Corrected January 6, 2025—see https://doi.org/10.1123/jab.2024-0347

Journal of Applied Biomechanics, 2025, 41, 37-46
https://doi.org/10.1123/jab.2023-0276

Human Kinetics
ORIGINAL RESEARCH

Effects of Different Inertial Measurement Unit Sensor-to-Segment Calibrations on Clinical 3-Dimensional Humerothoracic Joint Angles Estimation

Alessandro Bonfiglio,[1,2,3] Elisabetta Farella,[3] David Tacconi,[1] and Raoul M. Bongers[4]

[1]Euleria Health, Rovereto, Italy; [2]DISI—Information Engineering and Computer Science Department, University of Trento, Trento, Italy; [3]E3DA—Energy Efficient Embedded Digital Architectures, Fondazione Bruno Kessler, Trento, Italy; [4]Department of Human Movement Sciences, University Medical Center Groningen, Groningen, The Netherlands

Calibrating inertial measurement units (IMUs) involves converting orientation data from a local reference frame into a clinically meaningful reference system. Several solutions exist but little work has been done to compare different calibration methods with each other and an optical motion capture system. Thirteen healthy subjects with no signs of upper limb injury were recruited for this study and instrumented with IMU sensors and optical markers. Three IMU calibration methods were compared: N-pose calibration, functional calibration, and manual alignment. Subjects executed simple single-plane single-joint tasks for each upper limb joint as well as more complex multijoint tasks. We performed a 3-way analysis of variance on range of motion error, root mean squared error, and offset to assess differences between calibrations, tasks, and anatomical axes. Differences in the 3 IMU calibrations are minor and not statistically significant for most tasks and anatomical axes, with the exception of the offset interaction calibration × axes ($P < .001$, $\eta^2_G = .056$). Specifically, manual alignment gives the best offset estimation on the abduction/adduction and internal/external rotation axes. Therefore, we recommend the use of a static N-pose calibration procedure as the preferred IMU calibration method to model the humerothoracic joint, as this setup is the simplest as it only requires accurate positioning of the trunk sensor.

Keywords: Euler angles, alignment, orientation

Inertial measurement units (IMUs) have been widely used in rehabilitation due to their portability, ease of use, and low cost, especially when compared with alternative technologies such as optical motion capture. IMUs allow health care professionals to calculate movement parameters, such as range of motion (ROM), to evaluate a patient's performance and monitor their rehabilitation progress.[1] However, among all joints in the human body, modeling the shoulder joint is particularly challenging due to the number of joints involved (scapulothoracic, sternoclavicular, glenohumeral, and acromioclavicular), which yields a large ROM and 3 rotational degrees of freedom.

To be effective in a clinical setting, IMUs require an initial sensor-to-segment calibration to align their internal reference frame to the bone anatomical frame on which they are positioned. Importantly, sensor to segment (for brevity referred to as "calibration") mentioned in this work is not to be confused with sensor calibration, which is specific to the sensing elements within the IMU.[2] As calibration can be time consuming[1,3] and its accuracy can vary from subject to subject,[4] it is vital to quantify the reliability of IMU calibration techniques to allow the diagnosis and treatment of movement disorders.[5] Current literature has presented several calibration procedures that rely on holding a known pose (static calibration),[1,6–10] the execution of a strict movement (functional calibration [FC]),[11–14] accurate sensor positioning (manual sensor alignment), or a combination of these.[15,16] Although these have been extensively validated individually against optical reference systems, comparing the performance of different calibration techniques is challenging due to several elements that influence the accuracy of IMU joint angle estimation, namely: sensor placement,[17] subject-to-subject variability,[18] sensor orientation estimation error (bias, noise, and drift[19]), and anatomical sensor-to-segment alignment. To best single out the effect of sensor-to-segment calibration from all other variables that affect the computation of IMU joint angles, several calibrations should be applied to the same movement data of the same subject. Interestingly, Zhu et al[20] quantified the influence of sensor orientation estimation error and sensor-to-segment alignment on the overall humerothoracic joint angle estimation. To the best of our knowledge, and within an already limited body of literature on upper limbs[3,17,20] compared with lower limbs,[21,22] only Bouvier et al[17] followed this strict methodology when comparing several sensor-to-segment calibrations against optical motion capture.

Another important element that allows IMU-based joint angles to be employed for clinical applications is accuracy on both main joint axis (joint angle where movement occurs) and nonmain joint axis. Most of the existing literature on shoulder modeling focus on estimating the validity of their calibration methods only on the main joint axes[3,11,15] (ie, shoulder flexion axis during a shoulder flexion movement). However, accurately modeling nonmain joint axes can further enhance the applicability of IMU in disciplines such as rehabilitation where full 3-dimensional (3D) motion needs to be accurately tracked.[19]

Farella https://orcid.org/0000-0001-9047-9868
Bongers https://orcid.org/0000-0002-3518-7464
Bonfiglio (alessandro.bonfiglio1@gmail.com) is corresponding author, https://orcid.org/0009-0003-2098-1870

This paper focuses on the humerothoracic joint among all joints involved in the shoulder complex and aims to compare and validate 3 calibration methods to estimate its 3D angles from raw IMU orientation data during a variety of movement tasks. Therefore, our goals are to (1) assess accuracy and precision of each calibration against an optical motion capture system and (2) compare the 3 calibration methods against one another to provide recommendations on the most optimal shoulder joint IMU calibration.

Methods

Subject Recruitment

Thirteen healthy subjects (age 27.6 [6.1] y, weight 64.0 [13.3] kg, height 171.2 [6.1] cm) with no sign of upper limb injury, pain, or movement restrictions or disorders were recruited. This study received ethical approval from the ethical board of the University Medical Center Groningen, Groningen, The Netherlands (nr RR10982) and was performed in accordance with the Declaration of Helsinki.

Equipment and Materials

Five IMU sensors (Xsens DOT) were placed on each participant's sternum, scapula, upper arm, lower arm, and hand unilaterally on the right side of the body. Five 3-marker clusters were placed on top of each sensor to enable data acquisition with an active motion capture camera system (OptoTrak Certus) that served as a reference system.[19] Figure 1A depicts the full setup of both marker clusters and IMU sensors on a subject. Each combination of IMU sensor and marker cluster was firmly attached to the patient's skin by means of Kinesio Tape (Take Sport, Italy) while maintaining full marker visibility.

Experimental Protocol

The experimental protocol consisted of 3 phases: landmark digitization, sensor calibration, and task execution. Figure 1 shows details about the sensor setup and the data flowchart (Figure 1C) after acquisition for both Optotrak and IMU sensors. Details on the data acquisition protocol are provided next and in Table 1.

1. Landmark digitization: A 4-marker rigid probe was used to acquire static positional data of each body landmark presented in Wu et al.[23] These data were subsequently related to the position of each segment's cluster to compute the position of virtual body landmarks that were used to calculate joint angle data for the Optotrak system (optical motion tracking system [OPTO]); see "Optotrak Reference Model" section.
2. Calibration tasks: Subjects were required to hold a static N-pose (Figure 1) for 3 seconds and perform a series of single-plane movements for each joint, as described in Table 1, to estimate the rotation axis relative to each joint and degree of freedom.
3. Exercise tasks: Subjects performed a series of simple single-plane movements and complex multijoint tasks resembling activities of daily living that could be commonly executed in a real-life setting, as described in Table 1.

Optotrak Reference Model

Each segment reference for the optical system was computed following the International Society of Biomechanics recommendations from Wu et al.[23] For the shoulder joint, the glenohumeral rotation joint center was obtained using the formula described in Rab et al,[5] as it was considered the least invasive and most accurate estimation method described in the literature for this joint.[6] For the computation of the humerus reference frame, the H1 humerus model from Wu et al[23] was chosen.

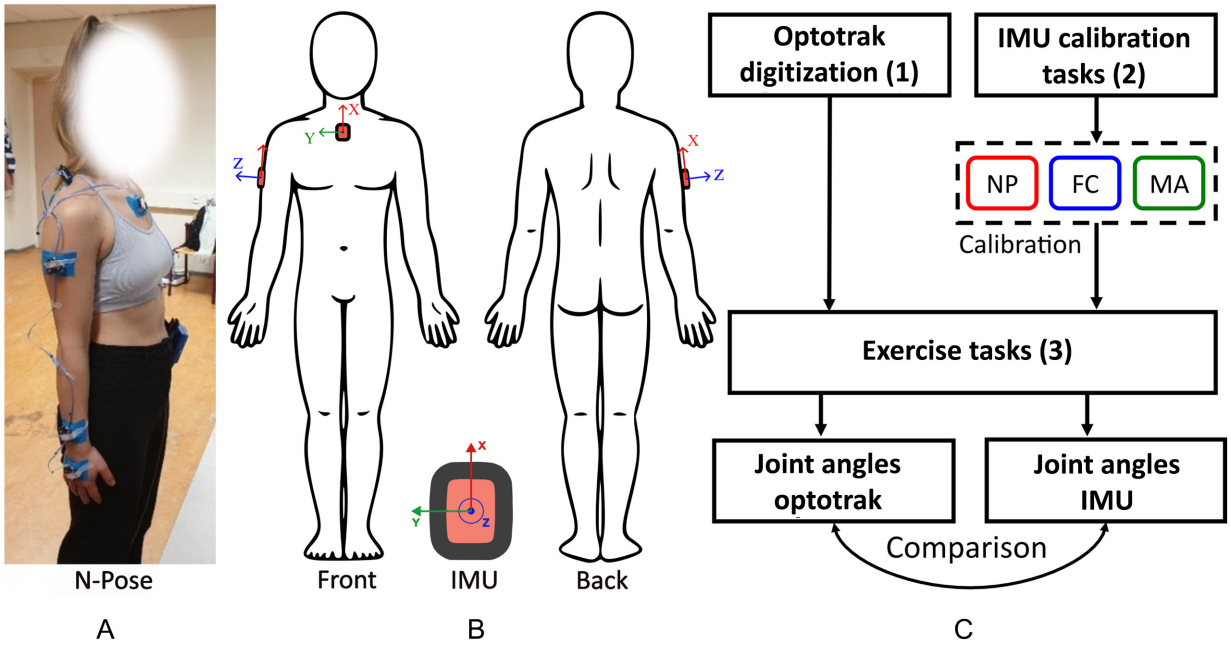

Figure 1 — (A) Subject setup with IMU and active marker clusters sensors placed on the thorax, scapula, upper arm, lower arm, and hand while the subject is standing in N-pose. Only the thorax and upper arm markers were used for this study. (B) Representation of the biomechanical model with IMU local axis orientation on the body: X-axis points cranially; Z-axis points forward and laterally for the trunk and upper arm sensor, respectively. (C) Data flow diagram of Optotrak and IMU data. Joint angle data from Optotrak is directly computed from positional markers. Joint angles from IMU are computed after calibration data recorded during calibration tasks is applied to the exercise data. FC indicates functional calibration; IMU, inertial measurement unit; MA, manual alignment; NP, N-pose calibration.

Table 1 Detailed Description of the Experimental Protocol, Including the IMU Calibration and Exercise Tasks

Body segment	Name	Calibration tasks
Multijoint	N-pose	The subject stands upright with their chest straight but not overextended. The arms are completely straight and kept alongside the body. Palms point inward and touch the side of the hip or upper leg. Feet are parallel and kept about 20 cm apart. The subject is required to keep this position for 3 s.
Shoulder	Shoulder flexion/extension	The subject stands upright with the arms hanging in front of the body and holds a stick with both hands at about shoulder width and thumbs pointing medially. From this position, the subject performs 5 shoulder flexion movements from 0° to about 90° by raising their upper arm in front of them while keeping their arms straight.
	Shoulder abduction/adduction	Starting from N-pose, the subject performs 5 shoulder abduction movements from 0° to about 90° by raising their upper arm laterally and keeping their arms straight.
	Shoulder internal/external rotation	The subject is seated with the olecranon supported on the table, elbow flexed at 90°, and the elbow touching the side of the body. From this position, the subject performs 5 shoulder internal/external rotation movements at full ROM without detaching their elbow from the body.
		Exercise tasks
Shoulder	Shoulder flexion/extension	The subject stands upright and holds a stick with their hands at about shoulder width and thumbs pointing medially. From this position, the subject performs 5 shoulder flexion movements from 0° to about 90° by raising their upper arm in front of them and keeping their arms straight.
	Shoulder abduction/adduction	Starting from an N-pose, the subject performs 5 shoulder abduction movements by lifting their arm laterally and keeping the elbow straight. The movement range is from about 0° to the participant's maximum shoulder abduction.
	Shoulder internal/external rotation	Starting with the elbow flexed at 90° and the elbow touching the side of the body, the subject performs 5 times a shoulder internal/external rotation movement at full ROM. Slight detachment of the elbow from the side of the body is allowed.
Multijoint	Drinking	A paper cup of water is placed on a shelf at about eye height (1.6 m). Starting from an N-pose, participants reach with their right arm toward the cup, grab it, bring it to their mouth, simulate drinking a sip of water, and put the cup back on the shelf. This movement is repeated 5 times.
	Box off shelf	Starting from an N-pose, participants stand in front of a shelf and move a shoe box from a higher shelf (1.48 m vertical height) to a lower shelf (0.96 m) and then put it back. The movement is repeated 5 times.
	Drawing	Participants are seated and required to draw imaginary circles anticlockwise with their right arm by sliding a pen on a table in front of them. This movement is repeated 5 times. Participants are instructed to involve some degrees of shoulder, elbow, and wrist motion without specifying the size of the circles.

Abbreviations: IMU, inertial measurement unit; ROM, range of motion. Note: The speed of each movement performed was self-selected and generally slow to mimic typical shoulder rehabilitation tasks or common activities of daily living.

IMU Calibration

Three IMU calibration methods were developed, which took advantage of different movements and poses to estimate individual segment references. Despite a full upper body setup, only the trunk and upper arm sensors were used for joint angle analysis. After sensor placement in static N-pose, the trunk sensor reference frames corresponded to X: pointing up, Y: pointing right, and Z: pointing forward, whereas for the upper arm, X: pointing up, Y: pointing backwards, and Z: pointing right. Subsequently, vector reference axes were computed for each calibration to define a sensor-to-segment quaternion (^{SI}q) to describe a unique rotation that translated the sensor's reference frame into the bone reference frame:

- Static N-pose (NP): When performing the N-pose, the trunk sensor and the gravity vector served as a reference to align each sensor. The advantage of this method was that the final segment reference frames solely relied on the trunk sensors' orientation.
- Functional calibration (FC): The joint rotation axis associated with each single-plane motion was computed for each body segment and degree of freedom. During each FC task, the axis of rotation was extracted from the gyroscope data by following the procedure described by Stančin and Tomažič.[9] For the upper arm, the shoulder flexion and abduction tasks were used to estimate the shoulder flexion and abduction axes, respectively. These estimations were then used to construct individual segment reference frames. Further information about the construction of the individual references is provided in Table 2. The trunk segment was calibrated in a static fashion as Cottam et al[8] showed no added benefit in performing a FC.
- Manual alignment (MA): Each sensor was carefully positioned on the body to ensure that the trunk and upper arm anatomical references were manually aligned with their respective sensor's reference frame. These references were then rotated to match as closely as possible the ISB conventions (Table 2). The N-pose was taken as a reference pose for this method to compare data easily. Although this method was computationally simple, consistent positioning and alignment of the sensor to the underlying segment over several sessions could be challenging and, therefore, influence joint angle accuracy.

Detailed specifications of each calibration frame can be found in Table 2. Note that the rotation matrices described in Table 2 specify the orientation of the sensor with respect to the underlying body segment; in other words, they represent the rotation needed to translate the orientation data from the IMU sensor to the orientation of the underlying bone-embedded reference frame, as defined by the ISB conventions presented in Wu et al,[23] while the subject stands in N-pose. Reference axes from Table 2 were used to compute a sensor-to-segment rotation matrix, as shown in Equation 1. Each rotation matrix was converted into quaternions (Equation 2), and the calibration data

Table 2 Detailed Description and Definition of Reference Frames Associated With Each Segment for Each IMU Calibration Method, Which Are Used to Compute a Sensor-to-Segment Rotation (^{SI}R)

Segment	Axis	Operations
N-pose		
Thorax	Y	$\bar{y} = [0,0,1]$: cranial
	Z	$\bar{z} = \frac{^{GI}TH[z]\times\bar{y}}{\|^{GI}TH[z]\times\bar{y}\|}$: lateral
	X	$\bar{x} = \frac{\bar{y}\times\bar{z}}{\|\bar{y}\times\bar{z}\|}$: forward
Upper arm	X	$\bar{x} = \bar{x}_{Thorax}$
	Y	$\bar{y} = \bar{y}_{Thorax}$
	Z	$\bar{z} = \bar{z}_{Thorax}$
Functional		
Thorax	X	$\bar{x} = \bar{x}_{Thorax-NP}$
	Y	$\bar{y} = \bar{y}_{Thorax-NP}$
	Z	$\bar{z} = \bar{z}_{Thorax-NP}$
Upper arm	Y	$y = {}^{UA}LA_{I/E}$: cranial
	Z	$\bar{z} = \frac{\bar{y}\times^{UA}UA_{F/E}}{\|\bar{y}\times^{UA}UA_{F/E}\|}$: lateral
	X	$\bar{x} = \frac{\bar{y}\times\bar{z}}{\|\bar{y}\times\bar{z}\|}$: forward
Manual		
Thorax	X	$\bar{x} = \bar{x}_{Thorax-NP}$
	Y	$\bar{y} = \bar{y}_{Thorax-NP}$
	Z	$\bar{z} = \bar{z}_{Thorax-NP}$
Upper arm	X	$\bar{x} = {}^{GI}UA[y]$
	Y	$\bar{y} = {}^{GI}UA[x]$
	Z	$\bar{z} = {}^{GI}UA[z]$

Abbreviations: IMU, inertial measurement unit; NP, N-pose calibration; SI, segment-to-inertial; UA, upper arm. Note: The upper left superscript represents the reference frame that each segment refers to: GI = global to inertial. Details on the IMU reference frame are shown in Figure 1.

were applied to the run time data (Equation 3) for each sensor. Finally, the joint angle between trunk and upper arm segments was calculated through quaternion multiplication, as shown in Equation 4, and then converted into Euler angles, as recommended by Wu et al.[23] Superscripts and symbols refer to G = global; S = sensor; I = inertial (sensor); R = rotation matrix; q = quaternion; ⊗ = Hamilton's quaternion product; and ()* = quaternion conjugate.

$$R = [\bar{x} \quad \bar{y} \quad \bar{z}] = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}, \tag{1}$$

$$^{SI}R \rightarrow {}^{SI}q, \tag{2}$$

$$^{GS}q_{runtime} = {}^{GI}q_{runtime} \otimes \left({}^{SI}q_0\right)^*, \tag{3}$$

$$q_{joint} = \left({}^{GS}q_{proximal}\right)^* \otimes {}^{GS}q_{distal}. \tag{4}$$

Data Analysis

Data analysis was performed in MATLAB (The MathWorks Inc, version R2022b). Quaternion joint angles from Optotrak and IMU were converted into Euler angles by choosing the following Euler sequences: *ZXY* for shoulder flexion, rotation, and all multi-joint tasks and *XYZ* for shoulder abduction. This choice was adopted to match the first rotation axis with the main joint angle of the movement when possible. Following joint angle decomposition, repetitions were singled out by identifying 2 subsequent joint angle minima along the joint axis where movement occurs. Finally, joint angles obtained with each IMU model were compared against the Optotrak optical system by computing the following quantities: ROM, ROM error (ROM ε), root mean squared error (RMSE), and offset. These quantities were calculated in similar papers in the field, such as Vitali and Perkins,[1] Filippeschi et al,[3] Cutti et al,[10] and Rigoni et al.[11]

- ROM was calculated as a difference between the maximum and minimum joint displacement (Equation 5) where θ represents the joint angle measured by each system. This was applied to both OPTO and IMU joint angles.
- RMSE was calculated according to Equation 6 where θ is the mean joint angle of each system during a single repetition. This formula was chosen to decouple the offset component from the RMSE calculation as it was analyzed separately.
- Offset was calculated as described in Equation 7.
- ROM ε in degrees (Equation 8) was also computed.

$$\text{ROM } [^\circ] = \theta_{MAX} - \theta_{MIN}, \tag{5}$$

$$\text{RMSE } [^\circ] = \sqrt{\frac{\sum_{i=1}^{N} |(\theta_{OPTO}(i) - \theta_{OPTO}) - (\theta_{IMU}(i) - \theta_{IMU})|^2}{N}}, \tag{6}$$

$$\text{Offset } [^\circ] = \theta_{OPTO} - \theta_{IMU}, \tag{7}$$

$$\text{ROM } \varepsilon \ [^\circ] = \text{ROM}_{OPTO} - \text{ROM}_{IMU}. \tag{8}$$

Outlier Removal

Due to a significant variability in joint angle data, outlier removal was performed through interquartile range analysis. Each IMU model performance index was analyzed individually, and data points that exhibited a value larger than 1.5 interquartile range from the 25th and 75th percent quartiles were identified and removed from the data pool.

Statistical Analysis

To evaluate differences between models, we performed a 3-way repeated-measures analysis of variance with $\alpha = .05$ chosen as significant level. We performed separate analyses for dependent variable (ROM ε, RMSE, and offset) with calibration methods (NP, FC, and MA); movement tasks (shoulder flexion, abduction, rotation, drinking, box off shelf, and circles); and anatomical planes (flexion/extension, abduction/adduction, and internal/external rotation) as independent variables. The 3-way repeated-measures analysis of variance allowed the exploration of interaction effects among different levels/factors while limiting type II error. Generalized eta squared (η^2_G) was chosen to calculate effect size,[24] and it was interpreted according to $\eta^2_G = .02$ as a small effect, $\eta^2_G = .13$ as a medium effect, and $\eta^2_G = .26$ as a large effect.[25] The software JASP (JASP Team 2023, version 0.17.2.1) was used to perform the statistical analysis. We used Mauchly test to ensure sphericity in the data and performed a Greenhouse–Geisser correction when necessary.

Results

Figure 2 depicts the joint angles computed with OPTO and the 3 IMU calibrations during all the tasks for one subject chosen randomly, displaying shoulder joint angles on the 3 anatomical planes: flexion/extension, abduction/adduction, and internal/external rotation. Qualitatively, the performance of all IMU calibration methods was comparable with one another for most tasks and axes, and differences between the methods were minor. However, more noticeable differences could be found with respect to the optical reference system on specific movements and axes, such as the internal/external rotation axis during flexion or the abduction/adduction axis during abduction (Figure 2).

NP calibration performed well on the flexion/extension axis during rotation (ROM ε; Figure 3A1) and on the internal/external rotation axis during abduction (ROM ε; Figure 3A3), but it showed large errors on the flexion/extension axis during abduction (ROM ε; Figure 3A1). FC performed well on the abduction/adduction axis during rotation (ROM ε and RMSE; Figure 3A2 and 3B2) but showed large errors on the flexion/extension axis during abduction (ROM ε, RMSE, and offset; Figure 3A1–3C1). MA showed the least offset on the flexion/extension axis during abduction (Figure 3C1) but the largest offset on the abduction/adduction axis during rotation (Figure 3C2). With the exception of these specific cases, the 3 calibration methods were comparable in terms of ROM ε, RMSE, and offset for the large majority of movements and joint axes. This was especially noticeable in joint tasks, where the performance of the 3 calibration methods was similar across all movements and joint axes (Figure 3).

All calibrations generally showed ROM error, RMSE, and offset values closer to 0, indicating better performance, and more contained variability between participants on main joint axes (eg, shoulder flexion task on the flexion/extension axis) compared with nonmain joint axes (eg, shoulder flexion task on the abduction/adduction axis; Figure 3). However, 2 exceptions were the abduction and rotation tasks. Shoulder abduction on the abduction/adduction axis showed worse performance on the main joint axis compared with nonmain joint axes for all metrics considered (Figure 3A1–3C3). Similarly, shoulder rotation displayed worse offset on main joint axes compared with nonmain joint axes (Figure 3C1–3C3).

No statistical differences were found between calibration methods in ROM ε and RMSE (Figure 3). However, the offset main effect and its interaction effect calibration × axes (Table 3; Figure 4) were statistically significant. In particular, the interaction effect (Figure 4B) highlighted most of the differences in performance across joint axes. Such an interaction (Figure 4B) indicated that FC performed worse than NP on the abduction/adduction axis (offset = $-14.06° \pm 6.01°$) and internal/external rotation axis (offset = $34.78° \pm 6.01°$), whereas MA performed better than NP and FC (MA offset abduction/adduction = $-7.74° \pm 6.00°$; FC offset internal/external rotation = $15.09° \pm 6.01°$).

Discussion

We have presented an in-depth analysis of the 3 most commonly employed IMU calibration methods for the humerothoracic joint during a variety of single-joint and multi-joint tasks by validating the performance of each calibration against an optical motion capture system and by comparing them among each other. As the main advantage of IMUs is their ease of use in real-life scenarios, we investigate their accuracy during tasks such as reaching and drinking to clarify the best approach to model the humerothoracic joint in a known task or, more generally, when the movement is unknown a priori. Our ultimate goal is to enable the widespread use of IMU technology in various fields, including rehabilitation, where accurate and precise joint angle measurement is key to diagnosing movement disorders or assessing rehabilitation progress.[13,26]

We analyzed ROM ε, RMSE, and offset as performance metrics as these are most commonly used in biomechanics research[3,27] and can facilitate comparison with similar studies. ROM is commonly monitored in rehabilitation as one of the primary recovery indicators for most musculoskeletal conditions,[28–30] and its measurement is often a feature of IMU rehabilitation devices[26,31–34]; therefore, quantifying its measurement accuracy is vital for clinical applications. In biomechanics studies, standard RMSE assesses curve similarity and offset between 2 motion signals simultaneously and can be found in most studies in the field.[1,3,10,18,30,32,35,36] However, the sensor-to-segment calibration required by the IMU could introduce a significant amount of offset due to holding a known pose, which might not affect curve similarity during motion. Thus, we presented these 2 parameters separately to facilitate the development of novel IMU calibration procedures that can compensate for the offset component from the joint angle estimation. More accurate joint angles could be translated to rehabilitation practice and applied to feedback control in video games,[32] character animation,[30] or movement sonification,[36] to name a few. Furthermore, this study investigated several simple and complex movements across all 3D joint angle axes, whereas the largest body of literature focuses on the main anatomical axis where movement occurs[15,35,37] (ie, flexion/extension axis during a shoulder flexion movement). Although this is suitable for general rehabilitation applications, there are scenarios wherein high accuracy on nonmain joint axes might be warranted. For instance, stroke patients commonly display excessive shoulder abduction movements during reaching that mainly involve shoulder flexion.[38,39] Alternatively, accurate 3D joint angles could be implemented in biofeedback control loops of video games to prevent compensatory movements.[40,41]

Besides minor differences among calibration methods in extremely specific combinations of movements and joint axes, our results show, overall, no significant differences in performance among calibrations for the ROM ε and RMSE variables and significant differences for offset main effect and calibration × axes interaction with a small effect size. Specifically, the main effect of calibrations is mediated by its larger interaction effect, which indicates significant differences in calibration performance as the movement axis changes (Figure 4B). In the calibration × axes interaction (Figure 4B), FC showed the worst performance on the abduction/adduction and internal/external rotation axes, whereas MA performed best on the same joint axes. When considering the overall performance of the 3 calibrations, NP shows similar or smaller ROM error, RMSE, and offset than FC and MA for the majority of the tasks analyzed. One exception to this is MA, which displays marginally better performance than NP during movements occurring mainly on the abduction/adduction or internal/external rotation axes. Nevertheless, practical considerations are necessary when choosing an IMU calibration, especially in a clinical setting, as the procedure required to perform NP, FC, and MA is largely different. NP requires holding an N-pose for a few seconds; FC requires holding a known pose for trunk alignment and performing several humeral elevation and abduction motions, therefore requiring at least 5 minutes per subject; and MA requires holding an N-pose as well as careful positioning and aligning the sensors

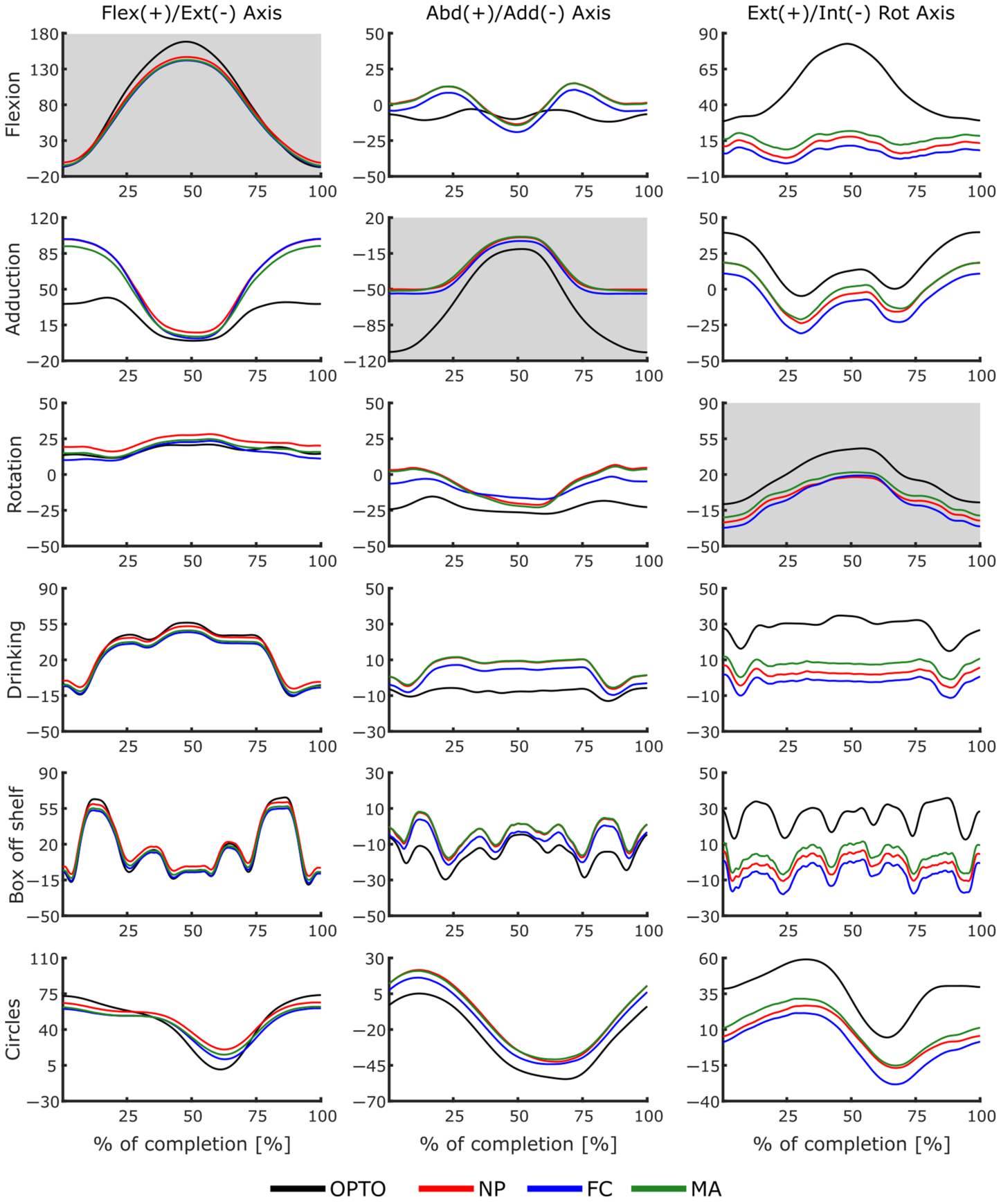

Figure 2 — Joint angles relative to one subject of OPTO and the 3 IMU calibrations for all tasks (rows) and all joint axes (columns). The *x*-axis represents the percentage of movement completion (0%–100%) from the N-pose resting position. Charts with a gray background show the main joint angle axis where movement occurs. Abd/Add indicates abduction/adduction; Ext/Int Rot, external/internal rotation; FC, functional calibration; Flex/Ext, flexion/extension; IMU, inertial measurement unit; MA, manual alignment; NP, N-pose calibration.

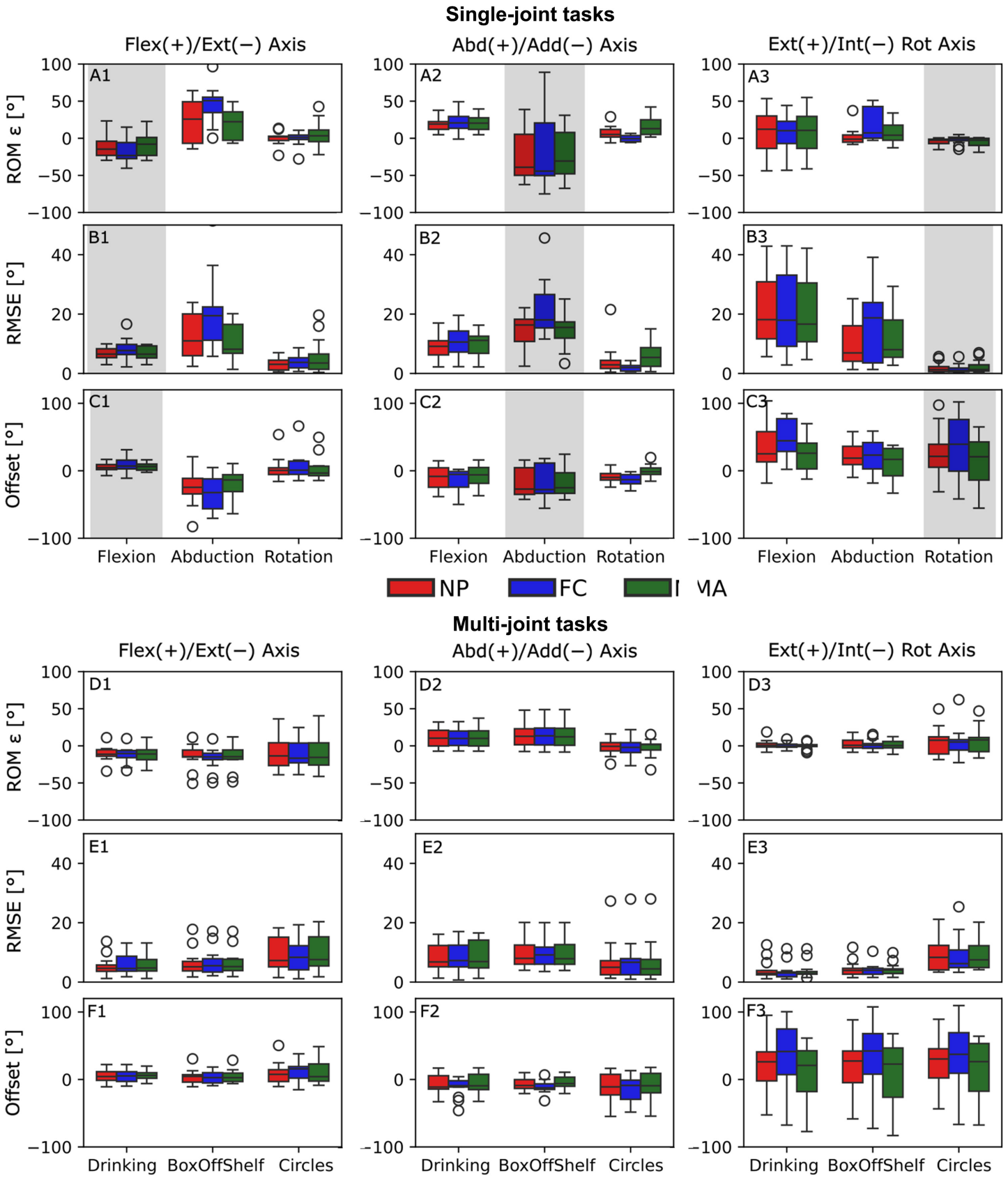

Figure 3 — Box-and-whisker plots of ROM error, RMSE, and offset calculated across all movement axes (Flex/Ext, Abd/Add, and Ext/Int Rot), displaying the performance of NP, FC, and MA during all movements. Gray-shaded areas correspond to the main joint angle axis where movement occurs. White dots represent outliers removed from the analysis. Abd/Add indicates abduction/adduction; Ext/Int Rot, external/internal rotation; FC, functional calibration; Flex/Ext, flexion/extension; MA, manual alignment; NP, N-pose calibration; RMSE, root mean squared error; ROM, range of motion.

Table 3 Results of 3-Way ANOVA Computed for Each Dependent Variable (ROM Error, RMSE, and Offset)

		ROM error			RMSE			Offset		
Cases	**DF[a]**	***F***	***P***	η^2_G	***F***	***P***	η^2_G	***F***	***P***	η^2_G
Calibration	2	0.510	.601	.001	2.248	.107	.005	**6.064**	**.004**	**.011**
Calibration × tasks	10	1.294	.233	.018	**0.216**	**.001**	**.014**	0.718	.693	.007
Calibration × axes	4	1.503	.201	.008	0.838	.502	.004	**16.044**	**<.001**	**.056**

Abbreviations: ANOVA, analysis of variance; DF, degrees of freedom; RMSE, root mean squared error; ROM, range of motion. Note: The table displays DF, *F* ratio, *P* value, and effect size (η^2_G) of the main calibration effect as well as its interaction effects with axes and tasks. Bold values indicate significantly different conditions.
[a]DF values only for *offset* are 1.814, 9.071, and 3.629 for calibration, calibration × tasks, and calibration × axes, respectively. DF values only for offset are 1.814, 9.071, and 3.629 for calibration, calibration × tasks, and calibration × axes, respectively.

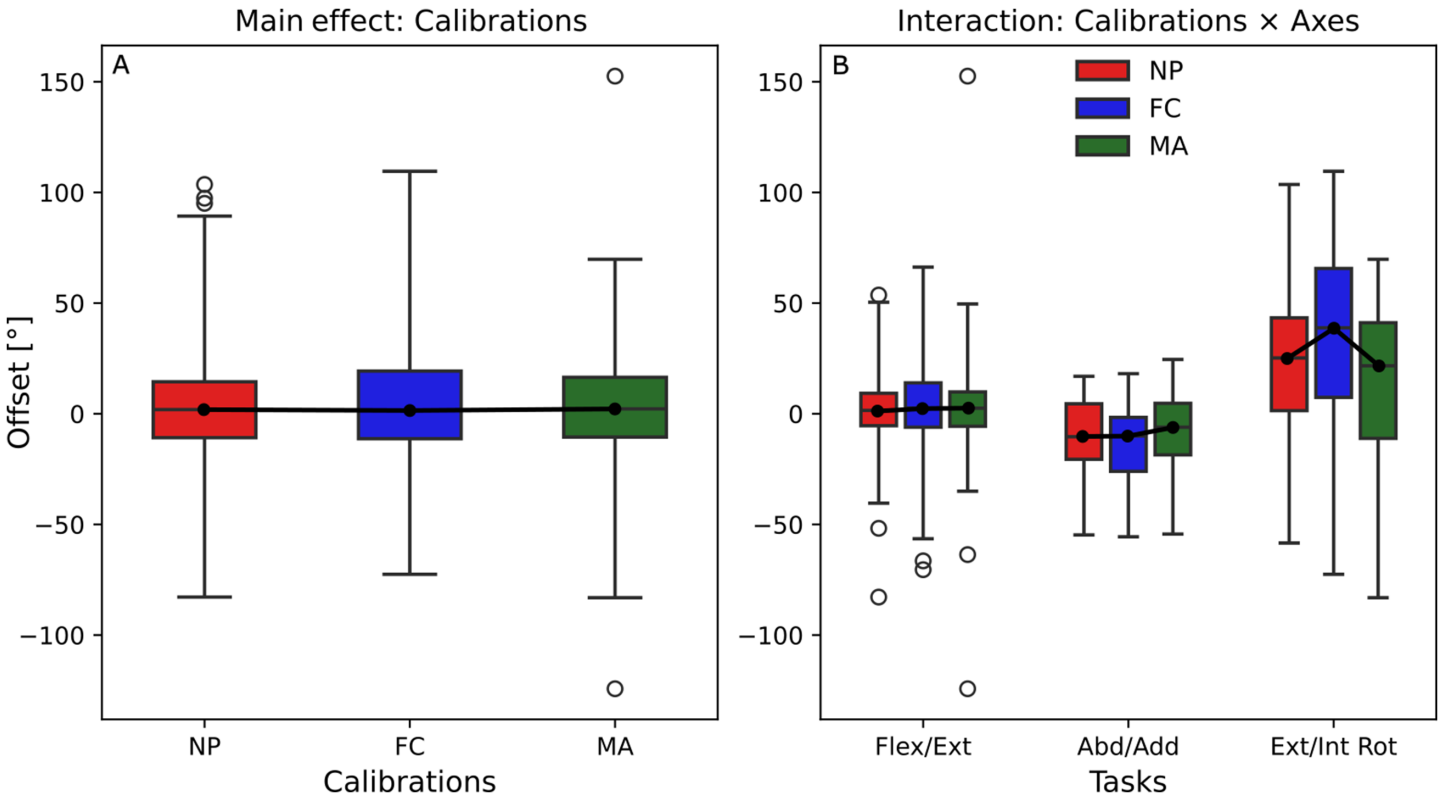

Figure 4 — Offset for main effect (calibrations) and interaction effect (calibration × tasks) of 3-way repeated-measures analysis of variance. The 2 graphs represent the offset as the average computed over tasks and axes (left) and offset as the average over tasks (right) for each joint axis, respectively. Abd/Add indicates abduction/adduction; Ext/Int Rot, external/internal rotation; FC, functional calibration; Flex/Ext, flexion/extension; MA, manual alignment; NP, N-pose calibration.

on the body, which largely depends on the skills and training of the operator and last approximately 5 minutes. As, in a clinical setting, a 5-minute setup could be meaningful in day-to-day practice and could increase the demands on the patient, FC is not recommended for the humerothoracic joint as results have shown no additional benefit over a simpler static calibration. Similarly, accurate sensor positioning using the MA calibration does not show any additional benefit over the NP calibration, except for extremely specific cases and minor performance improvements. Therefore, the extra time invested in performing FC or MA calibration, if compared with the simpler NP calibration, does not translate to increased joint angle accuracy when compared with OPTO. This applies to both main and nonmain joint axes. Our results show that ROM error and RMSE on nonmain joint axes are 10° to 15° larger than main joint axes. This phenomenon is also displayed on other joints,[42,43] but it is exacerbated on the shoulder due to its increased anatomical complexity as well as the large ROM available. We attribute the larger errors on nonmain joint axes to a misalignment of the body segments during the reference pose and skin artifacts that are predominant on the upper arm segment[44] and can severely corrupt measurements.

These results on main joint axes compare similarly with other studies in the field. For instance, Bravi et al[15] developed an IMU shoulder assessment tool for patients with cervical spinal cord injury, showing similar ROM error ranges to our NP calibration

(−12.9° ± 12.8° for shoulder flexion, −23.4° ± 34.8° for shoulder abduction, and −3.1° ± 3.5° for shoulder rotation). In addition, NP shows RMSE in the range of 6.7° ± 2.3° during shoulder flexion, 13.9° ± 5.7° during abduction, and 1.3° ± 1.2° during rotation, which is in a similar range to those identified by Proietti et al[35] and Chan et al[37] during similar shoulder motions. Two other papers[11,16] validated commercial IMU-based products for ROM assessments that employ static calibration and found differences in ROM within 10°,[11] depending on the movement and axis, and RMSE in the range of 10° to 20°.[16]

Finally, the limitations of this work need to be addressed as the humerothoracic joint is notoriously difficult to model due to its anatomy (ball-and-socket joint) and large potential ROM.[13] First, choosing the Euler sequence to convert quaternion data into anatomical angles plays a large role in interpreting the data from the IMUs. In this work, Euler sequences were chosen so that the first rotation within the sequence matches the axes where movement occurs. However, different Euler sequences might lead to different results. In addition, when modeling the shoulder joint without a priori knowledge of the movement, caution is warranted when choosing the Euler sequence. Second, sensor positioning was performed by trained operators, and no time restrictions were imposed. These circumstances might differ from a clinical setting wherein operators may have limited time available and may not always be trained. Third, participants recruited were healthy and showed no sign of anatomical deformities or musculoskeletal or neurological conditions. Therefore, these results might not generalize to unhealthy populations.

In summary, considering both joint angle estimation performance as well as setup time, a simpler static N-pose calibration yields, overall, the best estimates in joint angles. We, therefore, advice the use of NP as the preferred calibration method for clinical shoulder joint angle estimation as it only requires accurate placement of the trunk sensor, and its performance is similar, or at times better, than FC and MA. The only exception to this recommendation is when the focus of the application is on accurate offset estimation of shoulder motions on the abduction/adduction or internal/external rotation axes; in this case, the MA calibration could be a more suitable option, keeping in mind that this choice might be sensitive to accurate sensor placement and alignment; if users are not trained on the correct anatomical IMU placement, MA could result in less accurate joint angle estimations.[4,45] Finally, we provided a set of metrics (Figure 3) that quantify ROM error, RMSE, and offset of the most common IMU calibration methods used in literature to date during single-joint and multijoint shoulder movements.

Acknowledgments

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 956003. **Additional Material:** A detailed explanation of the data recording protocol is provided in the public repository: https://doi.org/10.5281/zenodo.11150444. The folder also contains raw data from IMU sensors and the optical motion capture system.

References

1. Vitali RV, Perkins NC. Determining anatomical frames via inertial motion capture: a survey of methods. *J Biomech*. 2020;106:109832. doi:10.1016/j.jbiomech.2020.109832
2. Poddar S, Kumar V, Kumar A. A comprehensive overview of inertial sensor calibration techniques. *J Dyn Syst Meas Control*. 2017;139(1):11006. doi:10.1115/1.4034419
3. Filippeschi A, Schmitz N, Miezal M, Bleser G, Ruffaldi E, Stricker D. Survey of motion tracking methods based on inertial sensors: a focus on upper limb human motion. *Sensors*. 2017;17(6):1257. doi:10.3390/s17061257
4. Höglund G, Grip H, Öhberg F. The importance of inertial measurement unit placement in assessing upper limb motion. *Med Eng Phys*. 2021;92:10. doi:10.1016/j.medengphy.2021.03.010
5. Rab G, Petuskey K, Bagley A. A method for determination of upper extremity kinematics. *Gait Posture*. 2002;15(2):113–119. doi:10.1016/S0966-6362(01)00155-2
6. Michaud B, Jackson M, Arndt A, Lundberg A, Begon M. Determining in vivo sternoclavicular, acromioclavicular and glenohumeral joint centre locations from skin markers, CT-scans and intracortical pins: a comparison study. *Med Eng Phys*. 2016;38(3):290–296. doi:10.1016/j.medengphy.2015.12.004
7. An KN, Morrey BF, Chao EYS. Carrying angle of the human elbow joint. *J Orthop Res*. 1983;1(4):369–378. doi:10.1002/jor.1100010405
8. Cottam DS, Campbell AC, Davey PC, Kent P, Elliott BC, Alderson JA. Functional calibration does not improve the concurrent validity of magneto-inertial wearable sensor-based thorax and lumbar angle measurements when compared with retro-reflective motion capture. *Med Biol Eng Comput*. 2021;59(11–12):2253–2262. doi:10.1007/s11517-021-02440-9
9. Stančin S, Tomažič S. Angle estimation of simultaneous orthogonal rotations from 3D gyroscope measurements. *Sensors*. 2011;11(9):8536–8549. doi:10.3390/s110908536
10. Cutti AG, Giovanardi A, Rocchi L, Davalli A, Sacchetti R. Ambulatory measurement of shoulder and elbow kinematics through inertial and magnetic sensors. *Med Bio Eng Comput*. 2008;46(2):169–178. doi:10.1007/s11517-007-0296-5
11. Rigoni M, Gill S, Babazadeh S, et al. Assessment of shoulder range of motion using a wireless inertial motion capture device—a validation study. *Sensors*. 2019;19(8):1781. doi:10.3390/s19081781
12. Dinno A. Nonparametric pairwise multiple comparisons in independent groups using dunn's test. *The Stata Journal*. 2015;15(1):292–300. doi:10.1177/1536867X1501500117
13. Gates DH, Walters LS, Cowley J, Wilken JM, Resnik L. Range of motion requirements for upper-limb activities of daily living. *Am J Occup Ther*. 2016;70(1):5487. doi:10.5014/ajot.2016.015487
14. Sato TO, Hansson GÅ, Coury HJCG. Goniometer crosstalk compensation for knee joint applications. *Sensors*. 2010;10(11):9994–10005. doi:10.3390/s101109994
15. Bravi R, Caputo S, Jayousi S, et al. An inertial measurement unit-based wireless system for shoulder motion assessment in patients with cervical spinal cord injury: a validation pilot study in a clinical setting. *Sensors*. 2021;21(4):1057. doi:10.3390/s21041057
16. Henschke J, Kaplick H, Wochatz M, Engel T. Assessing the validity of inertial measurement units for shoulder kinematics using a commercial sensor-software system: a validation study. *Health Sci Rep*. 2022;5(5):772. doi:10.1002/hsr2.772
17. Bouvier B, Duprey S, Claudon L, Dumas R, Savescu A. Upper limb kinematics using inertial and magnetic sensors: comparison of sensor-to-segment calibrations. *Sensors*. 2015;15(8):18813–18833. doi:10.3390/s150818813
18. Ligorio G, Zanotto D, Sabatini AM, Agrawal SK. A novel functional calibration method for real-time elbow joint angles estimation with magnetic-inertial sensors. *J Biomech*. 2017;54:106–110. doi:10.1016/j.jbiomech.2017.01.024

19. van Andel CJ, Wolterbeek N, Doorenbosch CAM, Veeger D, Harlaar J. Complete 3D kinematics of upper extremity functional tasks. *Gait Posture*. 2008;27(1):120–127. doi:10.1016/j.gaitpost.2007.03.002
20. Zhu K, Li J, Li D, Fan B, Shull PB. IMU shoulder angle estimation: effects of sensor-to-segment misalignment and sensor orientation error. *IEEE Trans Neural Syst Rehabil Eng*. 2023;31:4481–4491. doi:10.1109/TNSRE.2023.3331238
21. Pacher L, Chatellier C, Vauzelle R, Fradet L. Sensor-to-segment calibration methodologies for lower-body kinematic analysis with inertial sensors: a systematic review. *Sensors*. 2020;20(11):3322. doi:10.3390/s20113322
22. Weygers I, Kok M, Konings M, Hallez H, De Vroey H, Claeys K. Inertial sensor-based lower limb joint kinematics: a methodological systematic review. *Sensors*. 2020;20(3):673. doi:10.3390/s20030673
23. Wu G, van der Helm FCT, Veeger HEJ, et al. ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand. *J Biomech*. 2005;38(5):981–992. doi:10.1016/j.jbiomech.2004.05.042
24. Bakeman R. Recommended effect size statistics for repeated measures designs. *Behav Res Methods*. 2005;37(3):379–384. doi:10.3758/BF03192707
25. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. L. Erlbaum Associates; 1988.
26. Schiefer C, Kraus T, Ellegast RP, Ochsmann E. A technical support tool for joint range of motion determination in functional diagnostics - an inter-rater study. *J Occup Med Toxicol*. 2015;10:58. doi:10.1186/s12995-015-0058-5
27. Carnevale A, Longo UG, Schena E, et al. Wearable systems for shoulder kinematics assessment: a systematic review. *BMC Musculoskelet Disord*. 2019;20(1):546. doi:10.1186/s12891-019-2930-4
28. Lee GX, Low KS, Taher T. Unrestrained measurement of arm motion based on a wearable wireless sensor network. *IEEE Trans Instrum Meas*. 2010;59(5):1309–1317. doi:10.1109/TIM.2010.2043974
29. Timmermans AAA, Spooren AIF, Kingma H, Seelen HAM. Influence of task-oriented training content on skilled arm-hand performance in stroke: a systematic review. *Neurorehabil Neural Repair*. 2010;24(9):858–870. doi:10.1177/1545968310368963
30. Bleser G, Steffen D, Weber M, et al. A personalized exercise trainer for the elderly. *J Ambient Intell Smart Environ*. 2013;5(6):547–562. doi:10.3233/AIS-130234
31. Cinnera AM. Effects on balance skills and patient compliance of biofeedback training with inertial measurement units and exergaming in subacute stroke: a pilot randomized controlled trial. *Funct Neurol*. 2018;33:131–136.
32. Giggins OM, Persson UM, Caulfield B. Biofeedback in rehabilitation. *J Neuroeng Rehabil*. 2013;10:60. https://www.jneuroengrehab.com/content/10/1/60
33. Tsilomitrou O, Gkountas K, Evangeliou N, Dermatas E. Wireless motion capture system for upper limb rehabilitation. *ASI*. 2021;4(1):14. doi:10.3390/asi4010014
34. Schwarz A, Bhagubai MMC, Wolterink G, Held JPO, Luft AR, Veltink PH. Assessment of upper limb movement impairments after stroke using wearable inertial sensing. *Sensors*. 2020;20(17):4770. doi:10.3390/s20174770
35. Proietti T, O'Neill C, Hohimer CJ, et al. Sensing and control of a multi-joint soft wearable robot for upper-limb assistance and rehabilitation. *IEEE Robot Autom Lett*. 2021;6(2):2381–2388. doi:10.1109/LRA.2021.3061061
36. Newbold JW, Bianchi-Berthouze N, Gold NE, Tajadura-Jiménez A, Williams AC. Musically informed sonification for chronic pain rehabilitation: facilitating progress & avoiding over-doing. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. San Jose California USA: ACM; 2016:5698–5703. doi:10.1145/2858036.2858302
37. Chan LYT, Chua CS, Chou SM, et al. Assessment of shoulder range of motion using a commercially available wearable sensor—a validation study. *Mhealth*. 2022;8:30. doi:10.21037/mhealth-22-7
38. Ellis MD, Sukal T, DeMott T, Dewald JPA. Augmenting clinical evaluation of hemiparetic arm movement with a laboratory-based quantitative measurement of kinematics as a function of limb loading. *Neurorehabil Neural Repair*. 2008;22(4):321–329. doi:10.1177/1545968307313509
39. Levin MF, Liebermann DG, Parmet Y, Berman S. Compensatory versus noncompensatory shoulder movements used for reaching in stroke. *Neurorehabil Neural Repair*. 2016;30(7):635–646. doi:10.1177/1545968315613863
40. Kim GJ, Parnandi A, Eva S, Schambra H. The use of wearable sensors to assess and treat the upper extremity after stroke: a scoping review. *Disabil Rehabil*. 2022;44(20):6119–6138. doi:10.1080/09638288.2021.1957027
41. Bowman T, Gervasoni E, Arienti C, et al. Wearable devices for biofeedback rehabilitation: a systematic review and meta-analysis to design application rules and estimate the effectiveness on balance and gait outcomes in neurological diseases. *Sensors*. 2021;21(10):3444. doi:10.3390/s21103444
42. Cutti AG, Cappello A, Davalli A. In vivo validation of a new technique that compensates for soft tissue artefact in the upper-arm: preliminary results. *Clin Biomech*. 2006;21:S13–S19. doi:10.1016/j.clinbiomech.2005.09.018
43. Vargas-Valencia L, Elias A, Rocon E, Bastos-Filho T, Frizera A. An IMU-to-body alignment method applied to human gait analysis. *Sensors*. 2016;16(12):2090. doi:10.3390/s16122090
44. Cutti AG, Paolini G, Troncossi M, Cappello A, Davalli A. Soft tissue artefact assessment in humeral axial rotation. *Gait Posture*. 2005;21(3):341–349. doi:10.1016/j.gaitpost.2004.04.001
45. Donaldson B, Bezodis N, Bayne H. Within-subject repeatability and between-subject variability in posture during calibration of an inertial measurement unit system. 39th International Society of Biomechanics in Sport Conference, Canberra, Australia (Online): Sept 3-6, 2021. [Online]. Available: https://commons.nmu.edu/isbs/vol39/iss1/58.