Promoting Motor Learning in Squats through Visual Error Augmented Feedback: A Markerless Motion Capture Approach

BART SPRENKELS, University of Twente, the Netherlands

Error Augmentation (EA) is a method to benefit motor learning, proven effective in post-stroke rehabilitation. EA draws attention to errors in performance by intentionally amplifying the deviation between the performance and the ideal trajectory with the goal of promoting skill acquisition. So far EA has only been implemented on simple movements through advanced sensing and actuation tools. In this study, we explore the use of Markerless Motion Capture (MMC) to support EA for complex motor tasks in sports. This work contributes i) a low-cost, flexible, and easy-to-use MMC-based prototype to provide Error Augmented Feedback (EAF) in the field, and ii) an experiment (N=34) investigating the effectiveness of terminal visual EAF on Bodyweight (BW) squats for three pre-defined bio-mechanical features. Visual feedback delivered through our MMC-based prototype proved effective on one out of three selected bio-mechanical features. There is no evidence that indicates that participants that received EAF outperformed participants that received unmodified True Feedback (TF). Qualitative insights on using MMC to design visual feedback are promising but reveal inconsistencies introduced by variables such as clothing, long processing times, and inadequate camera quality. Observation on the effects of EA highlight the importance of a pre-defined ideal technique, the right gain, and the consideration to use deception or not.

AF Augmented Feedback
EA Error Augmentation
EAF Error Augmented Feedback
TF True Feedback
BW Bodyweight
ER Error Reduction
VE Vertical Error
BE Balance Error
KHE Knee-Heel Error
MMC Markerless Motion Capture
MoCap Motion Capture

1 INTRODUCTION

EA is a promising technique that can promote motor learning by identifying and amplifying errors in performance. Previous research in the field of rehabilitation showed that EA can be an effective way to learn or improve specific movements [32]. Insights from prior rehabilitation studies can be transferred to the field of Sports Interaction Technology (ITECH) as they share a similar goal: using advanced technology to teach specific movement patterns effectively. However, there are two main problems that need to be addressed: 1) EA has yet to be applied to complex sports movements and 2) there currently is no evidence of a lightweight and affordable way to provide EAF in the field. To address these problems, we explore the use of computer vision to provide EAF for complex motor tasks in sports. We focus on the case of squatting. By developing a MMC-based system that can record and analyze squats to identify and amplify any performance errors, this study aims to answer the following research question: **How might MMC be used designed to deliver reliable EAF for complex motor tasks in sports?** After development, an experiment among 34 subjects will provide insights to answer the sub-question: **What is the effectiveness of EAF in terms of skill acquisition using our prototype?**

Author's address: Bart Sprenkels, b.sprenkels@student.utwente.nl, University of Twente, Enschede, Overijssel, the Netherlands.

Artificial Intelligence Disclaimer: *In the development of the software prototype and statistical analysis, OpenAI's ChatGPT (GPT-4) was instrumental, translating pseudocode into executable MatLab, Python and R code for the prototype and statistical analysis, thereby facilitating a more efficient familiarization and a faster prototyping process. The use of ChatGPT for drafting textual content was deliberately minimal, serving primarily to refine paragraph structure and organize thoughts without influencing the thesis's intellectual direction. Grammatical accuracy and textual clarity were enhanced through Grammarly and Overleaf's autocorrect feature. This approach underscores a judicious application of AI tools to augment the research process, maintaining the integrity and originality of the work.*

The thesis is structured as follows: Section 2 covers the background of motor learning, followed by Section 3 on the background of Error Augmentation. This study focuses on squatting, of which the details are covered in Section 4. Section 5 outlines the development of the prototype and its implementation is documented in Section 6. The thesis ends with Section 7 on the general discussion and conclusion.

2 BACKGROUND OF MOTOR LEARNING

In order to call improvement in motor task performance "Motor Learning", the improvements need to be persistent over a longer period of time after the training period (formally called "Acquisition Phase") has ended. Any increase in performance displayed during or at the end of the Acquisition Phase is called Acquisition Performance. There are several mechanisms that help acquire the necessary neural changes to achieve motor learning, such as reward-based learning, error-based learning, observational learning, and use-dependent learning [22]. Qualitative feedback (right/wrong) is an essential factor in Knowledge of Results (KR) and forms the basis of reward-based learning. Quantitative feedback on the performance error forms the basis of error-based learning and leans more in the direction of Knowledge of Performance (KP). Both approaches can be an effective method of learning motor skills, but in a comparison study KP showed more potential [44]. This thesis focuses on error-based learning within a KP paradigm to potentially benefit motor learning.

2.1 Ideal Motor Learning Study Design

When researchers want to measure the impact of a designed intervention to promote Motor Learning, they need to follow a study design that includes **a pre-test, an acquisition phase, a post-test, a retention period, and a retention test**. By including these components, any recorded change in performance can accurate be identified and potentially be attributed to the tested intervention. In a hypothetical example, a novel simulator simulates speed-skating and researchers want to measure any potential persistent increase in performance after a retention period of two weeks. A simple test to measure performance is to measure the time it takes for a speed-skater to finish one lap of 400 meters. Such a test makes for an excellent pre-test, post-test and retention-test. A hypothetical acquisition phase might consist of four weeks of intense training in the developed simulator. After the acquisition phase, participants perform a post-test. The difference in performance between the pre-test (before acquisition phase) and the post-test (directly after acquisition phase) is called the Acquisition Performance. The difference between the pre-test and the retention-test (2 weeks after the post-test) can be considered persistent and could therefore be called Motor Learning.

If researchers want to compare a novel intervention to a common approach or known golden standard, they also need to include three groups with a varying Acquisition Phase: **a test group** that trains with the novel intervention, **a placebo group** that trains with the golden standard, and **a control group** that only performs the pre-test, post-test, and retention-test. Just participating in these tests can be considered practise, and practise is considered the most important factor that leads to increase in motor skill performance [20]. If all factors are kept constant, skill improvement is positively related to the amount of practise. Including these three groups forms a strong foundation to measure Motor Learning and attribute better performance to a newly designed intervention.

2.2 Feedback

An essential factor for acquiring new motor skills is the feedback provided to the subject [9][15]. Research focuses on designing ways to provide so called augmented feedback, also known as extrinsic feedback, which is defined as information that can only be made available to a person through an external source [46]. Jakus et al. describe augmented feedback as follows: "Augmented or extrinsic feedback is information about the implementation of a motion pattern, formed and transmitted to humans by means of an external source, i.e. separately and independently of the internal human perceptual processes." [24, p. 20409]. An external source can include, but is not limited to, a trainer, a monitor, sound or vibrations. Designing effective feedback mechanisms is a growing field among research as more advanced sensing and actuation technology becomes available. Current Motion Capture (MoCap) implementations struggle to design effective feedback based on factors like the nature of the movement, the complexity of the action, user context, and user experience [37]. Researchers can design different types of feedback for different movements. Another important distinction is the difference between concurrent and terminal feedback, which differ from each other in when the feedback is provided [42]. Concurrent feedback provides extrinsic information on performance in real time, as the movement is happening. Terminal feedback is only provided to the subject after the movement is completed.

3 ERROR AUGMENTATION

The implementation of complex sensing and actuating technology opens up new ways to generate extrinsic feedback. A promising example is EA, which modifies performance error in order to provide extrinsic feedback. This section covers the basics of EA and how it has been implemented in previous studies.

3.1 Background of EA

EA can be defined as intentionally drawing attention to errors in performance by artificially amplifying the deviation between the performance and the ideal trajectory. Isrealy et al. describe EA as intentionally magnifying or altering the (visual) feedback in order to emphasize visual and sensory feedback [23]. This definition is strengthened by Losey et al. who define EA as exaggerating errors in order to enhance feedback to the user, essentially making the subject's mistakes more pronounced [33]. Focusing on performance errors might seem counter intuitive if the goal is improvement. As mentioned by Abdollahi et al., "seldom does a therapist try to amplify a patient's mistake." [3, p. 121]. However, their statement is quickly followed up by a clarification: error-driven learning is central to neuroplasticity and (re)acquisition of movement. Using the identified error made while performing a motor task to artificially change feedback forms the foundation of EA, and this approach has a lot of similarities with Augmented Feedback (AF) as it requires an external source to provide.

The two main paradigms concerned with modulating error information are: Error Reduction (ER) and Error Augmentation (EA). Where the ER paradigm aims to reduce movement error during motor performance, the EA paradigm magnifies it. The attenuation of errors made during movements is often achieved through haptic forces and is mostly described in literature as Haptic Guidance. Sigrist et al. define haptic guidance as follows: "Haptic Guidance refers to physically guiding the subject through the ideal motion by a haptic interface." [46, p. 22]. Comparison studies between ER and EA showed that implementing EA gives slightly better results than ER [32]. Milot et al. found that EA provided better results in a subset of more skilled participants, whereas ER was better suited for less skilled subjects [36]. In a similar comparison study, Cesqui et al. found that EA is more suitable for less impaired stroke patients while ER is better suited for more impaired patients [13].

3.1.1 *Need for Ideal Technique.*

In order to design a successful application that implements EA to benefit motor learning, one needs to have a clear predefined ideal performance. If a participant performs a squat and the goal is to improve their performance by providing a modulated recording as terminal feedback, it has to be clear what aspects of the recording to augment. The targeted movement needs to be defined in advance. In the case of squatting, which is the focus of this thesis, one of the goals that will be identified in Section 4 is to descend until the thighs are parallel to the floor [30]. For many common closed movements, previous studies have accurately documented what combined bio-mechanical features together compose a complex movement. These studies can be used to understand the ideal execution of a movement. Once the ideal performance is clear and well documented, it is possible to identify deviations from the ideal technique, which in turn can be classified as performance errors. With an identified error signal, it is possible to artificially amplify the error and therefore implement EA.

EA is most promising for a non-novice target audience and can theoretically speed up learning by making (minor) mistakes more pronounced. This is especially true for complex motor tasks, as complete beginners tend to make a lot of mistakes and need time to understand a movement and its accompanying goals. Once a person is comfortable with the basics of a movement, EA can be used to make minor mistakes more pronounced. This type of feedback will make mistakes easier to spot and in turn allow someone to direct their focus to optimize their movement. Frameworks like the Challenge Point Framework have been developed around the effectiveness of Augmented Feedback based on task difficulty and a user's skill level [20]. When designing an EA application for a complex movement, keeping the skill level of the indented user in mind helps to ensure the degree of success.

3.1.2 Concurrent vs Terminal Error Augmented Feedback. Just like extrinsic feedback described in Section 2.2 , EAF can be either terminal or concurrent. In a concurrent scenario the user will experience feedback where the error is measured and amplified at any moment in time. In case of visual feedback, e.g. a Virtual Reality (VR) Headset or an animation, this means that EA is applied for every frame. If the feedback is haptic, the force amplifying the error is consistently applied throughout the motion. If EAF is given terminally, participants are free to perform the movement to the best of their ability. Once completed, AF will reveal how they did and potentially nudge them in the right direction.

Some scenarios allow for concurrent EAF, while other do not and thus require terminal EAF. A good example for this is a vertical jump. If a coach wants an athlete to jump to a certain height based on an athlete's weight and age, they could decide to implement visual EA. If you take a recording of the athlete performing the jump, you have no idea how high they will jump. This uncertainty prevents the implementation of concurrent EAF as it prevent researchers to calculate an error. It is only once the athlete lands that you know how high he jumped, which in turn allows you to calculate the deviation from the goal height (set based on age and weight) and their performance. This deviation, that can only be calculated terminally, is the error that you can use to design an implementation.

3.1.3 Mapped vs. Over-Time EA. There are two ways to apply EA: Over-Time EA and Mapped EA. A good example to illustrate the difference is to imagine a bow and arrow VR implementation. In an over-time approach, an archer will aim at a target and - before releasing the arrow - the deviation from the target and the cross hair will be amplified at every frame. This will force the user to have a steady hand, as any movement while aiming is amplified. Once released, the arrow will land exactly where the cross hair was on release. In a Mapped-EA approach, the user will aim and shoot an arrow without intervention. Once the arrow is released, the distance between bulls-eye (the target) and the true landing location of the error is measured and amplified. The result: if the archer shot too high, this will be accentuated by artificially increasing the distance between the bulls-eye and landing location. Over-Time and Mapped EA are different approaches. One is not necessarily better than the other but will lend them self better to various use cases. It is important to grasp that Over-Time EA is not synonymous with concurrent EAF, as a terminal EAF application can choose to implement Over-Time EA by amplifying performance error at every time point.

3.1.4 Gain and Deception. The gain is the chosen amplifier that determines the degree of amplification of the error signal. Selecting the right gain has a major impact on the degree of success of any EA implementation. Wei et al. tested three separate gains (1.5, 2, 3.1) and concluded that a gain that is too big might result in unstable learning [47]. Generally speaking, the gain can be larger for more advanced users. Professional athletes spend hundreds of hours perfecting complex motor movements. A great example is the basketball free-throw. This is a well-described closed motor task for which one could design an EA implementation. It is reasonable to assume that professional basketball athletes have perfected this movement and will therefore make smaller mistakes that are well suited for a larger gain to make them more insightful compared to a novice audience with large and more inconsistent angular errors.

It is up to the researcher whether or not they disclose the use of EA to their subjects. Theoretically, EA lends itself well to not disclosing its use. The implementation of EA itself should be sufficient to direct someone's attention to their mistakes and therefore to the "correct" performance; it is not necessary to inform them about EA. Causing someone to accept that EAF is unmodified feedback where in reality it is modified feedback is deception. The use of deception should be seriously considered as it can have potential ethical implications. Research that includes deception requires a detailed risk analysis, a debriefing to reveal the true nature of the experiment, and permission from an ethical committee. Besides the ethical aspect, there are other aspects that should be considered when choosing to include deception or not. For example, if users make large mistakes that are then amplified to absurd proportions, people could loose confidence in the feedback. Furthermore, it might be beneficial to explain EA as it could explain to the user any observed fluctuations in performance. The current state of the art has not looked into this, which is why it should be investigated in the future.

3.2 Evidence for the Use of EA From Rehabilitation

EA is a promising approach that could potentially promote motor learning in sports. Existing evidence of successful EA implementation can be found in the field of post-stroke rehabilitation. EA has been applied to re-learn movements such as arm reaching tasks, rotating limbs or functional gait rehabilitation. These are all relatively simple and well-defined motions that happen at a slower speed compared to complex sports techniques.

3.2.1 Visual Error Augmentation. Wei et al. have successfully implemented visual EA to improve the rate and extent of motor learning of visuomotor rotations in healthy subjects [47]. Wei et al. let participants follow a path with a cursor by moving their hands. The errors were amplified with a gain if the cursor deviated from the "ideal" path. The results show that EA can improve the rate and extent of motor learning in healthy subjects. The authors experiment with different gains (1, 2, 3.1). However, they warn that a gain that is set too high might result in unstable learning. In a study by Shum et al. visual augmented feedback has been applied to improve bimanual symmetry in a reaching task [45]. Participants who initially had an asymmetry between limbs were challenged to adapt by performing tasks in an immersive VR environment. In the VR environment, the perceived hand position deviated from the absolute hand position, which is a form of visual EA. After EA training, participants who completed the study achieved lower symmetry error after training with EA. A study conducted by Abdikkahi et al. is an example of multimodal EAF, where nineteen stroke survivors with chronic hemiparesis were asked to keep a cursor on a visual target [4]. The visual target was controlled by a therapist and the error is defined as the deviation between the cursor controlled by the patient and the visual target. The authors combined visual and haptic feedback, as the hand of the patient was attached to a robotic arm during the treatment. The study compared a treatment without EA (the control treatment phase) and a treatment with EA. During the EA treatment, the error was visually magnified by a gain of 1.5 and the robot arm applied an amplifying force of 100e (N/m) where e is the error and a maximum force of 4N. Although the results are not clinically meaningful, the findings indicate a real improvement of adding EA to simple massed practice treatment.

3.2.2 Haptic Error Augmentation. The work of Patton et al. is an example of successfully applying haptic EA and ER training that resulted in significant improvement among stroke subjects [40]. A robot-generated force field either guided/pulled the subject's hand towards the desired trajectory (haptic guidance) or provided resistance and imposed forces away from the desired trajectory (error amplification). Their findings showed improvement when applying EA training and detriment after applying ER training. In a study conducted by Givon-mayo et al. on stoke patients' reaching movements, a 5-week treatment with EA forces showed positive results [19]. In the experimental group, all velocity and directional deviations from a predefined ideal trajectory were amplified during reaching tasks. Kao et al. found EA to be more effective than ER when using EA and ER training to improve people's walking gait on a treadmill [25]. Only the EA group maintained a gait close to the target path once the feedback was removed. The authors describe a noteworthy step in determining the degree of force applied to the participant's hand. Subjects with an initial small average error need larger and faster forces than subjects with larger errors. Amplifying errors can be of great benefit to skilled subjects while retarding the learning process among novice subjects [36][34][19].

In a study by Marchal-Crespo et al., EA positively affected learning a locomotor task [34]. The authors also included a controller that applied random perturbing forces to the knee of their subjects to introduce noise force disturbance. They argue that the random noise benefit learning as they force their subject to continuously pay attention, as they could not anticipate the forces. This technique challenged subjects independently of their initial skill level, even if the motor task was simple for more skilled subjects. The study by Reisman et al. used a split-belt treadmill to improve step length asymmetry with the help of EA [41]. By running each treadmill belt at a different speed, they force subjects to compensate for the slower speed of their paretic leg by taking bigger steps. This form of EA shows positive results, but these results tend to be short-lived. To research long-term effects, the authors conducted a study where participants with a step length asymmetry of at least 5-cm trained on the treadmill 3 days per week for 4 weeks. Although not significant, their study is among the first to demonstrate that repetitive practice with EA can result in longer-term improvements in step length asymmetry after stroke. Some studies compared ER with EA with the goal of revealing what works best. Chen and Agrawal let participants drive an electric wheelchair with a force-feedback joystick along a predefined training path [14]. Assistive forces could reduce errors by pushing the joystick in the right direction while repelling forces could increase errors by pushing the joystick handle away from the desired direction. The results showed no significant difference between the assistive and repelling forces.

3.3 Error Augmentation in Sports Interaction Technology

Despite a wide body of research aimed at providing extrinsic feedback to athletes, there is close to no evidence of EAF in the field of sports. The only study that has implemented EA in a sports context is the work by Basalp et al., who studied the effect of visual EA on learning a complex rowing task [7]. More specifically, the task was trunk-arm sweep rowing and a virtual blue oar showed the reference trajectory on a screen. The angular deviation between the subject's oar and the reference oar was the defined performance error, which was amplified according to a gain along the principles of concurrent Over-Time EA. The authors specifically stated a certain degree of deception, as they note that subjects were not made aware that their performance error was visually augmented. In a within-subject study, researchers experimented with gradually increasing the gain to study which gain worked best. Tested gains are 1.0, 1.2, 1.4, 1.6, and 1.8. The study found that participants kept up with a gain as high as 1.4 but got confused at a gain of 1.6. Compared to reported gains of Patton et al., who reported significance performance increase at a gain of 2.0 [39], a gain of 1.6 is lower. The lower gain can be explained as participants in the rowing simulator are expected to keep up with concurrent feedback that bears the risk of becoming unstable at higher gains. The authors conclude that lower EA gains can be more suitable for more complex tasks but do not discuss the influence of their decision to rely on concurrent feedback, which could cause participants to not be able to keep up with amplified concurrent EAF.

3.4 Main Gaps in EA Research

By looking at existing EA applications in sports and rehabilitation, the two main challenges that are identified are: 1) EA has yet to be applied with affordable and flexible sensing and actuation tools, and 2) EA has yet to be adopted as a viable approach for complex motor tasks in sports. Although EA has shown to be effective in rehabilitation, all EA implementations on relatively complex tasks are realized through expensive equipment such as robotic arms and haptic actuators. Especially haptic actuators, such as the actuator that applied perturbation forces to a subject's knee [34], are incredibly complex and can only be designed within lab settings and extensive research trajectories. EA has been shown effective, so designing a low-cost and flexible way for researchers to conduct EA research could potentially be a well-received contribution to the field of Sports ITECH. The second challenge comes from the fact that EAF is complex sports movements is an highly underexposed field. With just one current prior study implementing EA on a sports motion, implementing EAF on further sport related topics bears a lot of potential for more effective technology-induced feedback applications. This thesis will try to bridge the gap and provide a proof-of-concept prototype that tackles the two identified challenges. The next section describes why BW squats are a suitable movement.

4 BODYWEIGHT SQUATS: A COMPLEX MOTOR MOVEMENT

A complex movement that is suitable for the purposes of this thesis is screened according to six pre-defined criteria, listed in Table 1. BW squats score well on all criteria. The movement is well described in prior research [30][38][17][30][43], can be broken down into bio-mechanical subfactors, can be executed indoors in a controlled environment, has a low barrier to entry as most people can do squats, uses the full-body and is a great indicator of overall health [18], is a sports movement, and is dynamic to a point where the temporal order of muscle contractions determines the degree of success of the performance. A BW squat is an excellent movement to proceed with for the purposes of this thesis.

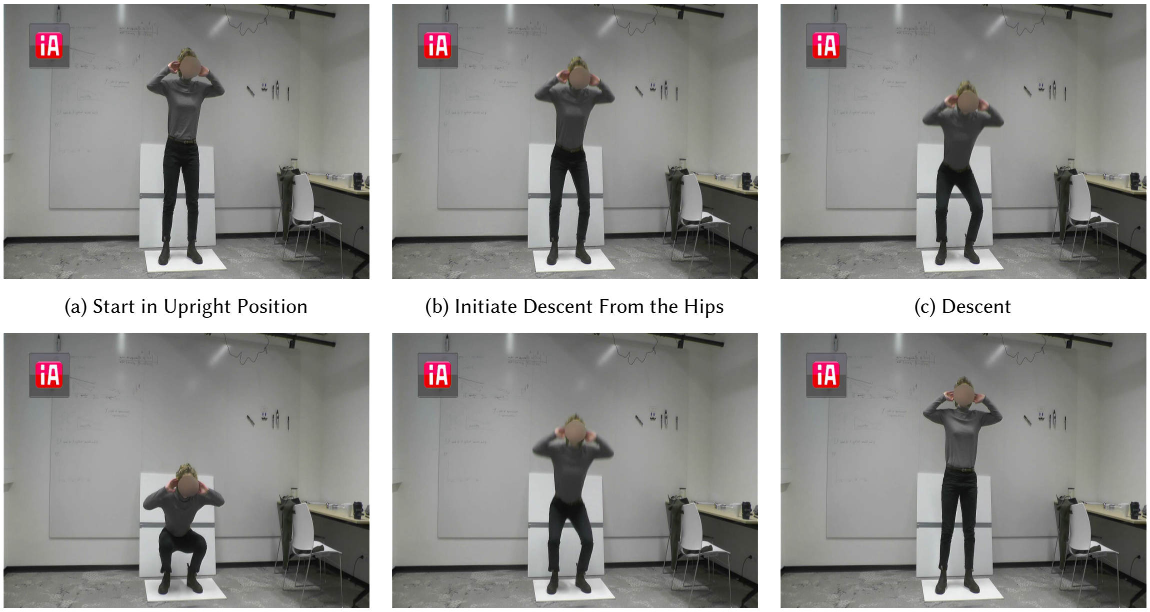
	Criterion	**Reasoning**
1	Well-described in prior research	EA can only be applied if the ideal technique is well described: angles, temporal, spatial instructions.
2	Sports Movement	Movement must be a sports-related movement.
3	Uses full-body	Although movements that use only a part of the body can be considered complex, a full-body movement was preferred.
4	Ability to be performed in controlled environment	Eliminating unnecessary variables was preferred. A motor learning study should be consistent and repetitive.
5	Easily broken down into bio-mechanical factors	Clear sub factors would help apply EA on multiple factors within the same complex movement.
6	Low barrier to entry	A foreseen challenge is subject recruitment, so an accessible complex movement makes recruiting easier.
7	Dynamic	A more dynamic movement is desired.

Table 1. Selection Criteria for Complex Movement

4.1 Bodyweight Squats

A BW squat starts in a standing position and is initialized by flexing the hip and knee joints to move the hips backwards. A person descends until the top part of the thigh at the hip joint is lower than the knee joints. After that, the movements are reversed to start ascending to return to the original standing position [30]. A more visual explanation of the described ideal squat technique can be seen in Figure 1. During a squat, the force is transmitted through the body while the feet are fixed to the ground [16]. The movement is not easy to learn and involves complex movements as it demands the person to control their lower body (knee and lower back) [9]. Because of the focus on the lower limbs and the fact that BW squats can train several large muscle groups at once, it is a common movement used in lower body rehabilitation [17]. The ability to perform a BW squat where the thighs descend parallel to the ground with balance, symmetry and coordination is a great indicator over overall movement quality.

Previous research has been conducted where researchers designed feedback to improve BW squat performance. O'Reilly et al. trained a machine learning model to classify squat performance based on sensor data of 5 IMUs [38]. They gathered data from 77 subjects and can classify a squat as acceptable or aberrant with 98% accuracy. The works of Bonnette et al. describe a concurrent visual feedback approach where subjects got to see a gray rectangular shape



(a) Start in Upright Position (b) Initiate Descent From the Hips (c) Descent

(d) Descent Until Thighs Parallel to Ground (e) Start Ascending (f) Return to Original Standing Position

Fig. 1. Ideal BW Squat Technique.

that mapped biomechanical parameters to 6 points [11]. The objective was to maintain the shape while performing BW squats. In Malaysia, bin Ahmad Tajudin et al. compared the effects of visual and verbal feedback on acquisition of performing BW squats [9]. As part of future studies they recommend researchers to use software or high technologies gadgets to give tailored feedback to participants regarding squat performance. Prior works in the field of Sports ITECH shows that BW squats are of interest to research. Additionally, the work of this thesis will contribute providing EAF on BW squats through technology, in line with recommendations for future research by [9].

4.2 Bio-mechanical Feature Selection

4.2.1 Common BW Squat Mistakes. In order to implement EA, the proper execution of a BW squat needs to be well-defined. The ideal technique serves as a golden standard and any deviation can be considered as performance error. However, according to Bedard this is not possible. They claim that: "there is no widely accepted standard for proper execution of the squat." [8, p.3]. It is true that there are many variations on what is considered ’proper execution’. A lot of studies focus on common mistakes that might cause injuries. [38] classified a BW squat as acceptable or aberrant based on common deviations of BW squats outlined by the National Strenght and Conditioning Association (NSCA) [43]. The 5 identified deviations used by [38] are (i) *Knees coming together during downward phase*, (ii) *Knees coming apart during downward phase*, (iii) *Knees ahead of toes during downward phase*, (iv) *Heels raising off the ground during squat exercise* and (v) *Excessive flexion of hip and torse during squat exercise*. In another study, Connor et al. described a sonification implementation that provides real-time auditory feedback on BW squats focusing on 4 parameters of the squat: *foot placement*, *knee flexion angle*, *knee alignment*, and *weight shifting* [17].

4.2.2 Movement Competency Screening. A common way to analyze BW squats is through qualitative analysis based on a set of predefined critical features [29]. An example of such a screening approach is through the Movement Competency Screening (MCS) developed by Kritz et al. [30]. In the development of the MCS, Kritz et al. compared the screening criteria of the bilateral BW squat as described in the works of [5], [10], and [27] to arrive at a summarized and well-described understanding of the proper squatting mechanics. These requirements were translated to 8 screening criteria for BW squats that are listed in Table 2. The authors of [9] used a variation of the MCS to score BW squats and gave verbal feedback to a subgroup of participants in line with predefined queues depicted in Table 3. Similarly, Bedard rated BW squats using a 0 to 6 scale where each point represents one critical feature [8]. The authors based their approach on the works of Herrington and Munro who analysed single leg squats according to a similar qualitative approach [21]. The 6 features selected by Bedard are shown in Table 4. All these studies that have outlined critical features and common errors help us understand the approximate ideal technique that we need in order to design EAF.

4.2.3 Selected Bio-Mechanical Features. Based on the reviewed prior research, it is possible to select individual aspects of a BW squat that will be the focus. This thesis will focus on three bio-mechanical features: **thighs parallel to the ground** [30], **knees stay behind toes** [38], and **knees track in line with orientation of the feet** [38]. All three features are suitable for EA as they can be translated into ideal executions and the deviation is a clear error signal. The respective errors are named: **Vertical Error (VE)**, **Balance Error (BE)**, and **Knee-Heel Error (KHE)**. As the proof-of-concept prototype is aimed at complex tasks, and complex tasks are defined as tasks with two ore more competing goals, the choice to focus on three unique errors is sufficient. Exactly how these three errors are extracted from a BW squat recording will be explained in the next section on the development of the prototype.

Head	Held in a neutral position appears centrally aligned.
Shoulders	Held down and away from ears. Elbows appear in line with ears.
Lumbar	Lumbar Held in neutral curve position.
Hips	Horizontally aligned and mobile. Move back and down during flexion.
Knees	Aligned with hips and feet during flexion.
Ankles	Mobility allows adequate dorsi-flexion during knee and hip flexion.
Feet	Stable with heels grounded during lower limb flexion.
Balance	Evenly distributed.
Depth	Top of thighs appear parallel with floor.

Table 2. BW Squat Screening Criteria as Outlined by the MCS [30, p.100]

Head	Centered
Shoulders	Held down away from ears.
Elbows	Elbows behind ears throughout the squat.
Lumbar	Neutral throughout squat
Hips	Movement starts here, aligned and extension is obvious
Knees	Stable, aligned with hips and feet
Ankle/Feet	Aligned with the knees and hips. In contact with the ground especially the heels at the bottom of the squat and feet appear stable.
Depth	Thighs parallel to the ground
Balance	Maintained

Table 3. Verbal Feedback Used by [9] based on MCS Criteria

Stance	The subject will assume a roughly shoulder width stance. The toes will be oriented directly forward or canted externally up to 30 degrees.
Heels	The entire bottom of the foot will remain in full contact with the ground throughout the movement.
Knees	The knees will track in line with the orientation of the foot throughout the squat movement.
Hip Depth	The subject will descend until the crease formed by the torso and the upper thigh clearly descends below the uppermost portion of the knee joint when viewed from the sagittal plane.
Hip Extension	After squatting, the subject will extend the knees and hips until both joints reach full extension and the subject is standing fully erect.
Back	The back will remain in a neutral position throughout the movement, maintaining the natural curvature of the spine.

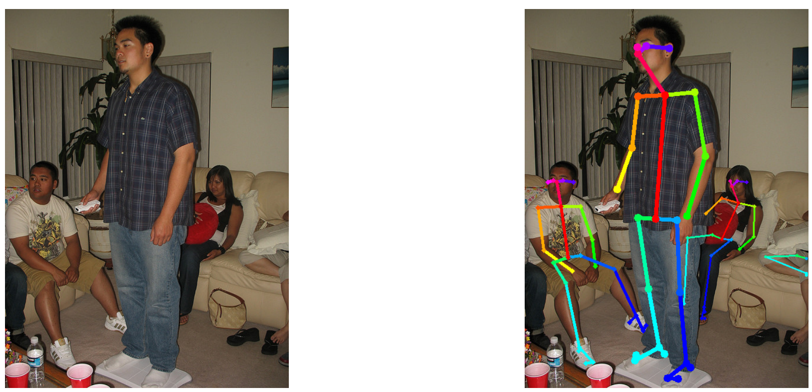
Table 4. BW Squat Screening Criteria as Outlined by Bedard [8, p.14]

5 PROTOTYPE DEVELOPMENT

One of the main contributions this thesis offers is a way to analyze movements and provide visual feedback at a low cost without the need for obtrusive sensors, making the system flexible to capture movement data in the field. A prototype is developed to implement the discovered principles of EA on a complex motor movement. This section covers the selected technology and the choices that were made to arrive at a prototype that can be used to gather movement data to analyse BW squats.

5.1 Markerless Motion Capture (MMC)

There are generally two ways to capture full-body data in sports: using marker/sensors and using cameras. MoCap systems that use markers/sensors are accurate but have downsides. Sensor-based systems can be expensive, often rely on inflexible lab environments, and require athletes to wear markers (often attached to a full-body suit) [28]. Markers can be uncomfortable and can be perceived as obtrusive. Because of these reasons, a system that is able to construct 3D images from two 2D images is preferred. MMC offers researchers the possibility to use two cameras to generate a 3D skeleton. This is done by first estimating the 2D skeleton in both 2D images. Figure 2 displays an example input and output generated through MMC. With the two 2D skeletons, the system calculates the 3D skeleton through triangulation. As this process relies on software and two cameras, it offers a low cost, unobtrusive, and flexible setup. MMC is selected to gather full-body data in order to provide EAF on BW squats.



(a) Input Image (b) Output Image with Skeleton Data

Fig. 2. MMC can estimate poses from a sample input image

BW squats are well-suited to be analyzed by a MMC system. A squat can be performed indoors and therefore is not impacted by weather and terrain. The movement is complex while still allowing the subject to stay in one place. The movement is short, which allows for quick terminal video analysis to provide feedback. Most importantly, all joints remain visible for both cameras throughout the motion of a BW squat. This in turn allows for complete data collection and does not rely heavily on interpolation or incomplete data.

5.1.1 No Need for Markers. MMC is able to track a person's movements without the need for intrusive sensors or reflective markers. An MMC-based system takes photos or videos as input, which are analysed by a trained machine learning model that can identify the location of important features on a picture. An example of such a machine learning model is OpenPose, which is used for this prototype. By extracting positional data from images, it is not necessary to have subjects wear any sensors. Not having to attach sensors to a subject when collecting data is less obtrusive and streamlines the data gathering process so it takes less time to complete. Non-intrusive movement tracking analysis offers a smoother workflow and therefore allows for convenient data collection. Placing two cameras pointed towards a complex movement is enough to gather 3D positional data that can be analysed.

5.1.2 Low Cost. MMC is a powerful tool to analyse movement due to its low implementation costs. For this specific prototype, the videos are captured by two Panasonic camcorders that are set up as webcams through capture cards that convert HDMI to USB signal. Although the specific camera models use in this thesis have been discontinued, the price of a similar setup costs around 500 euros. A high-end accelerometer-based tracking suit such as the MVN Link costs around 64.000 euros and an OptiTrack system for full body tracking costs around 40.000 euros. While these systems are able to deliver excellent results, they are not consumer grade products that might be used by athletes and coaches in sports. These price differences, together with the fact that MMC can work with a wide range of (consumer grade) cameras, are an excellent indication of the promising future of MMC-based MoCap systems in the field sports ITECH.

5.1.3 OpenPose. OpenPose has been selected to identify the location of keypoints based on input images. OpenPose is a real-time multi-person human pose detection library developed by Cao et al. [12] and is well-known for its multi-person detection performance. OpenPose is free for non-commercial use. The model used for this prototype is called BODY-25, which returns 2D positional data on 25 keypoints per person based on an input image. An overview of what keypoints are extracted by the BODY-25 model can be found in Table 5. OpenPose is a trained machine learning model that used the COCO dataset as its training data [31][28]. The COCO dataset is a large dataset that contains images of people annotated with 2D skeletons.

ID	Keypoint	ID	Keypoint
1	Nose	14	Left Knee
2	Neck	15	Left Ankle
3	Right Shoulder	16	Right Eye
4	Right Elbow	17	Left Eye
5	Right Wrist	18	Right Ear
6	Left Shoulder	19	Left Ear
7	Left Elbow	20	Left Big Toe
8	Left Wrist	21	Left Small Toe
9	MidHip	22	Left Heel
10	Right Hip	23	Right Big Toe
11	Right Knee	24	Right Small Toe
12	Right Ankle	25	Right Heel
13	Left Hip		

Table 5. Keypoints for OpenPose's BODY 25

A known limitation of OpenPose is the lack of unusual poses in the training data, which limits the ability to accurately identify skeleton data from unusual poses [28]. In sports, there are many poses that are underrepresented in normal life. Examples of sports movements that contain unusual poses are gymnastics, diving, and acrobatics. Because of this, MMC might not work for some sports, but the poses in BW squats are not a problem for OpenPose.

5.2 Calibration

To keep the prototype flexible, it is important that it can work with any off-the-shelf camera, as long as it is possible to easily measure its performance. For this thesis, two Panasonic HC-V520 camcorders are attached to a computer. A mini-HDMI to HDMI cable leads the camera output to a capture card, which essentially turns regular cameras into USB webcams. These specific cameras record Full HD footage at a 10.0 MP resolution. Different cameras have different focal lengths, resolutions and distortions. If the goal is to calculate accurate 3D positions, it is essential to understand the performance of the utilized camera. To calculate the unique properties of a camera, it needs to be calibrated. Once calibrated, it is possible to compensate for distortion and ensure that straight lines in real life appear straight on the captured images. As can be seen in Figure 3, distortion is not uniform and usually most prominent around the edges of the frame.

Fig. 3. Camera Distortion Examples

To map the performance of a camera across the whole frame, a calibration board with a checkerboard pattern is used. Specifically a 23x32 checkerboard with 25mm square sizes, printed on a Forex board. Forex is a durable, lightweight, and rigid PVC foam material. By moving the board in front of the camera and capturing frames in the meantime, it is possible to have a known golden standard (perfect squares with known dimensions) showing on all parts of a distorted frame. Figure 4 shows an aggregated representation of about twenty calibration frames all taken at a different position as part of the camera calibration procedure for the development of this prototype.

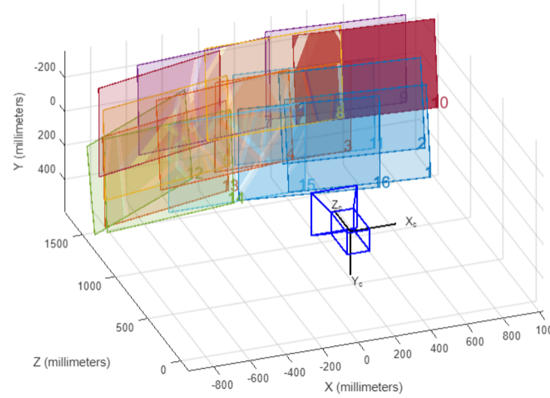


Fig. 4. Camera Calibrator App in MatLab

Straightforward camera calibration is one of the main ways that ensures this prototype is easy and flexible to use. With the help of nine buttons in a MatLab Graphical User Interface (GUI), it is possible to calibrate camera 1, camera 2 and the stereo pair. The buttons that form the GUI are appended in Appendix E. If the setup is moved to a new location, the cameras need to be re-calibrated to be able to determine 3D positions from two camera inputs. Relatively simple camera calibration also allows the user to use different cameras and does not force them to purchase specific equipment that they might not yet own to start conducting research. Calibrating the cameras in order to start collecting three dimensional positional data consists of three steps, namely i) calibrating camera one, ii) calibrating camera two, and iii) stereo calibrate the camera pair.

5.2.1 Single Camera Calibration Workflow. To calibrate camera 1, the user presses "Start Calibration Cam 1," which will start capturing frames on camera 1. The user will take the calibration board with the checkerboard pattern and start walking in front of camera 1. Once complete, the user presses "Stop Calibration Cam 1," which will save a selection of all collected frames to a directory. The amount of saved frames is determined by a variable called the saveRate, which is set to 1 FPS. If capturing calibration frames took 40 seconds, the prototype will save 40 frames to the directory. With all frames collected, the user now presses "Calibrate Camera 1," which will start the Camera Calibrator app within MatLab. The application will analyze all collected frames and identify the corners of the squares, which it in turn uses to calculate the camera parameters. Once all frames are analyzed, the user has a chance to manually remove any outliers and see if the calibration went well. If all went well, it should look something like the visualization in Figure 4. Once completed, the user will export the camera parameters from the app back to the workspace, where they will be stored.

5.2.2 Stereo Calibration Workflow. Once both individual cameras are calibrated, the next step is to capture a stereo pair and understand how the two cameras are positioned relative to each other. The button "Start Stereo Calibration" triggers both cameras simultaneously and starts capturing frames at a capture rate of 10FPS. The button "Stop Stereo Calibration" stops both cameras and saves the recorded frames to two separate directories within the "stereo calibration frames" sub-directory at a save rate of 1FPS.

Once all frames are collected, the user can press "Run Stereo Calibration" to open the Stereo Camera Calibrator app within MatLab. All collected calibration frames are loaded an analyzed automatically, after which the user can select and remove outliers, verify the validity of the calculated calibration and retrieve respective error rates. Figure 5 shows an example of a successful stereo calibration. It is important for the researcher to check if the distance between both cameras, as well as between the cameras and the panes, is feasible. To ensure an accurate calibration, a set of 20 accepted stereo pairs evenly spread over the whole frame should is sufficient. After a successful calibration the user can export the stereo parameters to the MatLab workspace, which will later be used for triangulation.

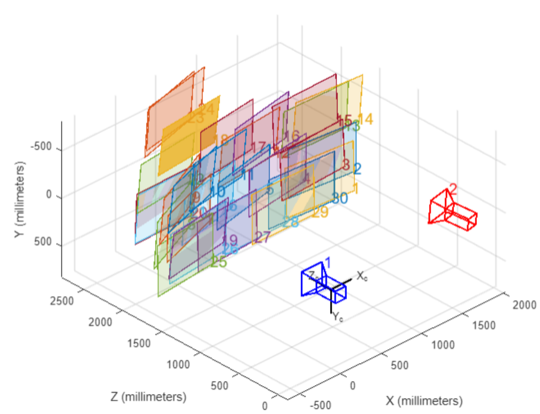


Fig. 5. Stereo Calibrator App in Matlab

5.2.3 Calibration Session File Structure. Each time the cameras need to be calibrated, the prototype creates a new calibration session. Calibrating is by far the most time consuming process when using the prototype, so to save time all completed calculations are stored. This also allows users to load an existing calibration session without having to capture new calibration frames. In Figure 6 the file structure is displayed, which also shows all "calibration_results" directories that will store a .mat MatLab file once all frames are analyzed and the calibration parameters are calculated. In short, the main directory is called "calibration sessions" and contains all individual calibration attempts. A new session directory will be named "Session001" and, when complete, will contain three sub-directories: "capture frames camera 1", "capture frames camera 2", and "stereo calibration frames". These folders store all captured calibration frames, as well as all calibration results once calculated.

```
calibration sessions
└── Session001
│   ├── capture frames camera 1
│   │   ├── frame1.png
│   │   ├── frame2.png
│   │   ├── ...
│   │   └── calibration_results
│   │       └── cameraParams.mat
│   ├── capture frames camera 2
│   │   ├── frame1.png
│   │   ├── frame2.png
│   │   ├── ...
│   │   └── calibration_results
│   │       └── cameraParams.mat
│   └── stereo calibration frames
│       ├── camera 1
│       │   ├── frame1.png
│       │   ├── frame2.png
│       │   ├── ...
│       │   └── calibration_results
│       │       └── cameraParams.mat
│       ├── camera 2
│       │   ├── frame1.png
│       │   ├── frame2.png
│       │   ├── ...
│       │   └── calibration_results
│       │       └── cameraParams.mat
│       └── stereo_results
│           └── stereoParams.mat
└── Session002
    └── ...
```

Fig. 6. Compact directory tree structure.

5.3 Video Capture and OpenPose Analysis

Once calibrated, video footage can be captured from two perspectives by starting and stopping the cameras at the same time. Triggering two cameras at exactly the same time requires a hardware trigger as MatLab cannot start capturing frames on two cameras with one command. To circumvent this problem, cameras are primed individually and triggered in two sequential lines of code. This approximation delivered adequate results. Once the squat movement is completed, both cameras are stopped and all captured frames are stored to a directory. Both video files are trimmed to make sure they both contain the same number of frames, which will be important when determining the 3D positional data.

```
command_to_execute = sprintf('bin\\OpenPoseDemo.exe --video "%s" --write_json "%s" --number_people_max
    1 --render_pose 0 --display 0', command_video_path, command_output_path);
```

Listing 1. OpenPose Analysis Command

To analyse both videos, the command outlined in Listing 1 will automatically run OpenPose. The command takes the path to a recorded video file as an input and will write all JSON files to a predefined output directory. The command is optimized for speed as it will always expect no more than one person in each frame (–number_people_max 1) and it will not display the progress of the analysis (–display 0 and –render_pose 0).

When OpenPose analysed all individual frames of both videos, the output directory contains two subdirectories called **openpose_output_1** and **openpose_output_2**. Both directories contain JSON files named **keypoint_data_frame_00001.json** where 00001 is the frame number of the input video. Each JSON file contains 2D information according to the BODY-25 model, which contains all keypoints shown in Table 5. Listing 2 shows the information contained in a single JSON file. Pose Keypoints 2D contain X, Y, Confidence data for all 25 keypoints. The confidence is a variable between 0 and 1 based on an OpenPose-determined confidence map. For this thesis, the confidence score is not utilized. The output directories with JSON files serve as the input of the Triangulation function as described in the next section.

```
{
  "version": 1.3,
  "people": [
    {
      "person_id": [-1],
      "pose_keypoints_2d": [
        611.361, 192.893, 0.964301,
        596.697, 245.308, 0.903299,
        557.005, 243.21, 0.87537,
        498.396, 224.269, 0.860914,
        (...),
        542.4, 701.3, 0.837754,
        550.675, 680.416, 0.706003
      ],
      "face_keypoints_2d": [],
      "hand_left_keypoints_2d": [],
      "hand_right_keypoints_2d": [],
      "pose_keypoints_3d": [],
      "face_keypoints_3d": [],
      "hand_left_keypoints_3d": [],
      "hand_right_keypoints_3d": []
    }
  ]
}
```

Listing 2. JSON Frame Data Format

The full file structure of a recorded and analysed squat can be seen in Appendix A, which shows that all information regarding to one individual squat is stored in a directory named "0001 - Session001" where 0001 is the name of a recorded trial (one participant folder can contain many trials) and "Session001" is the name of the camera calibration information necessary to later on determine the 3D positional data.

5.4 Triangulation

A pair of two calibrated cameras that can start and stop recording at exactly the same time allows the user to capture stereo pairs of images, which forms the basis for determining 3D positions through triangulation. The 'triangulate' function in MATLAB is used to estimate the three-dimensional positions of points from corresponding image points in a stereo pair of images [1]. It helps to reconstruct the 3D scene from two 2D images and takes the following variables as inputs: **matchedPoints1, matchedPoints2, stereoParams**, where matchedPoints one and two are the sets of worldpoints from each individual camera and stereoParams are the stereo camera parameters that result from a successful stereo calibration as described in Section 5.2.2. The function will output an Mx3 matrix where each row represents a 3D coordinate (X,Y,Z) of the corresponding stereo pair.

Although MatLab performs the calculations automatically, it is still valuable to understand how the triangulation process works. The function first constructs the projection matrices for each camera by combining the camera's intrinsic parameters with its extrinsic parameters. For the first camera, the extrinsic parameters are typically taken as identity and zero translation which makes camera one the reference camera. For the second camera the provided rotation and translation are used. Next, the triangulation function solves a geometric triangulation problem and finds the point in 3D space that best corresponds to the projections observed in the two images. The goal of the function is to minimize the reprojection error, which is the distance between the observed image points and the projected points of the estimated 3D point in pixels.

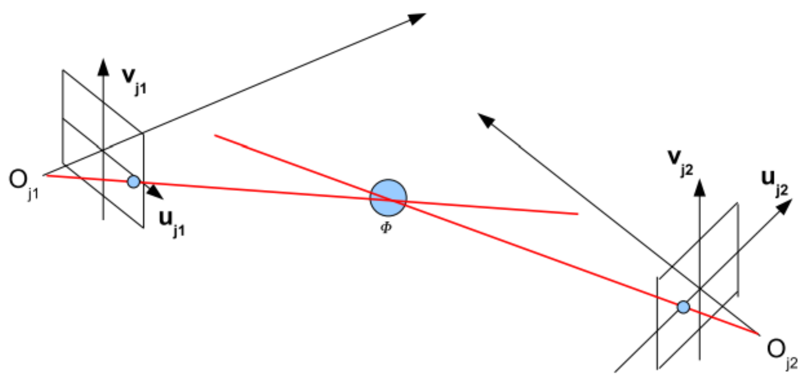


Fig. 7. Triangulation from two cameras as shown in [35, p.2]

The process of triangulation from two cameras is common the the field of computer vision and MoCap, and is well described by Masiero and Cenedese [35]. Their figure, as shown in Figure 7, shows how 3D position Φ can be determined once camera parameters from cameras Oj1 and Oj2 are known. The figure highlights how a line from camera sensor point O slices through a rectangular V by U sized pane that represents the recorded frame. One of the intrinsic camera parameter is the focal length, which in the figure can be seen as the distance from origin point O perpendicular to the rectangular pane. An essential extrinsic camera parameter is how exactly the two cameras relate to each other in physical space. It now is clear why camera calibration is so essential for MMC.

5.5 Data Preparation

Once triangulated, the data needs to be pre-processed in order to extract relevant artifacts later. The goal of pre-processing is to make the data more predictable and robust.

5.5.1 Filtering. All data is filtered using a forward-backward Butterworth filter with the following filter parameters: order = 2, sampling rate is equal to the framerate and a cutoff frequency of 3. The forward-backwards approach is used to avoid phase distortion that occurs with a single-pass Butterworth filter. This phase distortion can result in a 'loading effect' where it takes a couple of frames for the recorded spatial data to be accurate, as the filtered points depend on data entries before it. By first running a Butterworth filter in forward direction and then again in the reverse direction, any introduced phase shift is cancelled out. The forward-backward filtering approach is inherently terminal, meaning it requires the entire signal to be available upfront. Therefore, this approach would not work in a concurrent implementation. To minimize any edge effects, the signal is padded on either side to provide additional context for the filter at the signal boundaries. Right before filtering, any missing data is interpolated. Missing data is introduced to the signal if any joints were obstructed or OpenPose for whatever reason did not pick up on them during data collection. Interpolation is crucial as the Butterworth filter requires a continuous input signal to function properly.

5.5.2 Normalization. All data is normalized according to a predetermined height so that all stick figures are the same height. This is important when visualizing a recording, and makes the distance between the subject and the two cameras irrelevant. Obviously all joints need to be in frame to measure all keypoints at any time, but normalization played an important factor in providing consistent feedback to subjects during the experiment. In order to do so, the average vertical distance between the head keypoint and the two heel keypoints is measured over all recorded frames. A scaling factor is determined by dividing the desired vertical distance with calculated average vertical distance. In a second pass over all frames, all keypoints are scaled accordingly. Besides scaling for vertical height, all data is shifted along the horizontal X-Z plane. After shifting the recorded data, heel keypoints are at location (0,0,0) and the stick figure is therefore standing at the exact same location in each recording. Just like normalization, the stick figure is not shifted on every frame. Instead, the average position is determined in a first pass-through and the entire recording is shifted in a second pass-through. This way, if a subject were to take steps during a recording, this shift in position along the horizontal plane remains in the final data.

5.5.3 Rotate and Invert. During the development of the prototype it became clear that recordings were consistently upside down. The Y values needed to be flipped to obtain the desired result: consistent upright stick figures. After being inverted, the stick figure data is rotated around the Y-axis so that the line between the two measured hip keypoints is, on average throughout the recording, parallel to the Z-axis. By doing so, deviations in camera placement can be neglected when visualizing recorded data, as the stick figure will always face in the same direction.

5.6 Augmentation and Error Extraction

5.6.1 Vertical Error. One of the goals of a perfect squat is to squat down until your thighs are parallel to the floor. In this prototype, this state is achieved when the average vertical values of both hip keypoints are equal to the average vertical values of the knee keypoints at the deepest point of the squat. If these two values are not equal, there is a Vertical Error and the subject either squatted too deep or not deep enough. This deviation is the identified error and can be used to amplify according to the principles of EA. More specifically, the Vertical Error is an example of mapped-EA as described in Section 3.1.3 and will be amplified with a gain of 2. The prototype first extracts the error from the normalized data, which is doubled for augmentation purposes. In a second run, the prototype extracts the error in millimeters, which provide a better insight into the squat performance.

5.6.2 Balance Error. As described earlier, it is important that the center of mass does not move forward too much during the performance of a BW squat. This is often translated to: do not move your knees in front of your toes during the entire movement. In this prototype, the Error Augmented Feedback implements an over-time approach and amplifies this error as the movement progresses. The balance error is calculated along the X-Z plane and therefore excludes any vertical differences. A vector perpendicular to the line between both hip points determines the forward-facing direction. The balance error is the determined distance with which the knees are further along the perpendicular hip vector, averaged over both individual legs. The augmented data is altered over-time, which means that the error is amplified on each frame. The prototype stores the maximum measured balance error in millimeters for analysis purposes.

5.6.3 Knee-Heel Error. The Heel-Knee Error for one leg is defined as the angular deviation between the line from hip to knee and the line from the heel to big toe. Per this definition, the goal for the ideal squat form is to keep the knees pointed in the same direction as the feet throughout the squat motion. Similar to the Balance Error, this error too ignores vertical deviations by only including X-Z positional data. The Knee-Heel Error is calculated over time for each leg individually. Once calculated, the desired angle for each frame is calculated by multiplying the error with a gain of two, and the new knee position is calculated accordingly. The final Heel-Knee Error is the average absolute error from both legs. Unlike the Balance Error, the Knee-Heel Error takes the average angular error over the full movement.

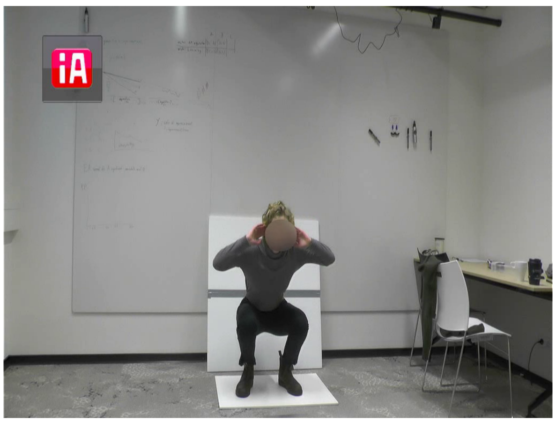
5.7 Data Visualization

The visual feedback is provided after every individual performed squat. Depending on the the participant's group, they receives no feedback (control group), TF (placebo group) or error augmented feedback (test group). For both the placebo and the test group, the feedback is generated based on recorded 3 dimensional positional data and consists of two parts: a stick figure and green target panes that showcase the target positions for a certain angle. The prepared BODY-25 keypoints, as listed in Table 5 are connected as shown in Table 6 to create an animated stick figure. The visual feedback is an animated video and is generated for three angles namely i) a front view, ii) a side view, and iii) a 3D view, which are displayed one after the other respectively. Figure 8 shows an example of generated TF (placebo group) output frames based on the recorded camera input.

5.7.1 Green Target Panes. On all three generated feedback views as shown in Figure 8c-e, there are additional green panes that help guide the subject towards their ideal performance. The horizontal pane that is visible from the front and 3D view is placed on knee-height at the detected deepest point of the recorded squat. The vertical pane visible from the side view is placed at the average X position of both big toe keypoints, which worked because the stick figure is rotated so that the line between both hips on average was parallel to the Z axis.

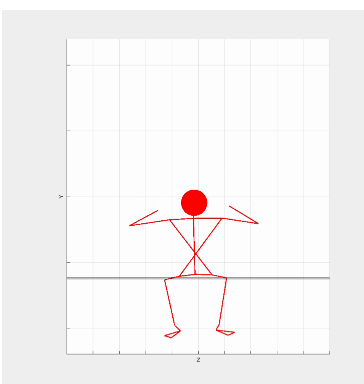
Connection	Description	Connection	Description
1-2	Nose to Neck	13-14	Left Hip to Left Knee
2-3	Neck to Right Shoulder	14-15	Left Knee to Left Ankle
3-4	Right Shoulder to Right Elbow	15-22	Left Ankle to Left Heel
4-5	Right Elbow to Right Wrist	22-20	Left Heel to Left Big Toe
2-6	Neck to Left Shoulder	20-21	Left Big Toe to Left Small Toe
6-7	Left Shoulder to Left Elbow	21-22	Left Small Toe to Left Heel
7-8	Left Elbow to Left Wrist	12-25	Right Ankle to Right Heel
2-9	Neck to MidHip	25-23	Right Heel to Right Big Toe
13-9	Left Hip to MidHip	23-24	Right Big Toe to Right Small Toe
9-10	MidHip to Right Hip	24-25	Right Small Toe to Right Heel
10-11	Right Hip to Right Knee	10-6	Right Hip to Left Shoulder
11-12	Right Knee to Right Ankle	13-3	Left Hip to Right Shoulder

Table 6. Stick Figure Connections for OpenPose's BODY 25

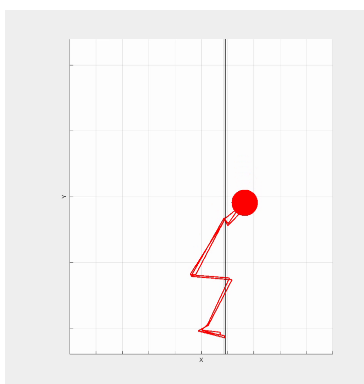


(a) Camera 1 Input

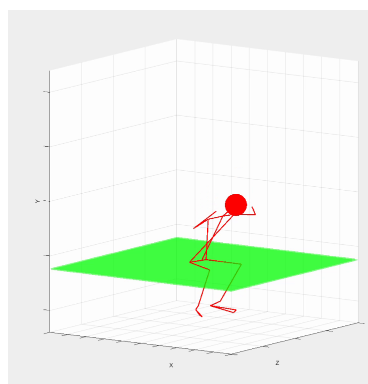
(b) Camera 2 Input



(c) Front View



(d) Side View



(e) 3D View

Fig. 8. Generated Visual Feedback Based on Camera Inputs

5.8 Materials

Table 7 provides an accurate list of materials used to realize this prototype. The cameras could be replaced by any consumer grade camera that outputs an color video signal. This prototype was built for a powerful computer in an attempt to speed up the analysis process in order to provide feedback within a reasonable time frame.

Item	Description
2 Cameras	Panasonic HC-V52, resolution of 1920x1080 at 10.0 MP. The system is flexible and is not limited to one specific type of camera.
2 Tripods	
2 Mini HDMI to HDMI Cables	
2 Capture Cards	Through the cables and capture cards, the cameras act like USB webcams and can be treated as such within the MATLAB software.
Calibration Board with Checkerboard Pattern	23x32 checkerboard pattern with 25mm checker size printed on 10mm forex PVC foam board, generated by www.calib.io [2].
Computer	Powerful PC (GeForce RTX 4080, AMD R7 7800X3D 8-Core, 32GB RAM) required to analyze video footage as fast as possible to provide timely terminal visual feedback.
1TB SSD	All data is written straight to encrypted and password protected external storage.
Monitor	
Speaker Set	Used for pre-programmed countdown sounds for consistent data collection.

Table 7. List of Materials and Descriptions

5.9 Prototype Validation

In order to gauge the accuracy of the prototype, it is important to compare it to a known ground truth of MoCap technology: the MVN Link. Five BW squats are recorded using both the developed MMC-based prototype and the MVN Link suit simultaneously, as seen in Figure 16. The squats are compared based on recorded knee angles for both legs. Figure 10 shows the recorded knee angle data for all five squats over time for both technologies. In the MMC-prototype, the extracted right knee angle is defined as the angle in the line from the Right Hip (keypoint 10) to the Right Knee (keypoint 11) to the Right Ankle (keypoint 12). The left knee angle is calculated from keypoints 13, 14, and 15. The MVN Knee Angles are extracted from the exported MVN Link data by looking at the columns "Right Knee Flexion/Extension" and "Left Knee Flexion/Extension". The MVN data is recorded at 120FPS, whereas the MMC-based prototype records data at 30FPS, which is why the MVN data is downsampled to accurately compare the differences.

(a) Frame capture using MMC (b) Frame capture using MVN (c) Frame capture from video

Fig. 9. Comparative analysis using MMC, MVN, and video.

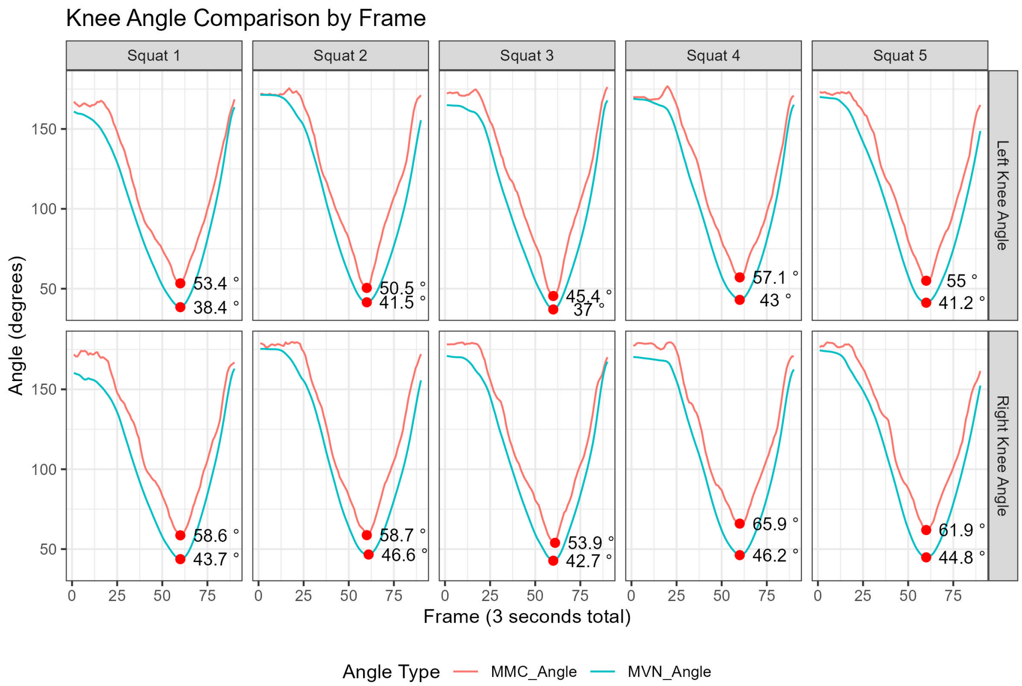


Fig. 10. Comparison of Knee Angles Between MMC and MVN

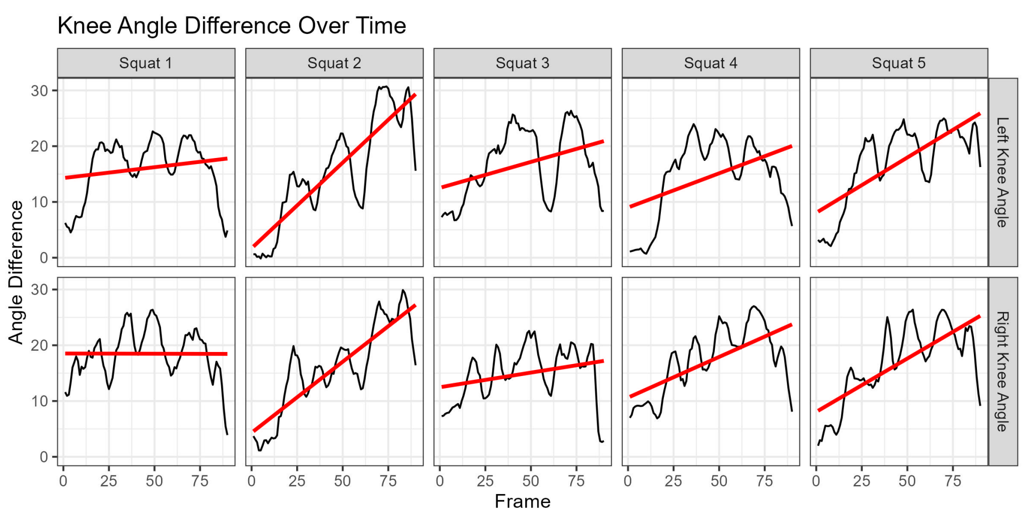


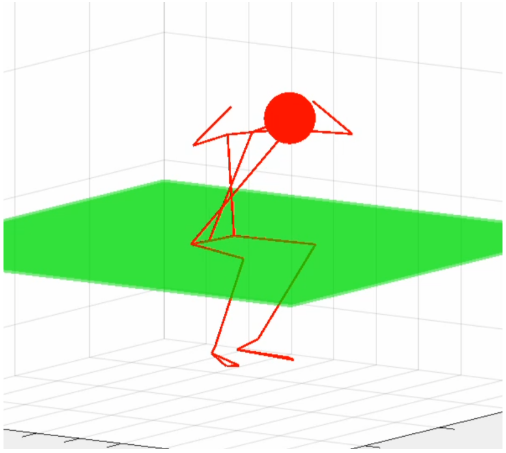
Fig. 11. Angle Difference Over Time

Overall, there is a mean deviation of 16 degrees between the MMC and MVN data. Figure 11 shows the angular deviation over time, with a fitted regression line for each facet. The figure shows that the angular deviation is inconsistent. Table 8 provides the descriptive statistics of the angular deviation per squat. Although the deviation seems constant, the total range of Squat 2 reveals a minimum of -0.15 degrees and a maximum of 30.71 degrees throughout the recording of the squat. This difference is noteworthy and needs to kept in mind when analyzing any recordings made by our prototype. Another important difference between both signals is the noticeable noise in the MMC signal compared to the much smoother MVN signal, even after an implemented Butterworth filter. Because of this noise it will be challenging to analyze minor mistakes in recorded movement data, which is something to take into account when drawing conclusions on recorded error signals.

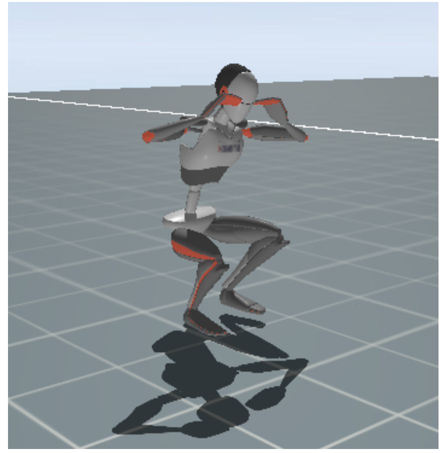
Squat	**Mean**	**Median**	**SD**	**Min**	**Max**
Squat 1	17.27	17.96	5.09	3.74	26.37
Squat 2	15.79	15.50	8.55	-0.15	30.71
Squat 3	15.81	16.22	5.64	2.66	26.36
Squat 4	15.92	17.42	6.62	0.71	27.01
Squat 5	16.92	18.56	6.85	1.98	26.39

Table 8. Descriptive Statistics for Squat Angle Differences Across Five Squats (degrees)

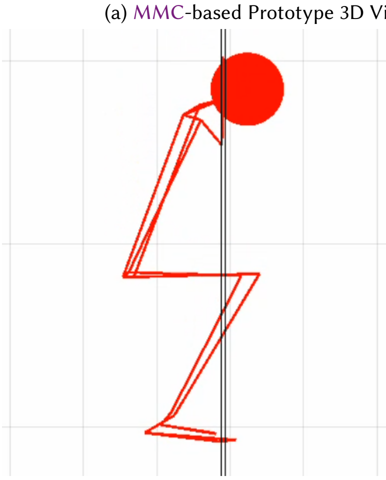
The main purpose of the MMC-based prototype is to provide visual feedback on a subject's squat performance. Therefore, a visual comparison as shown in Figure 12 displays the output of both MoCap technologies of the first validation trial at the deepest recorded point. A close visual inspection reveals great similarities between the body structure of the two avatars. These are excellent results and suggest that significant deviation in recorded angles could have a different cause. One of the underlying causes how the technology estimates keypoint position. Analysis of the OpenPose progress, as shown in Figure 13, that shows the fitted stick figure over any given image, shows that keypoints are estimated inaccurate due to various reasons. Identified reasons are **i) point of interest (partially) obstructed by bodyparts**, as can be seen at the right hip and right shoulder on Fig 13c, **ii) inaccurate estimation due to motion blur**, as can be seen at the shoulder keypoint in Figure 13b, **iii) low contrast due to clothes or lighting**, and **iv) OpenPose treats each frame separately**, which results in jittery and noisy signals. All in all, it is important to understand that MMC has the ability to generate convincing positional data but measurements can fluctuate, which can make drawing accurate conclusions based on extracted data unreliable.



(a) MMC-based Prototype 3D View

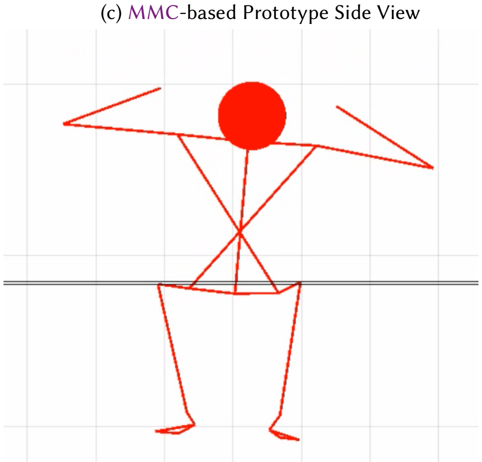


(b) MVN Link 3D View

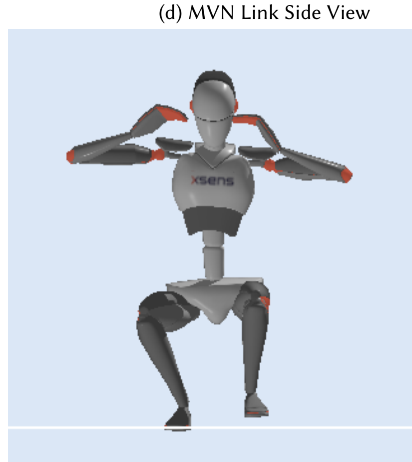


(c) MMC-based Prototype Side View

(d) MVN Link Side View

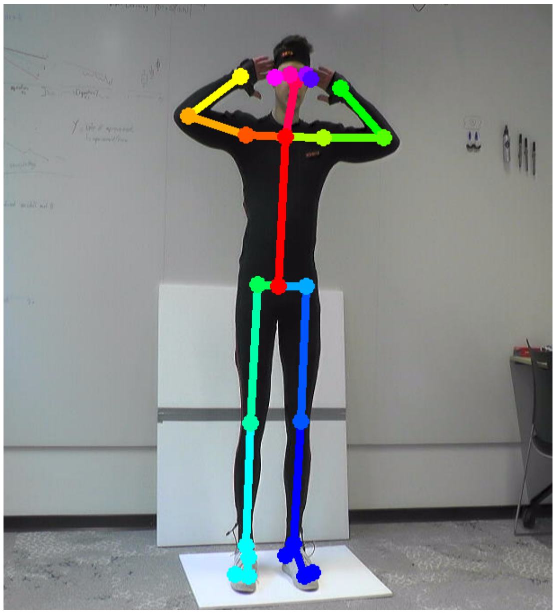


(e) MMC-based Prototype Front View

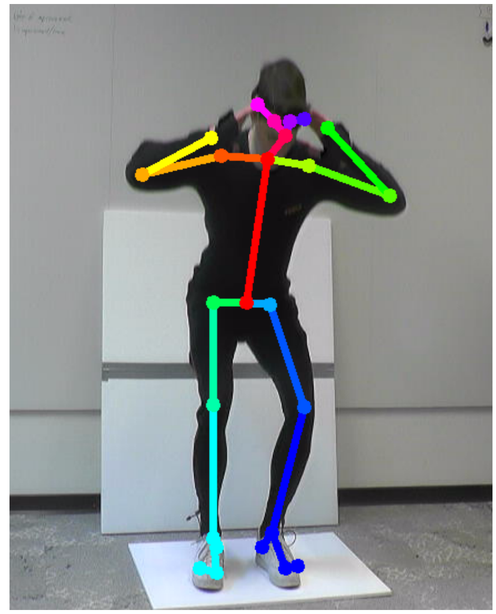


(f) MVN Link Front View

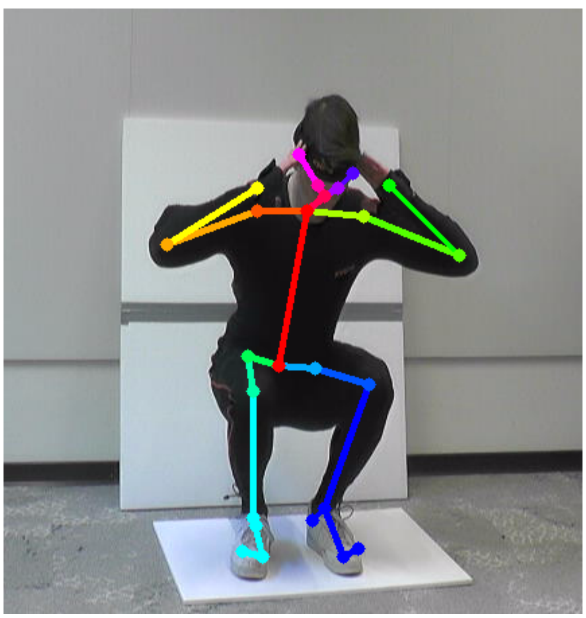
Fig. 12. Pose Comparison of Prototype and MVN Link



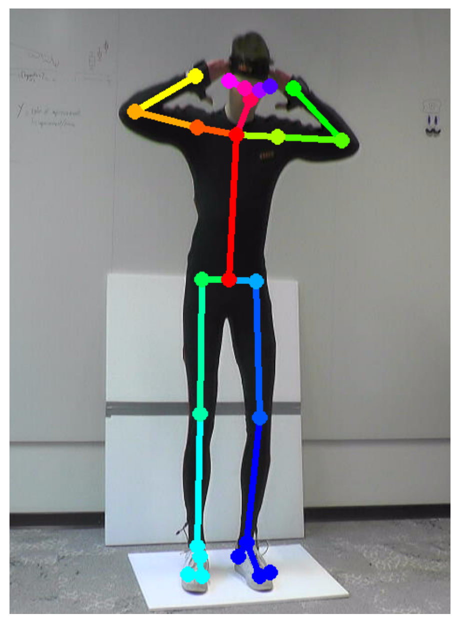
(a) Initial squat position with MMC overlay



(b) Descending phase of squat with MMC overlay



(c) Mid squat position with MMC overlay



(d) Ascent phase of squat with MMC overlay

Fig. 13. Sequential frames of a squat motion captured with MMC technology overlay.

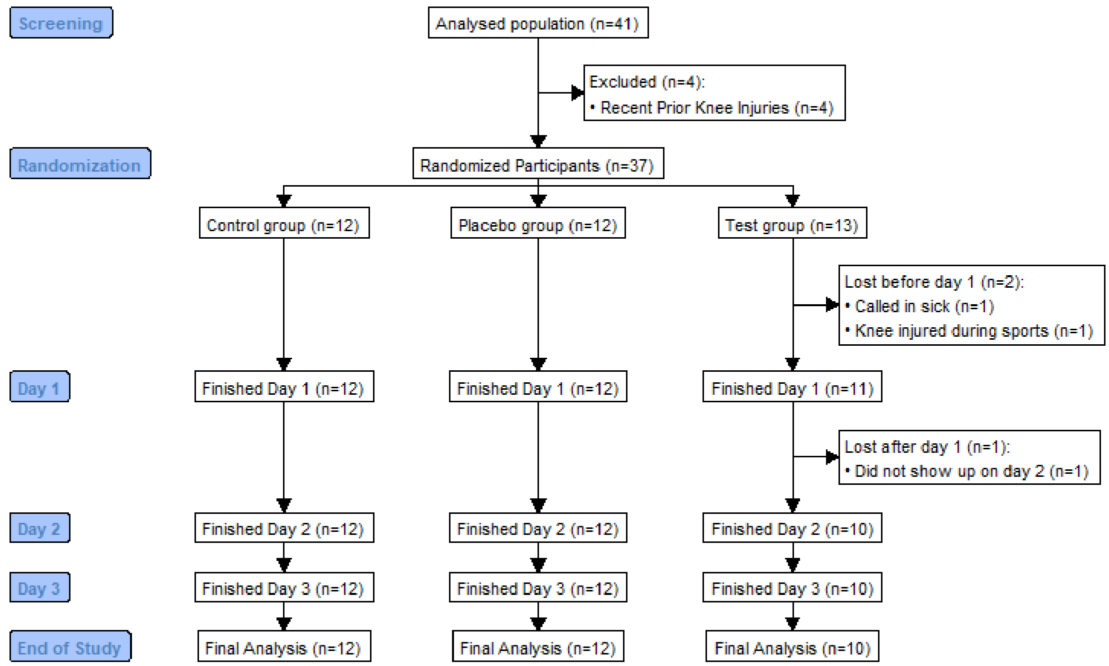


Fig. 14. Participant recruitment flow diagram.

6 MOTOR LEARNING STUDY: EFFECT OF EA ON SKILL ACQUISITION

In order to investigate the impact of EA on motor learning, as applied by our prototype, a motor learning study is conducted where participants receive terminal visual AF to improve their BW squat execution. Over a span of six weeks, participants were invited to perform 45 BW squats over a period of three days and perfect their technique through terminal visual AF. Comparing the progression of VE, BE, and KHE over the span of three days between three groups could potentially reveal a significant impact of EAF compared to TF and no feedback.

6.1 Subjects

A total of 34 subjects (mean age=22.5, SD=2.61, male=18, female=16) completed the experiment. The consort plot in Figure 14 shows that participants were divided into three groups at random. Participants were recruited through convenience sampling and were all students at the University of Twente. Participants were excluded from participating if they experienced knee injuries from which they were still recovering. A questionnaire among all participants revealed that the majority of the participants exercise 1-2 days per week and consider themselves to be moderately physically active.

6.2 Study Design

Participants are divided into three groups. Group 1: **test group** that received EAF, Group 2: **placebo group** that received unchanged TF and Group 3: **control group** that received no feedback in between squat repetitions. Table 9 shows the experiment schedule executed by all subjects on three consecutive days. Note that participants in Group C did not receive any visual feedback throughout the experiment.

6.2.1 Physical Setup. The experiment took place in a controlled environment at the University of Twente. Two cameras were placed on tripods 1.8 meters apart and 1.5 meters from the ground. The cameras were level and faced forwards, making sure they were not tilted upwards or downwards. The distance between both cameras and the marked location where the subjects perform BW squats was roughly 3.5 meters. Participants stood on a white PVC plate in front of a white background to ensure visual separation from the background. The experiment took place in a room with natural light coming in from wall, illuminating all participants on the right side of their body when facing the cameras.

6.2.2 Experiment Planning. Table 9 shows the three day planning of the experiment. The acquisition phase of this experiment consists of all squats where feedback is provided and is bolded in Table 9. The acquisition phase consists of 30 trials spread over the first two days. After a two day acquisition phase and a one day retention interval, subjects do a retention test on day three. This study design has been inspired by the works of Wulf et al. [49]. The session on day one and two lasted about 45 minutes for the groups that received feedback and 30 minutes for the control group. Day 3 took approximately 20 minutes for all groups. All three days started with a follow-along warming up and a pre-recorded video explaining the desired BW squat technique. On the first day all participants read an information letter, signed a consent form and watched a general instruction video explaining the purpose of the study. The test group was unaware they were going to receive EAF throughout the whole experiment and received a debriefing after completing the questionnaire on day 3.

Day 1	**Day 2**	**Day 3**
Signing Consent Form	Warming Up	Warming Up
Instruction Video	BW Squat Technique Video	BW Squat Technique Video
Warming Up	5 squats (no feedback)	5 squats (no feedback)
BW Squat Technique Video	**5 SQUATS (WITH FEEDBACK)**	Questionnaire
5 squats (no feedback)	BW Squat Technique Video	Debriefing
5 SQUATS (WITH FEEDBACK)	**5 SQUATS (WITH FEEDBACK)**	
BW Squat Technique Video	**5 SQUATS (WITH FEEDBACK)**	
5 SQUATS (WITH FEEDBACK)		
5 SQUATS (WITH FEEDBACK)		

Table 9. Experiment planning day by day (**CAPITALIZED** shows acquisition phase)

6.3 Ethical Considerations

As mentioned in Section 3.1.4, EA research lends itself well to not disclosing the modification of the EAF. Therefore, this experiment will use deception and make participants in the test group believe that their visual feedback is an accurate and unmodified representation of their BW squat performance. In order to conduct research that includes deception, researchers need to submit an application to the ethical committee. In this request for approval, all foreseen risks need to be outlined, as well as the envisioned solutions to mitigate any risk. To mitigate the risk of injuries caused by performing BW squats, participants with prior knee-related injuries are excluded from participating and a 4 minute warm up is added before each session. When it comes to EA, the deception in the study consists of passing off the motions of the animated avatar in the test group as exactly the motion of the subject, where in fact the movement is altered with a certain gain. While some participants may immediately spot that the avatar is moving differently from their own performance, other participants might not realize this. All participants in the test-group are debriefed on day 3. During the debriefing, they will be told about the principles of Error Augmentation and the fact that their movements were altered before passing them to the animated avatar.

6.4 Experiment Results

The collected data consists of 1530 recorded BW squats over 34 participants. For each trial, the prototype collected three error values: VE, BE, and KHE. From the collected data, obvious outliers are removed according to defined thresholds as shown in Table 10. The full data after the outlier removal process is plotted in Figure 15, with corresponding descriptive statistics in Table 11. The full data per participant is append in Appendix B, C, and D for all three error types. All data is split into two broad categories: **test data** and **acquisition data**. The *test data* consists of the five pre-test trials and 5 retention-test trials, and is used for motor learning analysis by conducting an Analysis of Variance (ANOVA) on the individual improvement scores. The *acquisition data* consists of all trials where participants in the test and placebo group received visual EAF, namely 15 trials on day 1 and 15 trials on day 2. The *acquisition data* is analysed by conducting three Linear Mixed Effects Regression (LMER) models as shown in Table 12. Appendix F shows plots used to validate the following assumptions: i) **Linearity**, ii) **Absence of collinearity**, iii) **Homoskedasticity**, iv) **Normality of residuals**, v) **Absence of influential data points**, and vi) **Independence** [48].

Error Type	Lower Limit	Upper Limit	# of Removed Outliers
Vertical Error	-500 mm	500 mm	1
Balance Error	0 mm	250 mm	28
Knee Heel Error	0 degrees	35 degrees	0

Table 10. Outlier Removal Thresholds per Error Type

Group	Day	VE Mean (mm)	VE SD	BE Mean (mm)	BE SD	KHE Mean (degrees)	KHE SD
1	1	47.55	67.10	52.77	37.27	9.96	3.55
1	2	36.71	33.54	62.08	46.59	11.35	4.70
1	3	15.84	40.98	53.47	42.98	11.57	3.57
2	1	65.09	63.71	48.41	40.19	8.56	3.77
2	2	34.39	61.12	53.55	40.43	8.94	2.93
2	3	20.13	47.90	42.89	36.70	9.17	2.82
3	1	72.50	82.92	48.52	33.30	10.00	4.78
3	2	70.45	73.78	47.55	37.63	10.76	5.00
3	3	72.28	67.96	38.41	25.00	10.03	4.04

Table 11. Descriptive Statistics per Error Type: VE, BE, KHE

Error Type	Model	Conditional R2	Marginal R2
Model 1	Vertical Error ~ Group + Day + Trial + Day*Trial + Group*Trial + Group*Day + Group*Trial*Day + (1 | ParticipantID)	0.741	0.064
Model 2	Balance Error ~ Group + Day + Trial + Day*Trial + Group*Trial + Group*Day + Group*Trial*Day + (1 | ParticipantID)	0.637	0.017
Model 3	Knee Heel Error ~ Group + Day + Trial + Day*Trial + Group*Trial + Group*Day + Group*Trial*Day + (Trial | ParticipantID)	0.702	0.051

Table 12. Linear Mixed Effects Models

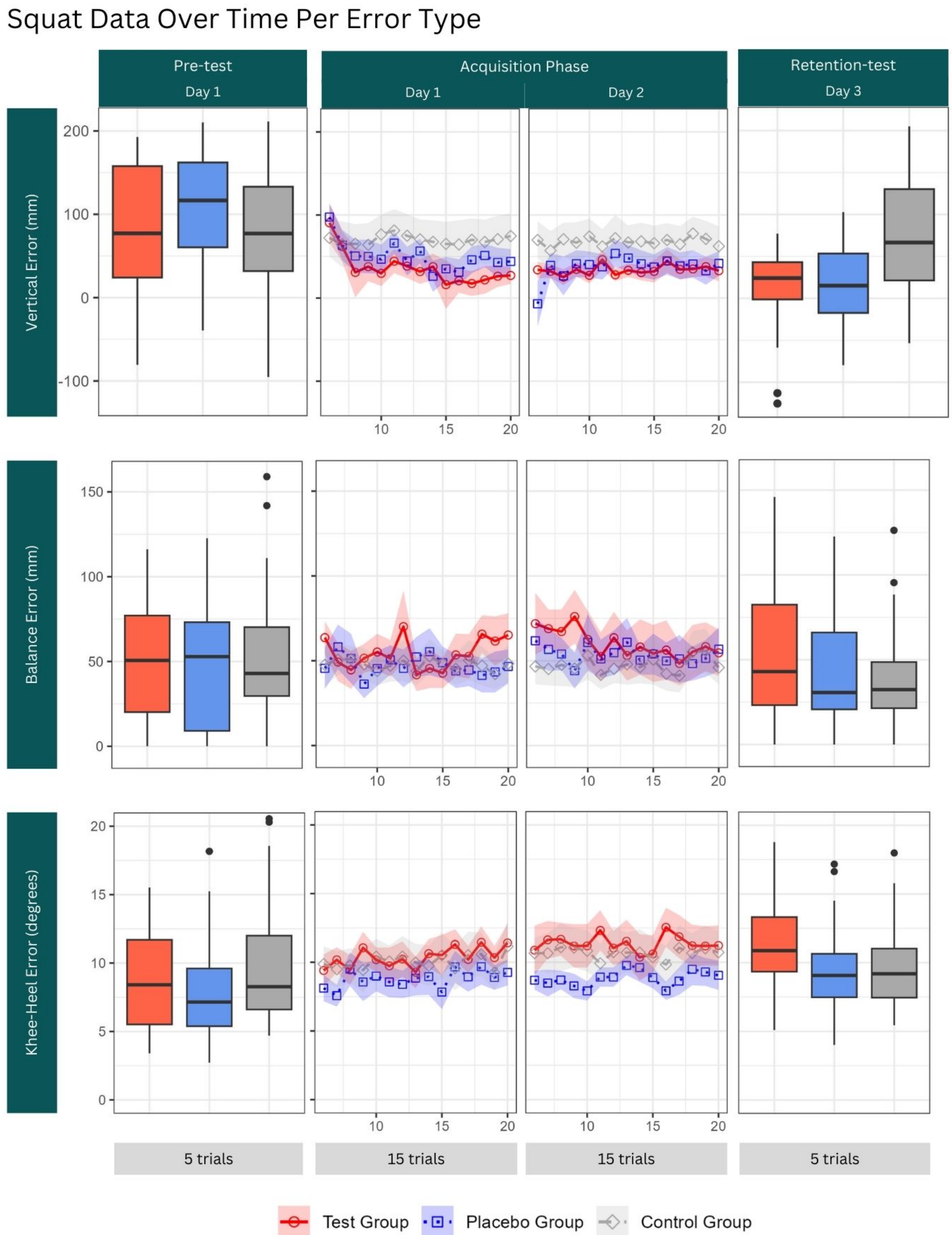


Fig. 15. Collected data split by error type (VE, BE, and KHE) and divided into three columns: pre-test, acquisition phase (Day 1 and Day 2), and retention-test. Pre-test and retention-test are aggregated box plots over 5 trials split by group. Acquisition data shows mean value with opaque standard errors.

6.4.1 Motor Learning - VE, BE, and KHE. The ANOVA conducted on the individual VE improvement score, defined as the difference between a person's average pre-test and retention-test VE performance, showed significant results, $F(2, 31) = 7.656$, $p = .002$. The test-group and the placebo group showed significant improvements in VE relative to the control group. The ANOVA for the BE ($F(2, 31) = 0.586$, $p = .563$) and KHE ($F(2, 31) = 1.837$, $p = .176$) are not significant. Figure 16 shows the individual performance difference between the pre-test and the retention-test.

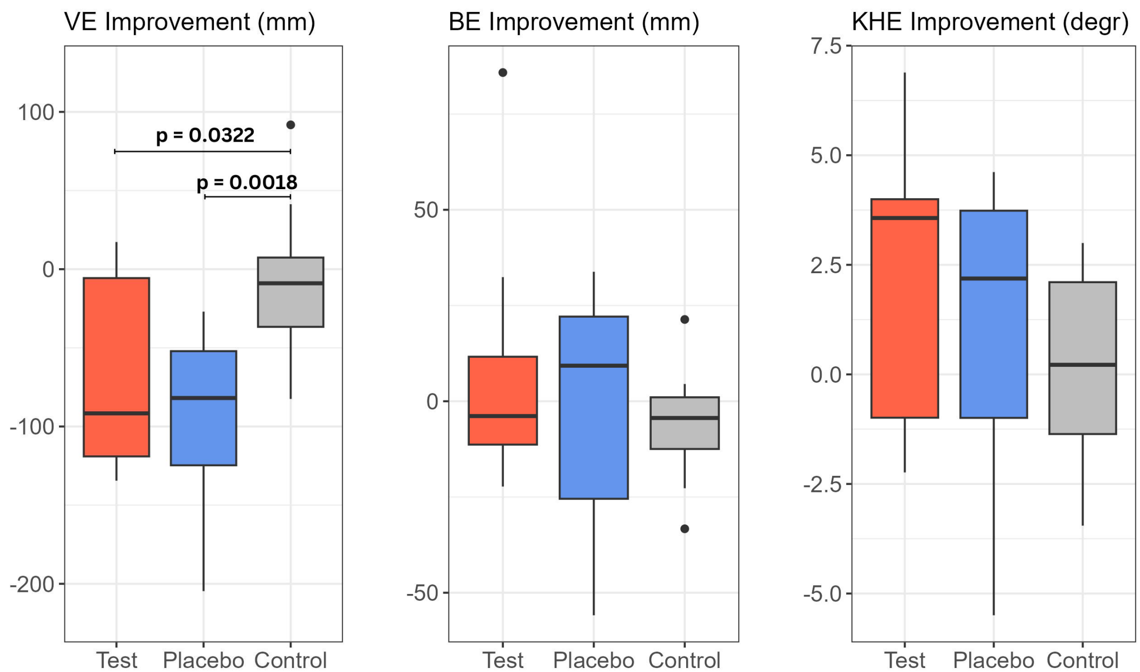


Fig. 16. Individual Improvement Scores of VE, BE, and KHE

Post hoc comparisons using the Tukey HSD test on the VE data, which controls for type I error across multiple comparisons at a 95% confidence level, further highlights these differences by revealing any differences between the individual three groups as shown in Table 13. The mean error for the control group was significantly larger than that for the test group (diff = 60.14, $p = .032$) and the placebo group (diff = 81.84, $p = .002$). No significant difference was found between the test and placebo group (diff = -21.70, $p = .608$). These results show that improvements seen in the test group and the placebo group, the groups that received visual feedback, were significantly greater than those in the control group. No significant differences are observed between the test group and the placebo group.

Comparison	diff	lwr	upr	p adj
placebo-test	-21.70	-77.41	34.01	0.61
control-test	60.14	4.43	115.85	0.03
control-placeo	81.84	28.72	134.96	0.00

Table 13. Tukey Multiple Comparisons of Means

6.4.2 Acquisition Performance - Vertical Error. Table 14 shows the effects obtained from running Model 1 (Table 12) on the acquisition data. Figure 17 shows the predicted VE values over time per group and day. There are six significant effects. **Day2** (SE = 12.2954, p < 0.001) has a -46.6mm error reduction on the intercept. **Trial** (SE = 0.6347, p < 0.001) reduces error with -3.09mm. The two-way interaction effect **Day2xTrial** (SE = 0.8975, p < 0.001) indicates that the increase of Trial has a different effect on Day 2 over Day 1. **Group3xTrial** (SE = 0.8593, p < 0.001) shows that Trial has a different effect on Group3 compared to Group 1. **Group3xDay2** (SE = 16.6481, p = 0.0102) shows that Day2 has a different effect on Group3 compared to Group1, and **Group3xDay2xTrial** (SE = 1.2153, p = 0.0063) is plotted in Figure 14 and tackles the difference in slope (progression of predicted error as Trial increases) on Day2 vs Day1, compared between Group3 and Group1.

Effect	Estimate	Std. Error	df	t value	Pr(>|t|)	Significance
Intercept	75.8301	19.2360	52.2477	3.942	0.000 241	***
Group2	4.6283	26.0499	52.2818	0.178	0.859 669	
Group3	−6.3413	26.0456	52.2477	−0.243	0.808 597	
Day2	−46.6214	12.2954	985.0006	−3.792	0.000 159	***
Trial	−3.0980	0.6347	985.0006	−4.881	1.23×10^{-6}	***
Day2:Trial	3.4577	0.8975	985.0006	3.852	0.000 125	***
Group2:Trial	0.7519	0.8608	985.0034	0.873	0.382 652	
Group3:Trial	3.1414	0.8593	985.0006	3.656	0.000 270	***
Group2:Day2	−13.4682	16.6548	985.0012	−0.809	0.418 902	
Group3:Day2	42.8431	16.6481	985.0006	2.573	0.010 214	*
Group2:Day2:Trial	0.1859	1.2163	985.0020	0.153	0.878 534	
Group3:Day2:Trial	−3.3239	1.2153	985.0006	−2.735	0.006 348	**

Significance levels: $^{***}\ p < 0.001$, $^{**}\ p < 0.01$, $^{*}\ p < 0.05$, $.\ p < 0.1$

Table 14. Vertical Error: Fixed Effects Results with Significance Levels

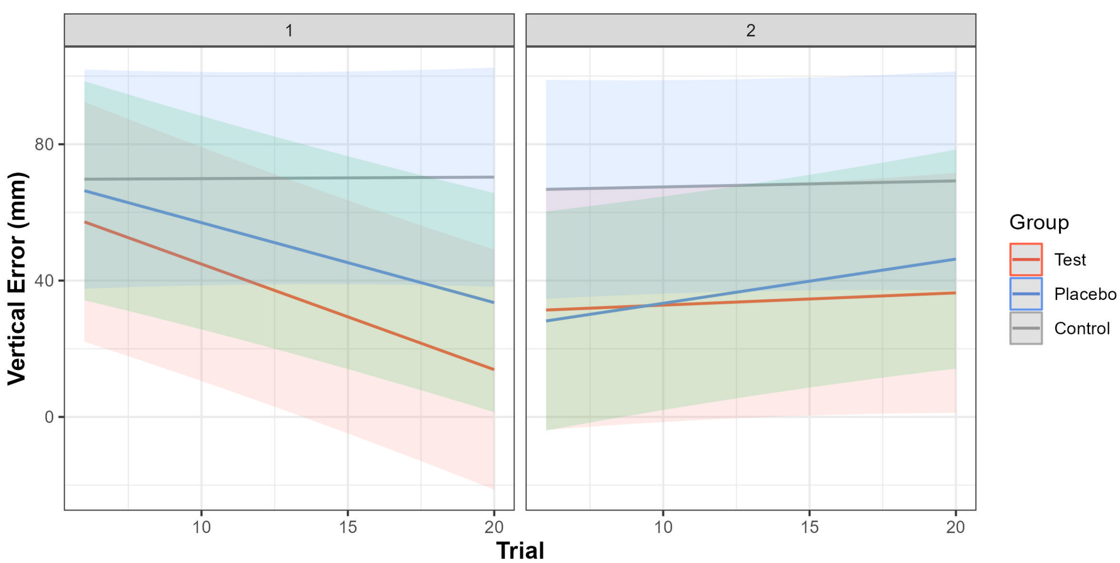


Fig. 17. Vertical Error TrialxGroupxDay Effects

6.4.3 Acquisition Performance - Balance Error. The BE acquisition data is analysed through M2 (Table 12) with results shown in Table 15. The reported significant effects are **Day2** (SE = 8.7224, p < 0.001) with an overall increase in predicted error of 28.88mm, **Day2xTrial** (SE = 0.6343, p = 0.004) that shows the effect of Trial on Day2 changes the predicted BE with -1.81mm compared to Day1. **Group2xDay2** (SE = 11.8333, p = 0.041) indicates that Group2 responds significantly different on the increase from Day1 to Day2 compared to Group1, with an estimate of -24.11mm. The same goes for **Group3xDay2** (SE = 11.7684, p = 0.002), with an estimate of -35.99mm estimated difference compared to Group1. The significant three way interaction effects **Group2xDay2xTrial** (SE = 0.8602, p = 0.026), and **Group3xDay2xTrial** (SE = 0.8565, p = 0.007) have an estimate change in error to their counterparts of 1.91mm and 2.30mm respectively per increase of Trial. Visualized progression of BE estimates by M2 are represented in Figure 15 that shows predicted BE prediction per group over time.

Effect	Estimate	Std. Error	df	t value	Pr(>|t|)	Significance
Intercept	47.5369	11.4424	63.0826	4.154	9.99	***
Group2	6.9476	15.5576	64.1203	0.447	0.656 69	
Group3	2.5955	15.5037	63.2536	0.167	0.867 58	
Day2	28.8895	8.7224	971.0509	3.312	0.000 96	***
Trial	0.5071	0.4439	971.0043	1.142	0.253 66	
Day2:Trial	-1.8058	0.6343	971.0350	−2.847	0.004 51	**
Group2:Trial	-1.0216	0.6079	971.0365	−1.681	0.093 15	0
Group3:Trial	-0.6886	0.6017	971.0064	−1.144	0.252 77	
Group2:Day2	-24.1092	11.8333	971.0403	−2.037	0.041 88	*
Group3:Day2	-35.9888	11.7684	971.0370	−3.058	0.002 29	**
Group2:Day2:Trial	1.9076	0.8602	971.0274	2.217	0.026 82	*
Group3:Day2:Trial	2.3023	0.8565	971.0258	2.688	0.007 31	**

Significance levels: $^{***}\ p < 0.001$, $^{**}\ p < 0.01$, $^{*}\ p < 0.05$, $.\ p < 0.1$

Table 15. Balance Error: Fixed Effects Results with Significance Levels

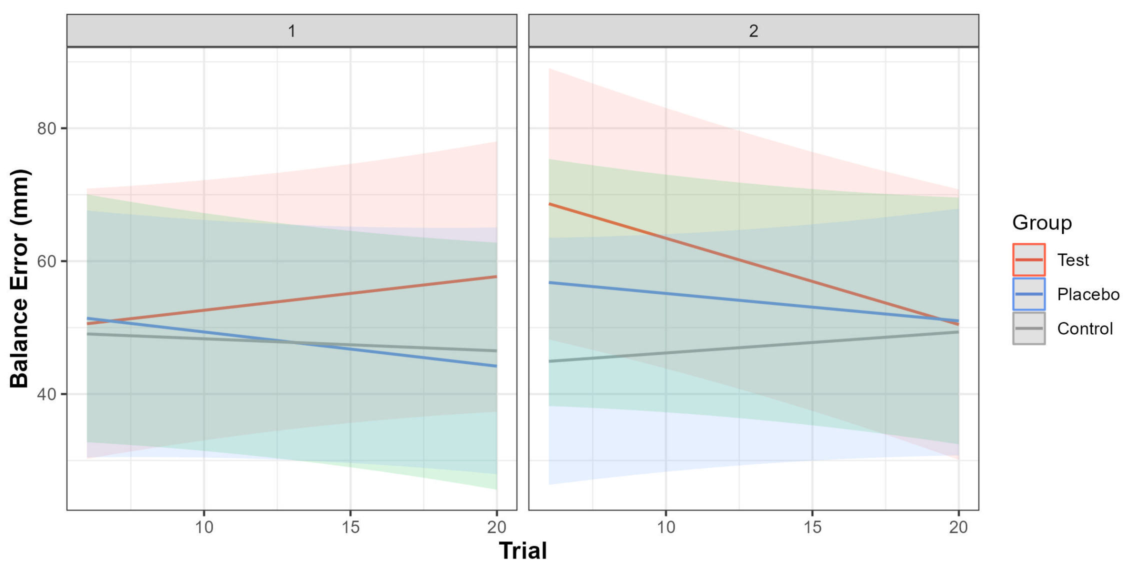


Fig. 18. Balance Error TrialxGroupxDay Effects

6.4.4 Acquisition Performance - Knee Heel Error. Analysis of KHE acquisition performance through M3 as shown in Table 12 provided the LMER results shown in Table 16. The fixed effect **Day2** (SE = 0.8614, p = 0.0089) is significant and shows that regardless of group and trial, Day2 increases the estimated intercept of the KHE by 2.238 degrees. The significant effect of Day is visualized in Figure 19, which shows the predicted KHE over trials per day.

Effect	**Estimate**	**Std. Error**	**df**	**t value**	**Pr(>|t|)**	**Significance**
Intercept	9.10555	1.202 75	44.406 23	7.571	1.59	***
Group2	-1.22233	1.628 53	44.406 23	−0.751	0.4569	
Group3	0.31289	1.628 53	44.406 23	0.192	0.8485	
Day2	2.25799	0.861 44	949.156 44	2.621	0.0089	**
Trial	0.09618	0.060 41	63.480 54	1.592	0.1163	
Day2:Trial	-0.09625	0.062 88	949.156 44	−1.531	0.1262	
Group2:Trial	-0.02678	0.081 80	63.480 54	−0.327	0.7445	
Group3:Trial	-0.04757	0.081 80	63.480 54	−0.582	0.5629	
Group2:Day2	-1.95106	1.169 19	949.502 40	−1.669	0.0955	0
Group3:Day2	-0.60579	1.169 42	949.501 15	−0.518	0.6046	
Group2:Day2:Trial	0.07639	0.085 30	949.421 40	0.896	0.3707	
Group3:Day2:Trial	0.02083	0.085 30	949.420 74	0.244	0.8071	

Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

Table 16. Knee Heel Error: Fixed Effects Results with Significance Levels

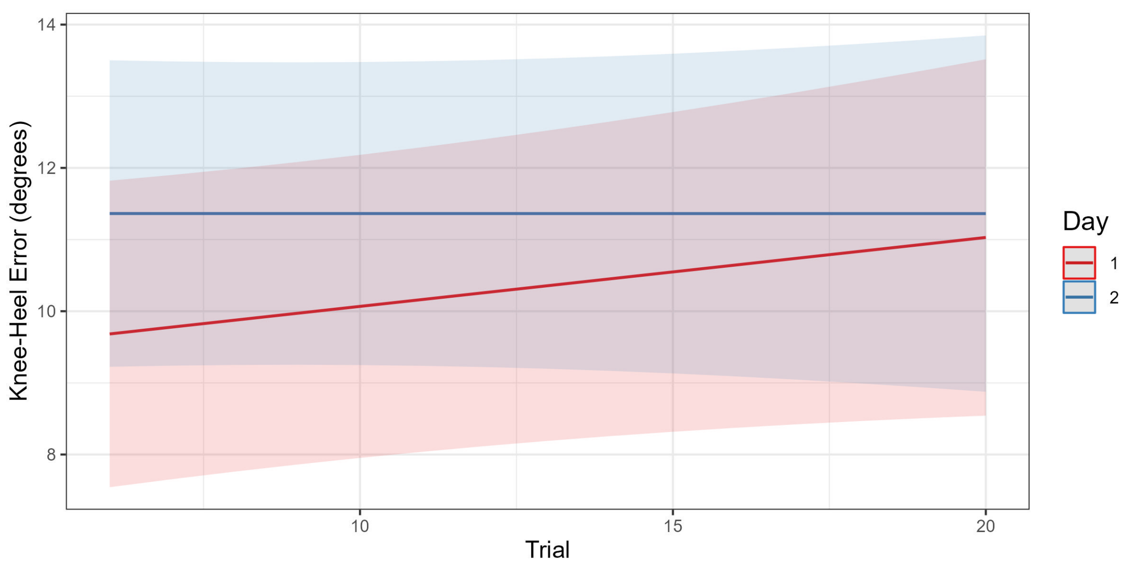


Fig. 19. Progression of Predicted KHE over Trial per Day

6.4.5 Experiment Observations. Observations while conducting the experiment gave the following results. First of all, participants show the tendency to join in when watching an explanation video on how to execute a BW squat. Next, the desire for a mirror was expressed several times throughout conducting the experiment, primarily by participants in the control group. A participant in the placebo group mentioned it felt unnatural to repeatedly perform a single squat and would have preferred groups of several squats back-to-back. Participants expressed confusion when missing data caused stick figure to jitter. The horizontal and vertical green panes were generally intuitively clear. One participant said: "*I want to go beyond the line!*", which indicates a potential counterproductive effective where participants view the line as a target to exceed instead.

In this current study design, observation is the only way to gather information on the participant's experience regarding EA. For all participants that received EAF, the first trials in the acquisition phase were important. At the start, the relatively large performance errors result in highly modified visual feedback. Remarks by participants in the test group, such as "*No way I went that deep! There's no way.*" and "*(translated) I just feel consistently bad actually*", indicate a potential effect of EA on trust and confidence of a participant. Especially at the start, when performance errors are large, the prototype generated feedback that at times did not align with the feelings and expectations of a participant.

6.4.6 Questionnaire Results. All participants completed a questionnaire on day 3, after completing the retention test. The questions asked are appended in Appendix G. The mean age of the participants is 22.5 with a 2.61 standard deviation. Out of the 34 participants, 18 were male and 16 female. Tables in Appendix H show the self-reported initial BW squat skill level, self-reported activity levels, and weekly exercise frequencies. As well as the categorical responses to perceived feedback ease of use and feedback helpfulness.

Through open questions in the questionnaire, participants were asked about their experience using the visual feedback. The single reported "Very difficult" was answered by a participant in the control group as the option "I did not receive feedback" did not yet exists. The option was added before the second participant in the control group completed the experiment. The same participant answered "Not helpful" in Table 21, which shows how helpful the visual feedback is perceived by participants. When asked about the initial impressions of the visual feedback, the participant answered the following. The majority of the people found the feedback helpful with statements saying "*it was clearly indicated how deep you had to squat*", "*really helpful visualisation of the body movement*", "*pretty clear, and easy to see the things to pay attention to*", and "*having the bit of feedback helped me make little adjustments to improve*". Participants mentioned the benefit of having the green panes as reference points, for example "*I found that the vertical and horizontal lines that were shown during the first two videos were rather helpful as it really gave me the ability to judge where I was landing.*".

The responses also reveal that the feedback on how deep participants needed to squat, represented by a horizontal green pane in the front- and 3D-view, was clearest. One participant summarized it as follows: "*Most of the visualisations were clear. Especially how deep the squat was was very clear for me. The other variables about whether I was performing the squat well were not so clear.*". Other people explicitly mentioned that it was clear to them how deep they were required to squat due to the horizontal line. A comment translated from Dutch states "*I was aware I needed to squat lower, as I ended up quite far above the line at the start*".

Besides the positive feedback, there were several points of critique and recommendations. The most common point repeated throughout the initial impressions focused on jittery animations due to poor interpolation and missing data. Participants mention "*Though the twitches or jittery was sometimes annoying*", "*sometimes the animation was a little bit weird*", "*sometimes the feedback was visibly squished or rotated*", and "*sometimes the video I was shown after my* *squats bugged a little so I couldn't decipher what I did wrong at some times*". Three people expressed that the 3D-view animation was not of value to them through comments like: "*The 3D variant of the feedback could be left out in my opinion*". Participants also expressed a desire for concurrent feedback by asking for live feedback, a mirror, or side-view cameras streaming the side profile.

When asked about the general flow of the experiment and the time between squat and visual feedback, participants mention the prolonged waiting periods. An answer that summarises the general consensus among participants is "*The overall flow of the experiment was good, that being said it took some time to receive the feedback, making it difficult to analyse a series of squats in a row.*". Another comment said: "*There was quite a long break between each squat. There might be a difference when doing singular squats or multiple squats with no pauses. The flow was fine and the experiment went quickly.*". The most critical comment among participants was: "*I found the wait to be a bit excessive, as unless I was doing a one rep max squat I doubt that I would be taking that long of a pause in between reps.*". People also mentioned the risk of getting distracted or bored, as well as the fact that quicker feedback would allow them to put the received feedback into practise more effectively.

The majority of the participants that received feedback reported they prefer squatting with visual feedback over squatting without visual feedback. In fact, among the 22 participants that completed the experiment and received visual feedback, 17 people prefer with feedback and 5 people prefer without feedback. Reasons submitted for preferring to squat without visual feedback are: "*because you can do many in a row without waiting*, *(translated) because it is easier and I do not have to think*", "*Now that I've seen many times what my form is while doing a squat, I am comfortable doing it without visual feedback.*", not being interested in further improving squat technique, and the last person demands an improved prototype in order to prefer with visual feedback, as became apparent by the response: "*With, but I would like to see improvements to the software, because right now it is not beating a mirror.*". To find potential improvements that align with the desires of the target users, participants were asked to give suggestions on improving or changing the prototype. Common factors include removing jitter and improving the animation (2x), improve the head placement to remove hunched back (4x), improve the look of the avatar (3x), faster feedback generation (2x), and participants again stressed the desire for concurrent feedback or a mirror (4x). Participants also suggested more accurate information on a potentially humped or arched back, more information on correct execution of the movement, overlaying a participant's previous performance in the form of a see-through avatar, and verbal instructions or the use of colors to provide tailored instruction to improve performance.

6.5 Experiment Discussion

6.5.1 Motor Learning. For one out of the three defined errors, namely VE, the results of the ANOVA on the *test data* suggest that Motor Learning took place. The improvement made by participants in the test and placebo group are significantly better than the control group. By comparing pre-test performance with retention-test performance, the study includes a 24 hour retention period. After this retention period, the increase in performance acquired during the acquisition phase is somewhat persistent and can therefore be classified as Motor Learning. This result shows that the visual feedback designed in Section 5.7 is effective and can promote Motor Learning in BW squats. The results of the ANOVA on the VE do not show any significant difference between the test group and the placebo group. This means that, in the *test data*, there is no evidence that suggests that EAF - as implemented by our prototype - outperforms TF. Finally, the conducted ANOVAs on BE and KHE improvement do not show any significant differences in performance between groups, indicating that the intervention (EAF or TF) had no impact on learning..

6.5.2 *Acquisition Performance.* The conducted LMER models on the *acquisition data* show a consistent significant effect of Day2 on all three defined errors. However, VE is the only error type where the fixed effect of Day2 reduces the performance error. The increase in predicted BE and KHE performance over time indicates a potential counterproductive effect of the designed visual feedback. The fixed effect of Trial significantly reduces VE, which is to be expected as an increase in Trial means an increase in practise. Generally speaking, the acquisition phase effectively reduced VE for the test and no effect on the control group. The acquisition phase increase BE for the test and placebo group, while slightly reducing BE performance in the control group. The KHE increased for the test and control group, but had a small effect on the placebo group.

Vertical Error - Visual feedback has effectively reduced VE during the acquisition phase among participants in the test and placebo group. Performance among the control group had a horizontal slope for both days. The effect of the visual feedback was most effective on the first day, as both the test and placebo group show downwards slopes on day 1 and slight upward slopes on day 2. The results show a clear significant difference in the effect of practise between the groups that receive feedback and the control group.

Balance Error - The effect of Trial on BE is significantly different on Day2 compared to Day1, with an estimate reduction of -1.805mm, shown by the interaction effect Day2xTrial. The difference in effect of Trial on Day2 of Group 1 compared to Group 2 is significant, as represented by the three-way interaction effect Group2xDay2xTrial. Group 1 slopes upwards on Day1 and downwards on Day2, whereas Group 2 slopes downwards on both Day1 and Day2. This indicates a potential difference caused by EA. A potential explanation of the upwards slope on Day 1 among participants receiving EAF could be the fact that participants are focused on reducing their VE. Evidence for this explanation is twofold, namely i) qualitative results obtained from the questionnaire show that participants found feedback on their VE clearest, and ii) the front-view feedback on their VE was shown first after each squat. The horizontal slope on Day 2 of predicted VE in Group 1, combined with the steep downwards slope on Day 2 of predicted BE in Group 1 could indicate a shift in focus of Group 1 participants from reducing VE to reducing BE. The predicted VE on Day 1 is lower for Group 1 than Group 2, which potentially means that participants in Group 1 focused more on reducing VE, sacrificing BE performance in the process.

Khee-Heel Error - The predicted KHE values indicate a potentially counterproductive effect of the implemented visual EAF. However, closer inspection shows that the predicted KHE values among participants in the control group also increases during the acquisition phase. Therefore, the increase in KHE cannot fully be contributed the the (error augmented) visual feedback. The fact that the KHE increases over time for the control group on the first day could potentially indicate the effects of fatigue. Participants in the placebo group show the lowest KHE at the end of the acquisition phase.

6.5.3 *Individual Differences.* More effort to compensate and normalize for individual differences could have given more definitive results. The most prominent improvement to the data collection process would be to normalize the findings for the Vertical Error based on the height of the participant. Out of the 34 participants that took part in the experiment, at least one person was taller than 200cm and one person was around 150cm. Normalizing the vertical error would account for these height differences, minimizing the impact on the results. Similarly, the experiment in no way accounts for different body types. Especially for the balance error, a person's body type has a large impact of the overall Centre of Mass (CoM). Despite not tracking these individual characteristics, it is important to note that the Linear Mixed Effects Regression models do account for initial individual performance differences. The models for the VE and BE include a term "(1|ParticipantID)" which accounts for random intercepts of the error prediction lines over time. The model for KHE includes a term "(Trial|ParticipantID)" which accounts for randoms intercepts as well as for random slopes. Ideally all three models would account for random slopes, but trial and error during the VE and BE model building phase revealed that models that include random slopes did not converge.

Choice of clothing impacts MMC performance, as black clothing as well as baggy clothes can lead to potentially less accurate skeleton estimations. Participants received no specific instructions on what clothing to wear other than a recommendation to: "wear sports clothes". This lack of clear instructions resulted in two things: i) **participants did not wear the same clothes on all three days of the experiment**, and ii) **there was no coherence between different participants**. Clothing worn by participants ranged from oversized and baggy clothes to tighter clothes that followed the shape of a participants body. Intuitively this has a major impact on positional data as OpenPose will struggle to accurately estimate keypoints if baggy clothes conceals the shape of the body. Clothing also ranged in terms of color, where some participants wore light shirts and pants and other participants wore dark clothes. The difference in color results in differences in contract that is recorded by the cameras, and low contrast frames are potentially more difficult to accurately review by OpenPose. One participant wore shoes of the same color as the carpet, resulting in a stickfigure with missing feet keypoints due to lack of contract. This was fixed by adding a white PVC plastic board on which participants could stand. This board was removed for participants wearing white shoes to maintain contrast. Finally, clothing ranged from even fabrics to fabrics with complex patterns and/or texts. In one instance, the pattern on a subject's sports leggings prevented OpenPose from accurately estimating keypoints to a degree where the visual feedback was unusable due to missing data and unpredictable spasms of limbs. This problem was quickly addressed by changing into another pair of sweatpants brought by the researcher. Overall, it is clear that clothing impacts the accuracy of the pose estimation but future research should reveal what clothing works best.

6.5.4 Calibration Session Inconsistencies. Despite the experiment taking place in a controlled lab environment that allowed the setup to remain in the same position over a period of several weeks, the cameras were re-calibrated several times for various reasons. Reasons include accidentally bumping against one of the tripods, moving the camera when holding the power on/off button, and having to move one camera to make space for building repairs. Although the calibration process was consistent, the orientation of the stick figure in the visual feedback deviated between calibration sessions. Although the proportions of the stick figure looked consistent, it appeared slanted at times. This slight tilt of the data impacts the results. For example, if the stick figure is slightly titled forwards, the BE increases. Furthermore, a tilted stick figure invites participants to attempt to compensate, potentially substracting from the overall learning experience. Potential solutions include fixating both cameras on one rig, making regular re-calibration unnecessary, or measuring the tilt and compensating for it with software before showing any generated visual feedback. MoCap systems generally have no native indication of what is "level" and MMC relies on the cameras being perfectly level and facing directly forward without any tilt or rotation. The OptiTrack system includes a step in the calibration process that requires an angled rod with reflective markers to be placed on the floor. Future MMC research should be aware of this phenomenon and should potentially develop more robust calibration methods that give more consistent results. For this research, adding an additional variable for the calibration session to the statistical model could help explain some unexplained variance in measurements.

6.5.5 *Study Design.*

The study design is not consistent with an ideal motor learning study due to the absence of a "post-test" and true "control group". Looking at the experiment planning in Table 9, participants start with a block of 5 squats on day 1 before they enter the acquisition phase. This block is considered a pre-test. What is missing is a post-test, which should be similar to the pre-test but measured directly after completing the acquisition phase on day 2. The difference in performance between the pre-test and the post-test is called the acquisition performance. On day 3, after a roughly 24 hour retention period, participants perform another block of 5 squats which is considered to be the retention-test. The retention-test is necessary to determine if any gained acquisition performance measured in the post-test is persistent and can therefore be called motor learning. Without the post-test, one can look at improvement measured over all trials in the acquisition phase to estimate motor learning, but data to compare performance with the pre-test is missing.

Furthermore, the control group cannot be seen as a control group, because a true control group does not partake in the acquisition phase. In a motor learning study, the control group should only participate in the pre-test, the post-test, and the retention-test. Nevertheless, the fact that the control group did participate in the acquisition phase seems to be beneficial for this study, especially because the post-test is missing. In a scenario where the post-test was missing AND the control group did not participate in the acquisition phase, the data generated by participants in the control group would barely contribute to any insights. In the case of this thesis, the control group benefits the results as it shows the trajectory of the participants that perform trials without receiving visual feedback. In the case of the VE, data from the control group shows that the performance of participants in the control group displays a horizontal trend and lacks improvement in performance over time.

Finally, the current study design does not allow for systematic analysis of the experience of EA among participants in the test group. The questionnaire is conducted before the final debriefing and therefore does not allow participants in the test group to share their experience with EAF. Observation revealed some insights regarding trust, confusion, and motivation, but this information is limited. The focus of the debriefing in this experiment focused on dealing with ethical implication and provided participants with an additional opportunity to be excluded from the experiment. There was room for the opportunity for a short unstructured interview that would have provided more insights.

6.6 Experiment Conclusion

The main question to be answered through conducting the experiment is: "**What is the effectiveness of EAF in terms of skill acquisition using our prototype?**. The quantitative results do not provide evidence that shows that Error Augmented Feedback outperforms True Feedback on promoting skill acquisition using our developed prototype. The significant ANOVA results on VE improvement indicate that the designed visual feedback can lead to improved performance compared to the control group. However, there is no evidence that suggests a significant difference between EAF and TF. Additionally, EAF resulted in slightly poorer performance during the retention-test for both BE and KHE compared to TF and no feedback. For both BE and KHE, groups with feedback did not outperform the control group, suggesting that the visual feedback provided by our prototype did not enhance performance. Qualitative insights show that terminal EAF might be confusing when not disclosed in advance, as large performance errors generate unrealistic EAF, having a potential negative impact of overall effectiveness. Factors such as inadequate camera quality, the failure to account for and track variations in individual body types, proportions, and clothing, as well as inconsistent hardware calibration, introduce uncertainties and inconsistencies that render definitive conclusions on the effects of EAF unreliable and unfeasible.

7 DISCUSSION

7.1 Advantages and Limitations of MMC

The main advantages of an MMC-based approach to providing EAF are the low cost, the accessibility, and the ease of deployment in a field setting. The fact that our prototype used two simple camcorders with an HDMI output to capture motion data is proof that this technique is accessible to a wide audience, which in turn makes MoCap research more accessible. MMC has also shown to be extremely flexible and it does not depend on one location. Our prototype has been borrowed to capture 3D motion data of climbers, and the system was ready to go within fifteen minutes. Both the experiment and implementing the system on a climbing wall revealed limitations that need to be taken into account. MMC with two cameras cannot perfectly track a keypoint when partially obstructed for one or both cameras. Even for bodyweight squats, a motion where the participant faces the cameras throughout the whole motion, analysis in section 5.9 revealed that keypoints such as the hips, shoulders, wrists and heels are hard to accurately locate when obstructed by arms and legs. Additionally, any movement that requires turning on the spot is hard to capture as this means that limbs are obstructed by the torso during the rotation. MMC also relies on hardware and calibration for accurate results. Frames captured by an overall better camera with a higher frame rate, a faster shutter speed, and a higher resolution will result in more and sharper frames, which after processing by OpenPose result in smoother and more accurate data points.

The accuracy of a MMC system depends on several variables and can be inconsistent across different days, different hardware, and different software. Variables that impact the accuracy are: camera quality, movement complexity, data (pre-)processing, and hardware calibration. Besides better cameras, the specific movement will also impact the accuracy of an MMC based system. Complex movements where bodyparts are regularly obstructed for one or both cameras result in incomplete data and rely heavily on interpolation. Interpolation will work well on data with low amounts of missing frames but is no solution for extended periods of missing data. A system with more cameras is possible and can allow researchers to capture keypoints from multiple angles. However, more cameras make the system less flexible, it will raise expenses and produce more data that takes longer to process.

Another MMC limitation related to movement type has to do with the OpenPose model limitations. The OpenPose model is trained on biased data and performs worse on uncommon movements such as somersaults, cartwheels and hand stands [28]. These type of uncommon poses are not adequately represented in the OpenPose training data, which results in a problem where skeletons are sometimes mislabelled if a person is not standing in a typical upright position. Somersault researchers have worked around this limitation by iteratively rotating and analysing each frame of a somersault video recording to find the pose estimation with the highest overall confidence to build a 3D model through MMC. When designing the BW-squat prototype, testing revealed that OpenPose struggles to identify limbs when crossed in unnatural poses. This problem is specific to MMC as it relies on an accurately trained machine learning estimation model.

7.2 EA Parameters: Gain Selection

The effectiveness of EAF on skill acquisition and motor learning depends on the selected EA parameters. When the goal is to implement EA, the error definition, gain selection, and any implemented deception have consequences on the outcome. This thesis worked with a gain of 2 in the test group. With the results of the experiment, it is not possible to confidently claim the test group with EAF outperformed the placebo group with TF. For the VE and the BE, the recorded data of both feedback-receiving-groups shows similar trajectories. However, for the KHE there seems to be a significant difference in performance between the test- and placebo group that shows that the test group performs significantly worse than the placebo group. The KHE EAF followed the over-time EA approach where the error was amplified for every frame. Each error was amplified with a gain of 2, which is a rather high gain compared to the existing literature. Researchers warned that learning becomes unstable when the gain is set too high, especially for a non-expert target group. Existing studies experimented with gains ranging from 1 (no amplification) through 3.1 [47], and the only paper currently focusing on sports ITECH gradually increased gain from 1.0 to 1.8 and reported that subjects got confused at a gain of 1.6 [7]. While a gain of two was recommended by [40], a gain of 1.5 would have been more in line with existing research and could have provided different results. During the experiment, there were instances where the gain of 2 was too high. In these instances, the performance of a bodyweight squat deviated far from the ideal trajectory. In one example, a subject in the test group would squat daily to maintain flexibly and developed a technique that included squatting down to where the hips almost touch the ground. This large deviation from a parallel squat meant that the stick figure in their visual feedback, constructed based on their performance, squats to a point where their hips disappear deep into the floor. This is the result of a gain that is too large, combined with a simplistic prototype that is not built to look for -and correct- edge cases. Researchers should build in constraints that prevent human-like avatars in visual feedback to warp beyond realistic human capabilities.

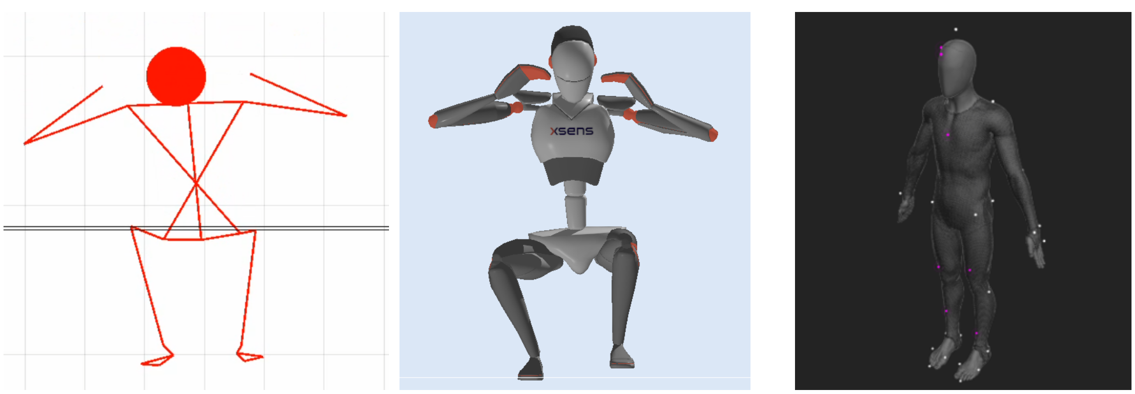
7.3 EA Parameters: Deception

The use of deception bears implications related to motivation and trust, which impacts the effectiveness of the EAF. The previous example describes one of the potential effects deception can have on a participant when the goal is to improve a motor task. Not disclosing the concept and implementation of EA could theoretically lead to rapid adaptation in performance as errors are exaggerated and easier to spot. Although the risk of potentially impacting trustworthiness and motivation were clear in advance, observing the effects of not disclosing EAF revealed that the impact needs to be taken seriously and explored further to document any effects. On each day, participants in groups one and two execute five BW squat repetitions before starting the acquisition phase in which they receive feedback. Looking at the VE, the first squat of the acquisition phase generally was not deep enough and therefore contained relatively large errors. Participants that received amplified feedback were unaware their feedback was modified and generally were confused when the stick figure they received as visual feedback barely descended during the animation. Continuing on the example raised in subsection7.2 where a subject is used to descending the hips as deep as possible, EAF, combined with the choice to decept the participants by passing off the visual feedback as if it was their own performance, could instantly dissolve any trust a subject has in the system and in the generated feedback. In another theoretical edge case the prototype could generate a stick figure that starts in an upright standing position and will ascent with the hips upwards, unrealistically elongating the legs. This could happen if a subject barely goes down with their hips during their squat, leaving a large deviation between knees and hips at the deepest recorded level (error), which is than doubled for the recording. In the edge cases where subjects received unrealistic visual feedback, or feedback that obviously did not resemble their own performance, it clearly shows that the effect of deception needs to be analysed as it could be likely that unrealistic visual feedback impacts the motivation and trust of a participant. Future work should study whether revealing the use of EA at the start of an experiment increases motivation to try and minimize errors. Disclosing how the feedback is generated could put a participant's mind at ease, boost their confidence, and increase immersion. Although the choice of deception makes sense, and the effect of EA can still be studied with deception as has been done in this thesis, disclosing the use of EA at the beginning of the experiment would be a more sensible option going forward, especially working with a non-expert target group.

7.4 Visual Feedback Design

7.4.1 Avatar. The effectiveness of EAF depends on the visual feedback design and the definition of the error signal. There are a lot of different ways that researchers are able to visualize 3D positional data. The prototype developed for this thesis utilized a simple stick figure, created by connecting the recorded keypoints with red lines. The positional data was filtered during the pre-processing stage to achieve a smooth and realistic animation that accurately represents the recorded movement. While accurate, one could argue this approach is too simplistic and does not accurately represent the subject's body. Limbs are represented by single lines and the subject's head is visualized by a featureless red sphere. Compared to the MVN Link and the OptiTrack system, the visualization indeed is much less complex, which can be seen in Figure 20. The MVN Link displays a simple humanoid robot with the option to overlay different pre-designed avatars. The OptiTrack system shows a mannequin-style avatar. When it comes to avatar design, other systems show more realistic approximations of a human and could potentially therefore provide better results when used as visual feedback to promote skill acquisition. However, it is important to acknowledge that there are effective representations of movement data that does not rely on realistic avatars. The work of Bonnette et al. mapped bio-mechanical parameters to 6 points to form an interactive gray rectangular shape [11]. While more realistic avatars could be considered more helpful, it is important to validate this assumption in the future.

One limitation of the MMC-based prototype when generating the avatar is the placement of the head. While OpenPose's BODY-25 model outputs positional data for both ears, these keypoints proved to be unreliable keypoints for two reasons: ears were often obstructed by a subject's hand during a squat, and both cameras struggled to capture the ear furthers away from the lens, resulting in incomplete and inaccurate data. Therefore, the red sphere that represents the head is placed at the nose-keypoint of the subject, which gave the impression of a hunched back or constant sloughing posture. While the overall animations are smooth and look realistic, the system can be improved by implementing a more realistic avatar. Newer pose estimation models, such as the trained model using MediaPipe [26], will provide more keypoints around the face that could be used for more accurate and realistic placement of the head and facial features. Future research can also use third party software such as Unity to overlay a complex avatar on the recorded positional data by the MMC implementation. A study should be conducted comparing several different avatars to find the most effective one to be used in future EA research.



(a) Avatar of MMC-Based Prototype (b) Avatar of MVN Link (c) Avatar of OptiTrack System

Fig. 20. Avatar Comparison

7.4.2 Bio-Mechanical Feature Selection. The bio-mechanical features and their respective defined errors could have large impact of the effectiveness of a MoCap implementation. For this thesis, the BW squat movement has been reduced to three critical bio-mechanical features based on existing literature: **squat until thighs are parallel with floor**, **knees do not exceed toes**, and **knees are pointing in line with feet**. These three features are selected because the corresponding error value could intuitively be defined, which resulted in the VE, BE, and the KHE. Current literature on the correct execution on BW squats did not differentiate between bio-mechanical features in terms of importance. By reducing to a selection of three bio-mechanical features, the prototype ignores other potentially more influential features such as a centered and neutral head position, a maintained natural curvature of the spine throughout the motion, grounded heels during lower limb flexion, or keeping elbows behind ears throughout the squat. Selecting just three features ensures that the prototype focused on a complex movement, but it did not provide feedback on the BW squat as a whole.

The authors of the Movement Competency Screen warn that it is critical to adapt for individual differences in body type and that it is less than ideal to force all subjects to perform a squat a certain way, regardless of their individual skill level and body type [30]. Looking specifically at the Balance Error, the goal should never be to keep the knees behind the toes. A deeper understanding of the complex movement shows that the goal is to keep a subject's Center of Mass (CoM) between the balls and heels of their feet. Generally speaking the CoM is too far forward when a subject's knees exceed their toes, but for different body proportions and body types this might not be the case. Looking at the features selected for this thesis, the defined Vertical Error combined with a mapped-EA approach caters well towards squatting until thighs are parallel to the floor. The Balance Error works well but should consider different body types and proportions. A more complex approach could instead estimate the subject's CoM over time, according to principles outlined in the works of Bai et al [6], and only apply EA when the CoM indicates any performance errors.

The third bio-mechanical feature aims to maintain a neutral stance where the knees point in line with the orientation of the foot throughout the squat movement. The implementation of the KHE in the code compares the vector between the heel and big toe with a vector between the hip and the knee and stores the angular deviation as the error. This error calculation is in line with prior work from Bedard and Kritz et al. [8][30] and also allows participants to place their feet in a way they felt comfortable. Results showed that applying EA on the KHE resulted in a counterproductive effect, as the KHE in the test group increased over time. A potential explanation for this could be because the error relies on the both the hip and heel keypoints, which are often partially obstructed by bodyparts. The hip keypoint is also subject to motion blur and therefore the position is estimated unreliably. A final explanation that could explain the gradual increase in error over time in the test group is consistently poor angle calculation due to the reduction of legs and hip to lines and (x,y,z) points. Take this exaggerated example: imagine that OpenPose estimates the right hip keypoint as the outer most part of part of the right hip and estimates the right knee at the inside of the knee. If a subject squats with their thigh facing exactly forward, a line between the right hip and right knee will point inwards. If this vector is compared to an accurately estimated foot-vector, the resulting error indicates that the subject has their knees pointed inwards. This angular deviation is amplified and the visual feedback will show an avatar with their knees pointing too far inwards. A pilot study in the development phase could have revealed persistent deviations in error definition.

7.5 Feedback Interval and Timing

The use of terminal feedback after every trial potentially reduced effectiveness of the MMC prototype. All described prior work where researchers implemented EA provided concurrent feedback. None of the researchers worked with terminal feedback. The main reason that our prototype developed in this thesis chose to deviate from prior work is to capture and analyze full-body positional data. All prior work focused on a specific movement such as reaching [45][19], tracing [47][4][40], or trunk-arm sweep rowing, except for gait rehabilitation on a (split-belt) threadmill [25][41], and driving an electric wheelchair [14]. The MMC prototype bridges the gap and applies EAF on a complex motor task and contributes a low-cost and accessible proof of concept. However, there currently is no way to use 3D MMC in a concurrent setup, which is a major limitation. OpenPose can analyse video in real-time, but converting the 2D data into 3D data requires both camera inputs and is done through triangulation. The triangulation process in MatLab cannot work real-time. Additionally, providing mapped-EA would not be possible in a concurrent setup, which in turn would have limited studying the effects of EAF on the Vertical Error. All in all, this means that terminal feedback can be justified. However, there is one major limitation of MMC that impacted the study design.

The MMC-based prototype as developed for this thesis is not fast enough, impacting the study design and results, as well as reducing the system's flexibility. Once the prototype recorded a video, several steps needed to happen before the subject received visual feedback, which all together took up too much time. Both videos are processed by OpenPose frame-by-frame at a rate of about 20FPS. Triangulation, error extraction, and augmentation for a ten second video was completed within two seconds. The final timely step was to produce three animations in MatLab. A high-end desktop computer was sourced to speed up the process. The OpenPose pose estimation analysis is a GPU-powered process which benefits using a more powerful GPU. However, MatLab's animation generation process is a CPU-powered process and does not allow for multi-threading, which turned out to be the primary bottleneck. The visual feedback is generated for three angles one by one, and watched three times per angle. This way, the subject can already review the front view animation while the other animations are still being generated. While it is believed beneficial to allow a few seconds between movement completion and terminal feedback to allow a subject to self-estimate their performance [46], the feedback should be provided within reasonable time. To maintain an effective cadence throughout the experiment, the maximum number of squats a subject could do before receiving feedback was one. Even with doing one squat, which lasts no longer than ten seconds, the time to seeing the first feedback approached thirty seconds. The long processing time combined with the inability to speed up the process with a powerful computer severely limited the study design. With a faster setup, the study design would provide feedback after sets of three squats. With a faster setup, a subject could have done up to fifty squats in one day instead of twenty, resulting in a longer acquisition phase with more practise.

7.6 Skill Level and Individual Differences

The system lacks flexibility to cater to varying skill levels and individual differences, therefore potentially limiting any beneficial effects of EAF. Prior research revealed that Augmented Feedback should be tailored based on task complexity and individual skill levels to reach maximum effectiveness. More specifically, EAF should be designed with a target skill level in mind in order to select the correct gain. The developed prototype ignored individual differences such as age, prior experience with the task, physical condition, and body type/proportions. The aim to find significant differences between the test group and the placebo group in this case compares modified feedback compared to TF. In sports, mirrors and video recordings are the golden standard to TF and are widely used. If the goal is to reveal the degree of effectiveness of EAF compared to TF, it is essential to cater to individual differences in order to make the terminal EAF as effective as possible. Not doing so theoretically does not reveal the true potential of applying well-designed EAF. This underestimated aspect during the prototype development phase could have influenced the results. Next iterations should include a degree of adaptability within the system that accounts for individual differences and optimizes the EAF to achieve the best results.

8 CONCLUSION

MMC is a promising tool to generate terminal error augmented visual feedback on complex full-body sports movements such as BW Squats. The significant improvements in VE using our proof-of-concept MMC-based prototype shows that MMC can successfully be implemented at low costs and is accessible to researchers interested in capturing and analyzing full-body data in complex dynamic sports tasks. To successfully design EAF for a specific movement, selecting the right bio-mechanical factors to augment is essential and an ideal technique needs to be well defined in advance. Movement where the performance error is unclear until completion, such as a vertical jump or a parallel squat require terminal mapped-EA. Continuously amplifying performance error is categorized as 'over-time EA' and theoretically could be implemented as concurrent feedback but current triangulation and augmentation technology inhibits concurrent generation of MMC-generated 3D positional data. Error Augmented skeleton data represented as an animated stick figure is promising, but feedback should be tailored to individual skill level and task complexity by carefully highlighting the ideal technique, as well as selecting a gain that highlights performance error while not confusing the athlete. Researchers can choose to not disclose the use of EA but qualitative evidence shows that novice athletes get confused by unrealistic augmented feedback. Overall, MMC-based EAF is promising and should be seriously considered by researchers in the Sports ITECH field. Applications designed for specific movements have the potential for widespread adoption in sports training and should be seen as a lightweight and affordable tool to analyze the performance of athletes in the field.

8.1 Future Work

To ensure a fruitful future of affordable and flexible MMC-driven EAF implementation, future research should tackle current uncertainties. Comparison studies should provide insights in most effective gains for various task complexity levels, the most accurate MMC model, and the difference in accuracy between MMC and other MoCap systems. Studies on the effect of camera resolution, the use of deception, various clothing options, and faster feedback generation could potentially unlock reliable and robust MMC systems that also allow for concurrent 3D EAF generation. Further analysis of the current state of the art will reveal how to design effective tailored feedback based on task complexity and skill level, which future researchers should not underestimate. If one was to improve on the current thesis, focus should be on a better avatar, more robust interpolation, and faster feedback generation mechanisms. This could potentially provide more definitive results. With less uncertainties, larger research projects can eventually start to work on different complex sports movements and conduct motor learning studies with longer acquisition phases. Future research will eventually lead to a new way of generating EAF that is robust and accurate, and therefore applicable to many types of complex movements, which allows for worldwide adoption and improved sports training methods.

REFERENCES

[1] [n. d.]. 3-D locations of undistorted matching points in stereo images - MATLAB triangulate - MathWorks Benelux — nl.mathworks.com. https://nl.mathworks.com/help/vision/ref/triangulate.html. [Accessed 15-03-2024].

[2] [n. d.]. Camera Calibration Boards, Targets and Software by Calib.io — calib.io. www.calib.io. [Accessed 15-03-2024].

[3] Farnaz Abdollahi, Emily D Case Lazarro, Molly Listenberger, Robert V Kenyon, Mark Kovic, Ross A Bogey, Donald Hedeker, Borko D Jovanovic, and James L Patton. 2014. Error augmentation enhancing arm recovery in individuals with chronic stroke: a randomized crossover design. *Neurorehabilitation and neural repair* 28, 2 (2014), 120–128.

[4] Farnaz Abdollahi, Sylvester V Rozario, Robert V Kenyon, James L Patton, Emily Case, Mark Kovic, and Molly Listenberger. 2011. Arm control recovery enhanced by error augmentation. In *2011 IEEE International Conference on Rehabilitation Robotics*. IEEE, 1–6.

[5] Thomas R Baechle and Roger W Earle. 2008. *Essentials of strength training and conditioning*. Human kinetics.

[6] Linge Bai and David Breen. 2008. Calculating center of mass in an unbounded 2D environment. *Journal of Graphics Tools* 13, 4 (2008), 53–60.

[7] Ekin Basalp, Nicolas Gerig, Laura Marchal-Crespo, Roland Sigrist, Robert Riener, and Peter Wolf. 2016. Visual augmentation of spatiotemporal errors in a rowing task. In *Human Movement and Technology: book of abstracts-11th Joint Conference on Motor Control & Learning, Biomechanics & Training*. Shaker Verlag GmbH.
[8] Alexander J Bedard. 2017. Augmented and intrinsic feedback in motor learning of the bodyweight squat.
[9] Ahmad Syariff bin Ahmad Tajudin, Mohamad Hamka Nizam bin Mohamad Rosni, and Jeffrey Low Fook Lee. 2016. Effects of augmented feedback on squat technique among eleven years old children. *Jurnal Sains Sukan & Pendidikan Jasmani* 5, 1 (2016), 1–8.
[10] J Bloomfield. 1998. Posture and proportionality in sport. In *Training in sport: Applying sport science*. John Wiley & Sons, 145–186.
[11] Scott Bonnette, Christopher A DiCesare, Adam W Kiefer, Michael A Riley, Kim D Barber Foss, Staci Thomas, Katie Kitchen, Jed A Diekfuss, and Gregory D Myer. 2019. Injury risk factors integrated into self-guided real-time biofeedback improves high-risk biomechanics. *Journal of sport rehabilitation* 28, 8 (2019), 831–839.
[12] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7291–7299.
[13] B Cesqui, S Aliboni, Stefano Mazzoleni, Maria Chiara Carrozza, F Posteraro, and Silvestro Micera. 2008. On the use of divergent force fields in robot-mediated neurorehabilitation. In *2008 2nd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*. IEEE, 854–861.
[14] Xi Chen and Sunil K Agrawal. 2013. Assisting versus repelling force-feedback for learning of a line following task in a wheelchair. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21, 6 (2013), 959–968.
[15] Ki-Hun Cho, Mi-Ran Hong, and Won-Kyung Song. 2022. Upper-limb robot-assisted therapy based on visual error augmentation in virtual reality for Motor recovery and kinematics after chronic hemiparetic stroke: a feasibility study. In *Healthcare*, Vol. 10. MDPI, 1186.
[16] Sungdeuk Chun. 2022. Using Augmented Reality for real-time feedback to enhance the execution of the squat. (2022).
[17] Cherelle Connor, Prithvi Kantan, and Stefania Serafin. 2022. A real-time movement sonification application for bodyweight squat training. In *Psychoacoustics in the Loop: Proceedings of the 7th Interactive Sonification Workshop*. 15–21.
[18] Gray Cook. 2003. *Athletic body in balance*. Human kinetics.
[19] Ronit Givon-Mayo, Esther Simons, Avi Ohry, Hana Karpin, Sharon Israely, and Eli Carmeli. 2014. A preliminary investigation of error enhancement of the velocity component in stroke patients' reaching movements. *International Journal of Therapy and Rehabilitation* 21, 4 (2014), 160–168.
[20] Mark A Guadagnoli and Timothy D Lee. 2004. Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of motor behavior* 36, 2 (2004), 212–224.
[21] Lee Herrington and Allan Munro. 2014. A preliminary investigation to establish the criterion validity of a qualitative scoring system of limb alignment during single-leg squat and landing. *J Exerc Sports Orthop* 1, 2 (2014), 1–6.
[22] Herbert Heuer and Jenna Luettgen. 2015. Robot assistance of motor learning: A neuro-cognitive perspective. *Neuroscience & Biobehavioral Reviews* 56 (2015), 222–240.
[23] Sharon Israely and Eli Carmeli. 2016. Error augmentation as a possible technique for improving upper extremity motor performance after a stroke–a systematic review. *Topics in stroke rehabilitation* 23, 2 (2016), 116–125.
[24] Grega Jakus, Kristina Stojmenova, Sašo Tomažič, and Jaka Sodnik. 2017. A system for efficient motor learning using multimodal augmented feedback. *Multimedia Tools and Applications* 76 (2017), 20409–20421.
[25] Pei-Chun Kao, Shraddha Srivastava, Sunil K Agrawal, and John P Scholz. 2013. Effect of robotic performance-based error-augmentation versus error-reduction training on the gait of healthy individuals. *Gait & posture* 37, 1 (2013), 113–120.
[26] Jong-Wook Kim, Jin-Young Choi, Eun-Ju Ha, and Jae-Ho Choi. 2023. Human pose estimation using mediapipe pose and optimization method based on a humanoid model. *Applied sciences* 13, 4 (2023), 2700.
[27] K Kinakin. 2004. Optimal muscle testing. *Champaign, IL: Human Kinetics* 122 (2004).
[28] Takumi Kitamura, Hitoshi Teshima, Diego Thomas, and Hiroshi Kawasaki. 2022. Refining OpenPose with a new sports dataset for robust 2D pose estimation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 672–681.
[29] Duane Knudson and Craig Morrison. 1996. An integrated qualitative analysis of overarm throwing. *Journal of Physical Education, Recreation & Dance* 67, 6 (1996), 31–36.
[30] Matthew Kritz. 2012. Development, reliability and effectiveness of the Movement Competency Screen (MCS). (2012).
[31] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. 2014. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 740–755.
[32] Le Yu Liu, Youlin Li, and Anouk Lamontagne. 2018. The effects of error-augmentation versus error-reduction paradigms in robotic therapy to enhance upper extremity performance and recovery post-stroke: a systematic review. *Journal of neuroengineering and rehabilitation* 15 (2018), 1–25.
[33] Dylan P Losey, Laura H Blumenschein, and Marcia K O'Malley. 2016. Improving the retention of motor skills after reward-based reinforcement by incorporating haptic guidance and error augmentation. In *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 857–863.
[34] Laura Marchal-Crespo, Jorge Lopez-Oloriz, Lukas Jaeger, and Robert Riener. 2014. Optimizing learning of a locomotor task: amplifying errors as needed. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 5304–5307.
[35] Andrea Masiero and Angelo Cenedese. 2012. Reconstruction error in a motion capture system. *arXiv preprint arXiv:1203.3230* (2012).

[36] Marie-Hélène Milot, Laura Marchal-Crespo, Christopher S Green, Steven C Cramer, and David J Reinkensmeyer. 2010. Comparison of error-amplification and haptic-guidance training techniques for learning of a timing-based motor task by healthy individuals. *Experimental brain research* 201 (2010), 119–131.

[37] David Monaghan, Freddie Honohan, Amin Ahmadi, Troy McDaniel, Ramin Tadayon, Ajay Karpur, Kieran Morran, Noel E O'Connor, and Sethuraman Panchanathan. 2016. A multimodal gamified platform for real-time user feedback in sports performance. In *Proceedings of the 24th ACM international conference on Multimedia*. 708–710.

[38] Martin A O'Reilly, Darragh F Whelan, Tomas E Ward, Eamonn Delahunt, and Brian M Caulfield. 2017. Technology in strength and conditioning: assessing bodyweight squat technique with wearable sensors. *The Journal of Strength & Conditioning Research* 31, 8 (2017), 2303–2312.

[39] JL Patton, YJ Wei, P Bajaj, and RA Scheidt. 2013. Visuomotor Learning Enhanced by Augmenting Instantaneous Trajectory Error Feedback during. (2013).

[40] James L Patton, Mary Ellen Stoykov, Mark Kovic, and Ferdinando A Mussa-Ivaldi. 2006. Evaluation of robotic training forces that either enhance or reduce error in chronic hemiparetic stroke survivors. *Experimental brain research* 168 (2006), 368–383.

[41] Darcy S Reisman, Heather McLean, Jennifer Keller, Kelly A Danks, and Amy J Bastian. 2013. Repeated split-belt treadmill training improves poststroke step length asymmetry. *Neurorehabilitation and neural repair* 27, 5 (2013), 460–468.

[42] Sean Sanford, Mingxiao Liu, and Raviraj Nataraj. 2021. Concurrent continuous versus bandwidth visual feedback with varying body representation for the 2-legged squat exercise. *Journal of Sport Rehabilitation* 30, 5 (2021), 794–803.

[43] Brad J Schoenfeld, Ronald L Snarr, et al. 2021. *NSCA's essentials of personal training*. Human Kinetics.

[44] Dhara A Sharma, Mohamed Faisal Chevidikunnan, Fayaz Rahman Khan, and Riziq Allah Gaowgzeh. 2016. Effectiveness of knowledge of result and knowledge of performance in the learning of a skilled motor activity by healthy young adults. *Journal of physical therapy science* 28, 5 (2016), 1482–1486.

[45] Leia C Shum, Bulmaro A Valdés, Nicola J Hodges, and HF Machiel Van der Loos. 2019. Error augmentation in immersive virtual reality for bimanual upper-limb rehabilitation in individuals with and without hemiplegic cerebral palsy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28, 2 (2019), 541–549.

[46] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review* 20 (2013), 21–53.

[47] Yejun Wei, Preeti Bajaj, Robert Scheidt, and James Patton. 2005. Visual error augmentation for enhancing motor learning and rehabilitative relearning. In *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005.* IEEE, 505–510.

[48] Bodo Winter. 2013. Linear models and linear mixed effects models in R with linguistic applications. arXiv, 1308.5499. *Online Tutorial. arXiv* (2013).

[49] Gabriele Wulf, Markus Höß, and Wolfgang Prinz. 1998. Instructions for motor learning: Differential effects of internal versus external focus of attention. *Journal of motor behavior* 30, 2 (1998), 169–179.

9 APPENDIX A

```
Matlab Measurements
├── calibration sessions
│   ├── Session001
│   ├── Session002
│   └── Session003
├── [Date Folder] (a new folder is created for each day)
│   ├── [Participant Folder] (e.g., John Doe 0001)
│   │   ├── [Trial Folder - Calibration Session] (e.g., 0001 - Session001)
│   │   │   ├── webcam_1_video.avi (video of cam 1)
│   │   │   ├── webcam_2_video.avi (video of cam 2)
│   │   │   ├── openpose_output_1
│   │   │   │   ├── keypoint_data_frame_00001.json
│   │   │   │   ├── keypoint_data_frame_00002.json
│   │   │   │   ├── keypoint_data_frame_00003.json
│   │   │   │   └── ...
│   │   │   ├── openpose_output_2
│   │   │   │   ├── keypoint_data_frame_00001.json
│   │   │   │   ├── keypoint_data_frame_00002.json
│   │   │   │   ├── keypoint_data_frame_00003.json
│   │   │   │   └── ...
│   │   │   ├── stereoResultsInfo.txt
│   │   │   └── animations_xx
│   │   │       ├── animation_subplot1.mp4
│   │   │       ├── animation_subplot2.mp4
│   │   │       └── animation_subplot3.mp4
│   │   ├── [ Trial Folder - Calibration Session] (e.g., 0002 - Session001)
│   │   │   └── ...
│   │   └── [ Trial Folder - Calibration Session] (e.g., 0003 - Session001)
│   │       └── ...
│   ├── [Second Participant Folder] (e.g., Jane Doe 0002)
│   │   └── ...
│   └── Recordings (all test recording that are not linked to a participant go here)
├── SessionData.xlsx
└── participantLookup.xlsx
```

Fig. 21. Compact directory tree structure.

10 APPENDIX B

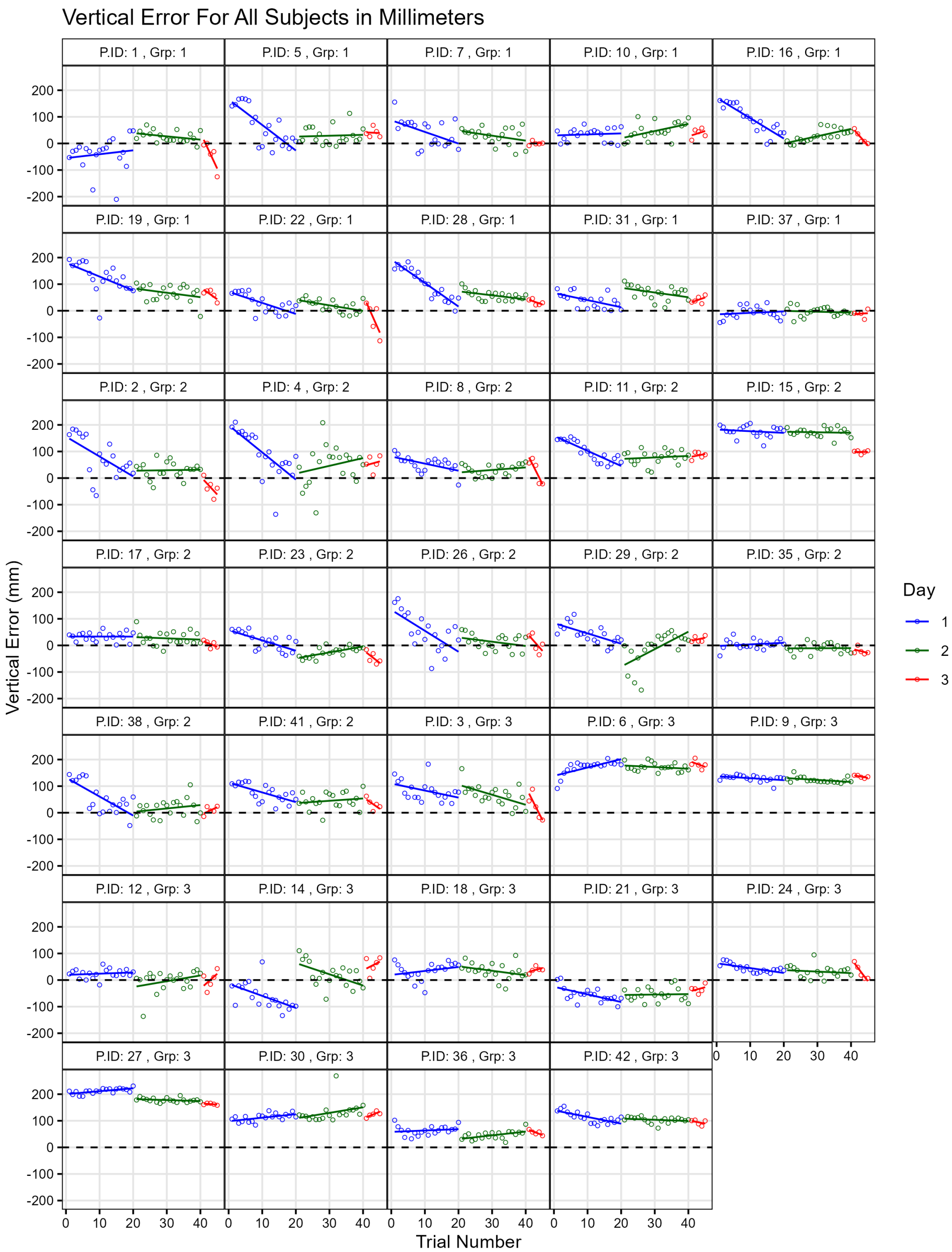


Fig. 22. Vertical Error For All Subjects in Millimeters

11 APPENDIX C

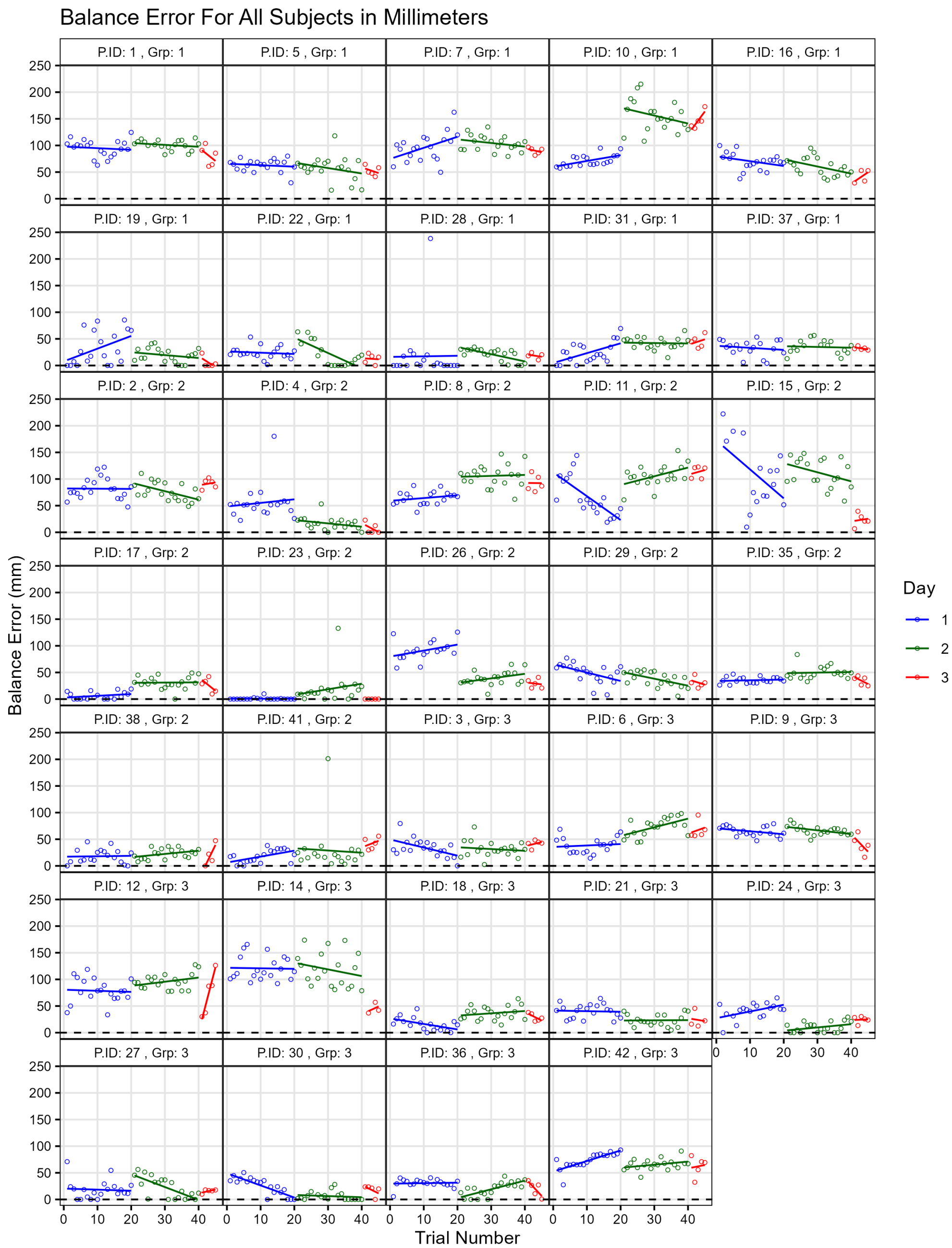


Fig. 23. Balance Error For All Subjects in Millimeters

12 APPENDIX D

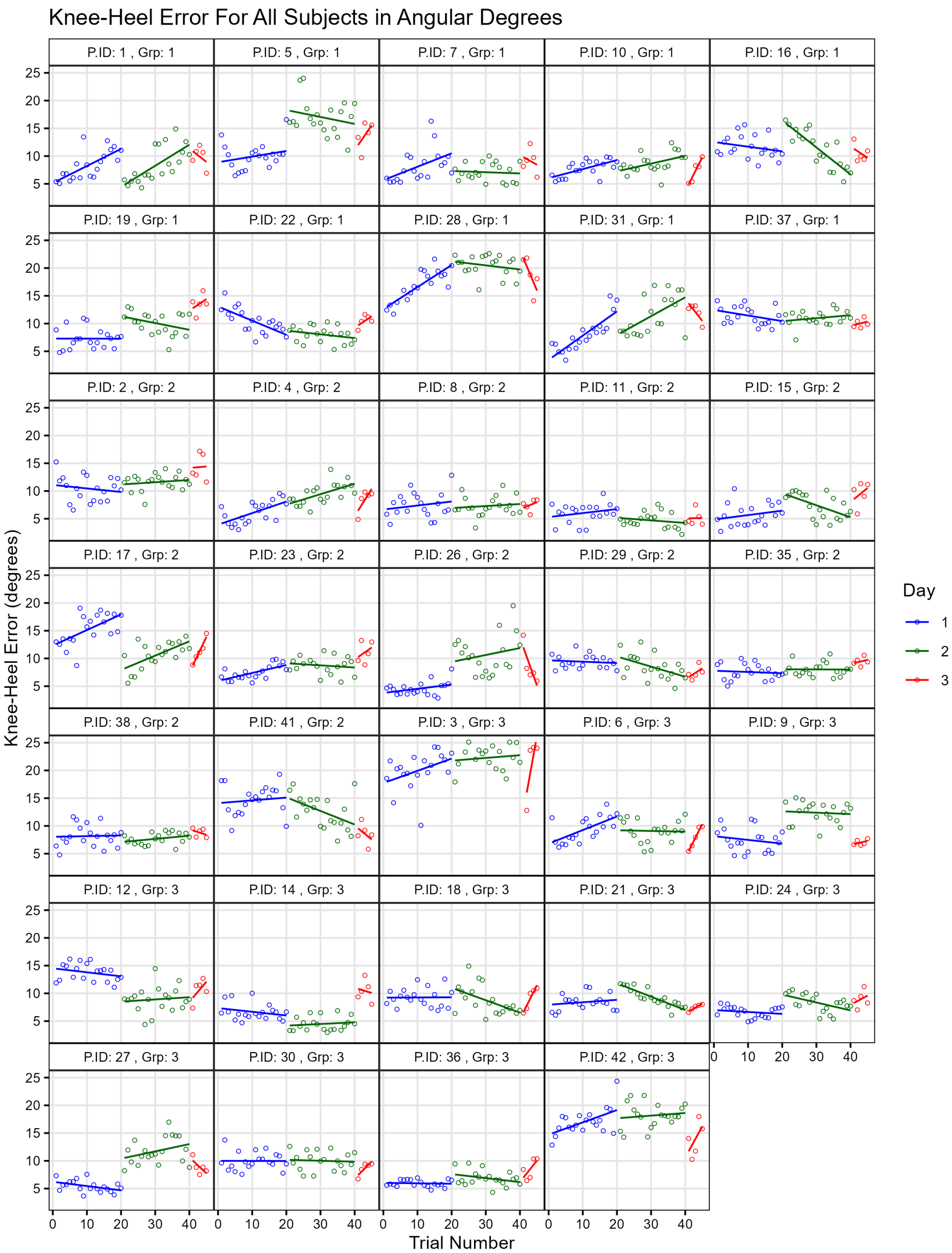


Fig. 24. Knee-Heel Error For All Subjects in Degrees

13 APPENDIX E: GUI OF THE PROTOTYPE

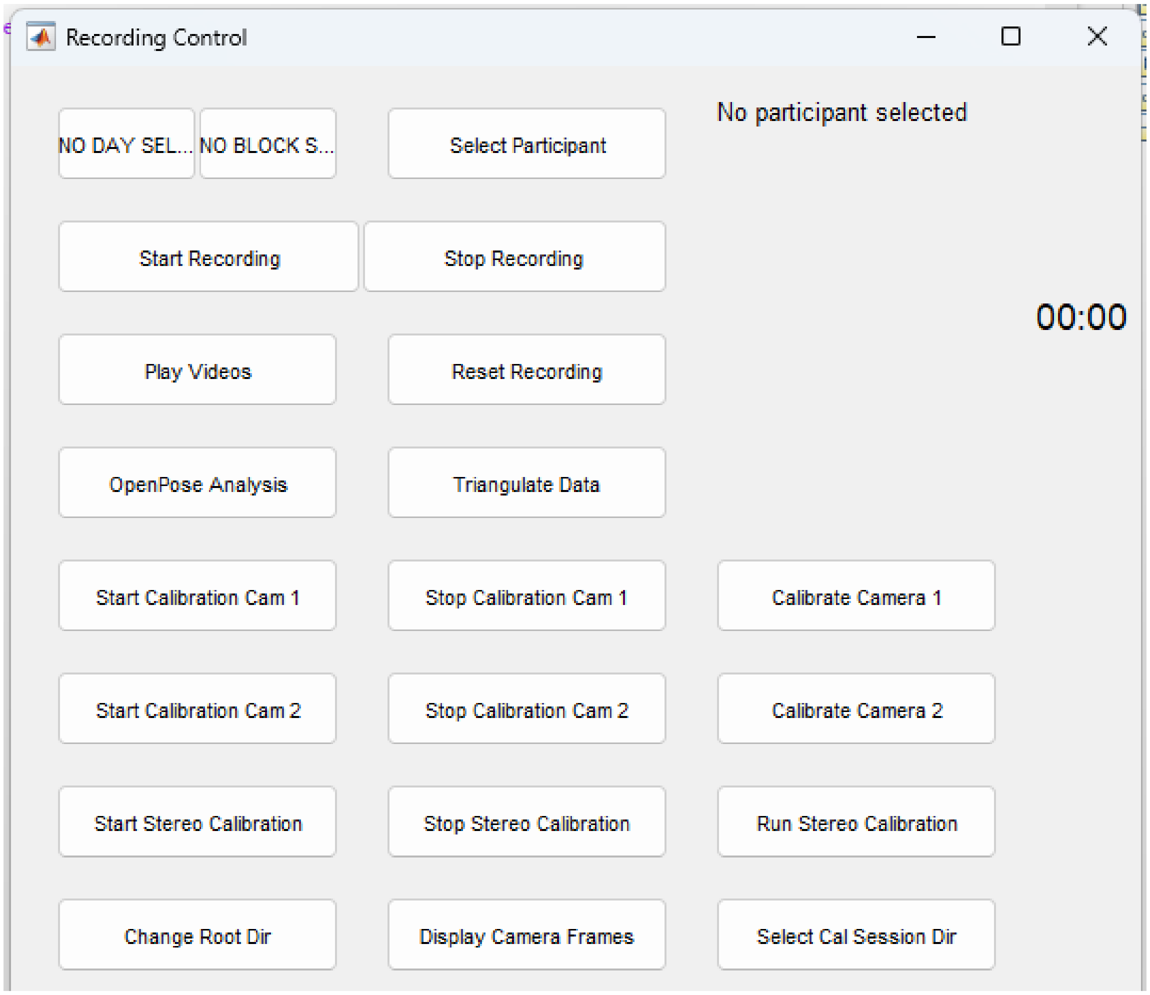
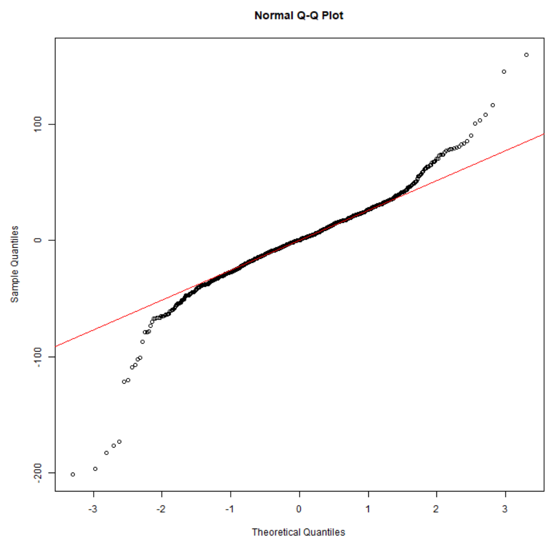
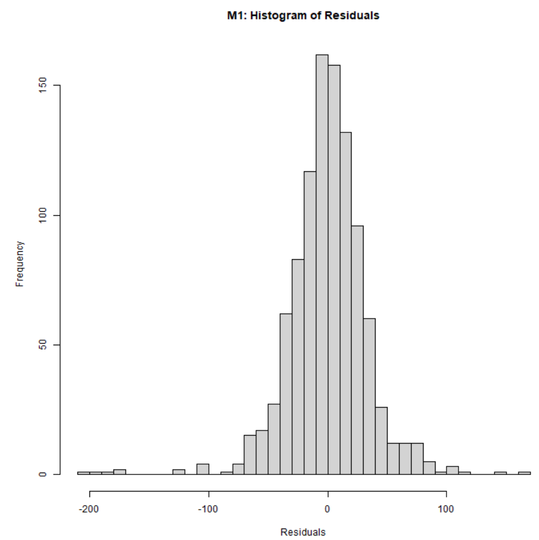


Fig. 25. GUI of the Prototype

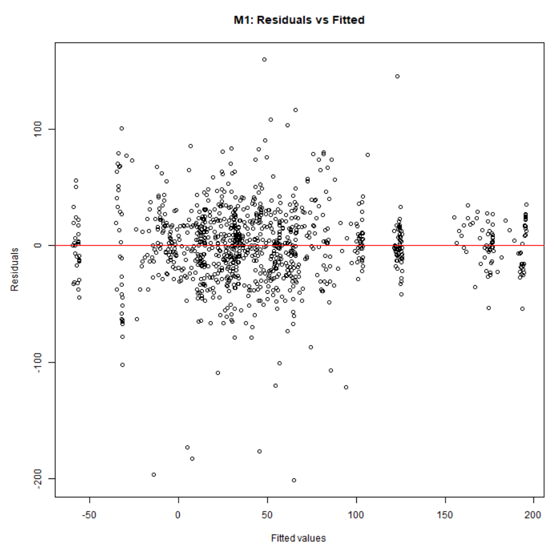
14 APPENDIX F: VE MODEL ASSUMPTION PLOTS



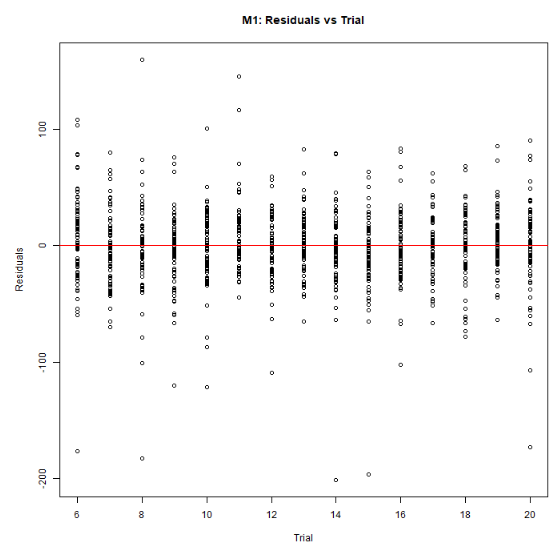
(a) Subcaption 1



(b) Subcaption 2



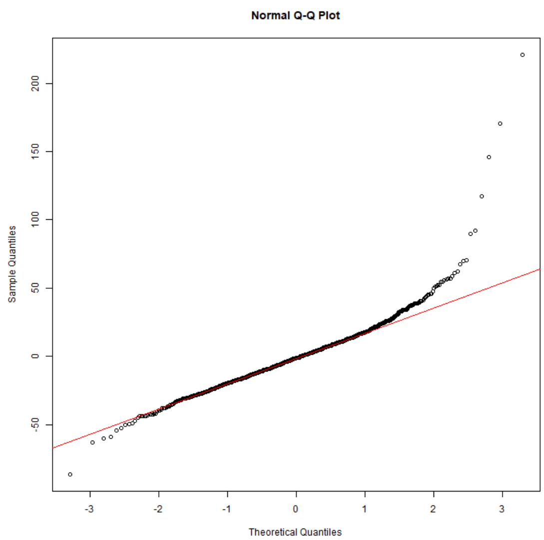
(c) Subcaption 3



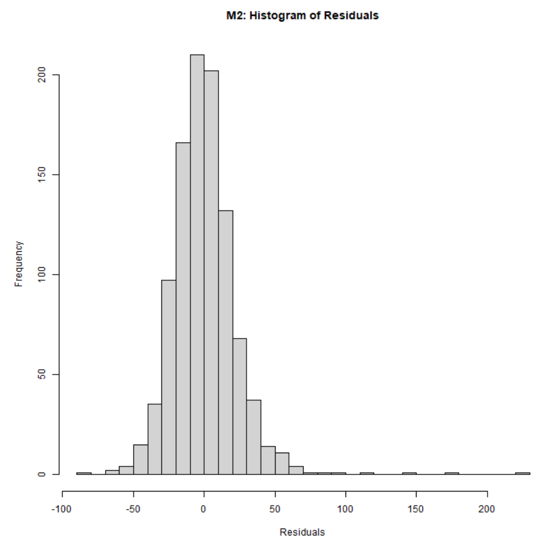
(d) Subcaption 4

Fig. 26. Assumptions Checking VE LMER

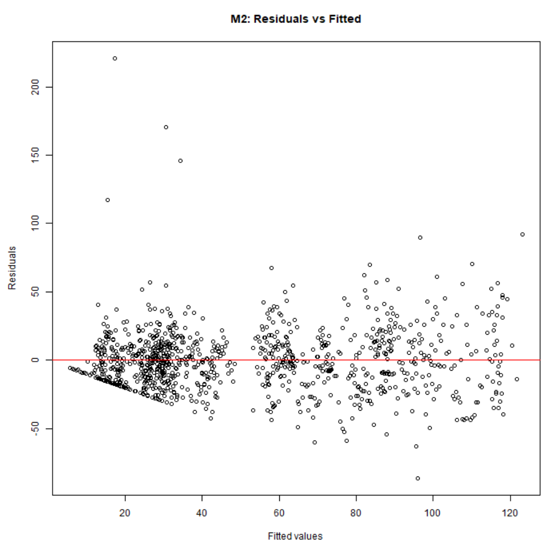
15 BE MODEL ASSUMPTION PLOTS



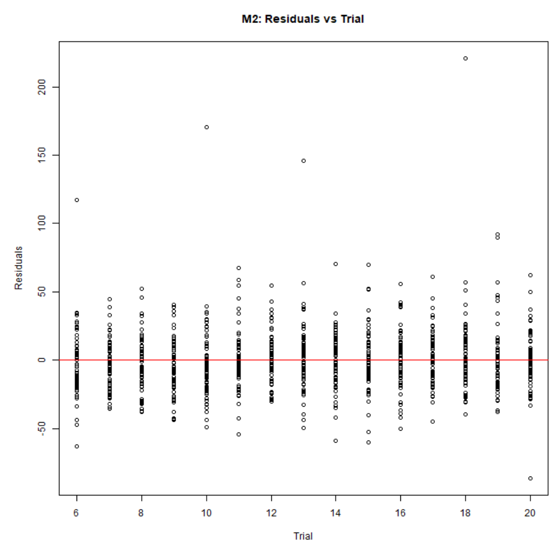
(a) Subcaption 1



(b) Subcaption 2



(c) Subcaption 3



(d) Subcaption 4

Fig. 27. Assumptions Checking BE LMER

16 APPENDIX G: QUESTIONNAIRE OVERVIEW

What is your name?
Text input field

What is your gender?
Male
Female
Non-binary
Prefer not to say

How old are you?
Text input field (The value must be a number)

How would you describe your usual level of physical activity?
Inactive
Lightly active
Moderately active
Very active

At the start of the experiment, how experienced were you with performing squats?
Beginner
Novice
Advanced
Expert

How often do you exercise per week?
0 days per week
1-2 days per week
3-4 days per week
5-6 days per week
Every day (7 days per week)

What were your initial impressions of the visual feedback that you received in between sets?
Text input field

What did you think of the time/duration between ending a squat and receiving the first feedback? How did you experience the overall flow of the experiments?
Text input field

How easy was it to understand and use the visual feedback?
Very easy
Somewhat easy
Neutral
Somewhat difficult
Very difficult
I did not receive feedback

How helpful was the visual feedback in improving your squat technique?
Not helpful
Slightly helpful
Helpful
Very helpful
I did not receive feedback

Do you prefer performing squats with or without visual feedback? Why?
Text input field

What improvements or changes would you suggest for the visual feedback system?
Text input field

Do you have any other comments or observations you would like to share about your experience?
Text input field

Are there other things you want to share that I haven't asked yet?
Text input field

17 APPENDIX H: QUESTIONNAIRE RESPONSES

Skill Level	Count	Percentage
Beginner	15	44.1%
Intermediate	15	44.1%
Advanced	4	11.8%
Expert	0	0%

Table 17. Distribution of Skill Levels Among Respondents

Activity Level	Count	Percentage
Inactive	2	5.9%
Lightly active	9	26.5%
Moderately active	20	58.8%
Very active	3	8.8%

Table 18. Distribution of Activity Levels Among Participants

Exercise Frequency	Count	Percentage
0 days per week	1	2.9%
1-2 days per week	16	47.1%
3-4 days per week	12	35.3%
5-6 days per week	5	14.7%

Table 19. Frequency of Exercise Days Per Week Among Participants

Feedback Category	Count	Percentage
Very easy	11	32.4%
Somewhat easy	9	26.5%
Neutral	2	5.9%
Somewhat difficult	0	0.0%
Very difficult	1	2.9%
I did not receive feedback	11	32.4%

Table 20. Feedback on Ease of Use

Feedback Category	Count	Percentage
Not helpful	1	2.9%
Slightly helpful	1	2.9%
Neutral	1	2.9%
Helpful	16	47.1%
Extremely helpful	4	11.8%
I did not receive feedback	11	32.4%

Table 21. Feedback on Helpfulness