à mesures répétées à un facteur a été utilisée pour tester l'effet des modalités. Les résultats (sphéricité vérifiée,  $p = .21$ ) montrent un effet principal significatif de la combinaison de modalités sur les erreurs d'estimation de distance ( $F_{(3,93)} = 10.56$ ,  $p < .001$ ,  $\eta^2$  partiel = .25; (figure 3, 1er graphe). Les comparaisons post-hoc avec correction de Bonferroni montrent que les erreurs moyennes étaient significativement plus faibles dans la condition Verbale-Visuelle-Haptique que dans les conditions Verbale-Visuelle ( $p = .001$ ) et Verbale-Haptique ( $p = .006$ ). Les erreurs moyennes étaient également significativement plus faibles dans la condition Visuelle-Haptique que dans les conditions Verbale-Visuelle ( $p = .008$ ) et Verbale-Haptique ( $p = .002$ ). Aucune autre différence significative n'a été observée.

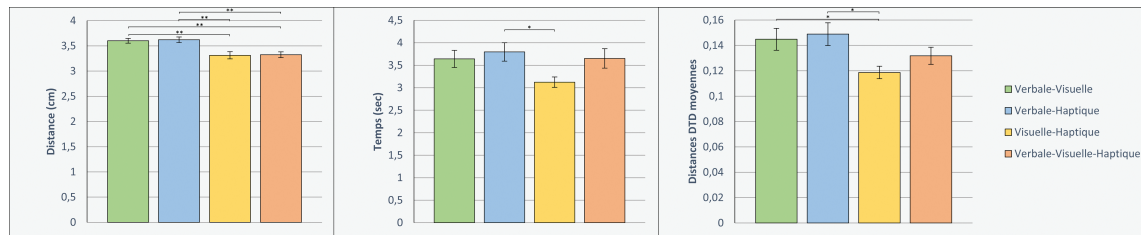


FIG. 3 – 1er graphe : Erreurs moyennes d'estimation de l'amplitude. 2ème graphe : Temps de manipulation moyens. 3ème graphe : Distance DTW moyenne. Les barres d'erreur représentent l'erreur standard (\*  $p < 0.05$ , \*\*  $p < 0.01$ ).

**4.1.2 Temps de manipulation.** Suite aux résultats des tests de normalité, une ANOVA à mesures répétées à un facteur a été utilisée pour tester l'effet des modalités. Les résultats (sphéricité vérifiée,  $p = .21$ ) montrent un effet principal significatif de la combinaison de modalités sur les temps de manipulation moyens ( $F_{(3,93)} = 3.50$ ,  $p < .019$ ,  $\eta^2$  partiel = .10; (figure 3, 2ème graphe). Les comparaisons post-hoc avec correction de Bonferroni montrent que les temps de manipulation moyens étaient significativement plus courts dans la condition Visuelle-Haptique que dans la condition Verbale-Haptique ( $p = .017$ ). Aucune autre différence significative n'a été observée.

**4.1.3 Trajectoires.** Suite aux tests de normalité, un test non-paramétrique de Friedman a été utilisé pour évaluer l'effet des modalités. Les résultats montrent un effet principal de la combinaison de modalités sur les distances moyennes DTD entre les trajectoires des participants et celles de l'instructeur ( $\chi^2 = 7.91$ ,  $p = .04$ ; (figure 3, 3ème graphe). Les tests par paires des rangs signés de Wilcoxon avec correction de Bonferroni montrent que les trajectoires des participants étaient significativement plus proches des trajectoires de référence dans la condition Visuelle-Haptique que dans les conditions Verbale-Visuelle ( $p = .02$ ) et Verbale-Haptique ( $p = .01$ ). Aucune autre différence significative n'a été observée.

## 4.2 Mesures subjectives

**4.2.1 Charge mentale perçue.** Les moyennes des scores bruts du NASA-TLX étaient de  $32.94 \pm 12.41$ ,  $34.66 \pm 13.56$ ,  $33.88 \pm 12.22$ ,  $34.34 \pm 13.13$ , respectivement pour les conditions verbale-visuelle, verbale-haptique, visuelle-haptique et verbale-visuelle-haptique. Le test de Friedman n'indique aucun effet principal significatif de la combinaison de modalités sur ces scores ( $\chi^2 = 3.94$ ,  $p = .26$ ).

**4.2.2 Questionnaire subjectif.** Chaque dimension du questionnaire a été étudiée séparément en calculant le score moyen des questions associées (figure 4). Les tests de Friedman montrent un effet principal des combinaisons de modalités sur les scores moyens pour la présence ( $\chi^2 = 15.33$ ,  $p = .002$ ), la présence sociale ( $\chi^2 = 10.96$ ,  $p = .012$ ), la coprésence ( $\chi^2 = 13.97$ ,  $p = .003$ ) et l'apprentissage ( $\chi^2 = 11.63$ ,  $p = .009$ ).

Le sentiment de présence des participants était significativement plus élevé dans la condition Verbale-Visuelle-Haptique que dans la condition Verbale-Haptique ( $p = .01$ ). Le score moyen était également plus élevé que dans les conditions Verbale-Visuelle ( $p = .07$ ) et Visuelle-Haptique ( $p = .07$ ), mais l'effet était marginal. Aucune autre différence significative n'a été trouvée.

Le sentiment de coprésence des participants était significativement plus élevé dans la condition Verbale-Visuelle-Haptique que dans les conditions Verbale-Visuelle ( $p = .03$ ) et Visuelle-Haptique ( $p = .006$ ). Le score moyen était également plus élevé dans la condition Verbale-Haptique que dans la condition Visuelle-Haptique ( $p = .06$ ), mais l'effet était marginal. Aucune autre différence significative n'a été trouvée.

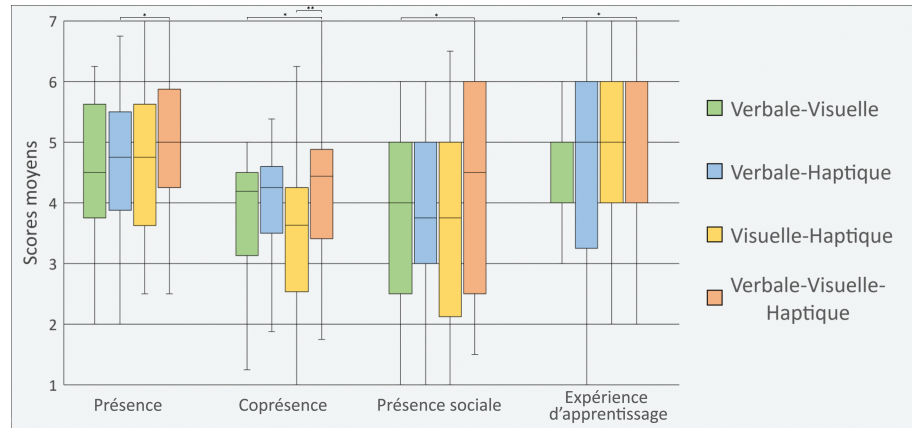


FIG. 4 – Scores moyens pour les quatre dimensions du questionnaire subjectif (\*  $p < 0.05$ , \*\*  $p < 0.01$ ).

TABLE 2 – Tests de corrélation de Spearman entre les quatre dimensions du questionnaire subjectif (\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ )

		Pr-CoPr	Pr-PrSo	Pr-ExAp	CoPr-PrSo	CoPr-ExAp	PrSo-ExAp
Verbale-Visuelle	$r(30)$	.36*	.40*	.41*	.70***	.32	.34
Verbale-Haptique	$r(30)$	.55**	.32	.53**	.49**	.59***	.58***
Visuelle-Haptique	$r(30)$	.35*	.32	.47**	.81***	.42*	.36*
Verbale-Visuelle-Haptique	$r(30)$	.41*	.39*	.54**	.76***	.39*	.33

Le sentiment de présence sociale des participants était significativement plus élevé dans la condition Verbale-Visuelle-Haptique que dans la condition Verbale-Visuelle ( $p = .03$ ). Aucune autre différence significative n'a été trouvée.

Enfin, les participants ont estimé avoir eu une meilleure expérience d'apprentissage dans la condition Verbale-Visuelle-Haptique que dans la condition Verbale-Visuelle ( $p = .04$ ). Aucune autre différence significative n'a été trouvée.

Des tests de corrélation de Spearman ont été réalisés pour étudier les relations entre ces dimensions. Les résultats sont affichés dans le tableau 2. A noter que toutes les corrélations significatives sont positives.

**4.2.3 Comparaison subjective.** Pour la comparaison subjective entre les conditions, chaque dimension a été examinée séparément en comparant le classement de chaque condition par rapport aux autres (figure 5). Les tests de Friedman (tableau 3) ont été utilisés pour évaluer l'effet des combinaisons de modalités sur ces classements pour chaque dimension.

Les comparaisons significatives par paires en utilisant les tests des rangs signés de Wilcoxon avec correction de Bonferroni pour chaque dimension sont rapportées ci-après.

Les participants ont classé la condition Verbale-Visuelle-Haptique comme la plus facile à comprendre (Q1), la plus appropriée (Q2), la plus précise (Q3) la plus agréable (Q4), et la plus efficace (Q10) pour recevoir des instructions ( $p < .01$  pour toutes les comparaisons). De plus, elle a été classée comme la plus facile pour mémoriser le mouvement (Q6), et pour reproduire le mouvement (Q7), comme la forme de communication la plus éducative (Q8) et la plus engageante (Q9) ( $p < .01$  pour toutes les comparaisons). Enfin, elle a été classée comme la méthode préférée de communication (Q11) par rapport aux autres conditions ( $p < .01$  pour toutes les comparaisons). D'autre part, elle a été classée comme significativement moins perturbante que la condition Visuelle-Haptique ( $p = .03$ ) et marginalement moins perturbante que la condition Verbale-Haptique ( $p = .09$ ). Aucune autre différence significative n'a été observée.

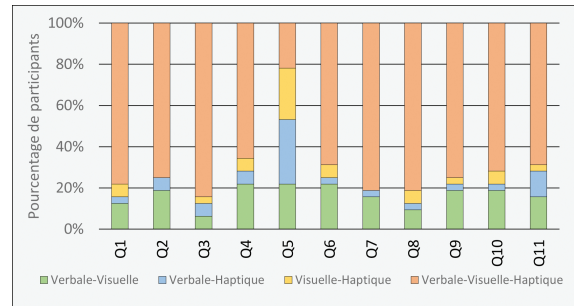


FIG. 5 – Pourcentage de participants ayant choisi chaque combinaison de modalités comme mode préféré par critère (Q1 : facile à comprendre; Q2 : adapté; Q3 : précis; Q4 : agréable; Q5 : déstabilisant; Q6 : facile à mémoriser; Q7 : facile à reproduire; Q8 : facilitant l'apprentissage; Q9 : engageant; Q10 : efficace; Q11 : mode préféré)

TABLE 3 – Tests de Friedman pour le questionnaire de comparaison (\*\*\*)= $p < .001$ )

Q#	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
$\chi^2$	28.27***	33.45***	25.42***	12.67***	30.93***	32.39***	32.39***	33.82***	27.37***	29.06***	24.56***

## 5 DISCUSSION

Cette recherche a examiné l'impact de différentes combinaisons de modalités de communication sur l'enseignement des mouvements de manipulation d'outils dans un EIP. L'étude a produit de nombreux résultats significatifs.

### 5.1 Performance

L'erreur d'estimation de la distance était le premier facteur pour évaluer la performance des apprenants après avoir reçu les instructions. Les résultats montrent que les combinaisons visuelle-haptique et verbale-visuelle-haptique sont les plus précises pour transmettre l'amplitude du mouvement, entraînant une diminution significative des erreurs lors de l'estimation de l'amplitude. Les résultats des travaux précédents [46], avaient suggéré que la modalité visuelle était la plus précise pour communiquer l'amplitude du mouvement alors que la modalité verbale était la moins précise des trois. Ainsi, il n'est pas surprenant que la combinaison visuelle-haptique se soit révélée comme la plus efficace dans l'étude actuelle. De plus, les instructions transmises avec les modalités haptique et visuelle sont similaires. En effet, elles fournissent des informations sur la forme de la trajectoire plutôt qu'une simple distance absolue (dans le cas des instructions verbales). La "règle spatio-temporelle" stipule que les stimulus présentés en proximité spatio-temporelle ont une probabilité plus élevée d'être combinés pour former une perception d'un événement physique [31]. Dans notre cerveau, les réponses des neurones multisensoriels (neurones intégrant des données provenant de différentes modalités) augmentent en cas de concordance spatio-temporelle et diminuent autrement [51]. Les instructions visuelles et haptiques sont davantage liées sur le plan spatio-temporel que les instructions verbales. La similarité spatio-temporelle entre les instructions visuelles et haptiques fournit une autre explication à l'amélioration de la précision obtenue par leur combinaison. La précision similaire obtenue avec la combinaison des trois modalités suggère que les informations reçues par la modalité verbale étaient moins utilisées par les participants pour reproduire la trajectoire que celles reçues par les deux autres modalités. Cette observation pourrait être attribuée au fait que les instructions verbales peuvent être soit excessives, soit insuffisantes, ne fournissant pas à l'apprenant les informations nécessaires pour accomplir le geste moteur [28]. Cela correspond également à des recherches antérieures indiquant que l'incorporation d'instructions

haptiques conduit à une amélioration de la performance des participants par rapport aux conditions n'incluant que des instructions visuelles et verbales [8]. D'un autre côté, lorsque les instructions verbales étaient combinées avec seulement une autre modalité, les participants étaient moins précis. Cela suggère que les instructions verbales pourraient interférer avec les informations reçues à travers l'autre modalité avec laquelle elle est combinée (visuelle ou haptique).

Le deuxième facteur pour évaluer la performance était le temps de réalisation du mouvement. Les résultats ont révélé que la reproduction du mouvement est plus rapide avec la combinaison visuelle-haptique. Dans les recherches antérieures [46], les instructions haptiques avaient entraîné un temps de réalisation du mouvement plus court. Il a également été constaté que les instructions visuelles réduisaient les temps de réalisation par rapport aux instructions verbales. Dans notre expérience, toutes les conditions incluant des instructions verbales ont augmenté les temps de réalisation. Ceci est en cohérence avec l'étude de Vitense et al. [50] montrant que l'introduction de feedback auditif peut captiver l'attention, entraînant ainsi une diminution de la performance lorsqu'il est associé à des retours visuels ou haptiques.

Les trajectoires du mouvement, le troisième critère d'évaluation, confirment davantage cette hypothèse. Effectivement, les résultats ont montré que les trajectoires des participants étaient plus proches de celles de référence avec la combinaison des modalités visuelle et haptique qu'avec les autres combinaisons. Comme mentionné précédemment, les instructions verbales ne fournissent pas d'informations sur les trajectoires des outils. Par conséquent, lorsque ces instructions étaient fournies (pendant trois des conditions expérimentales), les participants semblaient s'être davantage concentrés sur l'atteinte de la cible et moins sur la reproduction des trajectoires fournies par les autres modalités (visuelle ou haptique). D'autre part, lorsque les instructions verbales n'étaient pas fournies (condition visuelle-haptique), les participants ont tenté de reproduire la trajectoire du mouvement tout en atteignant la cible. Cela conduit à des trajectoires plus optimisées et à des mouvements plus précis, qui étaient également exécutés plus rapidement. Ces résultats valident notre hypothèse H1, suggérant que la combinaison d'instructions visuelles et haptiques conduit à la meilleure performance.

## 5.2 Expérience utilisateur

L'évaluation de l'expérience utilisateur impliquait trois mesures. La première mesure évaluait la charge mentale perçue par les participants. La deuxième mesure consistait en une évaluation subjective des interactions avec l'instructeur, comprenant le sentiment de présence, de coprésence, de présence sociale et l'expérience d'apprentissage. Enfin, la troisième mesure était axée sur la comparaison des quatre combinaisons de modalités.

En ce qui concerne la charge mentale, des recherches antérieures [45] ont suggéré que l'apprentissage multimodale réduit la charge mentale en raison de la distribution du traitement de l'information. Les résultats du NASA-TLX n'ont pas révélé de différences significatives dans les scores entre les combinaisons de modalités, ce qui suggère que les participants n'ont pas perçu de variation de charge mentale. Une méta-analyse récente comprenant 556 études a trouvé que le score moyen brut du NASA-TLX était de 42, tandis que le score moyen brut dans 72 études basées sur la RV (principalement des applications éducatives et de santé) était de 41 [24]. Sur la base d'instructions unimodales, le travail le plus proche de notre étude actuelle [46] a également signalé des valeurs similaires (une moyenne de score brut de 42). La moyenne des scores bruts du NASA-TLX dans notre étude (33.95) est nettement inférieure à ces valeurs rapportées. Bien qu'une comparaison systématique entre ces valeurs ne soit pas possible, nos résultats suggèrent que les instructions multimodales, quelles que soient leurs combinaisons, peuvent conduire à une diminution de la charge mentale. La différence non significative entre les combinaisons de modalités contredit quelque peu les résultats des performances des participants. Une explication possible pourrait être la nature simple des instructions fournies (une seule direction et une seule amplitude), qui n'ont pas nécessité un effort mental important pour être traitées. À noter également la grande

variabilité observée entre les participants, suggérant des différences inter-individuelles assez importantes et donc des profils différents. Des investigations supplémentaires utilisant des instructions plus complexes combinées avec des mesures plus objectives, telles que le suivi oculaire pour mesurer l'attention, seront nécessaires pour mieux comprendre l'impact potentiel des combinaisons de modalités sur la charge mentale. Par conséquent, l'hypothèse H2 est rejetée.

Les résultats du questionnaire subjectif indiquent que les participants ont ressenti des sentiments de présence, de coprésence, de présence sociale plus forts et une meilleure expérience d'apprentissage avec la combinaison des trois modalités. Ce résultat est conforme à des recherches antérieures indiquant que l'ajout de nouvelles modalités augmente la présence [43], et ainsi, plus il y a de modalités, plus la présence des participants est élevée. De plus, Slater et al. [47] expliquent que la présence et la coprésence ont souvent tendance à co-varier, de sorte que lorsque les utilisateurs ressentent un sentiment de présence plus fort, ils ont aussi tendance à ressentir un sentiment de coprésence plus intense. Nos résultats de corrélation sont en cohérence avec ces études et montrent une corrélation positive de la plupart des dimensions subjectives pour l'ensemble des conditions expérimentales. Dans le travail de Simon et al. [46], la modalité verbale a également augmenté le sentiment de coprésence par rapport aux modalités visuelles et haptiques seules. Cependant, notre étude diffère de ce précédent travail car l'instructeur était situé dans la même pièce alors que dans notre cas il était distant. Cela confirme que la modalité verbale est importante pour améliorer l'expérience utilisateur dans les EIP, même lorsque le partenaire d'interaction est situé à distance. Ceci est cohérent avec d'autres travaux [42] et vient corroborer notre hypothèse H3. Plus précisément, l'interaction entre les trois modalités est encore meilleure pour améliorer l'expérience utilisateur avec un instructeur distant. Cela suggère que la communication multimodale améliore les interactions interpersonnelles entre l'instructeur et l'apprenant.

Ceci est confirmé par les résultats du questionnaire de comparaison, indiquant que les participants ont préféré les instructions à trois modalités. Ils ont perçu cette combinaison comme la plus appropriée, plaisante, engageante et éducative. Ces conclusions sont en accord avec l'observation selon laquelle l'utilisation des trois modalités contribue à une expérience d'apprentissage améliorée. De plus, les participants ont indiqué que le fait de recevoir des instructions en combinant les trois modalités était associé à une plus grande facilité de compréhension et à une efficacité accrue. Cette combinaison était également la plus précise, ce qui est cohérent avec les mesures de performance.

Il convient de noter que les tests de corrélations entre les mesures subjectives et les mesures de performance n'ont révélé aucun résultat significatif. Néanmoins, les performances de la condition visuelle-haptique sont en contradiction avec certaines mesures subjectives. En effet, la condition visuelle-haptique ayant conduit à une meilleure trajectoire et un temps de réalisation plus rapide, était moins préférée que la combinaison à trois modalités et était perçue comme plus perturbante. Cela suggère que la nature des instructions reçues est plus importante pour améliorer la performance d'un apprenant que le sentiment de coprésence avec l'instructeur. En outre, l'ajout d'instructions verbales est utile pour améliorer l'expérience utilisateur et a tendance à rassurer les apprenants. Cependant, ces instructions ne doivent pas interférer avec celles provenant des autres modalités au risque de détériorer la performance des apprenants.

## 6 CONCLUSIONS

Cette étude s'inscrit dans le cadre d'un projet de recherche visant à intégrer le concept de compagnonnage pour l'acquisition de compétences motrices dans les environnements immersifs. Le compagnonnage est un modèle largement utilisé dans différentes disciplines au début du processus d'apprentissage, fournissant aux apprenants des instructions et des rétroactions d'experts. Cette étude a évalué l'effet de quatre formes de communication (verbale-visuelle, verbale-haptique, visuelle-haptique, verbale-visuelle-haptique) sur le transfert d'informations spatiales aux apprenants lors de la manipulation d'outils. Bien que les résultats présentés soient préliminaires, cette étude a montré que la combinaison



visuelle-haptique et la combinaison verbale-visuelle-haptique étaient les plus efficaces pour augmenter la précision des mouvements. La combinaison visuelle-haptique a également permis une manipulation plus rapide et une meilleure trajectoire de mouvement. Cependant, les participants ont indiqué que la combinaison visuelle-haptique était plus complexe pour mémoriser les instructions spatiales et plus perturbante que la combinaison à trois modalités. De plus, la combinaison à trois modalités a augmenté le sentiment de présence, de coprésence, de présence sociale et la qualité perçue de l'expérience d'apprentissage des participants. Cette dernière combinaison a également été perçue comme la plus adaptée pour l'apprentissage et la mémorisation des instructions.

### 6.1 Implications pour la conception

Ces résultats offrent des perspectives précieuses pour la conception d'interactions collaboratives visant à améliorer l'acquisition des compétences de manipulation d'outils au sein d'un EIP. En effet, ils confirment davantage que chaque modalité apporte des avantages distincts pour améliorer le processus d'apprentissage et que l'utilisation d'une stratégie de communication multimodale est optimale. La combinaison visuelle-haptique est démontrée comme étant la plus appropriée pour améliorer les performances des apprenants lors de la réception d'instructions spatiales sur les mouvements. L'ajout d'instructions verbales peut être utile pour améliorer l'expérience utilisateur et les interactions avec un instructeur distant. Cependant, ces instructions peuvent dégrader les performances des apprenants en interférant avec les informations reçues par le biais des autres modalités. Par conséquent, nous suggérons d'ajouter des instructions verbales pour améliorer la communication avec l'instructeur tout en évitant de les utiliser pour fournir des instructions spécifiques sur les mouvements des outils. Par exemple, elles peuvent être utilisées avant ou après la transmission des instructions sur le mouvement afin de rassurer les apprenants, attirer leur attention ou leur donner des conseils spécifiques sur la réalisation de la tâche.

### 6.2 Limitations et perspectives

Notre étude visait à obtenir des informations sur la communication entre les instructeurs et les apprenants au sein des environnements immersifs pour l'apprentissage. Cependant, certaines limitations doivent être notées. Premièrement, nous n'avons pas interrogé les participants après l'expérience sur leur perception de la présence de l'instructeur distant. Il est crucial de prendre en compte cette variable dans les prochaines études afin d'analyser si la simulation de la présence de l'instructeur a eu un impact sur l'expérience utilisateur. Deuxièmement, notre étude n'a pas examiné l'influence des modalités de communication sur le processus d'apprentissage, mettant en évidence la nécessité d'une étude longitudinale englobant des évaluations de type pré-apprentissage et post-apprentissage. Notre objectif principal dans cette étude était d'acquiescer une compréhension plus approfondie des forces et des faiblesses de chaque combinaison de modalités de communication avant de réaliser une étude aussi complète. Compte tenu des résultats obtenus, nos travaux futurs incluront la réalisation d'une étude longitudinale pour examiner les effets de chaque combinaison de modalités sur l'apprentissage. Troisièmement, l'étude s'est concentré exclusivement sur la perspective de l'apprenant lors de l'évaluation de l'influence des modalités de communication. Bien que cette approche ait été essentielle pour contrôler l'étude expérimentale, il est crucial, dans les futures recherches, d'explorer comment ces combinaisons de modalités affectent également les instructeurs et comment la technologie peut faciliter leur capacité à transmettre efficacement leurs compétences. Par ailleurs, nous envisageons également d'examiner comment la combinaison de modalités affecte l'apprentissage lors de l'acquisition de compétences motrices plus complexes dans un EIP. Ainsi, l'utilisation d'EIP plus sophistiqués qui reproduisent étroitement des scénarios du monde réel (l'apprentissage d'un geste médical, par exemple) peut augmenter la complexité des tâches et améliorer la pertinence des résultats de l'étude.

Ces investigations nous permettront de concevoir des interfaces utilisateurs plus efficaces pour supporter le modèle de compagnonnage et faciliter le transfert de compétences motrices d'un instructeur vers un apprenant dans ce type d'environnements.

## REMERCIEMENTS

Les auteurs souhaitent remercier tous les volontaires qui ont participé à l'étude. Ce travail a été soutenu par une subvention de l'Agence National de la Recherche (ANR-20-CE33-0010 Show-me).

## RÉFÉRENCES

- [1] Mortaja AlQassab and David Wang. 2016. Human Performance and Cognitive Workload in Multi-sensory Virtual Environments. In *Universal Access in Human-Computer Interaction. Interaction Techniques and Environments : 10th International Conference, UAHCI 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings, Part II 10*. Springer, 219–230. [https://doi.org/10.1007/978-3-319-40244-4\\_21](https://doi.org/10.1007/978-3-319-40244-4_21)
- [2] Michael J Barrington, Daniel M Wong, Ben Slater, Jason J Ivanusic, and Matthew Ovens. 2012. Ultrasound-guided regional anesthesia : how much practice do novices require before achieving competency in ultrasound needle visualization using a cadaver model. *Regional Anesthesia & Pain Medicine* 37, 3 (2012), 334–339. <https://doi.org/10.1097/AAP.0b013e3182475fba>
- [3] C. Basdogan, C. Ho, M. A. Srinivasan, and M. Slater. 2000. An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)* 7 (2000), 443–460. Issue 4. <https://doi.org/10.1145/365058.365082>
- [4] David Brickler, Matias Volonte, Jeffrey W Bertrand, Andrew T Duchowski, and Sabarish V Babu. 2019. Effects of stereoscopic viewing and haptic feedback, sensory-motor congruence and calibration on near-field fine motor perception-action coordination in virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 28–37. <https://doi.org/10.1109/VR.2019.8797744>
- [5] Dario Cecilio-Fernandes, Fokie Cnossen, Jenifer Coster, AD (Debbie) C Jaarsma, and René A Tio. 2020. The effects of expert and augmented feedback on learning a complex medical skill. *Perceptual and motor skills* 127, 4 (2020), 766–784. <https://doi.org/10.1177/0031512520914680>
- [6] Amine Chellali, Woojin Ahn, Ganesh Sankaranarayanan, JT Flinn, Steven D Schwartzberg, Daniel B Jones, Suvranu De, and Caroline GL Cao. 2015. Preliminary evaluation of the pattern cutting and the ligating loop virtual laparoscopic trainers. *Surgical endoscopy* 29 (2015), 815–821. <https://doi.org/10.1007/s00464-014-3764-7>
- [7] Amine Chellali, Cédric Dumas, and Isabelle Milleville-Pennel. 2011. Influences of haptic communication on a shared manual task. *Interacting with Computers* 23, 4 (2011), 317–328. <https://doi.org/10.1016/j.intcom.2011.05.002>
- [8] Amine Chellali, Cédric Dumas, and Isabelle Milleville-Pennel. 2012. Haptic communication to support biopsy procedures learning in virtual environments. *Presence : Teleoperators and Virtual Environments* 21, 4 (2012), 470–489. [https://doi.org/10.1162/PRES\\_a\\_00128](https://doi.org/10.1162/PRES_a_00128)
- [9] Amine Chellali, Isabelle Milleville-Pennel, and Cédric Dumas. 2013. Influence of contextual objects on spatial interactions and viewpoints sharing in virtual environments. *Virtual Reality* 17, 1 (2013), 1–15. <https://doi.org/10.1007/s10055-012-0214-5>
- [10] Elizabeth F Churchill and Dave Snowdon. 1998. Collaborative virtual environments : an introductory review of issues and systems. *virtual reality* 3 (1998), 3–15. <https://doi.org/10.1007/BF01409793>
- [11] Gregorio Convertino, Helena M Mentis, Aleksandra Slavkovic, Mary Beth Rosson, and John M Carroll. 2011. Supporting common ground and awareness in emergency management planning : A design research project. *ACM Transactions on Computer-Human Interaction (TOCHI)* 18, 4 (2011), 1–34. <https://doi.org/10.1145/2063231.2063236>
- [12] Cédric Dabo and Sylvain Connac. 2022. Les représentations des étudiants tuteurs en masso-kinésithérapie sur le dispositif de tutorat entre pairs. *Recherches en éducation* 46 (2022). <https://doi.org/10.4000/ree.10150>
- [13] Michel Desmurget. 2006. *Imitation et apprentissages moteurs : des neurones miroirs à la pédagogie du geste sportif*. Groupe de Boeck.
- [14] Oliver Ehmer. 2021. Synchronization in demonstrations. Multimodal practices for instructing body knowledge. *Linguistics Vanguard* 7, s4 (2021), 20200038. <https://doi.org/10.1515/lingvan-2020-0038>
- [15] Oliver Ehmer and Geert Brône. 2021. Instructing embodied knowledge : multimodal approaches to interactive practices for knowledge constitution. *Linguistics Vanguard* 7, s4 (2021), 20210012. <https://doi.org/10.1515/lingvan-2021-0012>
- [16] Jesse Fox, Sun Joo Ahn, Joris H Janssen, Leo Yeykelis, Kathryn Y Segovia, and Jeremy N Bailenson. 2015. Avatars versus agents : a meta-analysis quantifying the effect of agency on social influence. *Human-Computer Interaction* 30, 5 (2015), 401–432. <https://doi.org/10.1080/07370024.2014.921494>
- [17] R Brent Gillespie, M Sile O'Modhrain, Philip Tang, David Zaretsky, and Cuong Pham. 1998. The virtual teacher. In *ASME International Mechanical Engineering Congress and Exposition*, Vol. 15861. American Society of Mechanical Engineers, 171–178. <https://doi.org/10.1115/IMECE1998-0250>
- [18] Noah J Glaser and Matthew Schmidt. 2020. Usage considerations of 3D collaborative virtual learning environments to promote development and transfer of knowledge and skills for individuals with autism. *Technology, Knowledge and Learning* 25, 2 (2020), 315–322. <https://doi.org/10.1007/s10758-018-9369-9>
- [19] Erving Goffman. 1974. *Frame analysis : An essay on the organization of experience*. Harvard University Press. <https://doi.org/10.2307/2064022>
- [20] Carl Gutwin, Oliver Schneider, Robert Xiao, and Stephen Brewster. 2011. Chalk sounds : the effects of dynamic synthesized audio on workspace awareness in distributed groupware. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. 85–94. <https://doi.org/10.1145/1959167.1959200>

- 1145/1958824.1958838
- [21] William S Halsted. 1904. The training of the surgeon. *Bull Johns Hop Hosp* (1904), 267–275. <https://doi.org/10.1001/jama.1904.02500210043006>
- [22] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index) : Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183. [https://doi.org/10.1016/s0166-4115\(08\)62386-9](https://doi.org/10.1016/s0166-4115(08)62386-9)
- [23] John Hattie and Helen Timperley. 2007. The power of feedback. *Review of educational research* 77, 1 (2007), 81–112. <https://doi.org/10.3102/003465430298487>
- [24] Morten Hertzum. 2021. Reference values and subscale patterns for the task load index (TLX) : a meta-analytic review. *Ergonomics* 64, 7 (2021), 869–878. <https://doi.org/doi:10.1080/00140139.2021.1876927>.
- [25] Jon Hindmarsh, Mike Fraser, Christian Heath, Steve Benford, and Chris Greenhalgh. 1998. Fragmented interaction : establishing mutual orientation in virtual environments. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work*. 217–226. <https://doi.org/10.1145/289444.289496>
- [26] Adrian H Hoppe, Roland Reeb, Florian van de Camp, and Rainer Stiefelhagen. 2018. Interaction of distant and local users in a collaborative virtual environment. In *Virtual, Augmented and Mixed Reality : Interaction, Navigation, Visualization, Embodiment, and Simulation : 10th International Conference, VAMR 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part I 10*. Springer, 328–337. [https://doi.org/10.1007/978-3-319-91581-4\\_24](https://doi.org/10.1007/978-3-319-91581-4_24)
- [27] Keke Lu, Guanyang Liu, and Lingzhi Liu. 2013. A study on haptic collaborative game in shared virtual environment. In *Fifth International Conference on Machine Vision (ICMV 2012) : Computer Vision, Image Analysis and Processing*, Vol. 8783. SPIE, 425–431. <https://doi.org/10.1117/12.2021258>
- [28] Richard Magill and David I Anderson. 2016. *Motor learning and control* (11th ed.). McGraw-Hill Publishing New York.
- [29] Daniel Martin, Sandra Malpica, Diego Gutierrez, Belen Masia, and Ana Serrano. 2022. Multimodality in VR : A survey. *ACM Computing Surveys (CSUR)* 54, 10s (2022), 1–36. <https://doi.org/10.1145/3508361>
- [30] Helena M Mentis, Amine Chellali, and Steven Schwaitzberg. 2014. Learning to see the body : supporting instructional practices in laparoscopic surgical procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2113–2122. <https://doi.org/10.1145/2556288.2557387>
- [31] M Alex Meredith. 2002. On the neuronal basis for multisensory convergence : a brief overview. *Cognitive brain research* 14, 1 (2002), 31–40. [https://doi.org/10.1016/s0926-6410\(02\)00059-9](https://doi.org/10.1016/s0926-6410(02)00059-9)
- [32] Jonas Moll, Eva-Lotta Sallnäs Pysander, Kerstin Severinson Eklundh, and Sten-Olof Hellström. 2014. The effects of audio and haptic feedback on collaborative scanning and placing. *Interacting with computers* 26, 3 (2014), 177–195. <https://doi.org/10.1093/iwc/iwt031>
- [33] Kristine L. Nowak and Frank Biocca. 2003. The Effect of the Agency and Anthropomorphism of Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence : Teleoper. Virtual Environ.* 12 (2003), 481–494. <https://doi.org/10.1162/105474603322761289>
- [34] Catherine S Oh, Jeremy N Bailenson, and Gregory F Welch. 2018. A systematic review of social presence : Definition, antecedents, and implications. *Frontiers in Robotics and AI* 5 (2018), 409295. <https://doi.org/10.3389/frobt.2018.00114>
- [35] Sofianna Ojala, Joonas Sirola, Timo Nykopp, Heikki Kröger, and Henrik Nuutinen. 2022. The impact of teacher's presence on learning basic surgical tasks with virtual reality headset among medical students. *Medical education online* 27, 1 (2022), 2050345. <https://doi.org/10.1080/10872981.2022.2050345>
- [36] Vasiliki Papageorgiou and Petros Lamerias. 2017. Multimodal Teaching and Learning with the Use of Technology : Meanings, Practices and Discourses. *International Association for Development of the Information Society* (2017).
- [37] Stéphanie Philippe, Alexis D Souchet, Petros Lamerias, Panagiotis Petridis, Julien Caporal, Gildas Coldeboeuf, and Hadrien Duzan. 2020. Multimodal teaching, learning and training in virtual reality : a review and case study. *Virtual Reality & Intelligent Hardware* 2, 5 (2020), 421–442. <https://doi.org/10.1016/j.vrih.2020.07.008>
- [38] Mark C Porte, George Xeroulis, Richard K Reznick, and Adam Dubrowski. 2007. Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learning new surgical skills. *The American journal of surgery* 193, 1 (2007), 105–110. <https://doi.org/10.1016/j.amjsurg.2006.03.016>
- [39] Lauriane Pouliquen-Lardy, Isabelle Milleville-Pennel, Francois Guillaume, and Franck Mars. 2016. Remote collaboration in virtual reality : asymmetrical effects of task distribution on spatial processing and mental workload. *Virtual Reality* 20 (2016), 213–220. <https://doi.org/10.1007/s10055-016-0294-8>
- [40] J. Rasmussen. 1983. Skills, rules, knowledge : signals, signs and symbols and other distinctions in human performance models. *IEEE Transactions in Systems, Man, and Cybernetics* 13 (1983), 257–266. Issue 3. <https://doi.org/10.1109/TSMC.1983.6313160>
- [41] Aylen Ricca, Amine Chellali, and Samir Otrnane. 2021. The influence of hand visualization in tool-based motor-skills training, a longitudinal study. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 103–112. <https://doi.org/10.1109/VR50410.2021.00031>
- [42] Eva-Lotta Sallnäs. 2004. *The effect of modality on social presence, presence and performance in collaborative virtual environments*. Ph.D. Dissertation. KTH.
- [43] Eva-Lotta Sallnäs, Kirsten Rasmus-Gröhn, and Calle Sjöström. 2000. Supporting presence in collaborative environments by haptic force feedback. *ACM Transactions on Computer-Human Interaction (TOCHI)* 7, 4 (2000), 461–476. <https://doi.org/10.1145/365058.365086>
- [44] Richard A Schmidt, Timothy D Lee, Carolee Winstein, Gabriele Wulf, and Howard N Zelaznik. 2018. *Motor control and learning : A behavioral emphasis* (5th ed.). Human kinetics.
- [45] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning : a review. *Psychonomic bulletin & review* 20 (2013), 21–53. <https://doi.org/10.3758/s13423-012-0333-8>

- [46] Cassandre Simon, Manel Boukli Hacene, Samir Otmame, and Amine Chellali. 2023. Study of communication modalities to support teaching tool manipulation skills in a shared immersive environment. *Computers and Graphics* 117 (2023), 31–41. <https://doi.org/10.1016/j.cag.2023.09.011>
- [47] Mel Slater, Amela Sadagic, Martin Usoh, and Ralph Schroeder. 2000. Small-group behavior in a virtual and real environment : A comparative study. *Presence* 9, 1 (2000), 37–51. <https://doi.org/10.1162/105474600566600>
- [48] Jeanett Strandbygaard, Flemming Bjerrum, Mathilde Maagaard, Per Winkel, Christian Ribbjerg Larsen, Charlotte Ringsted, Christian Gluud, Teodor Grantcharov, Bent Ottesen, and Jette Led Sorensen. 2013. Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator : a randomized trial. <https://doi.org/10.1097/SLA.0b013e31827eee6e>.
- [49] Ankeet D Udani, T Edward Kim, Steven K Howard, and Edward R Mariano. 2015. Simulation in teaching regional anesthesia : current perspectives. *Local and regional anesthesia* (2015), 33–43. <https://doi.org/10.2147/LRA.S68223>
- [50] Holly S Vitense, Julie A Jacko, and V Kathlene Emery. 2003. Multimodal feedback : an assessment of performance and mental workload. *Ergonomics* 46, 1-3 (2003), 68–87. <https://doi.org/10.1080/00140130303534>
- [51] Mark T Wallace, M Alex Meredith, and Barry E Stein. 1993. Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *Journal of neurophysiology* 69, 6 (1993), 1797–1809. <https://doi.org/10.1152/jn.1993.69.6.1797>
- [52] Jinling Wang, Amine Chellali, and Caroline GL Cao. 2016. Haptic communication in collaborative virtual environments. *Human factors* 58, 3 (2016), 496–508. <https://doi.org/10.1177/0018720815618808>
- [53] Fei Wu, Jerald Thomas, Shreyas Chinnola, and Evan Suma Rosenberg. 2020. Exploring communication modalities to support collaborative guidance in virtual reality. In *2020 IEEE conference on virtual reality and 3d user interfaces abstracts and workshops (VRW)*. IEEE, 79–86. <https://doi.org/10.1109/VRW50115.2020.00021>
- [54] Geoff Wyvill, Craig McPheeters, and Brian Wyvill. 1986. Data Structure for soft Objects. *The Visual Computer* 2, 4 (Aug. 1986), 227–234. <https://doi.org/10.1007/BF01900346>
- [55] Eugenia Yiannakopoulou, Nikolaos Nikiteas, Despina Perrea, and Christos Tsigris. 2015. Virtual reality simulators and training in laparoscopic surgery. *International Journal of Surgery* 13 (2015), 60–64. <https://doi.org/10.1016/j.ijssu.2014.11.014>