



**HAL**  
open science

## Validation de la pertinence d'un simulateur d'arthroscopie en réalité virtuelle pour caractériser les chirurgiens expérimentés

A. Tronchot, J. Berthelemy, H. Thomazeau, Arnaud Huaulmé, P. Walbron, F.  
Sirveaux, Pierre Jannin

### ► To cite this version:

A. Tronchot, J. Berthelemy, H. Thomazeau, Arnaud Huaulmé, P. Walbron, et al.. Validation de la pertinence d'un simulateur d'arthroscopie en réalité virtuelle pour caractériser les chirurgiens expérimentés. *Revue de Chirurgie Orthopédique et Traumatologique*, 2021, 107 (8), pp.966-973. 10.1016/j.rcot.2021.09.033 . hal-03480972

**HAL Id: hal-03480972**

**<https://hal.science/hal-03480972>**

Submitted on 5 Jan 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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## Mémoire original

### Validation de la pertinence d'un simulateur d'arthroscopie en réalité virtuelle pour caractériser les chirurgiens expérimentés

### Validation of virtual reality arthroscopy simulator relevance in characterising experienced surgeons

Alexandre TRONCHOT, <sup>(1,2)</sup>; Julien BERTHELEMY, <sup>(1)</sup>; Hervé THOMAZEAU, <sup>(1,2)</sup>; Arnaud HUAULMÉ, <sup>(1)</sup>; Paul WALBRON, <sup>(3)</sup>; François SIRVEAUX, <sup>(3)</sup>; Pierre JANNIN, <sup>(1)</sup>

1. Univ Rennes, Inserm, LTSI – UMR 1099, F-35000 Rennes, France
2. Orthopaedics and Trauma Department, Rennes University Hospital, 2 rue Henri Le Guilloux, 35000, Rennes, France
3. Orthopaedics Department, Nancy University Hospital, Centre Chirurgical Emile Gallé, 49 rue Hermite, 54000 Nancy, France

#### Auteur correspondant :

Alexandre Tronchot, MSc

Univ Rennes, CHU Rennes, Inserm, LTSI, Equipe MediCIS- UMR 1099, F-35000, Rennes, France

Téléphone: Mobile +33 6 38 40 73 33 Work + 33 99 28 43 21

E-mail : Alexandre.TRONCHOT@etudiant.univ-rennes1.fr

ORCID : 0000-0002-3092-3282

Ne pas utiliser, pour citation, la référence française de cet article, mais celle de l'article original paru dans *Orthopaedics & Traumatology: Surgery & Research*, en utilisant le DOI ci-dessus.

#### Abstract

*Introduction* La simulation par réalité virtuelle (VR) est adaptée à l'apprentissage de l'arthroscopie. Malgré de nombreuses études, il reste néanmoins difficile de distinguer des niveaux de compétence (*Construct Validity*) parmi les chirurgiens expérimentés. Il semble donc adéquat de rechercher de nouvelles méthodes de mesure de compétence en utilisant l'analyse des trajectoires instrumentales au lieu des mesures couramment utilisées.

*Hypothèse* Nous avons émis l'hypothèse qu'une plus grande expérience en arthroscopie d'épaule serait corrélée à une meilleure performance sur simulateur VR d'arthroscopie d'épaule et que les opérateurs expérimentés partageraient des trajectoires instrumentales similaires.

*Matériels & Méthodes* Après réponse à un questionnaire standardisé, 104 trajectoires de 52 chirurgiens répartis en 2 cohortes (26 intermédiaires et 26 experts) ont été enregistrées sur simulateur d'arthroscopie. La procédure analysée était le retrait de corps étranger dans une articulation d'épaule droite. 10 mesures ont été calculées sur les trajectoires, notamment la durée de procédure, la distance instrumentale, l'économie de

mouvement et la fluidité. De plus, le Dynamic Time Warping (DTW) a été calculé sur les trajectoires pour une classification hiérarchique non supervisée des chirurgiens.

*Résultats* Les experts furent significativement plus rapides (Médiane 70.9s Écart Interquartile [56.4-86.3] vs. 116.1s [82.8-154.2],  $p < 0.01$ ), plus fluides ( $4.6 \cdot 10^5 \text{ mm} \cdot \text{s}^{-3}$  [ $3.1 \cdot 10^5$ - $7.2 \cdot 10^5$ ] vs.  $1.5 \cdot 10^6 \text{ mm} \cdot \text{s}^{-3}$  [ $2.6 \cdot 10^6$ - $3.5 \cdot 10^6$ ],  $p = 0.05$ ), et plus économes en mouvement ( $19.3 \text{ mm}^2$  [9.1-25.9] vs.  $33.8 \text{ mm}^2$  [14.8-50.5],  $p < 0.01$ ), mais il n'y avait pas de différence significative de performance sur la distance (671.4mm [503.8-846.1] vs 694.6mm [467.0-1090.1],  $p = 0.62$ ). Le clustering par DTW a différencié deux groupes de trajectoires similaires liées à l'expertise avec respectivement 48 trajectoires d'experts pour le premier groupe et 52 trajectoires d'intermédiaires et 4 d'experts pour le second groupe (Sensibilité de 92%, Spécificité de 100%). Le clustering hiérarchique avec DTW a identifié significativement les chirurgiens experts des intermédiaires en trouvant des trajectoires similaires pour 24/26 experts.

*Conclusion* Cette étude a démontré la *Construct Validity* du simulateur VR d'arthroscopie d'épaule au sein d'un groupe de chirurgiens expérimentés. Grâce à de nouveaux types de mesures simplement basées sur les trajectoires, il a été possible de distinguer de manière significative les niveaux d'expertise. Nous avons démontré que l'analyse par clustering avec Dynamic Time Warping est capable de distinguer de manière fiable les opérateurs experts des opérateurs intermédiaires.

*Pertinence clinique* Les résultats ont des implications pour l'avenir de la formation en chirurgie arthroscopique ou des programmes d'accréditation postuniversitaire utilisant la simulation par réalité virtuelle.

**Niveau de preuve** : III ; étude comparative prospective

**Mots-Clés** Simulation VR- Arthroscopie d'épaule – Entraînement – Compétences Arthroscopiques

## 1. Introduction

La simulation chirurgicale vise à enseigner des compétences pratiques en dehors du bloc opératoire. Elle est nécessaire pour des raisons de sécurité [1], d'organisation, liées au fonctionnement des établissements de santé [2] et aux contraintes de temps des enseignants et des étudiants. En France, la simulation procédurale est officiellement

incluse dans le programme national de formation médicale et chirurgicale des internes depuis 2017 [3].

La simulation par réalité virtuelle (VR) est un outil qui complète les techniques traditionnelles sur maquettes [4], ou sur sujets cadavériques [5]. Elle est particulièrement adaptée à l'apprentissage de l'arthroscopie [6-8]. De nombreuses recherches ont été consacrées au développement de cet outil et à son utilisation. Dans deux méta-analyses récentes [9,10], l'une des principales limites soulignées est la difficulté à distinguer les niveaux de compétences (*Construct Validity*) avec les métriques traditionnelles (telles que la durée et la distance instrumentale) dans des populations expérimentées. Par conséquent, il semble adéquat de rechercher de nouvelles métriques ou de nouvelles méthodes de mesure des performances.

En analyse de données, le classification hiérarchique (ou clustering) est une méthode d'analyse qui cherche à construire une hiérarchie de groupes avec critère de liaison. Pour l'analyse des séries temporelles, la déformation temporelle dynamique (*Dynamic Time Warping, DTW*) est un algorithme permettant de mesurer la similarité entre deux séquences temporelles dont la vitesse peut varier. Le clustering hiérarchique avec DTW a déjà été utilisé avec succès pour distinguer différents niveaux d'expertise dans des études portant sur la formation à la chirurgie robotique par simulation VR [11] mais n'a pas encore été analysé dans la formation à l'arthroscopie.

Nous avons émis l'hypothèse qu'une plus grande expérience en arthroscopie d'épaule serait corrélée à de meilleure performance sur simulateur VR d'arthroscopie d'épaule et que les opérateurs expérimentés partageraient des trajectoires instrumentales similaires. L'objectif de cette étude était d'analyser les trajectoires instrumentales et de déterminer la capacité du simulateur à distinguer les niveaux de performance individuels au sein de cohortes de chirurgiens expérimentés de différents niveaux d'expertise en utilisant des métriques innovantes et un clustering hiérarchique avec DTW.

## **2. Méthodes**

### ***2.1. Participants***

Les trajectoires de deux cohortes de chirurgiens ont été recueillies prospectivement lors de congrès internationaux de chirurgie orthopédique, celui de la

Société française d'arthroscopie (SFA) tenue à Strasbourg en décembre 2018, et celui de l'International Shoulder Course de Val d'Isère en janvier 2019. Chaque participant a répondu à un questionnaire standardisé sur son expertise en arthroscopie et ses loisirs (**tableau 1**). Le consentement éclairé a été obtenu pour tous les participants. Les questions concernant le nombre de procédures chirurgicales réelles déjà effectuées avant de réaliser l'exercice ont servi à établir les niveaux d'expérience de base pour chaque cohorte. La première cohorte appelée INTERMEDIAIRE est composée de 29 chirurgiens orthopédiques intéressés par l'arthroscopie (chirurgiens en fin d'internat, assistants spécialistes, jeunes installés), qui ont réalisé entre une et 99 chirurgies arthroscopiques en tant qu'opérateur principal éventuellement sous la supervision d'un chirurgien senior. La seconde cohorte appelée EXPERT est composée de 28 chirurgiens orthopédistes avec une surspécialisation en chirurgie arthroscopique d'épaule ayant réalisé plus de 100 arthroscopies en tant qu'opérateur principal. Ce seuil d'expérience a été précédemment utilisé dans une étude similaire [12]. Seule la pratique chirurgicale réelle a été évaluée pour déterminer le niveau d'expertise, la pratique antérieure sur simulateur chirurgical n'a pas été prise en compte. Les données de trois trajectoires INTERMEDIAIRE et de deux trajectoires EXPERT étaient incomplètes. Les 26 données restantes sur les trajectoires des participants INTERMEDIAIRE et 26 EXPERTS ont été analysées pour l'étude. Un calcul de puissance bilatéral a été effectué pour l'hypothèse nulle ("la performance des experts est égale à celle des intermédiaires") avec une erreur de type 1 acceptable,  $\alpha = 0,05$  et une erreur de type 2 acceptable  $\beta = 0,20$ , puissance = 0,8. Pour le calcul de la puissance, les données d'une précédente étude de validation d'un simulateur d'arthroscopie de l'épaule ont été utilisées [13]. Cela a permis d'obtenir un nombre minimum de 14 sujets dans chaque groupe, ce qui est conforme aux études de validation précédemment rapportées [12-15].

## **2.2. Méthodes d'évaluation**

Le simulateur d'arthroscopie ArthroS™ de VirtaMed AG (Zurich) a été utilisé pour recueillir les trajectoires. La *Face Validity* et la *Construct Validity* du simulateur ont été évaluées dans des études précédentes pour des populations de novices et d'experts [12,13,16]. Les modules *FAST* et *SHOULDER* ont été utilisés pour l'étude. L'exercice correspondait à une procédure de base en chirurgie de l'épaule : l'extraction de corps étrangers dans une articulation native d'épaule. "*Catch the Stars*

*Glenohumeral*" est un exercice semi-réaliste avec modélisation d'une articulation d'épaule indemne de pathologie où le but est de saisir et d'extraire cinq corps étrangers (e.g. cinq étoiles).

La récolte des données des participants à l'étude s'est faite selon un modèle standardisé (**Figure 1**). Premièrement, le participant a réalisé des exercices sur le module FAST (**1.1, 1.2 et 1.3, Figure 1**). Ensuite, le participant a effectué une identification vidéo-guidée des différentes structures de l'épaule droite dans le module SHOULDER, procédure effectuée en décubitus latéral pour tous les participants. (**2.1, Figure 1**). Troisièmement, les utilisateurs ont effectué l'exercice de « *Catch the Stars* » dans le module SHOULDER (**2.2, Figure 1**) en décubitus latéral, qu'ils ont effectué 3 fois. Le premier essai correspondant à la courbe d'apprentissage a été retiré de l'analyse.

#### **2.4. Analyse statistique**

Le programme TRACAS (*TRajjectory Analysis for Computer Assisted Surgery*) basé sur Matlab a été adapté aux données du simulateur et utilisé pour l'analyse des trajectoires. Dans un premier temps, les données extraites du simulateur ont été utilisées pour calculer dix métriques (définies dans la **Figure 2**) : *durée, longueur de la trajectoire, vitesse, accélération, tremblement, volume de travail, économie de mouvement, fluidité du mouvement, fluidité de la longueur de l'arc spectral et profondeur*. L'analyse a été effectuée en suivant les techniques publiées précédemment [17-20]. Le calcul de toutes les métriques et l'analyse des trajectoires ont été effectués sur un ordinateur dédié (Xeon E5-1650V4 @3.60GHz avec 32Go de RAM) en utilisant Matlab\_R2018b. Un test non paramétrique (test U bilatéral de Mann-Whitney) a été utilisé afin d'observer lesquelles des métriques présentaient des différences statistiques ( $p < 0,05$ ). Des diagrammes moustaches ont été modélisées. La deuxième étape de l'étude a consisté à regrouper les trajectoires des participants en utilisant la classification hiérarchique ascendante avec Dynamic Time Warping (DTW). Le DTW est pertinent pour l'indexeur de séries temporelles car il a la capacité de reconnaître des formes indépendamment de leur distorsion temporelle ou de leur déphasage [21].

### **3. Résultats**

#### **3.1. Analyse des Métriques**

L'analyse des trajectoires instrumentales par métriques a révélé des différences significatives entre les groupes INTERMEDIAIRE et EXPERT (**Tableau 2**). Les diagrammes moustaches sont représentés en (**Figure 3**) et (**Figure 4**). Le groupe EXPERT a obtenu des résultats significativement supérieurs à ceux du groupe INTERMEDIAIRE pour les métriques suivantes (**Figure 3**) : *Durée (s)*,  $p < 0,01$  (**Figure 3.1**) ; *Vitesse moyenne ( $mm.s^{-1}$ )*,  $p < 0,01$  (**Figure 3.2**) ; *Économie ( $mm^2$ )*,  $p < 0,01$  (**Figure 3.3**) ; *Fluidité de mouvement ( $mm.s^{-3}$ )*,  $p = 0,05$  (**Figure 3.4**) ; *Fluidité de longueur d'arc spectral*,  $p < 0,01$  (**Figure 3.5**) et *Volume de travail ( $mm^3$ )*,  $p < 0,01$  (**Figure 3.6**). Il n'y avait pas de différences significatives entre les groupes pour les métriques suivantes (**Figure 4**) : *Distance instrumentale (mm)*,  $p = 0,08$  (**Figure 4.1**) ; *Accélération moyenne ( $mm.s^{-2}$ )*,  $p = 0,39$  (**Figure 4.2**) ; *tremblement moyen ( $mm.s^{-3}$ )*,  $p = 0,41$  (**Figure 4.3**) et *Profondeur (mm)*,  $p = 0,11$  (**Figure 4.4**).

### 3.2. Analyse des Trajectoires

Les résultats du clustering non supervisé avec DTW sont donnés en (**Figure 5**). Sur ce dendrogramme, chaque donnée est représentée en abscisse pour se rejoindre toutes une à une jusqu'à former un seul cluster. Plus les regroupements de clusters sont faibles, plus la distance DTW entre les données agrégées est faible et plus les trajectoires sont similaires. Comme deux catégories sont étudiées ici, il a été décidé de ne comparer que les deux plus grands clusters. Le premier cluster était composé de 48 trajectoires similaires, 100% des données de la première catégorie provenaient de la cohorte EXPERT (24/26 experts) tandis que 93% du second cluster était composé de données de la cohorte INTERMEDIAIRE (26/26 intermédiaires) et 7% de la cohorte EXPERT (2 experts) pour un total de 56 trajectoires similaires (Sensibilité de 92%, Spécificité de 100%).

## 4. Discussion

À propos de cette procédure arthroscopique d'extraction de corps étrangers dans une épaule simulée VR, les résultats ont confirmé que six des dix métriques calculées sur les trajectoires étaient capables de distinguer les opérateurs experts des non-experts : les experts avaient un geste plus fluide (meilleur résultat pour la fluidité de la longueur d'arc spectral), plus économique (meilleurs résultats pour l'économie et le volume de

travail), et plus efficace (meilleurs résultats pour le temps et la vitesse). Cependant, l'accélération et la distance instrumentale ne différaient pas statistiquement entre les groupes. De la même manière, Rahm et al, [13] avaient remarqué que la métrique de distance instrumentale n'était pas aussi discriminante que les autres utilisées dans ce simulateur. Bien que souvent utilisée par les simulateurs VR, la distance instrumentale ne semble pas être corrélée à l'expertise. La mesure de fluidité du mouvement a été calculée avec différentes formules dans le passé [17,18,20]. Pour cette étude, trois formules de fluidité du mouvement ont été calculées. Le calcul selon Hofstad [18] n'était pas significatif et nous avons choisi de ne pas le représenter. Le calcul selon Cotin [17] était juste significatif pour cette étude. Plus intéressant encore, nos résultats sur la fluidité du mouvement via la longueur d'arc spectral étaient très significatifs et cohérents avec les résultats trouvés par Balasubramanian et al [20]. Avec ces nouvelles métriques simplement basées sur les trajectoires brutes du simulateur, il a donc été possible de distinguer de manière significative les niveaux d'expertise.

Nous avons démontré que l'analyse par cluster non supervisée avec DTW est capable de discriminer de manière fiable les opérateurs experts des opérateurs intermédiaires. Nous suggérons que cette approche basée sur la trajectoire complète pourrait être utilisée pour mieux classer le niveau de performance sur simulateur. Le cluster 2 est composé de l'ensemble de la cohorte intermédiaire et de deux experts. L'âge déclaré de ces deux experts est proche de l'âge médian de la cohorte intermédiaire (31 ans). Leur présence dans ce cluster pourrait s'expliquer par une expérience non liée à l'âge. Nous n'avons pas trouvé d'impact de la nationalité sur la performance en simulateur. Au contraire, la cohorte intermédiaire composée uniquement de participants français se distingue par l'hétérogénéité des trajectoires dans la classification DTW.

La *Construct Validity* du simulateur avait déjà été démontrée pour les novices et les experts [13], [22]. Mais, Stunt et al., [22] ont eu des difficultés à trouver des différences significatives de performance pour le groupe intermédiaire (moins de soixante chirurgies arthroscopiques en tant qu'opérateur principal). Roberts et al, [16] n'ont pas distingué l'expérience chirurgicale pour un exercice similaire d'arthroscopie de l'épaule avec les métriques de performance proposées par le simulateur, remettant en cause son intérêt pour les chirurgiens de niveau intermédiaire. Notre étude a montré



qu'il est possible pour un simulateur VR d'arthroscopie de distinguer plusieurs niveaux d'expertise en utilisant de nouvelles métriques ou une analyse par clustering avec DTW.

La majorité de la cohorte intermédiaire a bénéficié d'une formation théorique similaire en arthroscopie (84% titulaires du diplôme national français d'arthroscopie). Près de la moitié d'entre eux avaient déjà reçu une formation sur simulateur (46% vs 27% pour les experts). Cependant, ce groupe se distingue des experts par l'hétérogénéité de leurs trajectoires et leur moindre performance. Il est donc essentiel de combiner une formation pratique sur simulateur à une formation théorique. Or, Manoharan et al, [23] ont récemment démontré la difficulté au Royaume-Uni de créer un programme national de formation à l'arthroscopie. Walbron et al, [6,24] ont rapporté les efforts du Collège Français de Chirurgie Orthopédique et de la SFA pour proposer un programme national d'apprentissage de l'arthroscopie avec simulateur VR. Les premiers résultats furent intéressants mais avec certaines limites dues à la présence nécessaire mais compliquée d'un formateur auprès de l'étudiant. Ainsi, la création d'un programme national standardisé avec suivi régulier des performances de l'étudiant reste complexe en pratique avec seulement quelques exemples réussis dans le monde [25]. Un retour d'information immédiat et direct sur les performances semble obligatoire plutôt que de laisser l'étudiant travailler de manière autonome avec un retour d'information occasionnel lors de sessions espacées avec le formateur.

Il nous reste à démontrer que les résultats de cette méthode peuvent être extrapolés à d'autres exercices et simulateurs. En effet, l'exercice *Catch the Stars* est identique pour chaque tentative : la position ainsi que l'ordre d'apparition des étoiles sont les mêmes ; de même, l'anatomie et les dimensions de l'épaule simulée sont toujours les mêmes. Cette invariance de l'exercice rend la mesure du DTW très efficace car chaque participant va utiliser le même parcours. Le niveau d'expérience a été évalué uniquement par le nombre d'arthroscopies réalisées en tant qu'opérateur principal, en accord avec les seuils d'expérience des études précédentes [13,22]. Bien que l'exercice ait été réalisé dans des conditions similaires pour les deux cohortes (lors d'un congrès et non dans un environnement de bloc opératoire), les résultats ont pu être faussés par les distractions inhérentes au congrès. Cela peut expliquer la présence de certaines valeurs aberrantes dans l'étude. Également mentionné par Fucentese et al, [12] la collecte de données pendant un congrès peut entraîner un biais de sélection en attirant des

participants très intéressés par l'enseignement et la formation VR. L'analyse en sous-groupes pour les pratiques de sports, de jeux vidéo et d'instruments de musique étant limitée par la petite taille de l'échantillon et l'hétérogénéité des sous-groupes, nous avons choisi de ne pas représenter les résultats. Enfin, cette étude ne porte que sur les compétences techniques du chirurgien. L'étude des compétences non techniques, y compris les compétences cognitives et interpersonnelles, revêt également une grande importance [26].

## 5. Conclusion

Cette étude a démontré la *Construct Validity* du simulateur CR d'arthroscopie de l'épaule au sein d'un groupe de chirurgiens expérimentés. Grâce à de nouveaux types de mesures simplement basées sur les trajectoires du simulateur, il a été possible de distinguer de manière significative les niveaux d'expertise. Nous avons démontré que l'analyse par cluster avec Dynamic Time Warping était capable de distinguer de manière fiable les opérateurs experts des opérateurs intermédiaires.

## 6. Remerciements

Nous remercions tous les participants de l'étude, ainsi que l'équipe de VirtaMed, Martina Vitz et Benjamin Grauer pour le support technique.

**Conflits d'intérêt :** Les auteurs n'ont aucun conflit à déclarer. F. Sirveaux est consultant pour Tornier Wright Medical en dehors de ce travail.

**Source de financement:** aucune

### Contribution des auteurs:

Troncot: conception de la méthodologie, collecte, analyse et visualisation des données, écriture du manuscrit, superviseur principal du projet.

J. Berthelemy: analyse et visualisation des données, co-auteur du manuscrit.

H. Thomazeau: Conception et supervision de l'étude, conception de la méthodologie, révision de l'article.

A. Huault: conception de la méthodologie, aide au sujet de l'analyse des dossiers collectés, révision de l'article.

P. Walbron: réalisation de l'enquête, collecte les données, révision l'article.

F. Sirveaux: Conception de l'étude, réalisation de l'enquête, révision de l'article.

P. Jannin: Conception et supervision de l'étude, conception de la méthodologie, révision de l'article.

## References

- [1] Barger LK, Rosner B. Extended Work Shifts and the Risk of Motor Vehicle Crashes among Interns. *N Engl J Med* 2005;10.
- [2] Erivan R, Chaput T, Villatte G, Ollivier M, Descamps S, Boisgard S. Ten-year epidemiological study in an orthopaedic and trauma surgery centre: Are there risks involved in increasing scheduled arthroplasty volume without increasing resources? *Orthop Traumatol Surg Res* 2018;104:1283–9. <https://doi.org/10.1016/j.otsr.2018.08.009>.
- [3] Bonnafeous S, Amiot-Chanal H, Rouanet P. Arrêté du 12 avril 2017 portant organisation du troisième cycle des études de médecine. NOR: MENS1708241A, n.d.
- [4] Banaszek D, You D, Chang J, Pickell M, Hesse D, Hopman WM, et al. Virtual Reality Compared with Bench-Top Simulation in the Acquisition of Arthroscopic Skill: A Randomized Controlled Trial. *J Bone Jt Surg* 2017;99:e34. <https://doi.org/10.2106/JBJS.16.00324>.
- [5] Henn RF, Shah N, Warner JJP, Gomoll AH. Shoulder Arthroscopy Simulator Training Improves Shoulder Arthroscopy Performance in a Cadaveric Model. *Arthrosc J Arthrosc Relat Surg* 2013;29:982–5. <https://doi.org/10.1016/j.arthro.2013.02.013>.
- [6] Walbron P, Common H, Thomazeau H, Hosseini K, Peduzzi L, Bulaid Y, et al. Virtual reality simulator improves the acquisition of basic arthroscopy skills in first-year orthopedic surgery residents. *Orthop Traumatol Surg Res* 2020:S187705682030116X. <https://doi.org/10.1016/j.otsr.2020.03.009>.
- [7] Marcheix P-S, Vergnenegre G, Dalmay F, Mabit C, Charissoux J-L. Learning the skills needed to perform shoulder arthroscopy by simulation. *Orthop Traumatol Surg Res* 2017;103:483–8. <https://doi.org/10.1016/j.otsr.2017.02.008>.
- [8] Baumann Q, Hardy A, Courage O, Lacombes P, Accadbled F. Lessons taught by a knee arthroscopy simulator about participants in a European arthroscopy training programme. *Orthop Traumatol Surg Res* 2019;105:S287–91. <https://doi.org/10.1016/j.otsr.2019.09.008>.
- [9] Frank RM, Wang KC, Davey A, Cotter EJ, Cole BJ, Romeo AA, et al. Utility of Modern Arthroscopic Simulator Training Models: A Meta-analysis and Updated Systematic Review. *Arthrosc J Arthrosc Relat Surg* 2018;34:1650–77. <https://doi.org/10.1016/j.arthro.2017.10.048>.
- [10] Aim F, Lonjon G, Hannouche D, Nizard R. Effectiveness of Virtual Reality Training in Orthopaedic Surgery. *Arthrosc J Arthrosc Relat Surg* 2016;32:224–32. <https://doi.org/10.1016/j.arthro.2015.07.023>.
- [11] Despinoy F, Zemiti N, Forestier G, Sánchez A, Jannin P, Poignet P. Evaluation of contactless human-machine interface for robotic surgical training. *Int J Comput Assist Radiol Surg* 2018;13:13–24. <https://doi.org/10.1007/s11548-017-1666-6>.
- [12] Fucetese SF, Rahm S, Wieser K, Spillmann J, Harders M, Koch PP. Evaluation of a virtual-reality-based simulator using passive haptic feedback for knee arthroscopy. *Knee Surg Sports Traumatol Arthrosc* 2015;23:1077–85. <https://doi.org/10.1007/s00167-014-2888-6>.
- [13] Rahm S, Germann M, Hingsammer A, Wieser K, Gerber C. Validation of a

virtual reality-based simulator for shoulder arthroscopy. *Knee Surg Sports Traumatol Arthrosc* 2016;24:1730–7. <https://doi.org/10.1007/s00167-016-4022-4>.

[14] Bayona S, Fernández-Arroyo JM, Martín I, Bayona P. Assessment study of insight ARTHRO VR ® arthroscopy virtual training simulator: face, content, and construct validities. *J Robot Surg* 2008;2:151–8. <https://doi.org/10.1007/s11701-008-0101-y>.

[15] Gomoll AH, O’Toole RV, Czarnecki J, Warner JJP. Surgical Experience Correlates with Performance on a Virtual Reality Simulator for Shoulder Arthroscopy. *Am J Sports Med* 2007;35:883–8. <https://doi.org/10.1177/0363546506296521>.

[16] Roberts PG, Guyver P, Baldwin M, Akhtar A, Alvand A, Price AJ, et al. Validation of the updated ArthroS simulator: face and construct validity of a passive haptic virtual reality simulator with novel performance metrics. *Knee Surg Sports Traumatol Arthrosc* 2017;25(2):616–25. <https://doi.org/10.1007/s00167-016-4114-1>.

[17] Cotin S, Stylopoulos N, Ottensmeyer M, Neumann P, Rattner D, Dawson S. Metrics for Laparoscopic Skills Trainers: The Weakest Link! In: Dohi T, Kikinis R, editors. *Med. Image Comput. Comput.-Assist. Interv. — MICCAI 2002*, vol. 2488, Berlin, Heidelberg: Springer Berlin Heidelberg; 2002, p. 35–43. [https://doi.org/10.1007/3-540-45786-0\\_5](https://doi.org/10.1007/3-540-45786-0_5).

[18] Hofstad EF, Våpenstad C, Chmarra MK, Langø T, Kuhry E, Mårvik R. A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance. *Surg Endosc* 2013;27:854–63. <https://doi.org/10.1007/s00464-012-2524-9>.

[19] Oropesa I, Sánchez-González P, Lamata P, Chmarra MK, Pagador JB, Sánchez-Margallo JA, et al. Methods and Tools for Objective Assessment of Psychomotor Skills in Laparoscopic Surgery. *J Surg Res* 2011;171:e81–95. <https://doi.org/10.1016/j.jss.2011.06.034>.

[20] Balasubramanian S, Melendez-Calderon A, Burdet E. A Robust and Sensitive Metric for Quantifying Movement Smoothness. *IEEE Trans Biomed Eng* 2012;59:2126–36. <https://doi.org/10.1109/TBME.2011.2179545>.

[21] Keogh E, Ratanamahatana CA. Exact indexing of dynamic time warping. *Knowl Inf Syst* 2005;7:358–86. <https://doi.org/10.1007/s10115-004-0154-9>.

[22] Stunt JJ, Kerkhoffs GMMJ, van Dijk CN, Tuijthof GJM. Validation of the ArthroS virtual reality simulator for arthroscopic skills. *Knee Surg Sports Traumatol Arthrosc* 2015;23:3436–42. <https://doi.org/10.1007/s00167-014-3101-7>.

[23] Manoharan G, Sharma N, Gallacher P. Competence in using the arthroscopy stack system: a national survey of orthopaedic trainees in the UK. *Ann R Coll Surg Engl* 2020;102:149–52. <https://doi.org/10.1308/rcsann.2019.0131>.

[24] Walbron P, Common H, Thomazeau H, Sirveaux F. Evaluation of arthroscopic skills with a virtual reality simulator in first-year orthopaedic residents. *Int Orthop* 2020. <https://doi.org/10.1007/s00264-020-04520-1>.

[25] Strandbygaard J, Bjerrum F, Maagaard M, Rifbjerg Larsen C, Ottesen B, Sorensen JL. A structured four-step curriculum in basic laparoscopy: development and validation. *Acta Obstet Gynecol Scand* 2014;93:359–66. <https://doi.org/10.1111/aogs.12330>.

[26] Yule S, Flin R, Maran N, Rowley D, Youngson G, Paterson-Brown S. Surgeons’ Non-technical Skills in the Operating Room: Reliability Testing of the NOTSS Behavior Rating System. *World J Surg* 2008;32:548–56. <https://doi.org/10.1007/s00268-007-9320-z>.

Tableaux

	INTERMEDIAIRE (n=26)	EXPERT (n=26)	TOTAL (n=52)	P-VALUE
<b>AGE, an</b>	30 [28-31]	41 [36-46]	33 [29-41]	< 0.001*
<b>SEXE</b>				1 <sup>†</sup>
<i>Homme</i>	23 (88.5%)	24 (92.3%)	47 (90.4%)	
<i>Femme</i>	3 (11.5%)	2 (7.7%)	5 (9.6%)	
<b>COTE DOMINANT</b>				0.6 <sup>†</sup>
<i>Droit</i>	24 (92.4%)	25 (96.2%)	49 (94.3%)	
<i>Gauche</i>	1 (3.8%)	0 (0.0%)	1 (1.9%)	
<i>Ambidextre</i>	1 (3.8%)	1 (3.8%)	2 (3.8%)	
<b>PAYS</b>				< 0.001 <sup>†</sup>
<i>France</i>	26 (100.0%)	5 (19.2%)	31 (59.6%)	
<i>Autre</i>	0 (0%)	21 (80.8%)	21 (40.4%)	
<b>PRATIQUE SPORTIVE</b>				0.01 <sup>†</sup>
<i>Oui</i>	15 (57.7%)	24 (92.3%)	39 (75.0%)	-
<i>Non</i>	11 (42.3%)	2 (7.7%)	13 (25.0%)	
<b>PRATIQUE JEUX VIDEOS</b>				0.22 <sup>†</sup>
<i>Oui</i>	10 (38.5%)	5 (19.2%)	15 (28.8%)	
<i>Non</i>	16 (61.5%)	21 (80.8%)	37 (71.2%)	
<b>PRATIQUE INSTRUMENT MUSIQUE</b>				1 <sup>†</sup>
<i>Oui</i>	7 (26.9%)	6 (23.1%)	13 (25%)	
<i>Non</i>	19 (73.1%)	20 (76.9%)	39 (75.0%)	
<b>EXPERIENCE SIMULATEUR CHIRURGICAL</b>				0.25 <sup>†</sup>
<i>Oui</i>	12 (46.2%)	7 (26.9%)	19 (36.6%)	
<i>Non</i>	14 (53.8%)	19 (73.1%)	33 (63.4%)	

**Tableau 1.** Caractéristiques de base des participants

NOTE. Les données sont présentées sous forme de médiane [écart interquartile] ou de nombre (pourcentage).

\* *Test t de Student sur échantillons indépendants*

<sup>†</sup> *Test Chi<sup>2</sup> de Pearson*

METRIQUES	Valeurs (Médiane [Écart Interquartile (IQR), 25% - 75%])		$p^\dagger$
	INTERMEDIAIRE	EXPERT	
Distance_Instrumentale <i>PL</i>	695 mm [IQR, 467-1090]	671 mm [504-846]	0.62
Durée <i>DUR</i> *	116.1 s [82.8-154.2]	70.9 s [56.4-86.3]	<0.01
Vitesse moyenne <i>AV</i> *	6.9 mm.s <sup>-1</sup> [4.7-9.5]	10.0 mm.s <sup>-1</sup> [7.4-11.6]	<0.01
Accélération moyenne <i>AA</i>	92.3 mm.s <sup>-2</sup> [61.0-141.4]	116.6 mm.s <sup>-2</sup> [79.2-144.2]	0.09
Tremblement moyen <i>AJ</i>	3717 mm.s <sup>-3</sup> [2431-6087]	4211 mm.s <sup>-3</sup> [3017-5604]	0.31
Economie <i>ECO</i> *	33.8 mm <sup>2</sup> [14.8-50.5]	19.3 mm <sup>2</sup> [9.1-25.6]	<0.01
Fluidité du Mouvement <i>HMS</i> *	1.5.10 <sup>6</sup> mm.s <sup>-3</sup> [2.6E+5-3.5E+6]	4.6.10 <sup>6</sup> mm.s <sup>-3</sup> [3.1E+5-7.2E+5]	0.05
Fluidité de Longueur d'Arc Spectral <i>SPARC</i> *	-37.8 [-60.7 - -26.5]	-21.2 [-31.0 - -16.8]	<0.01
Profondeur <i>D</i>	123.8 mm [97.1-155.0]	116.6 mm [92.1-138.8]	0.11
Volume de Travail <i>WV</i> *	28236 mm <sup>3</sup> [7193-52229]	12687 mm <sup>3</sup> [5753-22309]	<0.01

**Tableau 2.** Résultats des métriques pour les catégories INTERMEDIAIRE et EXPERT, p-values

Les p-values significatives sont surlignées en gris et marquées d'un \*.

<sup>†</sup> p-value U Mann-Whitney

## **Figure**

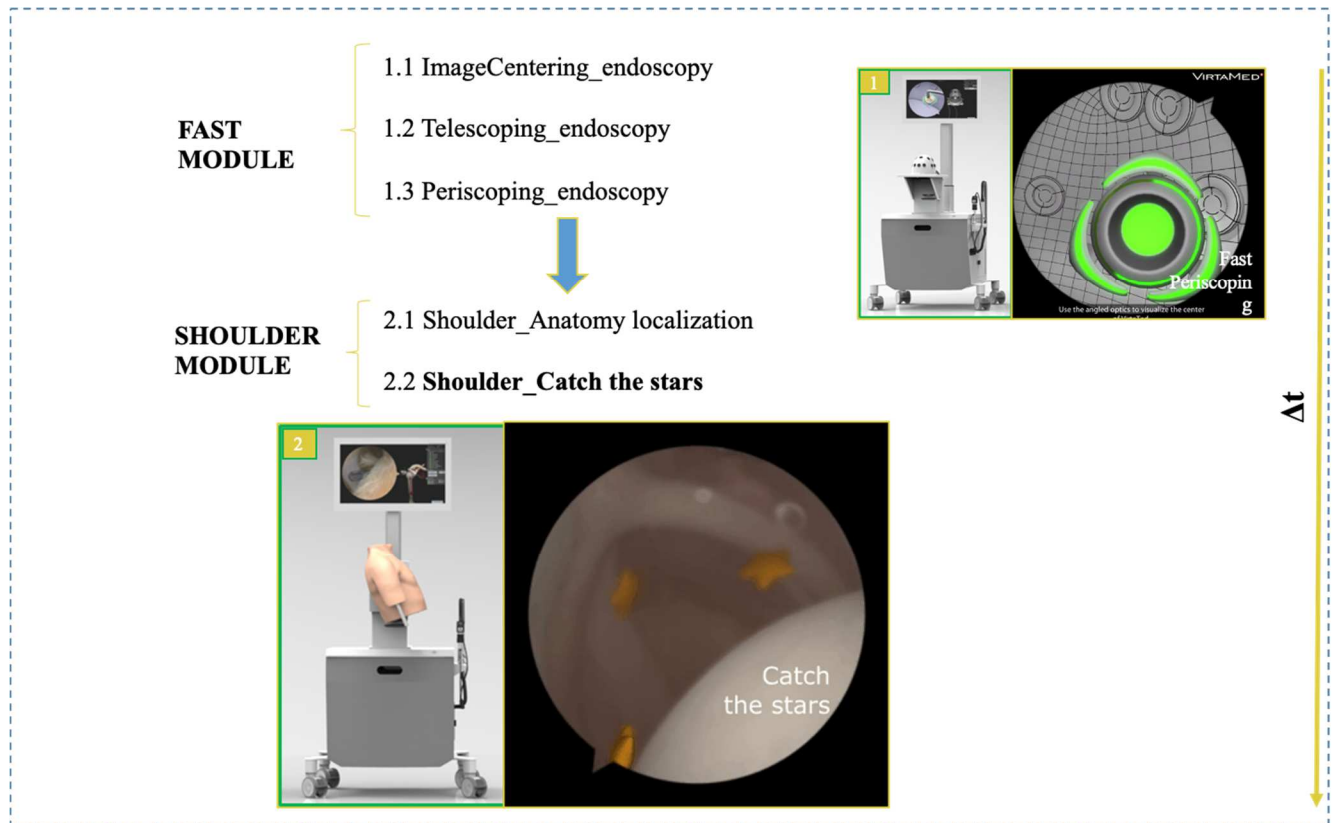
**Figure 1.** Protocole Standardisé (1. Le participant est initié au simulateur avec le module FAST (Fundamentals of Arthroscopy Surgery Training) avec 3 exercices différents 1.1, 1.2 et 1.3 ; 2.1 Le participant est initié à l'anatomie de l'épaule du module SHOULDER en décubitus latéral, 2.2. Le participant réalise l'exercice 3 fois en décubitus latéral)

**Figure 2.** Définitions des métriques

**Figure 3.** Diagrammes moustaches des métriques avec différences significatives (3.1. Durée, 3.2. Vitesse, 3.3. Volume de Travail, 3.4. Economie, 3.5. Fluidité du Mouvement, et 3.6. Fluidité de Longueur d'Arc Spectral)

**Figure 4.** Diagrammes moustaches des métriques sans différences significatives (4.1. Distance Instrumentale, 4.2. Accélération, 4.3. Tremblement, et 4.4. Profondeur)

**Figure 5.** Dendrogramme du Clustering Hiérarchique avec Dynamic Time Warping (En abscisse les trajectoires et en ordonnées la distance entre clusters. Les deux principaux clusters sont représentés, le cluster 1 est uniquement composé d'experts avec une faible distance DTW, le cluster 2 est composé des intermédiaires et de 2 experts avec une distance DTW plus hétérogène.)



**Figure 1.** Protocole Standardisé (1. Le participant est initié au simulateur avec le module FAST (Fundamentals of Arthroscopy Surgery Training) avec 3 exercices différents 1.1, 1.2 et 1.3 ; 2.1 Le participant est initié à l'anatomie de l'épaule du module SHOULDER en décubitus latéral, 2.2. Le participant réalise l'exercice 3 fois en décubitus latéral)



---

**DEFINITION DES METRIQUES**

---

<b>Durée (<i>DUR</i>)</b>	Temps d'exécution entre la première fois que la main déplace la sonde, jusqu'à ce que la sonde soit relâchée à la fin de la tâche. Mesurée en secondes (s).
<b>Distance Instrumentale (<i>PL</i>)</b>	Somme des distances totales parcourues par les sondes (caméra et pince) pendant l'exécution de la tâche. Mesurée en millimètres (mm).
<b>Vitesse Moyenne (<i>AV</i>)</b>	Vitesse linéaire moyenne de la sonde pendant la tâche. Mesurée en millimètres par seconde. ( $\text{mm}\cdot\text{s}^{-1}$ ).
<b>Accélération Moyenne (<i>AA</i>)</b>	Accélération instantanée moyenne de la sonde pendant la tâche. Mesurée en millimètres par seconde au carré. ( $\text{mm}\cdot\text{s}^{-2}$ ).
<b>Tremblement Moyen (<i>AJ</i>)</b>	Tremblement moyen (dérivée de l'accélération) pendant la tâche, également appelée "fluidité". Mesuré en millimètres par seconde au cube. ( $\text{mm}\cdot\text{s}^{-3}$ ).
<b>Volume de Travail (<i>WV</i>)</b>	Volume de la coque convexe pour chaque trajectoire. Mesuré en millimètres cubes. ( $\text{mm}^3$ ).
<b>Economie de Mouvement (<i>ECO</i>)</b>	Rapport entre le volume de Travail et la Distance Instrumentale. Mesuré en $\text{mm}^2$ .
<b>Fluidité du Mouvement (<i>HMS</i>)</b>	Variation totale de l'accélération de la pointe de l'instrument. Mesurée en millimètres par seconde au cube. ( $\text{mm}\cdot\text{s}^{-3}$ ) et normalisé par la durée de procédure.
<b>Fluidité de Longueur d'Arc Spectral (<i>SPARC</i>)</b>	Fluidité du profil de vitesse du mouvement linéaire par la méthode de la longueur d'arc spectrale. Mesure sans unité.
<b>Profondeur (<i>D</i>)</b>	Distance totale parcourue par la pointe de l'instrument le long de l'axe z. Mesurée en millimètres (mm).

---

**Figure 2.** Définitions des métriques

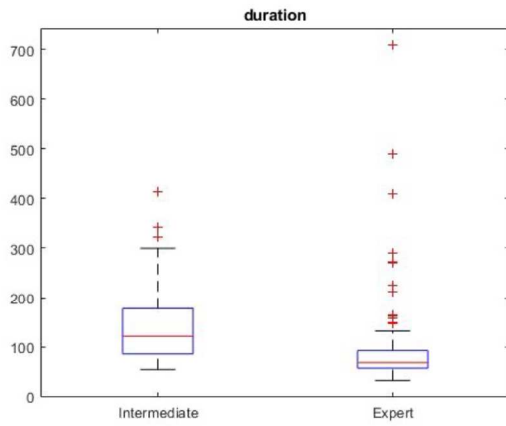


Figure 5.1. Durée

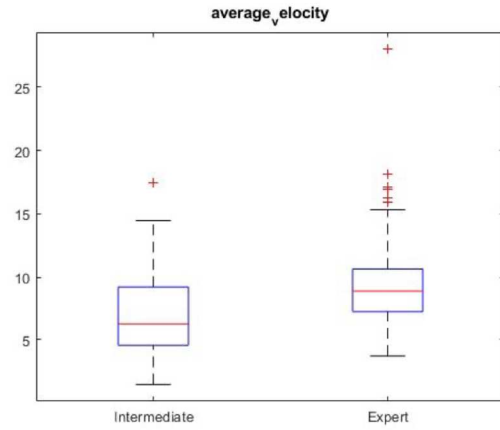


Figure 5.2. Vitesse

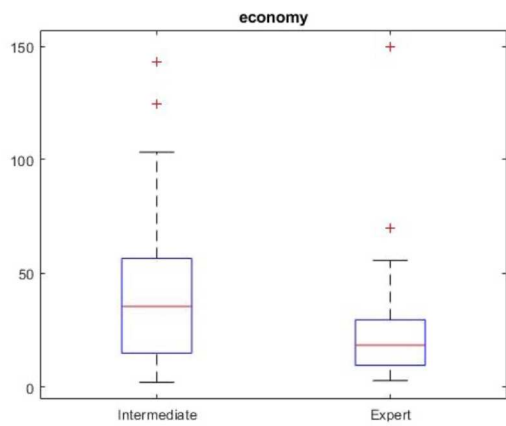


Figure 5.3. Economie

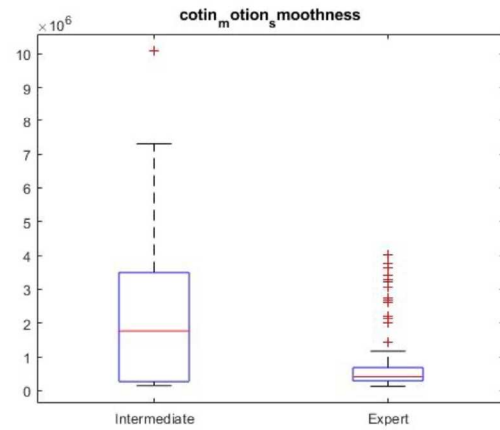


Figure 5.4. Fluidité

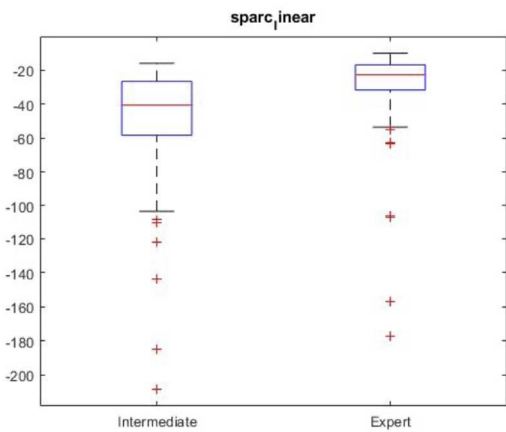


Figure 5.5. Fluidité de longueur d'Arc Spectral

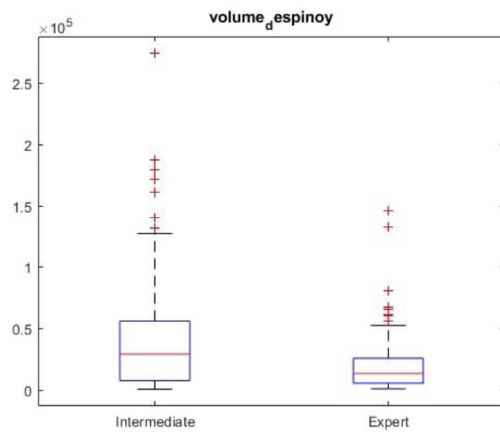
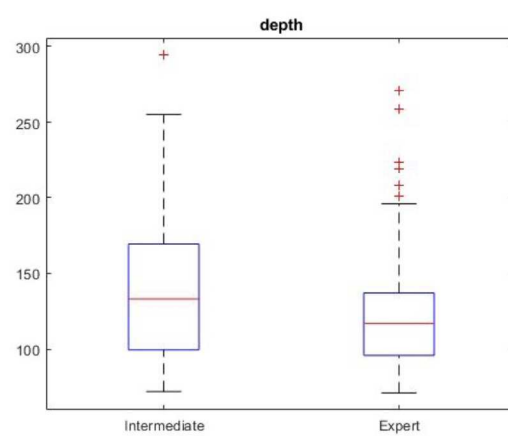
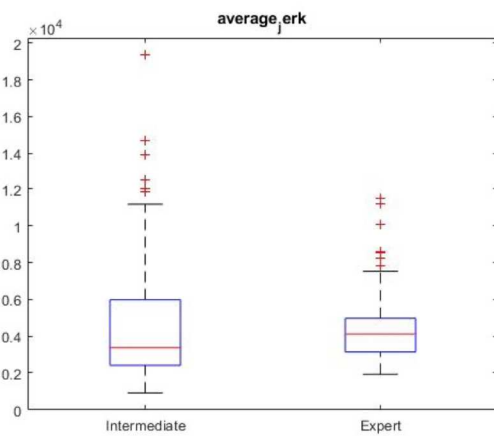
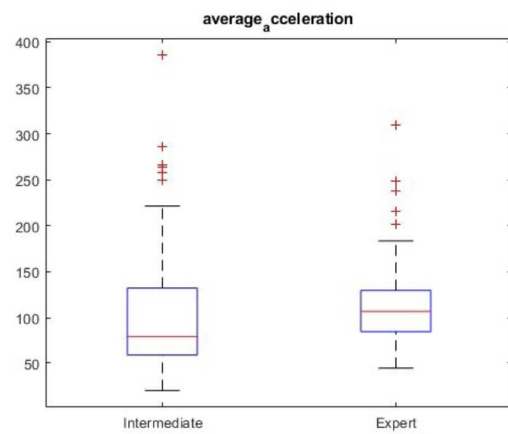
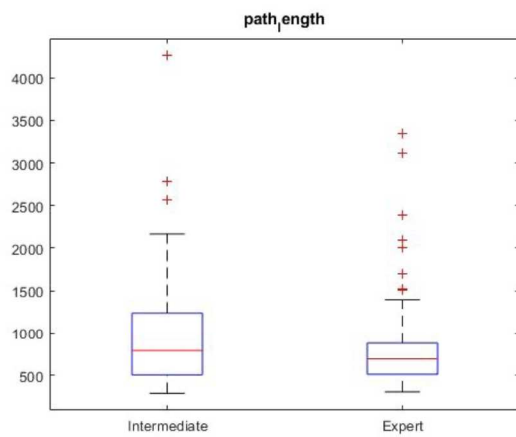
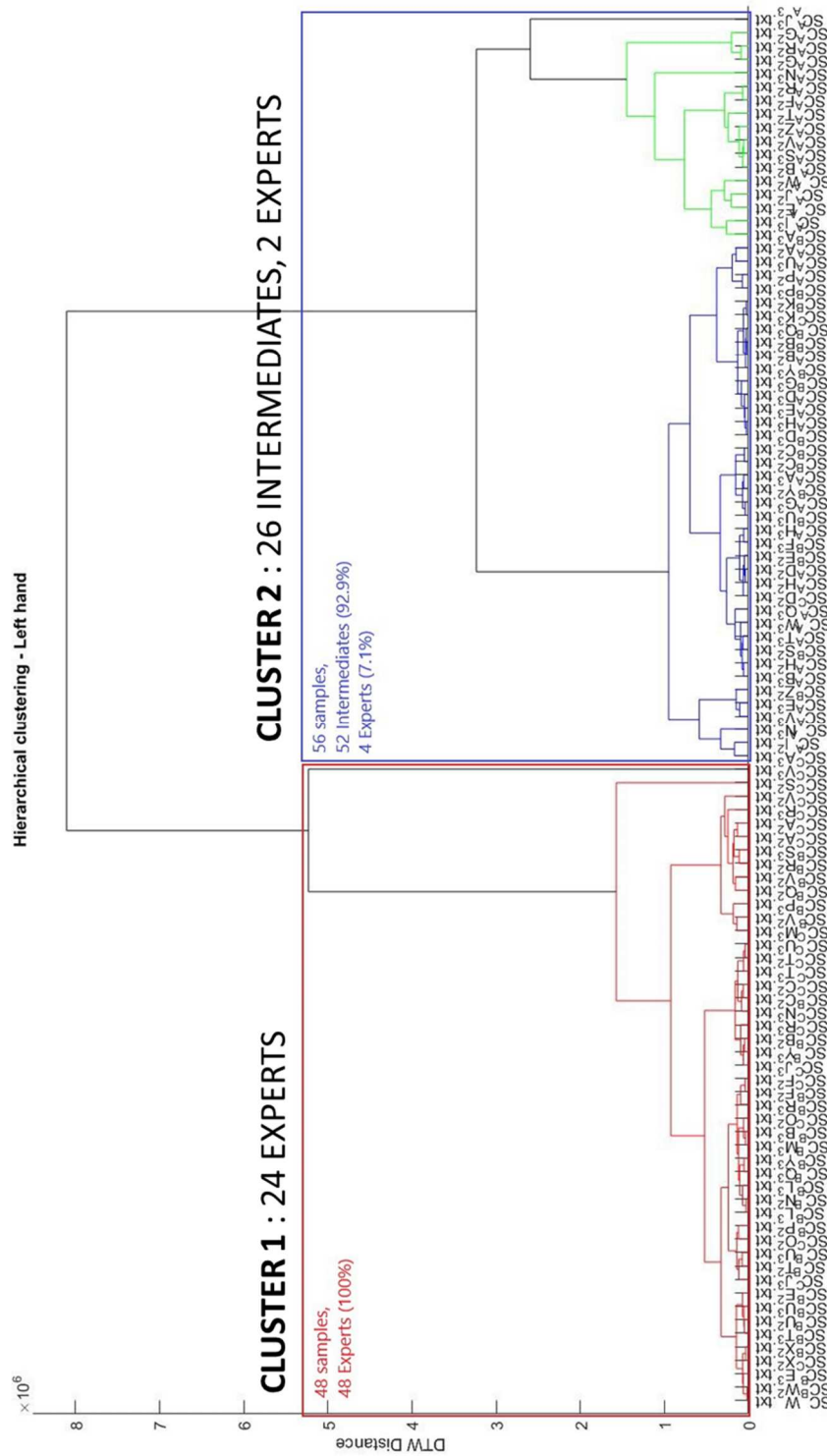


Figure 5.6. Volume de Travail

**Figure 3.** Diagrammes moustaches des métriques avec différences significatives (3.1. *Durée*, 3.2. *Vitesse*, 3.3. *Volume de Travail*, 3.4. *Economie*, 3.5. *Fluidité du Mouvement*, et 3.6. *Fluidité de Longueur d'Arc Spectral*)



**Figure 4.** Diagrammes moustaches des métriques sans différences significatives (4.1. Distance Instrumentale, 4.2. Accélération, 4.3. Tremblement, et 4.4. Profondeur)



**Figure 5.** Dendrogramme du Clustering Hiérarchique avec Dynamic Time Warping (En abscisse les trajectoires et en ordonnées la distance entre clusters. Les deux principaux clusters sont représentés, le cluster 1 est uniquement composé d'experts avec une faible distance DTW, le cluster 2 est composé des intermédiaires et de 2 experts avec une distance DTW plus hétérogène.)