Revisiting Consumed Endurance: A NICE Way to Quantify Shoulder Fatigue in Virtual Reality

Yi Li
Monash University
Melbourne, Australia
yi.li5@monash.edu

Robert Crowther
Australian Catholic University
Melbourne, Australia
robert.crowther@acu.edu.au

Jim Smiley
Monash University
Melbourne, Australia
Jim.Smiley@monash.edu

Tim Dwyer
Monash University
Melbourne, Australia
Tim.Dwyer@monash.edu

Benjamin Tag
Monash University
Melbourne, Australia
Benjamin.Tag@monash.edu

Pourang Irani
University of British Columbia
British Columbia, Canada
pourang.irani@ubc.ca

Barrett Ens
Monash University
Melbourne, Australia
barrett.ens@monash.edu

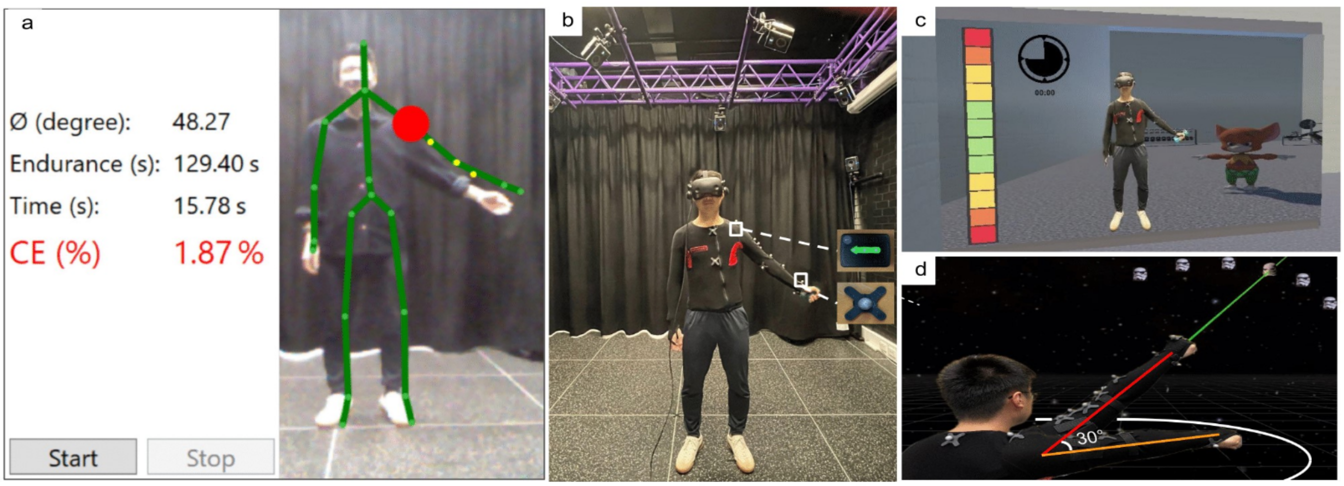

Figure 1: (a) The original Consumed Endurance (CE) interface that uses the Microsoft Kinect tracking system to estimate the CE level for a performed arm gesture. (b) The marker-based Vicon motion tracking system in our study setup. We use 14 14mm X-base reflective markers to track arm movement and four Delsys Trigno Avanti Sensors to collect analog EMG signals. (c) Illustration of the VR environment of Study 1 and the weight-lifting task performed by participants. The colour gradation bar provides direct visual feedback on the shoulder elevation angle. (d) Illustration of the VR setup of Study 2 while a participant performs a target-pointing task. In the actual study trial, only one target is visible to participants at once

ABSTRACT

Virtual Reality (VR) is increasingly being adopted in fitness, gaming, and workplace productivity applications for its natural interaction with body movement. A widely accepted method for quantifying the physical fatigue caused by VR interactions is through metrics such as Consumed Endurance (CE). Proposed in 2014, CE calculates the shoulder torque to infer endurance time (ET)—i.e. the maximum amount of time a pose can be maintained—during mid-air interactions. This model remains widely cited but has not been closely examined beyond its initial evaluation, leaving untested assumptions about exertion from low-intensity interactions and its basis on torque. In this paper, we present two VR studies where we (1) collect a baseline dataset that replicates the foundation of CE and (2) extend the initial evaluation in a pointing task from a two-dimensional (2D) screen to a three-dimensional (3D) immersive environment. Our baseline dataset collected from a high-precision tracking system found that the CE model overestimates ET for low-exertion interactions. Further, our studies reveal that a biomechanical model based on only torque cannot account for additional exertion measured when the shoulder angle exceeds 90° elevation. Based on

VRST 2023, October 09–11, 2023, Christchurch, New Zealand

ACM ISBN 979-8-4007-0328-7/23/10…$15.00
https://doi.org/10.1145/3611659.3615702

these findings, we propose a revised formulation of CE to highlight the need for a hybrid approach in future fatigue modelling.

CCS CONCEPTS

• **Human-centered computing** → **Usability testing**.

KEYWORDS

VR interactions, Interaction design, Endurance, Consumed Endurance, Ergonomics

ACM Reference Format:
Yi Li, Robert Crowther, Jim Smiley, Tim Dwyer, Benjamin Tag, Pourang Irani, and Barrett Ens. 2023. Revisiting Consumed Endurance: A NICE Way to Quantify Shoulder Fatigue in Virtual Reality. In *29th ACM Symposium on Virtual Reality Software and Technology (VRST 2023), October 09–11, 2023, Christchurch, New Zealand.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3611659.3615702

1 INTRODUCTION

Immersive technologies such as Virtual Reality (VR) offer exciting new ways of interacting with computing systems using natural, embodied input, such as the motion of one's head, hands, legs, or even the whole body [7]. As such, VR is increasingly popular for applications in fitness [1], gaming[36, 38], and workplace productivity [30]. As these applications increasingly lead to prolonged usage, supporting the design of VR experiences that manage physical *fatigue*, a cumulative effect of exertion over time is becoming essential. For instance, a productivity application requiring excessive motion will quickly cause fatigue, presenting a barrier to its adoption. Conversely, a fitness application that requires high exertion may aim to induce fatigue after a predictable amount of time. Further, a game design may need direct control in adjusting the physical intensity to maintain user engagement while adapting to individual users' fatigue threshold. Consequently, VR designers will benefit from practical guidance on controlling fatigue across various applications.

There are two established models that might be suitable for modelling fatigue in VR interactions. The first model is the Consumed Endurance (CE) metric [20], which estimates the expended exertion of shoulder muscles during mid-air arm gestures in real time. As endurance time (ET) is the maximum duration a muscle contraction can be maintained before fatigue sets in, CE is the ratio of a given interaction time to the total predicted ET. Inputs to the CE model are the positions of the shoulder, elbow, wrist, and hand, which can be inferred via the head and hand positions tracked by head-worn displays (HWD) [9]. As such, CE is being adopted in designing 3D user interfaces. For instance, Xia et al. [40] recommend using CE to evaluate shoulder fatigue in HCI gesture designs, and Belo et al. developed a tool called XRgonomics [11] that applies CE to Augmented Reality (AR) interactions. The second established model is Cumulative Fatigue (CF) [22, 37], which addresses a fundamental limitation of CE by modelling muscle transition between active, fatigue, and rest periods. Unlike CE, CF does not predict endurance time but rather aims to predict the rating of perceived exertion (RPE).

The current work aims to achieve a model that addresses the limitations of both CE and CF for use in VR applications. We begin with a close inspection of the CE model, as it is currently applied in tools such as XRgonomics [11] due to its publicly available code base. We later plan to expand our investigation to CF, with a longer-term goal of generalising these models beyond the shoulder, to complex uses with other body segments.

Despite its continued adoption, CE has not been fully validated. Following its initial evaluation with a large 2D screen and a Kinect tracking camera [19], the outputs were shown to correlate with subjective measures of RPE. However, this correlation does not confirm the validity of the CE in an extensive 3D space in VR. CE makes an assumption that low-exertion interactions can last indefinitely, which is contested by prior studies [14, 25]. Further, CE relies on shoulder torque estimations, which assume that exertion is symmetric around 90°. However, recent subjective data [11] and muscle contraction data [2] have indicated otherwise.

In this paper, we conduct a thorough investigation of the CE metric for use in VR. In an initial study, we collect a baseline dataset of shoulder exertion from a diverse group of participants. From the dataset, we reconstructed the ET function that forms the foundation for CE. Whereas the CE formulation is based on data collected from high-intensity tasks over a wide range of body joints, our study includes low-intensity tasks with shoulder muscles akin to typical interaction with 3D UIs. Consequently, our findings challenge the assumption that shoulder exertion can be sustained indefinitely when the exerted force is below the minimum threshold of 15% of its maximum.

Second, we validate CE with a dynamic pointing task in VR. This task includes the 120° range of horizontal shoulder motion to add external validity for VR applications. In addition, we compare three vertical motion ranges versus only two levels in the original study. By doing so, we provide direct evidence that exertion is not symmetric at equal angles above and below 90° as assumed in the original CE formulation, which is based solely on calculated torque.

Ultimately, we propose an alternative formulation of CE *New & Improved Consumed Endurance (NICE)* that uses our reconstructed ET function and introduces a correction term to account for additional exertion required by the shoulder movement when the arm elevates above the shoulder.

In summary, the contributions of this paper include:

- A new ET-exertion curve constructed using a wide range of low to high exertion of the shoulder joint.
- The first comparison of the widely accepted fatigue metric, CE, with objective measures of muscle contraction in VR interactions.
- A proposed formulation, NICE correcting substantial overestimation of predicted ET in the traditional CE and accounting for additional exertion of shoulder elevation above 90°.

2 RELATED WORK

Fatigue is a cumulative effect that occurs after a persistent period of physical tasks. Conversely, ET is a temporal concept that measures the sustained period of an exercise until failure [16]. This section introduces literature on direct fatigue measurements and studies that indirectly model fatigue via ET. In this paper, we solely refer to physical fatigue, i.e., muscle fatigue, not mental fatigue.

2.1 Subjective Fatigue Measurement

Levels of fatigue can be described by subjective approaches. Two commonly used methods are the National Aeronautics and Space Administration Task Load Index (NASA-TLX) [17] and Borg CR10 RPE [6]. The NASA-TLX evaluates perceived workload from dimensions like mental demand, physical demand, and frustration on a 21-gradation scale. Similarly, Borg CR10 uses discrete levels between 0 and 10 to classify physical exertion. Though a rough estimation of fatigue can be obtained from subjective fatigue measurement, two significant drawbacks exist. Firstly, it is not easy for subjective measurements to reveal small but vital differences due to limited scales [20]. The second drawback is that individuals may perceive the scales differently due to the uniqueness of their backgrounds and bring bias to study findings [6, 23].

2.2 Objective Fatigue Measurement

Objective approaches to quantifying fatigue focus on changes in physiological characteristics, including heart rate [5], and muscle oxygenation [12]. Though objective measurements eliminate cognitive bias between individuals, these methods are insufficient to properly understand muscle fatigue in real time and potentially add confounding factors to the fatigue assessment. For example, heart rate can be collected with lightweight wearable devices like smartwatches or fitness trackers. However, changes in heart rate are subtle during low-intensity physical tasks [5].

A non-invasive method estimates fatigue based on the participant's muscular activity [18]. It starts with collecting participants' Maximum Voluntary Contraction (MVC) through readings of the force transducer, which defines each participant's maximum muscle contraction capacity. The ratio between the current exerted force or torque and the known maximum quantifies the momentary %MVC. Decreasing MVC over time can be considered evidence of muscle fatigue. However, these measurements are post-trial assessments and cannot provide real-time feedback on muscle fatigue during the study.

Wireless surface electromyography (EMG) sensors allow for objective and non-invasive monitoring of real-time muscle activity. However, the occurrence of muscle fatigue is indicated by the increased magnitudes in muscle contraction [29]. It is unclear how cumulative fatigue can be estimated from instantaneous muscle activation.

2.3 Modeling Fatigue in Mid-air Interaction

While no previous research has specifically modelled fatigue for VR interaction, there has been an extensive effort to improve the ergonomics of upper-body interaction. Plantard et al. [34] developed a software tool to compute the most well-known ergonomic metric: Rapid Upper Limb Assessment (RULA) [27] based on a real-time markerless motion capture camera. Similarly, Bachynskyi et al. [3, 4] designed a novel protocol that estimates the physical workload of arm postures based on clustering biomechanical simulated muscle activation to reduce arm fatigue. Recently, Evangelista Belo et al. [11] integrated the above ergonomic metrics in an AR toolkit to provide guidance on 3D UIs placement. However, ergonomics-based metrics are mainly focused on static postures and cannot account for the cumulative fatigue effect from prolonged dynamic interaction.

There are three relevant models for broader usage in dynamic mid-air interaction. The Rohmert's ET model[35] was the first study that quantifies the relationship between the ET and the exertion level (%MVC) with a power curve, as can be seen in Equation (1). However, the ET model assumes that in calculating the exertion level (%MVC), the maximum exerted force (Max_Force) and the current force ($Force$) — are comparable.

$$ET = \frac{1236.5}{(\frac{Force}{Max\_Force} * 100 - 15)^{0.618}} - 72.5 \tag{1}$$

The CE model [20] improved Rohmert's ET model by estimating the exertion using shoulder torque of the right arm, as in Equation (2). The cumulative average exertion level is obtained by normalizing the average torque ($Torque$) with the maximum shoulder torque (Max_Torque) and is represented in the unit of *%MVC*. The CE model is able to quantify fatigue indirectly by representing the expended physical effort as the ratio between the spent time and the estimated ET, as can be seen in Equation (3). This one single value, CE, will guide VR designers to choose interaction gestures that have the desired fatigue level. A visualization of Rohmert's ET model and an example of CE can be seen in Figure 2.

$$ET = \frac{1236.5}{(\frac{Torque}{Max\_Torque} * 100 - 15)^{0.618}} - 72.5 \tag{2}$$

$$CE = \frac{SpentTime}{ET} * 100 \tag{3}$$

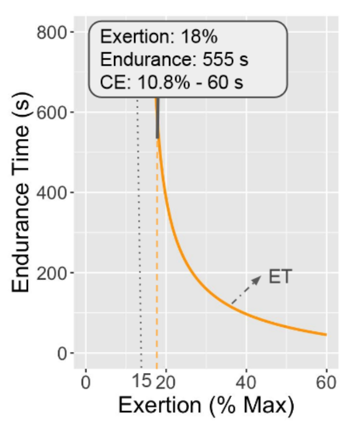

Figure 2: Rohmert's ET model and CE [20].

CE [20] was the first to implement a body-tracking system for real-time torque computation, which freed the participants' hands from the necessity to hold devices. In the original study, participants were asked to hold their dominant arms at different angles for different periods of time. CE showed a high correlation ($R^2 = 0.716$) with Borg CR10 in the linear regression analysis and agreed with Borg CR10 in revealing the main effects in angle, time, and the interaction effects between them. In demonstrating using CE as a design parameter [20], a novel text-entry layout-SEATO was created by allocating the most frequently used characters to the positions that achieve the lowest CE scores. Compared with the traditional QWERTY layout, SEATO obtains a lower Borg CR10 on average. Potentially, CE can help decision-making over selection methods, haptic feedback, and object placements in VR interactions. However, CE strongly relies on the assumption that the exertion level of arm movement will never be lower than 15% MVC (i.e., lift a 1 kg dumbbell to a shoulder angle of 30°). Thus, the model will predict infinite ET for any such interactions.

Another relevant work is the Cumulative Fatigue (CF) model developed by Jang et al. [22]. It applies skeleton-based body-tracking cameras to calculate shoulder torques, similar to the CE model. However, the main difference between CF and CE models is that the CF model uses the transition cycle of motor units, the basic functional units of muscles [8], to predict perceived subjective fatigue: Borg CR10 RPE. In practice, the proportion of fatiguing, activating, and resting motor units will be estimated from the arm gestures performed in real time by the fixed fatiguing and recovery factors. The constant fatiguing and recovery factors were updated later to functions of exertion in the improved CF model [37]. Hence, the CF model can work with tasks of any exertion level. The previous CF evaluation conducts a 2×2 within-participant study investigating different target placements and orders of resting periods. Participants performed a periodic mid-air pointing task on a 2D plane while reporting their perceived fatigue through Borg CR10 every 20 s. Model parameters, including fatiguing and recovery factors, need to be pre-trained with collected Borg CR10 scores. Cross-validation was applied by testing the model on the data collected under different conditions. An overall RMSE of 1.33 was achieved between CF and Borg CR10. Furthermore, a complex mid-air docking task was implemented to evaluate the improved CF model. Similar RMSE results confirmed the flexibility of model performance.

While CE and CF positively correlate with Borg CR10, CF is found to underestimate the fatigue during multi-touch interaction [24], and CE may overestimate the ET by applying Rohmert's ET curve [15]. Importantly, neither metric has been investigated using objective measures, for example, by comparing CE with the empirical data of ET and comparing CF with direct measurement of muscle activities via EMG, nor have they been validated in VR.

3 REVISITING CONSUMED ENDURANCE

Our work aims to examine the validity and reliability of CE for mid-air interactions in VR. We supplement the prior subjective evaluations with objective data, including empirical observation of ET under low to moderate exertions and measured activity of four different muscle groups involved in the shoulder-arm motion.

In the first of two user studies (Section 4), we closely inspect the CE model from its theoretical foundation, Rohmert's ET curve, by reproducing this curve with a diverse participant pool conducting shoulder-motion tasks over a wide range of task intensities. Our results counter Rohmert's assumption on infinite ET below 15 %MVC (Section 4.1) and reveal a substantial overestimation of exertion in the CE model.

In the second study (Section 5), we implement a design similar to the initial CE evaluation [20] but include an additional height level to investigate arm angles both above and below horizontal. Our objective measures demonstrate that intensity increases above the 90° angle for straight arm tasks. This observation confirms that CE underestimates the exertion of arm movements above 90° and highlights the limitation of using shoulder torque only in quantifying shoulder exertion.

3.1 Shoulder Motion Data Collection

Figure 1-a shows an early interface of CE measurement integrated with a Microsoft Kinect V1 camera to capture the body movement [20]. The Kinect camera is a markerless motion-tracking system using a default 32-joint skeleton. Among all 32 joints, only the shoulder, elbow, wrist, and hand joint centres are used as the input of CE to calculate the applied shoulder torque. In the current study, we use 10 Vicon Vantage cameras that capture 14 reflective markers placed on the upper body and right arm (see Figure 1-b) at 100 Hz using VICON Nexus software (v2.12, VICON, Oxford UK). All captured trials are reconstructed, labelled, and gaps filled in VICON Nexus. They are later exported to Visual 3D (v6.0, c-motion, MD, USA) for body modelling (shoulder, upper arm, lower arm, hand, and joint centres). Finally, we calculate the CE measurements in Matlab [26].

3.2 EMG Signal Collection

In our second study, Delsys Trigno wireless EMG sensors[1] are used to collect the surface EMG signal of four muscles at the shoulder: Upper Trapezius (TR), Middle Deltoid (MD), Anterior Deltoid (AD), and Infraspinatus (IF). The skin preparation for sensor placement followes the SEMIAM recommendations[2] and is shown in Figure 3. The EMG signals are synchronised with the VICON Nexus software at 2000 Hz and captured during motion trials.

Figure 3: EMG sensors placement of four investigated muscles that contribute to shoulder range of motion.

In the signal processing, we remove the DC offset from the EMG signal with a high pass filter of 50 Hz and compute the EMG linear envelope using full-wave rectify and a low pass filter of 250 Hz. Finally, a moving RMS with a window size of 500 will be applied before the MVC normalization, following the recommendation of Visual3D[3].

The MVC normalization is needed to make EMG signals comparable between participants. All captured EMG signals will then be able to be represented in the unit of the percentage of the maximum muscle contraction (%MVC). Therefore, we collect the MVCs of the desired muscles of each participant before the study trials. Details about the MVC collection will be explained in the Supplementary Materials.

4 STUDY 1 - ENDURANCE TIME FOR SHOULDER EXERTION

The CE model relies on the ET curve developed by Rohmert [35] (see Figure 2) to estimate ET from the current shoulder torque. This dependency is followed by a strong assumption that the ET will be

[1]https://delsys.com/trigno-avanti/
[2]http://seniam.org/sensor_location.htm
[3]https://c-motion.com/v3dwiki/index.php?title=Tutorial_EMG

approximated to infinity when the task intensity is below 15% of an individual's maximum strength.

Furthermore, Rohmert's curve was created from high-intensity exertions of various arm and leg muscle groups. This approach overlooks the low-intensity exertion and is also invalid for modelling shoulder movement since ET is recognised as joint-specific [14]. Thus, we test the validity of this assumption by reproducing a similar curve on our baseline dataset collected from static shoulder elevation of varying exertion levels.

To collect a sufficiently large set of data points, we include a broader set of task intensities than would be possible from free-hand arm motion alone by providing participants with a set of small dumbbells (with weights ranging from 1-3 kg).

Apparatus. The study environment was built for virtual reality with the Unity engine. There is a precedent to studying exercise in VR, showing it to improve user concentration and endurance [28]. Thus, we design a virtual gym in VR, wherein virtual dumbbells depict those held in the real world. Participants wear an HP Reverb G2 headset. Their precise right-arm movement is streamed in VR, having Unity and Vicon Nexus synchronized in real time. As such, we can monitor the shoulder elevation and elbow extension during study trials. Participants receive visual feedback on their arm position to keep them within the target range (colours between the two red scales) as shown in Figure 1-c.

Participants. We recruited 12 right-handed volunteers (five female and seven male), aged 18-39 years, height 1.50-1.83 m, and weight 50-90 kg. Only two participants had no prior experience with VR. Since the study task involves purely physical activities, we designed a pre-study questionnaire about the participants' workout routine and caffeine consumption to avoid confounding factors.

Task. After the MVC collection, participants were introduced to the study environment in VR. The study task is side-lifting a dumbbell with the required weight at the desired arm angle for as long as possible up to a maximum of 5 minutes. Trials will also be ended if the participants fail to maintain their arms within ± 5° of the target angle. At the end of each trial, participants were provided with fully explained Borg CR10 ratings to self-report their perceived fatigue.

Design. Two independent factors are implemented in the study: `Arm_Angle` and `Arm_Weight` to produce various torque levels, including low to high exertion. We consider four levels of `Arm_Weight` to stimulate the perceived shoulder fatigue from a heavier arm.

- **`Arm_Weight`**: the weight of the dumbbell held by the participant. Our design included four levels: 0, 1, 2, and 3 kg.
- **`Arm_Angle`**: the angle between the participant's arm and their torsos. There are five levels: 30°, 60°, 90°, 120°, and 150°.

Due to a large number of combinations ($4 \times 5 = 20$), we avoid a full within-participants design to mitigate participant fatigue. Following initial MVC collection, each participant is asked to continue with conditions from a random list of four repetitions of all 20 conditions ($4 \times 20 = 80$ in total). The number of study conditions varied between participants. Participants can choose to complete four to seven conditions based on their physical capacity and relax their muscles with a 5 minutes break between conditions. The maximum study duration is one hour. The total study duration, including MVC collection, is approximately 90 minutes. Across all participants, we collect data from four repetitions for each `Arm_Angle`−`Arm_Weight` pair, resulting in 80 data points.

Measures. **`Duration (s)`**: Duration of the trial in seconds. Trials will be terminated by the end of 5 minutes or by any failures to maintain arms at the target `Arm_Angle`. **`True_Torque (%MVC)`**: The average physical intensity represented in shoulder torque, and it is normalized to the corresponding participant's Max_Torque as measured at the MVC collection.

In total, we have $5 \times 4 \times 4 = 80$ measurements for `Duration`, `True_Torque` (%MVC).

4.1 Revised Endurance Function

We begin by investigating the asymptotic term in the original ET model, as introduced in Section 2.3. The ET model used in the original CE study (Equation (2) was improved from Rohmert's (Equation (1)). As shown in Equation (2), this curve has two additional terms (introduced in a formula from Freivalds [13]) not found in an ordinary power curve $y = a * x^b$: one of these is the asymptotic term (15) and the other is a constant term (-72.5). While the constant term has no biomechanical implication, the asymptotic term reveals the maximum workload individuals can take without getting fatigued (see the black dotted line in Figure 2). In other words, individuals can sustain any physical activity below the threshold for an indefinite period of time.

To produce a similar curve and further validate the previous ET model, we first need to determine the asymptotic term to use in our function. Literature has an extensive discussion regarding including the asymptotic term and its potential impact. In Monod and Scherrer's study [31], the asymptotic term is recognized as the critical force for physical tasks without exhaustion. Their ideas were supported by Morton [21] and others [32] who consider the asymptotic term as the buffer muscle capacity and argue that spending all non-buffer physical effort does not lead to exhaustion and argue that spending all physical exertion does not lead to exhaustion. Meanwhile, Garg et al. argue that the asymptotic term will overestimate the ET of low to medium exertion [15]. Frey et al. [14] conclude that a regular 2-parameter power curve can best describe their observations. Given that the minimum observation of torque in our data collection is 5.8 %MVC, we decided to compare the goodness of fit with power curves with different values for the asymptotic parameter, between 0-5, from our measurements of `Duration` and `True_Torque`.

As shown in Table 1, when the asymptotic term is taken as 0, i.e. a regular power curve, the curve best aligns with our collected data. The plot in Figure 4 shows the varying endurance capability among participants in our study. Nonetheless, the power curve has a reasonable fit with $R^2 = 0.57$, meaning that the physical exertion explains 57% of the variation in ET. Our conclusion of the asymptotic term is consistent with the prior study done with natural arm weight at 90°, where empirical evidence indicates that the ET is not infinite for exertion lower than 15 %MVC [25].

Therefore, we propose ET^+ as a revised Endurance Time curve for shoulder movements:

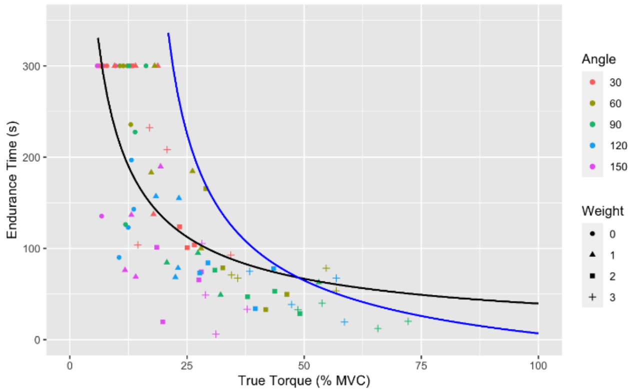

Figure 4: Black line: The new ET model fits with 2-parameter power curve; Blue line: The Rohmert's ET model used in CE.

$$ET^{+} = \frac{1277.9066}{(\frac{Torque}{Max\_Torque} * 100)^{0.7546}} \quad (4)$$

To our knowledge, this is the first complete ET-exertion function constructed from such coverage of low to high exertion (6 - 72 %MVC) at the shoulder joint. This established formulation will be valuable for future fatigue studies in biomechanics and HCI.

5 STUDY 2 - EVALUATION OF CE IN VR

To evaluate the validity of CE for use over a wide range of motion in VR interactions, we designed a study that closely follows a design from the initial evaluation of CE by Hincapié-Ramos et al. [20]. This evaluation study (hereafter referred to as the 'original' study) evaluated CE with a task requiring participants to touch a 2D array of targets on a large screen. The study used a 2×2 factorial design with two height levels and two distance levels designed to coerce participants into using either an extended or bent elbow.

Our study design includes several changes from this design. First, we replace the original task using direct touch on a large screen with a task with greater external validity in VR. Arguably the most common VR interaction, and the most feasible over extended durations, is pointing using a raycast metaphor. Second, we mimic the extended and bent arm conditions of the original study but allow a free and relaxed posture to be used in the bent arm condition, without any artificial constraint. Third, whereas the original study included two 'height' conditions for target locations, we include a third condition. This addition allows us to investigate differences in exertion in angles as equal distances above and below 90°, which was overlooked in prior studies [20, 22, 37].

Table 1: ET curve coefficients for asymptotic terms, where shoulder torque (%MVC) is between 0-100; ET is Endurance Time in seconds.

c: Asymptote	a	b	R^2
Power: $ET(s) = a * (\%MVC + c)^b$			
-5	355.1425	-0.3704	0.4823
-4	482.2708	-0.4709	0.5206
-3	627.1240	-0.5518	0.5415
-2	801.3887	-0.6240	0.5552
-1	1014.6980	-0.6910	0.5651
0	1277.9066	-0.7546	0.5725

Participants. We recruited 12 volunteers (six female, six male), ages 18-39 years, height 1.51-1.95m, and weight 48-82kg. The number of participants in study 2 was based on a power analysis based on pilot data suggesting a minimum sample size of 9 will have 95%

Task. Following the MVC collection, participants start the study trials in an immersive virtual galaxy in VR seen on Figure 1-d where they are told to fight against the "stormtroopers" with the "lightsaber" in their hands (from the fictional Star Wars universe). In detail, participants must destroy a set of 20 targets one by one, presented in random order, by selecting them with a laser extended from their right hand. The direction of the laser will align with the lower arm direction. 20 target positions are evenly distributed within a 120° arc horizontally, which is a proper functional Range-Of-Motion for shoulder-scapula movement [33], and within a vertical range of ± 5° of the given height condition. In the end, targets will only show on an invisible spherical surface. During the study, an inelastic buckle belt is used to constrain participants' backs in a chair to limit trunk movement. To mitigate variance in task completion time between participants, we specify a 1.5 s dwell time, during which the laser must continuously touch the target before it is destroyed. We encourage participants to finish tasks as quickly as possible. Same as in the implementation in User Study 1, a detailed explanation of the Borg CR10 is provided at the end of each trial to collect the self-reported fatigue.

Design. There are two independent variables applied in the study: `Arm_Extension` and `Target_Group`. We used a 2×3 within-subject design to compare `CE`, `EMGs`, and `Borg CR10` in each condition. In detail, we have two levels of `Arm_Extension` in controlling the interaction:

- **Extended**: The laser will only be available when the elbow angle is greater than 145°.
- **Bent**: The laser will only be available when the elbow angle is lower than 145°.

Meanwhile, we chose three levels of `Target_Group` for monitoring the target location:

- **120**: Targets will be placed at the upper 25°-35° relative to the participant's right shoulder to obtain an average shoulder torque at roughly 120° from vertical (see the red line in Figure 1-d).
- **90**: Targets will be placed in between the upper 5° and the lower 5° relative to the participant's right shoulder to obtain an average shoulder torque at 90° (see the orange line in Figure 1-d).
- **60**: Targets will be placed at the lower 25°-35° relative to the participant's right shoulder to obtain an average shoulder torque at roughly 60° from vertical.

Participants must point at 20 targets for one repetition and complete three repetitions for each condition. The presentation order of the conditions within and between participants is balanced by a Latin square design. Each trial takes roughly one minute to complete, and the entire study, including the MVC collection, requires roughly 90 minutes.

Measures. **CE (%)**: output of the CE model. **Borg CR10**: Self-report perceived fatigue between 0-10. The average physical intensity represented in muscle strength of each muscle group: **EMG_TR (%MVC)**: upper trapezius; **EMG_MD (%MVC)**: middle deltoid; **EMG_AD (%MVC)**: anterior deltoid; **EMG_IF (%MVC)**: infraspinatus.

In total, we have $2 \times 3 \times 3 = 18$ CE ratings, Borg CR10 and $2\times3\times3\times4 = 72$ measure for EMG (%MVC) across the four different muscles for each participant.

5.1 Results

In our analysis, we excluded one participant due to a technical issue with the sensor, as among all 18 trials, 10 of them contained faulty EMG readings. In addition, we removed four data points (2% of the remaining data) because they are out of three Standard Deviations of the completion time.

We used a multi-factor ANOVA with aligned rank transform [39] to compare means for the measured variables, CE, Borg CR10 scores, and EMG readings. Results are reported in Table 2 and visualised in Figure 5 and Figure 6. In cases where an interaction effect between Arm_Extension and Target_Group was found, a separate one-way ANOVA was conducted for each Arm_Extension (see Table 3). For brevity, we show post hoc pairwise comparisons with Bonferroni correction using brackets in all the following figures (Significance Level: '*' 0.05, '**' 0.01, '***' 0.001).

5.2 Discussion

Disagreement between CE, Borg CR10, and EMG. Inspection of the Extended arm conditions reveals interesting patterns across the Target_Group. Because CE quantifies the shoulder exertion based on the torque at the shoulder joint, we hypothesise that the CE scores will increase for Target_Group between 60° and 90° and decrease for Target_Group between 90° and 120°. As seen in Figure 5, the pattern in CE is as we expected. Though the dynamic torque is affected by the movement speed in Study 2, CE at 60° and 120° are roughly equal (with small variations from the ±5° target range) and both lower than 90.

Whereas results of the Borg CR10 scores of perceived exertion appear to disagree with this trend; although a high variance in scores prevented us from finding significant differences, there appears to be an increasing pattern of Borg CR10 with increasing Target_Group in the Extended arm condition.

This mismatching inspires us to compare the objective exertion measured by EMG. As shown in Figure 6, all four muscle groups show significant differences between the 60° and 120° conditions in Extended arms. Furthermore, patterns found in EMG_TR, EMG_MD, EMG_AD, EMG_IF clearly show that the intensity of muscle contractions increases with Target_Group for the Extended arm.

Since CE is a function of physical intensity and time, these observations of muscle exertion in EMG_TR, EMG_MD, EMG_AD, and EMG_IF imply that CE should similarly increase with shoulder elevation over a fixed duration. However, as discussed above, the reliance of CE on estimated torque assumes a symmetrical distribution of CE peaks at 90°, which does not agree with EMG.

The disagreement between torque and EMGs leads us to an exploration where we replace the Torque (%MVC) in Equation (4) with EMG (%MVC) measured from one of the muscles, for example, EMG_MD (%MVC), we see the expected increasing pattern in the resulting CE_MD values in Figure 6 with significant difference

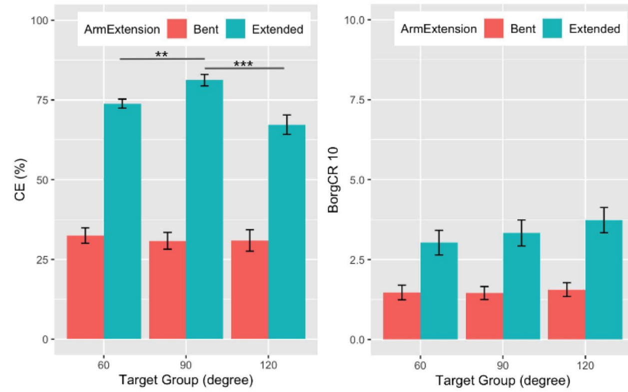

Figure 5: Mean values of each condition in Study 2. From left to right are CE and Borg CR10. Bars represent ±1 SE.

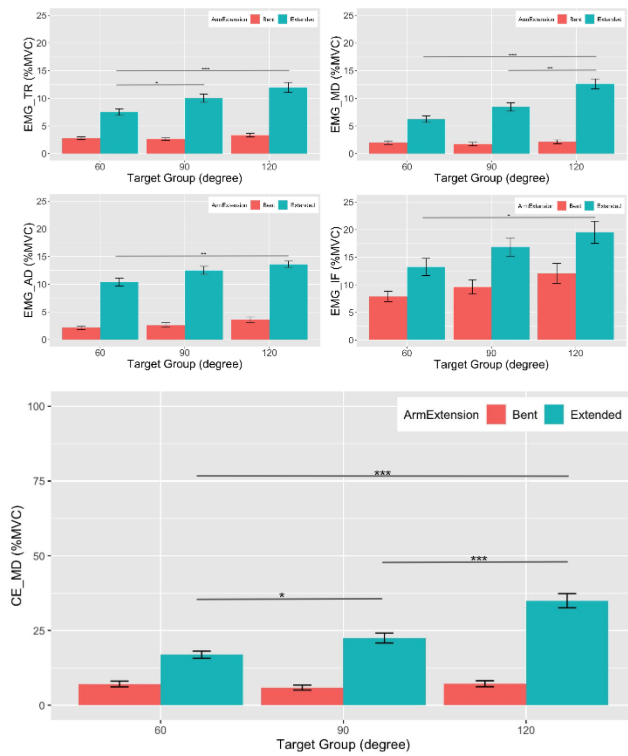

Figure 6: Mean values of each condition in Study 2. From top to bottom, left to right are EMG_TR, EMG_MD, EMG_AD, EMG_IF, and CE_MD. Bars represent ±1 SE.

Table 2: Results of the ART ANOVA on each of our dependent measurements.

Measurements	Main Effect of Arm_Extension	Main Effect of Target_Group	Interaction Effect
CE	$F_{1,194} = 453.11$, $\mathbf{p < 0.001}$	$F_{2,194} = 7.84$, $\mathbf{p < 0.001}$	$F_{2,194} = 6.02$, $\mathbf{p < 0.01}$
Borg CR10	$F_{1,194} = 151.10$, $\mathbf{p < 0.001}$	p = 0.13	p = 0.32
EMG_TR	$F_{1,194} = 477.10$, $\mathbf{p < 0.001}$	$F_{2,194} = 20.44$, $\mathbf{p < 0.001}$	$F_{2,194} = 13.21$, $\mathbf{p < 0.001}$
EMG_MD	$F_{1,194} = 1099.15$, $\mathbf{p < 0.001}$	$F_{2,194} = 48.42$, $\mathbf{p < 0.001}$	$F_{2,194} = 48.75$, $\mathbf{p < 0.001}$
EMG_AD	$F_{1,194} = 1028.59$, $\mathbf{p < 0.001}$	$F_{2,194} = 19.70$, $\mathbf{p < 0.001}$	$F_{2,194} = 3.17$, $\mathbf{p < 0.05}$
EMG_IF	$F_{1,194} = 146.30$, $\mathbf{p < 0.001}$	$F_{2,194} = 20.92$, $\mathbf{p < 0.001}$	$F_{2,194} = 6.42$, $\mathbf{p < 0.01}$

Table 3: Results of the 1-way ANOVA on our dependent measurements, conducted separately under each `Arm_Extension` condition.

Measurements	Extended Arm	Bent Arm
CE	$F_{2,194} = 7.56$, $\mathbf{p < 0.001}$	p = 0.87
Borg CR10	-	-
EMG_TR	$F_{2,194} = 15.49$, $\mathbf{p < 0.001}$	p = 0.66
EMG_MD	$F_{2,194} = 30.95$, $\mathbf{p < 0.001}$	p = 0.88
EMG_AD	$F_{2,194} = 8.51$, $\mathbf{p < 0.001}$	p = 0.17
EMG_IF	$F_{2,194} = 3.96$, $\mathbf{p < 0.05}$	p = 0.17

between conditions in Extended arms. At this point, we conclude that torque fails to capture the observed physical intensity when the arm elevation is above the shoulder height, which highlights the limitation of using torque only to estimate fatigue.

Unlike the clear patterns seen between `Target_Group` conditions with an Extended arm, results in conditions with Bent arms show no significant difference between Target_Group conditions. This is different from the original CE study, which showed a significant difference between low and high target groups, resulting from participants compensating for high target locations with adjustments in their arm postures.

Correction term in torque calculation. Based on the preceding observations, we propose a correction term C for CE to correctly predict physical intensity when the arm elevation is above 90°. As described in Equation (5), the updated torque calculation $h(torque)$ will add the correction term C to $g(torque)$, which is the torque calculation in the original CE formulation. We expect this additional term to explain the mechanism of shoulder elevation [2] as $f(\theta)$, where θ is (the angle between the upper arm and the torso) - 90 while correcting the bias in estimating exertion using torque only.

$$h(torque) = g(torque) + C \tag{5}$$

Inspired by the study done by Crosbie [10], where the relationship between the shoulder joint and upper arm is well explained by the logarithmic function, we hypothesize that the correction term $f(\theta)$ is a variant of $log(\theta)$. We start with $log(\theta + 1)$ by adding a constant term +1 to left shift the log curve across the origin to avoid decreasing the physical intensity at 90°. The correction term C is defined in Equation (6).

$$C = w * log(\theta + 1) \tag{6}$$

Table 4: The Pearson Correlation Coefficients (r) of torque and four muscle groups show that our proposed torque function $h(torque)$ is more strongly correlated than $f(torque)$ from the original CE formulation.

r	EMG_TR	EMG_MD	EMG_AD	EMG_IF
$g(torque)$	0.53	0.62	0.73	0.40
$h(torque)$	**0.69**	**0.80**	**0.84**	**0.46**

In terms of the unknown coefficient: w, our initial attempt is based on the idea of getting the exertion at 120° greater than the exertion at 90°. We use Equation (7) below to obtain the $w_{threshold} = 1.8$ that makes $Torque_{90}$ equal to $Torque_{120}$.

$$w_{threshold} = \frac{\Delta Torque(\%MVC)}{log(\Delta TargetGroup)} = \frac{Torque_{90} - Torque_{120}}{log(30 + 1)} \tag{7}$$

Then we increment the value by 0.1 and observe that 1.9 is the minimum value to make $Torque_{120}$ significantly greater than $Torque_{90}$ in our empirical data. We decide to round it to 2 to add more confidence for future evaluation. As seen in Table 4, the corrected torque calculation $h(Torque)$ shows a higher correlation with the four muscle groups compared with the pre-correction torque $g(Torque)$. The above observation suggests that $h(Torque)$ successfully follows the same pattern as using EMG to estimate the exertion of the three levels in `Target_Group`.

6 LESSONS LEARNED

Below we summarise several takeaway lessons from our study.

CE overestimates ET for low to moderate physical exertion. Our first study identified the reason for this: CE follows Rohmert's assumption that physical exertion can be extended indefinitely when the %MVC is below 15%. This implies that participants can engage in VR interactions with bare hands forever without taking a break. However, our empirical data confirms that the maximum duration is finite for VR interactions that require shoulder movement with an extended arm.

A torque-based model fails to capture additional exertion for arm movement above 90°. Our second study demonstrates the inconsistency between CE, Borg CR10, and EMG for above-shoulder movement. Our measurements of muscle contractions of four different muscles in the shoulder show that the intensity of contractions increases with arm angle. The use of only calculated torque is an attractive feature of CE as it allows fatigue to be predicted using minimal information about the arm pose. However, this attempt risks underestimating the exertion of interacting with objects above shoulder height in VR.

A revised CE formulation: New & Improved Consumed Endurance (NICE).. Based on the revised ET function established in Section 4.1 and the correction term introduced in Section 5.2, we propose a revised CE formulation: *NICE* as defined in the Equation 8.

$Duration$ is the spent interaction time from the start, θ is (the angle between the upper arm and the torso) - 90, and $Torque$ is the torque at the shoulder joint. In practice, θ and $Torque$ can be estimated from arm pose inferred from commodity headset and controller positions. This improved metric may be validated in a study using a similar design to the current study. We discuss our plans for such a study along with other future work in Section 7.1 below. An open-source package, including scripts of the revised formulation and data, will be published with the paper.

$$NICE = \frac{Duration * (\frac{Torque}{Max\_Torque} * 100 + 2 * log(\theta + 1))^{0.7546}}{1277.9066} * 10 \tag{8}$$

A hybrid approach using torque and muscle activation addresses current limitations. Currently, there are three approaches in the design of fatigue models for VR interactions. The first is a pure torque-based approach, which the original CE model follows. This approach suffers from inaccurate estimations of the over-shoulder movement, as verified by our data. The second approach is supervised learning, which is the direction CF takes. Though CF has been shown to accurately predict subjective RPE, it requires pre-training the model on a specific type of activity, this approach is potentially prone to overfitting and may lose generalisability when applied in varying interaction tasks. The last approach is based on muscle activation. An example of this approach is the clustering model in studies [3, 4], where biomechanical simulation maps the arm posture to energy cost in recruited muscle groups. However, how to estimate cumulative fatigue from instantaneous muscle activation remains unclear.

The improved CE model: NICE, was constructed from the revised ET curve that was fit to our empirical data (see Figure 4), along with the proposed correction term, which increases correlation between torque calculation and muscle contraction (see Table 4). Therefore, NICE takes a hybrid approach by combining torque and muscle activation. As such, NICE can estimate the maximum interaction duration and account for additional exertion in over-shoulder activities. Furthermore, the implementation of NICE does not require model training or pre-calibration, which makes it more user-friendly than CF.

7 CONCLUSION

We conducted two studies to investigate the CE metric, previously evaluated by subjective approaches only, using objective measures. The ground truth data identified the causes for overestimates in ET predicted by the original CE model. In a second study, we examined the validity of CE in VR through a typical VR pointing task to further explore the relationship between CE, torque, and muscle intensity in an extensive 3D space. Our findings presented that the CE metric requires the further addition of a new term in order to correctly reflect asymmetric physical exertion for VR interactions above the shoulder height. Lastly, We proposed an improved metric, NICE, aimed at correcting the limitations we identified.

7.1 Limitations and Future Work

Like previous studies in estimating the ET curve, our first study is influenced by the limited duration of trials (capped at 5 minutes). This was done to prevent physical exhaustion or potential injury to participants. However, it artificially constrains the range of data available for a better fit with the ET-exertion function. Furthermore, for each study, we recruited 12 participants. While this is similar to the sample size from similar studies ([20] and validation data sets from [22, 37]), a future study with a greater diversity of participants may lead to results that better generalise to the broader population.

As inspired by the last point in Section 6, our next step is to investigate muscle-based fatigue indicators under different interaction conditions. We propose slicing the interaction area around the arm into a 2D grid with horizontal and vertical lines. Participants will continue each trial within a particular block until they can no longer hold their arms. By observing the corresponding muscle activation, we can explore various approaches for deriving the correct term, potentially through supervised learning. It is expected that the NICE model with the revised correction term will achieve significant improvements by accounting for the exertion that cannot be explained by the torque. This will allow its use in a wide variety of applications with varying requirements for inducing fatigue.

Due to the limited time frame in the current study, we did not include a direct comparison with CF. However, since CF also relies on estimated torque, it is expected that CF will underestimate RPE when the shoulder angle exceeds 90°. We will investigate this assumption in a follow-up study. Our current study, nevertheless, brings new insights into improving CF. In the longer term, we hope our efforts will contribute toward a universal model that addresses the current limitations of both CE and CF, enabling designers to predict both endurance time and RPE in a diverse range of tasks in VR.

More importantly, current literature only develops the model to target upper body movement. Yet, VR interaction has become more complex, from direct manipulation to full-body gaming. demanding consideration of more body segments and individual variations in future model improvement. Our longer-term goal is to generalize these models beyond the shoulder and arm for use with other body segments. We believe our work will help lead toward future models based on muscle activation rather than torque solely. This may prove useful for modelling muscle groups such as those used for neck motion that may be less influenced by torque than the shoulder.

REFERENCES

[1] [n. d.]. Four apps to try if you're new to VR fitness. *Washington Post* ([n. d.]). https://www.washingtonpost.com/lifestyle/2021/11/11/vr-fitness-apps-to-try/

[2] Maximilian Aurbach, Jan Spicka, Franz Süß, and Sebastian Dendorfer. 2020. Evaluation of musculoskeletal modelling parameters of the shoulder complex during humeral abduction above 90. *Journal of biomechanics* 106 (2020), 109817.

[3] Myroslav Bachynskyi, Antti Oulasvirta, Gregorio Palmas, and Tino Weinkauf. 2014. Is motion capture-based biomechanical simulation valid for hci studies? study and implications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3215–3224.

[4] Myroslav Bachynskyi, Gregorio Palmas, Antti Oulasvirta, and Tino Weinkauf. 2015. Informing the design of novel input methods with muscle coactivation

clustering. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 6 (2015), 1–25.

[5] Kayne Barclay, Danny Wei, Christof Lutteroth, and Robert Sheehan. 2011. A quantitative quality model for gesture based user interfaces. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*. 31–39.

[6] Gunnar Borg. 1998. *Borg's perceived exertion and pain scales.* Human kinetics.

[7] Doug A. Bowman, Ryan P. McMahan, and Eric D. Ragan. 2012. Questioning Naturalism in 3D User Interfaces. *Commun. ACM* 55, 9 (sep 2012), 78–88. https://doi.org/10.1145/2330667.2330687

[8] F Buchthal and H Schmalbruch. 1980. Motor unit of mammalian muscle. *Physiological reviews* 60, 1 (1980), 90–142.

[9] Gavin Buckingham. 2021. Hand tracking for immersive virtual reality: opportunities and challenges. *Frontiers in Virtual Reality* (2021), 140.

[10] Jack Crosbie, Sharon L Kilbreath, Luise Hollmann, and Sarah York. 2008. Scapulohumeral rhythm and associated spinal motion. *Clinical biomechanics* 23, 2 (2008), 184–192.

[11] João Marcelo Evangelista Belo, Anna Maria Feit, Tiare Feuchtner, and Kaj Grønbæk. 2021. XRgonomics: Facilitating the Creation of Ergonomic 3D Interfaces. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 290, 11 pages. https://doi.org/10.1145/3411764.3445349

[12] Sue A Ferguson, W Gary Allread, Peter Le, Joseph D Rose, and William S Marras. 2011. Shoulder muscle oxygenation during repetitive tasks. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 55. SAGE Publications Sage CA: Los Angeles, CA, 1039–1041.

[13] Andris Freivalds. 2011. *Biomechanics of the upper limbs: mechanics, modeling and musculoskeletal injuries.* CRC press.

[14] Laura A Frey Law and Keith G Avin. 2010. Endurance time is joint-specific: a modelling and meta-analysis investigation. *Ergonomics* 53, 1 (2010), 109–129.

[15] A Garg, KT Hegmann, BJ Schwoerer, and JM Kapellusch. 2002. The effect of maximum voluntary contraction on endurance times for the shoulder girdle. *International journal of industrial ergonomics* 30, 2 (2002), 103–113.

[16] M Hagberg. 1981. Muscular endurance and surface electromyogram in isometric and dynamic exercise. *Journal of Applied Physiology* 51, 1 (1981), 1–7.

[17] 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.

[18] Kimberley Hayes, Judie R Walton, Zoltan L Szomor, and George AC Murrell. 2002. Reliability of 3 methods for assessing shoulder strength. *Journal of Shoulder and Elbow surgery* 11, 1 (2002), 33–39.

[19] Juan David Hincapié-Ramos, Xiang Guo, and Pourang Irani. 2014. The Consumed Endurance Workbench: A Tool to Assess Arm Fatigue during Mid-Air Interactions. In *Proceedings of the 2014 Companion Publication on Designing Interactive Systems* (Vancouver, BC, Canada) *(DIS Companion '14)*. Association for Computing Machinery, New York, NY, USA, 109–112. https://doi.org/10.1145/2598784.2602795

[20] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

[21] R Hugh Morton. 1996. A 3-parameter critical power model. *Ergonomics* 39, 4 (1996), 611–619.

[22] Sujin Jang, Wolfgang Stuerzlinger, Satyajit Ambike, and Karthik Ramani. 2017. Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 3328–3339.

[23] Thomas Kosch, Jakob Karolus, Johannes Zagermann, Harald Reiterer, Albrecht Schmidt, and Paweł W. Woźniak. 2023. A Survey on Measuring Cognitive Workload in Human-Computer Interaction. *Comput. Surveys* (1 2023). https://doi.org/10.1145/3582272

[24] Zhe Liu, Daniel Vogel, and James R Wallace. 2018. Applying the cumulative fatigue model to interaction on large, multi-touch displays. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays*. 1–9.

[25] Svend Erik Mathiassen. 1999. Prediction of shoulder exerion endurance from personal factors. *International Journal of Industrial Ergonomics* (1999).

[26] MATLAB. 2010. *version 7.10.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts.

[27] Lynn McAtamney and E Nigel Corlett. 1993. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied ergonomics* 24, 2 (1993), 91–99.

[28] COLEEN McCLURE and Damian Schofield. 2019. Running virtual: The effect of virtual reality on exercise. (2019).

[29] J Matt McCrary, Bronwen J Ackermann, and Mark Halaki. 2018. EMG amplitude, fatigue threshold, and time to task failure: A meta-analysis. *Journal of science and medicine in sport* 21, 7 (2018), 736–741.

[30] Brian Merchant. 2021. The Metaverse Has Always Been a Dystopian Idea. https://www.vice.com/en/article/v7eqbb/the-metaverse-has-always-been-a-dystopia

[31] H Monod and J Scherrer. 1965. The work capacity of a synergic muscular group. *Ergonomics* 8, 3 (1965), 329–338.

[32] Toshio Moritani, Akira Nagata, HERRFRT A DEVRIES, and MASUO MURO. 1981. Critical power as a measure of physical work capacity and anaerobic threshold. *Ergonomics* 24, 5 (1981), 339–350.

[33] Surena Namdari, Gautam Yagnik, D David Ebaugh, Sameer Nagda, Matthew L Ramsey, Gerald R Williams Jr, and Samir Mehta. 2012. Defining functional shoulder range of motion for activities of daily living. *Journal of shoulder and elbow surgery* 21, 9 (2012), 1177–1183.

[34] Pierre Plantard, Edouard Auvinet, Anne-Sophie Le Pierres, and Franck Multon. 2015. Pose estimation with a kinect for ergonomic studies: Evaluation of the accuracy using a virtual mannequin. *Sensors* 15, 1 (2015), 1785–1803.

[35] Walter Rohmert. 1960. Ermittlung von Erholungspausen für statische Arbeit des Menschen. *Internationale Zeitschrift für angewandte Physiologie einschließlich Arbeitsphysiologie* 18, 2 (1960), 123–164.

[36] David Vidal-Tomás. 2022. The new crypto niche: NFTs, play-to-earn, and metaverse tokens. *Finance Research Letters* 47 (2022), 102742. https://doi.org/10.1016/j.frl.2022.102742

[37] Ana Villanueva, Sujin Jang, Wolfgang Stuerzlinger, Satyajit Ambike, and Karthik Ramani. 2023. Advanced modeling method for quantifying cumulative subjective fatigue in mid-air interaction. *International Journal of Human-Computer Studies* 169 (2023), 102931.

[38] Jan Wöbbeking. 2022. Beat Saber generated more revenue in 2021 than the next five biggest apps combined. https://mixed-news.com/en/beat-saber-generated-more-revenue-in-2021-than-the-next-five-biggest-apps-combined/

[39] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[40] Haijun Xia, Michael Glueck, Michelle Annett, Michael Wang, and Daniel Wigdor. 2022. Iteratively Designing Gesture Vocabularies: A Survey and Analysis of Best Practices in the HCI Literature. *ACM Transactions on Computer-Human Interaction (TOCHI)* 29, 4 (2022), 1–54.