CTU

**CZECH TECHNICAL
UNIVERSITY
IN PRAGUE**

F3

**Faculty of Electrical Engineering
Department of Computer Graphics and Interaction**

Master's Thesis

Interactive application using motion data for the analysis of an individual's musculoskeletal system

Master's Thesis

Bc. Barbora Koudelková
Open Informatics, Human-Computer Interaction

**May 2024
Supervisor: Ing. David Sedláček, Ph.D.**

MASTER'S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Koudelková Barbora**   Personal ID number: **483651**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Computer Graphics and Interaction**

Study program: **Open Informatics**

Specialisation: **Human-Computer Interaction**

II. Master's thesis details

Master's thesis title in English:

Interactive application using motion data for the analysis of an individual's musculoskeletal system

Master's thesis title in Czech:

Interaktivní aplikace využívající pohybová data pro analýzu pohybového aparátu jedince

Guidelines:

Create an interactive application that playfully guides a child user in performing selected exercises to analyze his/her musculoskeletal system.
1) Familiarize yourself with the methodologies of musculoskeletal analysis in pediatric individuals [1, 2] and possible methods of motion recording.
2) According to (1), select appropriate methods, propose their implementation using projection, and select an appropriate method for recording movement data.
3) Design the selected methods in a visually and gamifically engaging way and implement them. Draw inspiration from interactive applications in museums and science centers, e.g. [4]. Design a way to analyze the recorded data and implement the necessary support for the analysis (e.g., export to appropriate software, basic statistics).
4) Test the developed implementation with the target group (children) and consult with experts (e.g., physiotherapists) during the work.

Bibliography / sources:

1] Kolá Pavel ; Má ek Miloš. Základy klinické rehabilitace. 2021.
2] Henderson, S. E., Sugden, D., & Barnett, A. L. (1992). Movement assessment battery for children-2. Research in Developmental Disabilities.
3] Joseph J. LaViola, Jr. et all. 3D User Interfaces: Theory and Practice, second edition. 2017. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA.
4] Initi Playground, https://www.initi.org/

Name and workplace of master's thesis supervisor:

Ing. David Sedlá ek, Ph.D.   Department of Computer Graphics and Interaction FEE

Name and workplace of second master's thesis supervisor or consultant:

Date of master's thesis assignment: **16.02.2024**   Deadline for master's thesis submission: **24.05.2024**

Assignment valid until: **21.09.2025**

_____   _____   _____
Ing. David Sedlá ek, Ph.D.   Head of department's signature   prof. Mgr. Petr Páta, Ph.D.
Supervisor's signature   Dean's signature

III. Assignment receipt

The student acknowledges that the master's thesis is an individual work. The student must produce her thesis without the assistance of others, with the exception of provided consultations. Within the master's thesis, the author must state the names of consultants and include a list of references.

_____._____   _____
Date of assignment receipt   Student's signature

Acknowledgement / Declaration

I want to express my gratitude to Ing. David Sedláček, PhD., for his guidance, advice, but also for procuring the necessary equipment and testing facilities. His optimism not only motivated me to complete this project but also made this collaboration immensely valuable to me. Thank you for all of it.

As this work could not have been accomplished without external cooperation, I extend my thanks to everyone who assisted me in the development process:

Special thanks to Mgr. Iva Štolová (and Amálka), without whom this project might not have been possible. Further appreciation goes to Mgr. Radka Beranová, DiS., for her expert consultation.

I am also immensely grateful to Mr. Otto Naderer, MSc., for his willingness to dedicate time to this project on such short notice and for his crucial revisions that have amazed me.

My thanks also go to Ing. Josef Kortan, who inspired me to undertake this project and who literally saved my user testing.

Lastly, I would like to thank Kittens, my partner, and especially my dad, who has shown endless support to me and my studies. Thank you.

v

Abstrakt / Abstract

Tato diplomová práce zkoumá převedení současných metod pro zachytávání pohybu bez snímačů (markerů) a principů gamifikace v pediatrické fyzioterapii a ergoterapii do moderní doby počítačového věku. Cílem výzkumu je zjednodušit proces testování a vyhodnocování pohybu testovaných osob tak, aby se zvýšila jejich angažovanost a bylo možné začlenit tyto metody do standardizovaných lékařských postupů.

Vývoj zahrnoval testování různých metod snímání pohybu, přičemž byl kladen důraz na metody snímání bez markerů, využívající pokročilé technologie počítačového vidění, jakými je například MediaPipe. Tento přístup byl doplněn vývojem aplikace pro vizuální projekci, která implementuje část testu MABC-2, která slouží jako intervenční metoda využívaná v ergoterapii.

Podle metodiky User-centered design byla provedena dvě kola testování, která měla za cíl zdokonalit a iterativně vylepšit navrženou aplikaci.

Výsledky testování s uživateli naznačují, že systém výrazně zlepšuje interakci s pacienty a podstatně zvyšuje efektivitu terapeutického cvičení.

Klíčová slova: snímání pohybu bez markerů, analýza pohybu, snímání pohybu, snímání pohybu dětí, ergoterapie, fyzioterapie, okupační terapie, MABC-2, interaktivní projekce, HCI

This thesis explores integrating markerless motion capture methods and gamification principles in pediatric physical and occupational therapy. The research aims to simplify the process of testing and evaluating the movement of test subjects to increase engagement and incorporate these methods into standardized medical procedures.

The development involved testing various motion capture methods, emphasizing markerless capture methods using advanced computer vision technologies such as MediaPipe. This approach was complemented by developing a visual projection application that implements a portion of the MABC-2 test, which serves as an intervention method used in occupational therapy.

Following a user-centered design methodology, two rounds of testing were conducted to refine and iteratively improve the proposed application.

The results of testing with users indicate that the system significantly improves interaction with patients and substantially increases the effectiveness of therapeutic exercise.

Keywords: markerless motion tracking, motion analysis, motion capture, motion capture of children, ergotherapy, occupational therapy, physiotherapy, MABC-2, interactive projection mapping, HCI

Contents /

Tables / Figures

Chapter **1**
Introduction

Recently, it has become quite common to see the blending of the medical world with modern technologies. Whether it involves the integration of computer vision methods and usage of artificial intelligence during medical procedures [1], or the application of gamification principles in patient examination and better subsequent recovery [2–3], the intersection of medicine and technology is evident. This integration of both worlds introduces new approaches to engage and motivate patients through their journey, and can greatly speed up and facilitate diagnosis for doctors.

This thesis explores the potential applications of motion-tracking technologies and gamification principles in movement therapy methods, such as physiotherapy and occupational therapy.

1.1 Project specification

This project builds upon previous work, specifically integrating movement data for the methodology developed by prof. PaedDr. Pavel Kolář, Ph.D. The methodology of Prof. Kolář is focused on observing simple exercises, according to which the future problem of the individual's locomotor system can be recognized. The goal was to create software and hardware support for this methodology that would allow trained personnel to operate it in the shortest possible time.

Several technological approaches were proposed for consideration, with a strong emphasis on using markerless motion detection - the main reason behind favoring the markerless systems was that the final product in the future iterations was expected to be used *en masse*, meaning that the idea was to establish physiotherapy centers or child screening clinics in every city, implementing the discussed methodology. Considering the limited time available for examining each child (typically 5 to 15 minutes), the markerless method would become crucial for prioritizing cases. The alternative, involving the application of trackers and markers, would consume a comparable amount of time, in addition to necessitating staff training or the recruitment of additional technical experts.

Despite a lack of direct collaboration between the author and the domain expert, Professor Kolář, as will be referenced in later chapters, the thesis retains and sets itself the following tasks:

1. To understand common physiotherapy and occupational therapy methods and to find a methodology that is standardizable, meaning it can be measured in a way that goes beyond simple visual assessments by medical professionals.
2. To select a segment of the methodology for feasibility verification and subsequent implementation, with the time constraint that was given previously.
3. To design a system for analyzing chosen exercises in a visually engaging and gamified way. This system should use markerless methods and projections to make the examination process quicker and simple enough for staff without technical skills to use.

Chapter **2**
Method research

This chapter explores the common diagnostic methods in physiotherapy and ergotherapy, with a specific focus on their relevance in the context of pediatric patients. By examining these foundational approaches, we aim to understand their limitations and strengths and their possible applications of motion tracking on standardized assessments, if possible.

2.1 Physiotherapy - common methods

This section is dedicated to exploring methods widely used in sports physiotherapy and placing them in the context of Professor Kolář's work.

2.1.1 Vojtova therapy

Vojta's therapy [4], also known as Vojta's Method, was developed by Czech neurologist Professor Václav Vojta. Rooted in pediatric physiotherapy, this therapeutic approach addresses motor disorders and developmental challenges in infants and young children. The method is grounded in the idea that innate movement patterns are genetically encoded in the human central nervous system, drawing inspiration from principles of child developmental physiology.

Central to Vojta's Method is the therapeutic principle of reflexive locomotion. This approach involves the activation or elicitation of motor responses (coordination) that are inherent to every individual. These reflexive responses manifest from birth and persist throughout one's life. Remarkably, Vojta discovered that even in patients with damage to the central nervous system, these innate motor responses can be triggered. This underscores the method's adaptability and its ability to influence motor coordination beyond typical developmental stages.

2.1.2 Method of Ludmila Mojžíšová

The Mojžíšová method [5] is a reflexive approach that utilizes mobilization, soft tissue techniques, passive movements, massages, and thermal procedures. The treatment focuses on the spine, pelvis, and muscles throughout the entire body and addresses functional infertility in women.

Part of the method are standardized exercises, focusing on the active involvement of the pelvic floor muscles.

2.2   Kolář´ s methodology

Professor Kolář is a world-renowned physiotherapist who integrates the approach of Ludmila Mojžíšová (as previously described in Section 2.1.2) with the Vojta methodology (see Section 2.1.1). Notably, he was a student of both of these influential figures, drawing from their combined expertise. The method itself is called Dynamic Neuromuscular Stabilization (DNS).

It is a comprehensive method that incorporates several principles and techniques. The entire DNS method is rooted in developmental kinesiology, where the maturation of the central nervous system (CNS) in a child reflects its ability to perceive the environment and move within it. Any deviation from the brain's natural development eventually manifests in the motor skills of either the child or an adult [6]. The DNS theory focuses on two main areas: the Integrated Spinal Stabilizing System (ISSS) and the role of intra-abdominal pressure (IAP). The disbalances between those areas can lead to postural deformities, as seen in Figure 2.1.

Optimální posturální vzor | Abnormální posturální vzory

OPTIMÁLNÍ POSTURA: páteř napřímená, hrudník umístěný nad pánví, horizontální postavení bránice, pánev v neutrální poloze.

SYNDROM ROZEVŘENÝCH NŮŽEK: elevace hrudníku, šikmé postavení bránice, anteverze pánve.

PŘEDSUNUTÉ DRŽENÍ: osa hrudníku a pánve je paralelní, v důsledku nedostatečného napřímení na kyčlích je celý trup předsunutý vpřed.

PŘEDSUNUTÉ POSTAVENÍ HRUDNÍKU: Hrudník je elevován a postaven ventrálně oproti pánvi.

POSTAVENÍ S HRUDNÍKEM ZA PÁNVÍ: Hrudník je postaven dorzálně vůči pánvi, rigidní hrudní kyfóza.

Figure 2.1. Optimal postural alignment on the left, where the pelvis is properly aligned, characterized by a neutral pelvis position, horizontal shoulder line, and the ribcage positioned directly over the pelvis. The subsequent figures illustrate abnormal postural patterns: 'Syndrome of Open Scissors,' indicating an elevated ribcage and skewed pelvis; 'Forward Shift,' showing a parallel alignment of the ribcage and pelvis but shifted forward; 'Forward Tilted Pelvis,' with the ribcage elevated and shifted forward ventrally compared to the pelvis; and 'Pelvis Pushed Back,' depicting the ribcage positioned dorsally to a rigidly extended pelvis (image taken from [7]).

A key premise of the DNS [8] approach is that every joint position depends on stabilizing muscle function and coordination of both the local and distant muscles to ensure the neutral or centered position of joints in the kinetic chain. DNS presents a set of functional tests to assess the ISSS and to assist in finding the "key link" of the dysfunction.

Figure 2.2. Infant motor milestones aligned with DNS, to reinforce proper movement and posture (image taken from [7]).

These postural defects can be addressed by employing basic postural functions derived from toddlers' and babies' natural movements in patients, as seen in Figure 2.2. What is known is how to get rid of the problems by employing these exercises, but not how to uncover them or measure them in a short amount of time, which is one of the main ideas behind this thesis, as was mentioned in the specifications of this project (see Chapter 1). Sadly, a lack of professional knowledge and inability to reach out to domain experts (physiotherapists of Kolář) in detecting and measuring these deformities led to a search for more standardized methods in the field of ergotherapy, mentioned further in Chapter 4.

2.3 Ergotherapy

Ergotherapy [9], also called occupational therapy, is a type of treatment that, through meaningful engagement, aims to preserve and utilize an individual's abilities for managing everyday, occupational, recreational, and leisure activities. It applies to individuals of any age with various types of disabilities—physical, sensory, psychological, mental, or social. Our primary focus in pediatric occupational therapy is addressing dyspraxia (praxia stands for dexterity).

In childhood, dyspraxia [10] (also known as developmental coordination disorder or DCD) usually refers to a disorder in which children do not develop the motor skills that are expected for their age. While some children outgrow the condition, most continue to experience movement difficulty as adolescents and adults. Treatment can help them function better as they grow.

We categorize dyspraxia into three distinct types:

- **Ideatic** - Senzoric and perception skills are impaired.
- **Motoric** - Can be characterized by selective movement disorders, postural adaptation disorder, impaired relaxation, and balance disorder [11].
- **Ideomotoric** - Combination of the above-mentioned types.

4

This thesis deals with the possible investigation of motoric dyspraxia and the possibilities of quick detection, exploring the so-called interventional methods of ergotherapy. The standardized ergotherapy measurement methods used to categorize patients' dyspraxia are described below.

2.3.1 Purdue Pegboard Test

Purdue Pegboard Test (shortly PPT) is purposed to evaluate fine motor skills, fingertip dexterity, and overall gross movement and coordination of arm, hand, and fingers. The idea is to place pins, washers, and collars into a pegboard. A pegboard is a perforated hardboard with evenly spaced holes for placing the pegs, approximately 20 centimeters long, and a drawn black line in the middle from top to bottom (as seen in Figure 2.3).

Figure 2.3. Purdue Pegboard Assessment. (image taken from [12]).

The test is divided into three subtests lasting exactly thirty seconds per hand. During each subtest, the patient is encouraged to insert as many pins as possible into the pegboard holes, starting from the top and progressing to the bottom. The patient is allowed three trials for each task:

1. **Unilateral manipulation** -The patient uses only one hand to insert pegs, beginning with the right hand and then the left hand. The task should commence in this correct order unless the patient is left-handed [13].

2. **Bilateral manipulation** - The individual uses both hands simultaneously to insert pegs into specified holes in the pegboard.

3. **Bimanual interplay** - Assembly - In this task, the individual must assemble a specified number of pins, collars, and washers into a particular configuration using both hands.

   The expected subtest score (the number of expected correctly inserted pins) is determined by a table-defined number minus the coefficient dependent on the patient's gender, multiplied by age [13].

   The PPT takes approximately 5 to 10 minutes to administer and score. Its low purchase price, minimal training requirements for staff, and quick evaluation make it an effective method for assessing pediatric patients. However, due to its focus on fine motor skills and complexity in the context of movement capture and evaluation, it represents a very complex task that is currently unsuitable for further exploration in this thesis.

2.3.2 Box and Block Test

The Box and Block Test (BBT) [14] is conducted using a wooden test box divided into two compartments, one equipped with 150 wooden colored cubes with an edge length of 25 mm, the second remaining empty. The right and left compartments are partitioned, as seen in Figure 2.4.
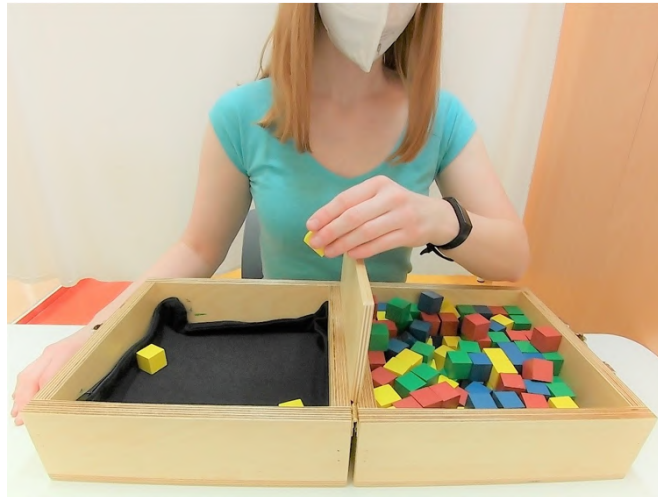
Figure 2.4. Transferring colored blocks from one compartment of a divided box to another, (image taken from [14]).

This test serves several purposes, including evaluating fine and gross motor skills in the upper limbs, assessing the speed and dexterity of the upper limbs, testing eye-hand coordination, and evaluating the ability to precisely follow instructions during an action. Specifically designed to measure unilateral gross manual dexterity, the BBT applies to children aged 6 to 12. It is particularly useful for this age group, as individuals younger than six generally have yet to have a distinguished laterality.

In each trial, the patient has a 60-second time limit to spill unilaterally as many dice/boxes as possible from the full compartment into the empty compartment. Each successfully transferred cube earns one point, and a higher overall score indicates better gross manual dexterity. The testing procedure consists of one trial attempt followed by three measured attempts for each limb, with the best result being considered for analysis.

This test examines the proficiency of both left and right limbs, analyzing how they perform in terms of division, speed, and skill. The goal is to determine if there is a sufficiently small difference in performance between the right and left hands.

2.3.3 MABC-2

Movement Assessment Battery for Children [15] is a standardized test employed to assess motor, cognitive, and ideomotor skills in children and to identify the degree and nature of motor difficulties or developmental disorders of movement coordination in children. This thesis will refer to the Movement Assessment Battery as MABC and its second version as MABC-2.

The MABC-2 is specifically designed for children aged 3 to 16 and is categorized into three distinct age groups: 3-6 years (AB1), 7-10 years (AB2), and 11-16 years (AB3). It is based on the behavioral concept of evaluating human motor functions and assumes that the level of functioning of the child's basic motor functions is reflected in the output performance and the observable way of performing sensorimotor tasks [16].

Each testing category (AB1, AB2, and AB3) contains eight exercises. Each item is designed to assess one of the following components of motor ability:
- The manual dexterity component (fine motor skills)
- The gross motor skills component
- The static and dynamic balance component

The testing process takes approximately 20-30 minutes, and completing the record sheet takes about 10 minutes. The patient must be given time beforehand to get familiar with the tasks. The assessment involves recording the best performance in each part of the test. The results are then compared to percentiles based on norms established for three age groups. Specific percentile thresholds are used to evaluate the record sheet. Consequently, the test and the record sheet results are expressed in percentiles, categorizing the child into one of three groups of motor abilities. Each group indicates the extent to which a particular child or adolescent may have motor skill difficulties or impairments.

The primary drawbacks of this method include the complete testing time - briefing the patient, preparing the necessary physical aids, performing the test, and completing the sheet, which amounts to approximately one hour per patient. The MABC-2 test is also freely available, but its purchase price is considerably high. The full package, inclusive of necessary objects, tables, and standards, is priced at around thirty thousand Czech Crowns on the local market[1].

2.3.4 ADL

ADL is a medical abbreviation for Activities of Daily Living and is broadly used in occupational therapy. The Functional Independence Measure (FIM) is an international assessment tool designed to evaluate a patient's independence and self-sufficiency in managing ADL activities.

The ADL activities are broken down into six categories, measured by FIM Tables [17]: Personal hygiene, Continence, Movement (transfers), Locomotion, Communication, and Social aspects, with each category having its subgroup of tasks. This assessment rates each FIM item using a scale ranging from 1 to 7, with 1 representing complete dependence and 7 indicating complete independence. The scale gauges the degree of assistance required for each activity and supplements this assessment with a percentage expression [18].

The main advantage of this standardized method is its excellent reliability and adequate validity [19–21]. The predictive ability of FIM items, including food intake, bathing, toileting, urinary control, dressing, and social interaction, has been established in previous studies [22, 19]

However, even though this method is commonly applied in pediatric cases, its utilization makes more sense in elderly patients who suffer from, e.g., dementia or after brain-related injuries. Also, awarding the points in each category relies on subjective observation of the patient by the occupational therapist, not the patient's objective performance, as other methods, e.g., 2.3.1, 2.3.2, or 2.3.3. Due to the challenge of achieving a comparable rating through an artificial solution, comparable to trained experts, this thesis will not further consider exploring this test.

---

[1] https://hogrefe.cz/mabc-2

Chapter 3

Movement capture methods

This chapter is dedicated to researching possible motion capture methods, their performance analysis in the context of movement assessments, and the possibilities of utilization for the thesis requirements.

3.1 Motion Capture Methods

Motion capture, also shortly called *mocap*, could be described as a process of capturing the motion of an existing object or person and transferring the captured data to its digital form, where it can be, e.g., subjected to subsequent analysis, altered, or used for various purposes.

3.1.1 Gait

Gait could be described as the manner of a person´s walking and is a standardized method for measuring normal and pathological walking patterns. It begins when the first foot contacts the ground and ends when the same foot contacts the ground again. The gait cycle is typically divided into two main phases: the stance and swing phases, as seen in Figure 3.1.
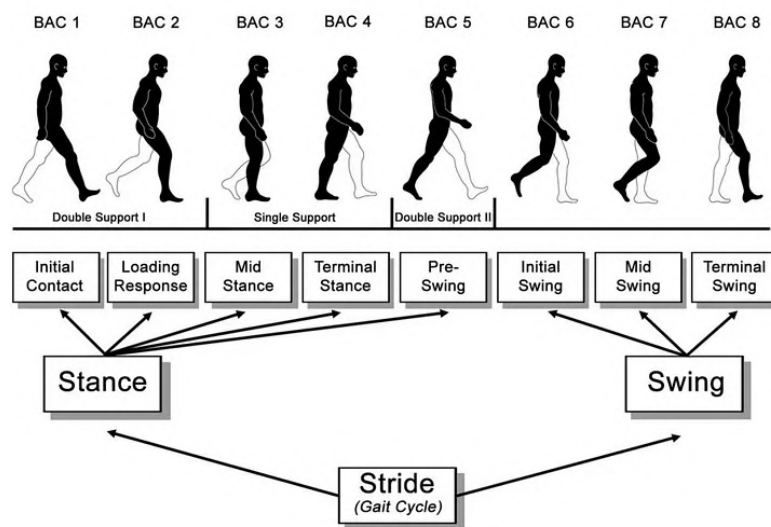
Figure 3.1. The image depicts the phases of the human gait cycle, which is divided into two main categories: Stance and Swing. The Stance phase includes 'Initial Contact,' 'Loading Response,' 'Mid Stance,' 'Terminal Stance,' and 'Pre-Swing.' The Swing phase consists of 'Initial Swing,' 'Mid Swing,' and 'Terminal Swing.' The cycle starts with one foot making contact with the ground (Initial Contact), progresses through the weight-bearing stance phase, and transitions into the swing phase where the foot is off the ground. The sequence ends with the same foot touching the ground again, completing one full gait cycle or stride (image taken from [23]).

8

Capturing the Gait cycle serves as an important reference point for comparing the performance of various motion capture systems. This method provides a valuable benchmark for assessing and contrasting the capabilities of different systems.

3.1.2 Planes of motion

From a clinical point of view, the measurements can be divided into three planes of motion [24], as depicted in Figure 3.2:
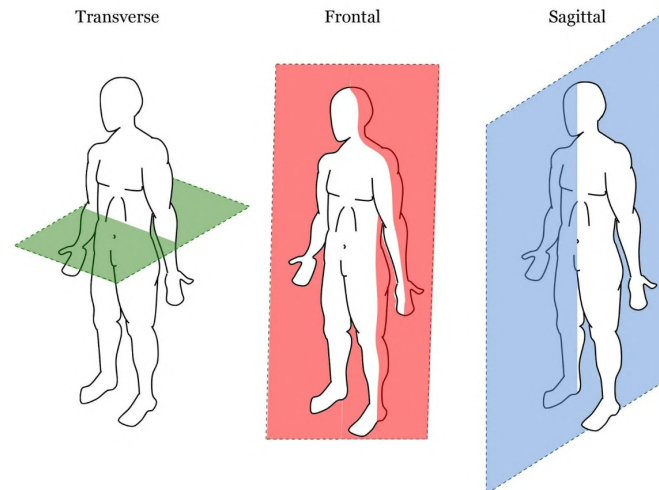
Figure 3.2. Three main planes of motion (image taken from [25]).

- **Sagittal plane of motion** - This plane divides the body into left and right halves. Movements that occur in the sagittal plane are typically forward and backward, such as flexion (bending) and extension (straightening).
- **Coronal plane of motion (Frontal)** - This plane divides the body into front and back halves. Movements that occur in the frontal plane are side-to-side movements, such as abduction (moving away from the midline) and adduction (moving toward the midline) of limbs.
- **Transverse plane of motion (Horizontal)** - This plane divides the body into upper and lower halves. Movements that occur in the transverse plane are rotational or twisting movements, such as internal and external rotation.

3.1.3 Position deviation

Despite motion capture being relatively accurate and used for many years as a good capturing method, from the clinical point of view the (marker-based) mocap yields an error load based on how the mocap principle works, in general. We assume the skeleton shape by placing markers alongside the patient, basically claiming that the bones start where the markers are. This consideration results in soft tissue artifacts (STA). Even though there exist various techniques that mitigate the STA problem [26–27], the problem itself is still not completely diminished.

Biplanar videoradiography (BVR), on the other hand, is a method that eliminates the STA problem. Biplanar videoradiography (biplanar fluoroscopy or biplane imaging), is a medical imaging technique that involves the use of two X-ray systems positioned at right angles to each other. This setup allows for simultaneous imaging from two different perspectives, typically in orthogonal planes, and the deviation of joints' position near zero. Its limitations lie in a narrow Field of View (FOV) only

where one or few joints can be monitored simultaneously [28], costly setup, and thus a limited range of real-world usability.

The STA artifact causes inaccuracy in precise joint positions during motion such as jump-cut movement. According to different studies focused on hip and ankle joint placement, mocap joint rotations and translations were observed to differ as much as 1,5° and 28mm [28], which is considerably high for precise, medically conducted experiments. Despite the observed differences, different studies performed on foot and ankle kinematics [29] suggest that mocap and BVR can yield similar results for standard gait analysis, especially in the sagittal plane (previously discussed in Section 3.1.2). It is safe to say that motion capture, in general, is a relatively accurate measurement tool, but it must be taken into account that sometimes the given joint position can be a rough approximation, not an exact data information.

3.2 Marker-based motion tracking

In this section, existing solutions utilizing markers are described. While there are, of course, numerous cameras and methods available, given the preferred focus of this thesis on markerless solutions, only the most relevant systems, often referred to as a *golden standard* in the field of motion capture and gait analysis [30], suitable for use in this work, are presented here. The main disadvantage of using marker-based solutions is their lengthy setup time as mentioned in Chapter 1, placing the markers correctly on the patient's joints and calibrating them in the camera software environment.

3.2.1 Vicon

Vicon [1] is a type of optical motion capture system, widely used in clinical trials. When Vicon cameras are set to their normal motion capture mode, they output the movement of Vicon markers, rather than the body to which the markers are attached. This is achieved with the combination of high-powered, narrow-band strobed illumination, retroreflective markers, and a corresponding filter that is tuned to the same wavelength as the strobe [31].

Vicon systems find extensive applications in other various fields, including biomechanics, sports science, and animation. Their technology plays a crucial role in enhancing the accuracy of motion capture data, facilitating detailed investigations into human and animal movement patterns for a wide range of scientific and industrial purposes.

3.2.2 Optitrack

OptiTrack [2] is a state-of-the-art optical motion capture system widely used in biomechanics and animation for precise and real-time tracking of objects and body movements. The Optitrack camera system works similarly to the Vicon (3.2.1) - employing a network of high-speed infrared cameras, OptiTrack captures the three-dimensional position of retro-reflective markers. This system facilitates detailed motion analysis by recording the trajectories of these markers.

The comparative study between Optitrack and Vicon motion analysis systems revealed close agreement in mean vector magnitudes, differing by only 1–3mm [32].

---

[1] `https://www.vicon.com/`
[2] `https://optitrack.com/`

Despite a higher number of trajectory gaps in Optitrack, the system demonstrates comparable accuracy and reliability to the more expensive Vicon system. This suggests that Optitrack could serve as a cost-effective solution for accurate 3D marker tracking in research and clinical applications.

3.2.3 Xsens

Movella Xsens is a Dutch technology company that specializes in 3D motion analysis, motion capture, and 3D motion tracking solutions [3]. One of Xsens' notable products is the MVN motion capture system, which consists of inertial motion sensors that can be worn on the body to capture and analyze movements in a 3D space without needing external cameras or markers. This makes their technology versatile and suitable for indoor and outdoor use, and also popular in physiotherapy and professional sports analysis.

The main disadvantage of using inertial sensors (in general) is generating drift from the space origin across the x, y, and z-axis. The Xsens solution appears to generate the smallest drifts among other inertial systems, according to studies - Xsens MotionGrid system's performance achieves a remarkably low 3D root-mean-square (RMSE) positioning error in challenging environments, including fast and dynamic movements, measuring deviation from origin approximately 0.065m and 0.10m per scenario, as has been proven by [33].

3.3 Markerless methods

Markerless motion capture methods do not rely on markers as reference points of joints necessary for obtaining motion data. Markerless mocap can be divided into appearance-based and model-based approaches [34].

The appearance-based approaches learn image features such as silhouettes and map them directly into a multidimensional space of poses. These approaches need training data closely resembling the motions to be analyzed and used, for example, deep neural networks for training (e.g., MediaPipe, discussed in Section 3.3.1). Model-based approaches, on the other hand, use primitive representations of the human anatomy with an underlying kinematic chain that is fitted to silhouette from each camera view to estimate the poses [35]. The markerless method can combine both approaches (for example, Azure Kinect, see Section 3.3.2).

The biggest advantage of using markerless systems is their relatively small cost compared to the purchase price for robust marker-based systems and their necessary hardware. However, in most cases, their affordable outcome comes with a trade-off, where, in contrast to measuring using marker-based systems, the inaccuracy in measuring motion and estimating poses can be high.

3.3.1 Mediapipe

Mediapipe[4] is an existing framework created and open-sourced by Google. It offers various methods of using a pre-trained network to build machine learning pipelines and process time-series data, e.g., estimating full body poses, hands, and facial tracking. The input to MediaPipe is typically a stream of RGB images, which could be framed from a video captured in real time from a camera or an existing video file.

---

[3] https://www.movella.com/products/xsens
[4] https://developers.google.com/mediapipe

11

Once the image is prepped, it is passed to a Convolutional neural network model for pose estimation.

Following pose estimation, the neural network outputs the positions of landmarks in either 2D or 3D coordinates, as illustrated in Figure 3.3. The data fed into Mediapipe are in 2D; therefore, the model employs a *scaled orthographic projection*, or weak perspective method, maintaining a constant average depth (z-coordinate) with the hip landmark at the origin (0,0,0). This hip landmark acts as a reference for the depth of other landmarks, which are normalized to the x and y coordinates through weak projection [36].

0. nose	17. left_pinky
1. left_eye_inner	18. right_pinky
2. left_eye	19. left_index
3. left_eye_outer	20. right_index
4. right_eye_inner	21. left_thumb
5. right_eye	22. right_thumb
6. right_eye_outer	23. left_hip
7. left_ear	24. right_hip
8. right_ear	25. left_knee
9. mouth_left	26. right_knee
10. mouth_right	27. left_ankle
11. left_shoulder	28. right_ankle
12. right_shoulder	29. left_heel
13. left_elbow	30. right_heel
14. right_elbow	31. left_foot_index
15. left_wrist	32. right_foot_index
16. right_wrist	

Figure 3.3. Pose landmarks in Mediapipe Pose Landmark Model. The model tracks 33 joints in Pose Landmark detection mode but can also track both hand and face in the Holistic Model, with a total of 543 landmarks (33 pose landmarks, 468 face landmarks, and 21 hand landmarks per hand [37]) (image taken from [38]).

Mediapipe has a high potential to be the next reliable tool for evaluating and identifying gait patterns. It shows almost accurate knee flexion/extension angles, with low RMSE [39] and average ankle trajectories [40]. However, compared to the performance of Vicon(3.2.1) and Kinovea posthoc analysis(3.4.2) in measuring hip movement, flexion and extension, and ankle angle extension, the results of MediaPose estimation can be described as highly problematic, mainly because the algorithm for estimation often mislabels joints [41].

3.3.2 Microsoft Kinect

Microsoft Kinect is a line of motion-sensing input devices developed by Microsoft for use with their Xbox gaming consoles and Windows PCs [42]. While Kinect was initially designed for gaming, its capabilities found applications beyond the gaming industry. Developers and researchers explored its potential in healthcare, education, and art.

It's worth noting that Microsoft discontinued the production of the Xbox Kinect with the Kinect v2 in 2017 and focused more on integrating the Kinect sensors into the cloud environment with Microsoft Azure services. The Azure Kinect is a spatial computing developer kit with computer vision and speech models and Software Development Kits (SDKs) that can be connected to Azure AI Services. It is the successor to the Microsoft Kinect line of sensors [43]. The Azure Kinect contains one depth camera, an inertial measurement unit, a 360-degree microphone array, and one RGB camera with up to Ultra-high-definition resolution (4K).

For a long time, Kinect aid measurement was considered unreliable for scientific purposes [44]. Since the release of v2 Kinect, studies found that the tracking demonstrated medium to high accuracy in measuring spatiotemporal gait parameters in the sagittal plane of motion, excluding the rotational parameters [45]. However, the RMSEs of rotation-related movements (hip rotations, ankle rotations) remained considerably high compared to Vicon [46]. Kinects are still less efficient for surveillance applications due to their small field of view and the low-quality outdoor performance of the sensor, especially in direct sunlight [47].

3.3.3 Leap Motion Controller

Ultraleap, formerly and widely known as Leap Motion, is a touchless sensor device using infrared light to track hand and finger motions without physical contact.

Its core components include two built-in cameras and infrared LEDs. The Leap Motion Service then processes streamed image data obtained from LEDs. It compensates for background and lighting, reconstructing a 3D representation. The tracking layer reconstructs a 3D representation and infers object positions with filtering techniques. The outcomes are communicated as frames through a transport protocol [48].

This technology could possibly be applied in exploring fine motoric skills of a child in ergotherapy-related assessments, for example, finger dexterity in the Purdue Pegboard Test (discussed in Section 2.3.1), or parts of MABC-2 assessment focused on manipulation of objects (discussed in Section 2.3.3)

3.3.4 Lidar

LiDAR [49], which stands for Light Detection and Ranging, is a remote sensing technology that uses laser light to measure distances and generate precise, three-dimensional information about the shape and characteristics of objects or surfaces in its line of sight.

LiDAR systems typically consist of a laser scanner, a GPS receiver, and an inertial navigation system. The LiDAR system emits laser pulses toward a target area, typically in rapid, short bursts of laser light. Pulses then interact with objects or surfaces in the target area. The laser light is reflected back to the LiDAR sensor, and the sensor detects the reflected laser pulses. The time the laser pulses travel to the target and back is measured precisely. Using the known speed of light, the LiDAR system calculates the distance to the object or surface based on the time it took for the laser pulses to make the round trip [50].

By repeating this process across different locations or from various angles, LiDAR generates a dense set of 3D coordinates, often referred to as a Point Cloud. This Point Cloud represents the spatial distribution of objects and surfaces in the scanned area.

Due to its point cloud nature, it is not expected to detect humane skeletal systems on the go or typically detect them in the form of obstacles (often used in automotive self-driving cars). Therefore, special computer vision algorithms for pose detection must be applied. Its clinical validity in gait analysis has been tested with results comparable to Optitrack(see Section 3.2.2) [51]. However, it still represents a very high-cost, high-effort solution to detect movement anomalies in children. However, we can leverage the 2D detection capability, which is suitable for identifying a child and the position of their legs in space during the measurement of spatial exercises,

13

such as the MABC-2 (as previously discussed in Section 2.3.3), and its capability is being used through this thesis, see the Chapter 7.

3.3.5 iKeleton

The iKeleton is an iPhone-designed application available on AppStore[5]. It sends tracked data in the form of Open Sound Control (OSC) messages, a common protocol in the creative industry, facilitating communication with visual programming frameworks like Touch Designer and Max For Live.

Given the extremely low purchase price (4.99 USD), it represents an extremely interesting solution. Notably, Google Play does not have an equivalent for sending tracked data through standardized protocols. The possible application of accurate motion tracking in the medical field has not yet been examined at all.

3.4 Market Overview and Solution Evaluation

This section explores the available solutions on the market that are suitable for addressing the needs in Chapter 1 and combines both the software solution and analytics together with hardware solvers or are widely used in sports and physiotherapy. However, it is important to note that the pricing information or the functionality principle on most of the products is not accessible to the public.

3.4.1 Eumotus

Eumotus BodyWatch[6] represents an analytics and markerless computer vision solution rooted in TensorFlow and MediaPipe (see Section 3.3.1). Its primary application lies in analyzing the positions of individual joints during the execution of conventional sports maneuvers such as squats, cleans, deadlifts, and other Olympic skills. Following the analysis, Eumotus transforms complex data into simplified diagrams and graphs. The platform promises a simplified learning curve for mastering these movements and finds extensive use in physiotherapeutic and sports clinics in the United States.

However, it's essential to note that this solution, while valuable in its domain, doesn't align with the central theme of this thesis, as it does not contribute to the development of the main concept.

3.4.2 Kinovea

Kinovea[7] is a free and open-source software designed for video analysis. It is widely used by sports professionals, physiotherapists, and biomechanics researchers. The solution does not detect poses in real-time but rather offers a posthoc analysis of the skeleton, trajectory, joint angles, and overall performance, as seen in 3.4. This solution would be roughly sufficient for the ways they currently analyze posture morbidities using DNS, as referred to in Section 4.1.1 and 2.2.

---

[5] https://apps.apple.com/cz/app/ikeleton-osc/id6448196217?platform=iphone

[6] https://www.eumotus.com/

[7] https://www.kinovea.org/

Figure 3.4. Overview of Kinovea user interface. (image taken from [52]).

3.4.3 Theia3D / Contemplas

Theia3D, an enterprise based in Canada, specializes in markerless motion capture software leveraging deep learning algorithms to estimate 3D human pose from synchronized 2D camera views[8]. Deep learning algorithms then operate frame-by-frame, identifying and tracking over 100 landmarks on visible subjects in the 2D videos.

The triangulation of landmark detections from multiple 2D videos into 3D positions, using camera calibration parameters, results in a scaled and optimized 3D skeletal model with 17 rigid segments [53]. This model accurately represents individual poses, requiring timecode-synchronized cameras, with frame rates tailored to the movement dynamics.

The efficacy of Theia3D's markerless motion capture solution is substantiated by measurements in angle differences on the sagittal plane, as referenced in [54], and comparable results in gait kinematic patterns, as indicated in [55]. With numerous usages in clinical trials, the Theia markerless solution states that it is the best widely-used markerless motion capture.

This technology has been adopted by the German company Contemplas, which integrates Theia3D algorithms and cameras into one of its solutions, TEMPLO[9]. Contemplas specializes in the analysis of movement, including gait and athletic performance in sports such as swimming, athletics, and running. The company provides a precise analysis of the kinematics of athletes' joints.

Within TEMPLO, users can define report templates and add parameters or entire pages, allowing for the creation of customized diagrams that suit specific user requirements. This adaptability ensures a comprehensive and User-centric approach to motion analysis, accommodating the diverse needs of users in different contexts. This solution can be applied from start to end to satisfy the needs of Prof. Kolář.

3.4.4 MoveAhead

MoveAhead is an Irish company specializing in developing motion-tracking engines specifically designed for pediatric patients. According to their website, their soft-

---

[8] https://www.theiamarkerless.ca/
[9] https://contemplas.com/en/motion-analysis/3d-analysis/

15

ware focuses on two key aspects: Children's movement data and data analysis and children's movement learning science. The company adopts a privacy-by-design approach, ensuring full compliance with GDPR standards for children's data protection.

While the specific workings of this solution are not detailed, MoveAhead provides a straightforward deployment of their SDK into platforms such as Unity and similar engines[10]. MoveAhead stands out as the sole market player with its narrow standardization exclusively tailored for children. Still, it's important to note that MoveAhead has not been utilized in any scientifically validated clinical trials, and its solution is not freely accessible for testing purposes.

---

[10] `https://moveahead.io/`

Chapter 4
Method selection

This chapter focuses on the technical analysis, procedure, and fitting of suitable potential methods and workflows in collaboration with experts on the given topic, as referenced in Chapter 1.

4.1 Consultations with domain experts

This section outlines the collaboration with physiotherapy and ergotherapy experts, which helped shape the research methods and assessments discussed in this thesis.

4.1.1 Physiotherapy

During the preparation of this thesis, a collaboration between the author and the commissioner (client), Professor Kolář, or his graduate students, was expected. It was anticipated that these would be simple, standardized coordination and motor exercises, adapted for a short examination time and the age of the participants. At the same time, no clear age limit was set for the child patients. Despite repeated urging, throughout the entire period of the creation and specification of activities necessary for identifying appropriate method proposals, the estimation was based solely on professional resources and books, as previously discussed in Section 2.2.

Actual contact was established only at the end of December with a Master's student from the Second Faculty of Medicine at Charles University in Prague, specializing in Pavel Kolář's DNS method, Bc. Klára Hudíková. The consultation revealed that there is no standardized procedure in the DNS measuring methods for detecting movement abnormalities at the moment; it all depends on the subjective judgment of the physiotherapist. The main difficulty is that if one focuses on a specific abnormality in the altered postural chain, it can sometimes represent only a compensatory mechanism, and it is necessary to perceive the overall manifestation.

However, measuring two axes in the sagittal plane between the diaphragm (chest) and pelvis could be considered a standardized assessment; for a better idea of how such a method would be measured, see Figure 2.1.

This assessment could be easily performed with a simple integration using MediaPipe. For a simple prototype, see Section 6.1. For more accurate results, a Convolutional network could be pre-trained. However, this requires a lot of annotated videos. Given the nature of the variables observed, using the sports annotation tool Kinovea (as already described in Section 3.4.2) is suggested.

In collaboration with the department head, it had been anticipated that at least a rough manual would be compiled by February or March 2024 on evaluating various DNS methods. However, this timeline for the so-called *cookbook* posed a significant challenge for the subsequent assembly of the thesis. Additionally, in March 2024, it became evident that Ms. Hudíková, due to serious reasons, would no longer continue in the project. As a result, no standardization occurred, and the potential for communication with the Second Faculty of Medicine was consequently diminished.

Although Prof. Kolář did reach out at the end of March, his communication came at a time when other methods had already been selected for this work.

4.1.2 Ergotheraphy

Due to the lack of data in the field of physiotherapy and the inability to find a physiotherapist specializing in DNS, through an acquaintance, Mgr. Ivana Štolová, a professional assistant at the Department of Occupational Therapy, Faculty of Health Studies, Jan Evangelista Purkyně University in Ústí nad Labem, was contacted in early November 2023. Mgr. Štolová provided insight into this field and helped select and explain standardized methods for pediatric patients (with references to occupational therapy and individual methods found in Section 2.3).

After expert consultations and clarifying potential technical difficulties associated with the translation of individual assessments, the MABC-2 method was chosen, for which she provided original materials and equipment, see further Section 4.2.

4.2 MABC-2: Analysis of possible implementation

As was mentioned beforehand in Section 4.1.2, the MABC-2 (2.3.3) was chosen, due to its versatility and clear metrics for scoring the patient's performance and also fast completion time per task.

It was decided that for the analysis, the most basic form of tasks, designated under the numerical category AB1 and designed for preschool-aged children, would be used. The tasks are divided into three classes:
- **MD** - Manual Dexterity Circle
- **AC** - Aim and Catch Circle
- **BAL** - Balance Circle

Below listed is an analysis for every assessment class:

4.2.1 MD1 - Coin Insertion

The task involves the child inserting six coins into a slot within a time limit (as seen in Figure 4.1), testing precision and coordination. The performance is observed and evaluated for efficiency of movement, technique, and preferred hand.
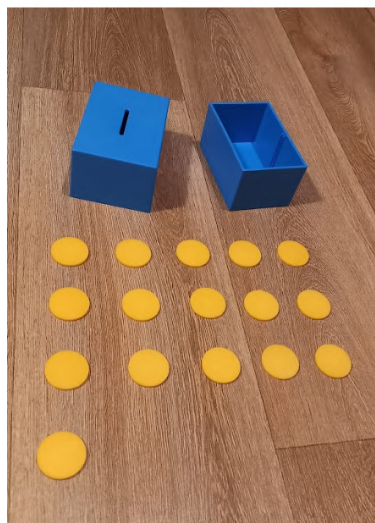
Figure 4.1. Coin box for MD1.

The child must:

1. Hold the coin box for the whole time.
2. Insert coins one at a time.
3. Use only one hand per trial

A Mediapipe's hand pose detection could be employed. However, it is important to keep in mind that there are still necessary physical objects, such as a box and coins, and that Mediapipe needs to detect a coin successfully dropped into the box.

4.2.2   MD2 - Stringing Beads

The task involves the child stringing 6 beads on a shoelace. The time between the start and all 6 beads string to a shoelace is measured.

The task largely involves qualitative observations such as holding the string too close or too far from the face, holding the string too close to the tip, and so on. This task is unsuitable for implementation; the only feasible creation would be virtual stopwatches that measure the start (the task begins by placing both hands of the child on the mat) and the end (the same conclusion) and then compare the time with the age-related reference norm.

4.2.3   MD3 - Drawing of a Path

In the MABC-2 assessment involving drawing a path, children are asked to follow a pre-determined route on paper with a pencil (as seen in Figure 4.2), testing their visuomotor control.
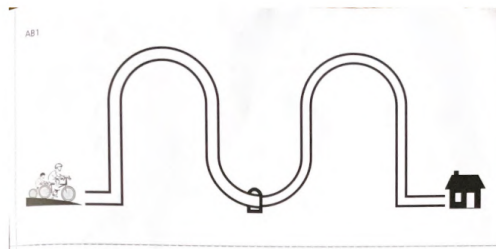
Figure 4.2. Official paper route for MD3.

To objectively measure performance, computer vision techniques can be applied to quantify the deviation between the child's tracing and the indicated path, but it is worth noting, that developing such a system would require careful machine learning models, calibration, and testing.

This can be reached by image preprocessing and path detection (using, for example, the Canny edge detector algorithm [56]. This method detects high-intensity changes in an image by tracing edges connected through gradients and reduces noise with steps that include Gaussian filtering and hysteresis thresholding.), aligning the reference path with the child's drawing aligning the start and end points, and then calculating the deviation. The problem imposes the analysis and scoring. Again, the problem with qualitative observations (*Putting too much pressure on the pen, etc..*) arises.

4.2.4   AC1 - Beanbag Catching

The task involves the child catching the beanbag thrown at a distance 1.8 meters in front of him. The child must:

1. Stand in a designated place for the whole time, moving outside the circle is forbidden.
2. Catch the bag using only hands, not other parts of the body.
3. Catch the beanbag at least once in ten attempts.

19

The task of catching a beanbag could be feasible for implementation. Since it requires two (objectively quite heavy) non-slip mats, a certain distance between them, and two circles. it can easily be simulated using projection. As for the catching aspect - this presents a more complex problem, where a combination of hand gesture detection and tracking of the beanbag object could be utilized. Also, it is worth noting that the child should follow the beanbag with its eyes for the whole time of executing this task. This could potentially be employed by using eye-tracking abilities[1] included in Mediapipe solutions and lenses with a short focal length.

4.2.5 AC2 - Throwing a Beanbag onto the Mat

The task involves the child throwing a beanbag onto a mat placed 1.8 meters away. The child must:

1. Stand within a designated area throughout the task, with stepping outside the marked zone being prohibited.
2. Throw the beanbag using only one hand, with an underhand grip.
3. Successfully land the beanbag on the mat in at least one out of ten attempts.
4. The child may, during the trials, change the hands used.

The task requires again two non-slip mats set at a specific distance apart, a requirement that can be effectively imitated using projection technology, as referenced in 4.2.4. Since it is known there exist solutions to detecting objects thrown on a mat (see 10) using a 2-D LiDAR, this task is easily implementable.

4.2.6 BAL1 - Standing on One Leg

The task measures the child's ability to maintain balance on one leg, a key indicator of motor control and stability. The attempt is considered complete when the child loses balance or the time limit is reached, whichever comes first.

The child must:

1. Stand on one leg, raising the non-supporting leg for up to 30 seconds. The second leg should be held in a specific position without touching the supporting leg.
2. Keep balance without swaying or hopping. If the balance is lost, the attempt is stopped, and the duration is recorded.
3. Perform the task twice for each leg, allowing a maximum of two attempts per leg within the 30-second time frame.

The assessment involves careful observation of the child's posture and balance. The examiner should record the duration of each successful attempt and note any deviations from the expected posture. A maximum of two trials are given for each leg to achieve the 30-second balance goal.

Since the child must stand in one designated place, the mat can be replaced by projection mapping again. An attempt to demonstrate the functionality via the Mediapipe has been made, as shown in the prototype further discussed in Section 6.2.

---

[1] `https://github.com/google/mediapipe/blob/master/docs/solutions/iris.md`

4.2.7   BAL2 - Walking on Tiptoes

This assessment examines the child's balance and coordination while walking on tiptoes, a skill that requires both fine motor control and stability.

The child is required to:

1. Walk forward on tiptoes, starting with the toe on the edge of a designated line and proceeding to walk along it.
2. Keep both hands lifted and maintain balance without swaying or stepping off the line.

The task is to be done without any help, so that the heels do not touch the ground.

To measure performance using markerless methods, we could employ a LiDAR and a projector to project a line on the floor that the child must follow, using LiDAR to measure the elevation of the feet above the ground The measurement of the feet elevation itself could be tricky. Using a MediaPipe for this task is not recommended since measuring the ankle rotations and elevations in study [41] due to frequent ankle joint mislabeling and inaccurate measurement.

Each correct performance is logged, and the number of accurately completed steps is counted. Errors are acknowledged by the child verbally (by saying *oops* for example) or with a hand raise, and the task is considered complete upon reaching the end line or after fifteen steps, whichever comes first. The attempt should be repeated if the child achieves less than the required number of tiptoe steps.

4.2.8   BAL3 - Jumping on Mats

This exercise assesses the child's ability to jump accurately and with coordination across a series of six mats placed alongside each other.

The child is instructed to:

1. Jump with both feet onto the first mat and then proceed by jumping onto the subsequent mats, which are arranged in sequence, ending with a target mat marked with a circle.
2. Perform the jumps without slipping or stepping off the mats.
3. Successfully land with both feet on each mat, avoiding any part of the feet landing outside the mat edges.
4. Complete a set number of jumps in sequence, typically five, which aligns with the age-appropriate expectations.

For markerless measurement, as mentioned in 4.2.6 4.2.7, a projection mapping and markerless measuring hardware could be employed.

4.3 Possible issues to take into consideration

This section discusses challenges, such as accurately capturing qualitative observations and maintaining exact measurements in automated systems and the variability in children's performance on tests like the MABC-2.

4.3.1 Qualitative Observations

The MABC-2 manual can be divided into two parts: the manual itself and the evaluation manual. As has been discovered, the manual includes not only a quantitative section but also a subcategory for qualitative observation, which the therapist checks off, as can be seen in Image 4.3.

Figure 4.3. Evaluation Manual of MABC-2. The form assesses single-leg balance with two main sections: one for timing the balance duration on each leg and another for checking qualitative aspects of posture and control.

Some qualitative observations could be captured using clever heuristics of computer vision; however, columns such as *Body appears tense/stiff,* etc., are extremely difficult to evaluate. This problem is most evident, especially in MD tasks, like *Holds the string too far from the tip* - detecting such deviations is almost impossible.

The evaluation also contains a subcategory Psychological Factors, where the therapist checks off boxes like *Impulsive, Anxious, Overestimates their abilities, Does not show pleasure from success....* Again, the same problem arises.

4.3.2 Exact measures

Mgr. Štolová has frequently mentioned that even though the solution would try to simulate the physical MABC-2 assessments, it is vital to keep the exact dimensions and sizes of objects, namely in assessments such as AC1, MD1, or mats for BAL3. It might be due to the effect of Fitts's law (Fitts's Law predicts the time required to rapidly move to a target area, such as dragging coins to a slot or stepping on a mat. It suggests that the smaller and further away the target, the longer the movement will take.) Therefore, it is crucial to keep the dimensions of projected objects the same as in the toolkit for

clinical validity. During the loan period of the toolkit, precise measurements of all components were documented and photographed.

4.3.3 Working with children

Mgr. Štolová has frequently mentioned that the MABC-2 score does not always capture the extent of motor difficulties. Despite the test's ease of use and interpretation, its effectiveness for diagnosing developmental coordination disorder has been questioned due to certain underlying problems. A common concern is that the MABC-2 score may not accurately represent the motor challenges observed and reported by individuals, as the test's non-holistic nature might fail to consider daily fluctuations in a child's physical and mental condition.

A study conducted by Kathryn J. Hadwin on the strengths and weaknesses of the MABC2 test is summarized very nicely with the following quote:

"I feel that a child's performance is dependent on many factors on the day. For example, their attention and concentration, emotional wellbeing, fatigue, and routine change as they will have often missed school…" [57].

4.4 Summary

Since the MABC-2 is both standardized *and* allows for ongoing consultations with experts during the development of the application, this method was selected for further implementation (refer to Chapter 7). Due to the challenges associated with implementing the Manual Dexterity Circle assessments, which require highly precise hand-tracking models (for example, exercise MD1, discussed in Section 4.2.1), it was decided to focus solely on the Aim-Catch Circle (hereafter referred to as AC) and Balance Circle (hereafter referred to as BAL). These will be implemented using projection technology to create a visually engaging way to administer these tests.

Chapter 5
Application Design and Requirements

This section is dedicated to defining the user group and their functional requirements and selecting appropriate resources for the completion of the project. This includes software, hardware, and visual considerations necessary to meet the project's objectives.

5.1 Characteristics of the Children User Group

The primary user group for this study consists of children aged 4-7. This age range is critical for identifying and addressing developmental concerns such as dyspraxia [58]. While the Movement Assessment Battery for Children is often used as an intervention method, children participating in the tests may vary in motor abilities. Some may exhibit visuomotor impairments, affecting their coordination and spatial understanding, while others might experience mild hindrances that do not signify a specific motor disorder.

Children of this age group are developing rapidly in physical, cognitive, and social domains. They are honing essential motor skills, both gross (like hopping and running) and fine (such as drawing and manipulating objects). Cognitively, they are beginning to grasp more complex instructions. They can interact with technology, but it is necessary to mention that their attention spans tend to be brief, requiring tasks to be engaging and potentially necessitating frequent guidance, as stated by Mgr.Štolová.

5.2 Requirements

The functional and non-functional requirements outlined below are derived from the thesis assignment's content and subsequent research conducted in the previous chapter on method selection (refer to Chapter 4).

5.2.1 Functional requirements

- **FR1 - Standardized Testing Protocol:** The application implements the MABC-2 method, focusing on standardized test administration to ensure consistency and reliability across uses (refer to Chapter 7).
- **FR2 - Aim-Catch and Balance Circle Assessments:** The application exclusively supports Aim-Catch Circle (AC) and Balance Circle (BAL) assessments, avoiding the more complex Manual Dexterity Circle due to its precision requirements (as discussed in Section 4.4).
- **FR3 - Projection Technology Integration:** Utilizes projection technology to visually display the test areas for AC and BAL, enhancing the engagement and clarity of the testing process.
- **FR4 - Dimensional Accuracy of Projected Objects:** Ensures that the dimensions of projected objects in assessments are maintained accurately per the physical MABC-2 toolkit, to adhere to ensure clinical validity (as referred in Section 4.3.2).

- **FR5 - Motion Data Management and Export:** The application will record, manage, and store motion data during user testing in a format that can be easily accessed and used for future analysis and replication of the tests.

■ 5.2.2 Non-functional requirements

- **NFR1 - Immediate feedback and rewards:** Immediate feedback and rewards should be provided to motivate continued play,
- **NFR2 - Expert Consultation Support:** The app design supports ongoing consultations with domain experts throughout the development phase, ensuring the implementation is aligned with current best practices and expert advice.
- **NFR3 - Visual Engagement:** The interface should be visually appealing and engaging to maintain the attention of the user group, using bright colors and clear, large visuals.
- **NFR4 - Measurement Accuracy:** Implements precise measurement capabilities to ensure that the distances and dimensions within the application reflect those of the physical objects used in traditional MABC-2 assessments.

■ 5.3 User-Centric Interface Design for Children

This section describes the User-centric design approach used for developing an interactive application tailored to children. It highlights using simple, engaging visuals and designs to ensure the app is accessible and enjoyable for young users [59].

■ 5.3.1 Interaction Design

In designing the interface for this application, it's crucial to acknowledge that the primary users may not yet be able to read and often do not understand abstract concepts such as geometric shapes like squares and rectangles, which are frequently used in the test. Based on the best practices outlined by Inclusive School Principles [60], this application uses familiar objects and scenarios in explanations to make interactions more intuitive and relatable for children.

Children should also be encouraged to explore and fully view the visual setup before beginning any tasks. This exploratory phase is designed to familiarize them with the interactive environment, reducing any potential anxiety and increasing their comfort level with the activities they will engage in.

Figure 5.1. This image illustrates an interaction design developed in consultation with Mgr. Štolová for the balance assessment. The design employs a visual metaphor - a well progressively filling with water - which is easily understandable for children who cannot yet comprehend time concepts. This feature, adapted from an exercise on the Balance Tutor[1] (a perturbation treadmill designed for postural control training, similar to what is measured in BAL1) engages patients by having them maintain balance while observing the filling well, simplifying the concept of duration without the need for clocks.

■ 5.3.2 Visual aesthetics

The visual aesthetics of the application are crucial for communicating effectively with young users. Following a suggestion from her supervisor, for which the author is very grateful, the author decided to design the interface in a fairy-tale, cartoon-like style, yet in a well-known environment that is appealing to children. This approach helps to create a friendly and engaging atmosphere.

The application features five distinctive visual environments, each brightly colored and hand-painted to attract and hold the children's attention. To ensure a uniform visual tone across these scenes and drawing from her previous experience, the author proposed and decided to employ color-grading and unified shaders. Color grading in this context involves adjusting the hues and saturation, strongly emphasizing the bloom effect to maintain a consistent mood and style. The bloom effect in image postprocessing creates a soft glow around bright areas in an image, magnifying the overall luminosity and thus giving a dreamy or ethereal touch to the scene.

---

[1] `https://www.btl.cz/produkty-pokrocile-rehabilitacni-systemy-balance-tutor`
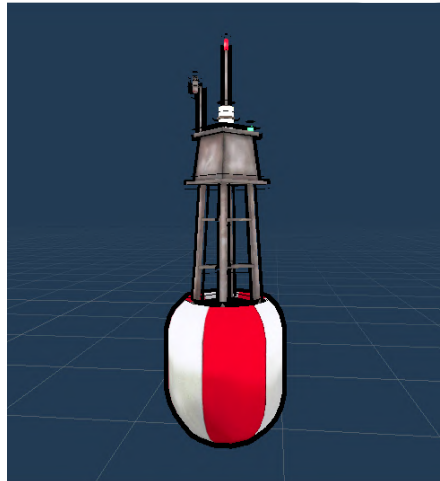
Figure 5.2. An example of an outline shader applied on a buoy model from BAL3 (taken from Asset Store [2]), that makes its edges stand out clearly.

For a cartoon-like appearance, all objects in the application are given bold outlines, similar to those seen in children's cartoons (as seen in Figure 5.2). This effect was achieved using a Simple Toon Shader[3], which simplifies shading and adds these outlines (shaders are programming scripts used in computer graphics to define the appearance of surfaces and lighting, thereby playing an important role in enhancing the dynamic effects of the scenes [61]). Additionally, a hand-drawn aesthetic was chosen for the visuals; every asset was either hand-painted by the author or carefully selected from a Unity Asset Store[4] to ensure it fits the overall design.

5.3.3 Feedback Mechanisms

During testing, as per MABC-2 guidelines (discussed in Section 4.2), the child receives minimal verbal support, mainly basic nods and sounds for confirmation or disagreement during the testing phase. Therefore, the application should notify the child when a trial has ended and use specific sounds to indicate if a task was performed incorrectly. Importantly, these sounds are designed to be gentle to avoid causing stress to the child.

---

[2] `https://assetstore.unity.com/packages/3d/props/exterior/simple-buoy-4734`

[3] `https://assetstore.unity.com/packages/vfx/shaders/simple-toon-185038`

[4] `https://assetstore.unity.com/`

5.4 Hardware selection

During the development phase, there was an opportunity to borrow several depth sensors for testing. These tests evaluated their suitability for the desired implementation mentioned in Section 4.2. The sensors included Azure Kinect (see Section 3.3.2), Intel RealSense D415[5], RealSense D435i[6] (the RealSense 400 series features depth cameras for applications such as robotics and augmented reality, using stereo vision to capture 3D data), and RealSense LiDAR(discussed in Section 3.3.4), as seen in Figure 5.3.

Figure 5.3. Available depth cameras.

Below is a brief description of each sensor, including its operating principle and testing setup, as seen in Table 5.1. Efforts were made to standardize the testing parameters across all devices as much as possible, for example, capturing at 30 Frames per Second (FPS) for every sensor.

List of Sensors	Depth-sensing Technology	IMU	Sensor Resolution
Azure Kinect (NFOV unbinned)	IR ToF	yes	1024 x 1024
Intel RealSense D415	IR Stereo	no	1280 x 720
Intel RealSense D4335i	IR Stereo	yes	1280 x 720
Intel RealSense L515	LiDAR	yes	1024 x 768

Table 5.1. Tested depth sensors specifications. IR stands for infrared light, and ToF stands for time-of-flight technology, where the sensor measures distances by timing how long it takes for an IR light signal to travel to an object and back to the sensor. The IMU stands for inertial measurement unit, where the sensor can measure acceleration, rotation, and velocity,

The performance was evaluated on a laptop with the following technical specifications, meeting the minimum hardware requirements for each sensor:

- PC: MSI GP63 Leopard 8RE
- CPU: Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz
- GPU: GeForce® GTX 1060 with 6GB GDDR5, 1,280 CUDA Cores
- OS: Windows 10 Home
- RAM: 32 GB DDR4-2400

---

[5] `https://www.intelrealsense.com/depth-camera-d415/`
[6] `https://www.intelrealsense.com/depth-camera-d435`

■ 5.4.1 Comparison test of depth sensing quality

The test was conducted in a room with four exercise mats placed at one-meter intervals, extending to a total distance of 4.5 meters from the sensor's position, the longest distance required for the MABC-2 assessment (as seen in Figure 5.4).

Figure 5.4. Test setup for sensor measurements.

A human subject assumed a T-pose on each mat, serving as a dynamic reference model. The positioning against a window allowed for the introduction of direct natural light, testing each sensor's ability to cope with potential backlighting and high-contrast scenarios.
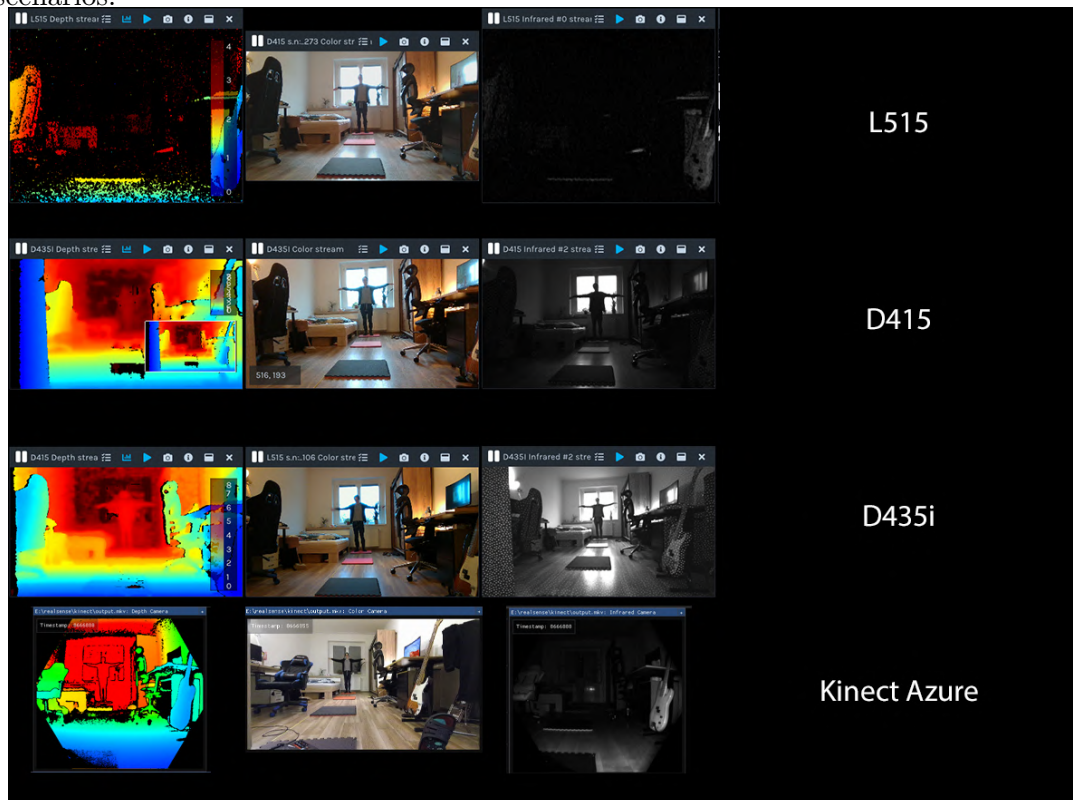
Figure 5.5. Comparison of output from sensors.

29

■ **Intel RealSense D415**

The D415 captured detailed depth and color data across the range, with reasonable accuracy and detail preservation, even at the furthest mat. The sensor's structured light technology provided reliable data in the indoor setting with ample lighting [62].

■ **Intel RealSense D435i**

With a wider field of view, the D435i offered better spatial coverage and captured the T-pose with a broader perspective. The depth data indicated some susceptibility to edge distortion, a common trait in wide-angle lenses, but maintained a good level of detail overall [62].

■ **Intel RealSense L515**

The L515, despite being a LiDAR-based sensor, underperformed significantly in this setup. At a distance of 4.5 meters, the depth data was sparse and noisy, possibly due to the sensor's range limitations or difficulty in handling the high-contrast lighting from direct sunlight. Another option might be, that the setup for the sensor was not done correctly [62].

■ **Kinect Azure**

The Kinect Azure demonstrated great performance, producing a robust and accurate 3D model of the subject and the environment. It managed the lighting conditions well, suggesting advanced capabilities in handling variations in illumination and contrast [62].
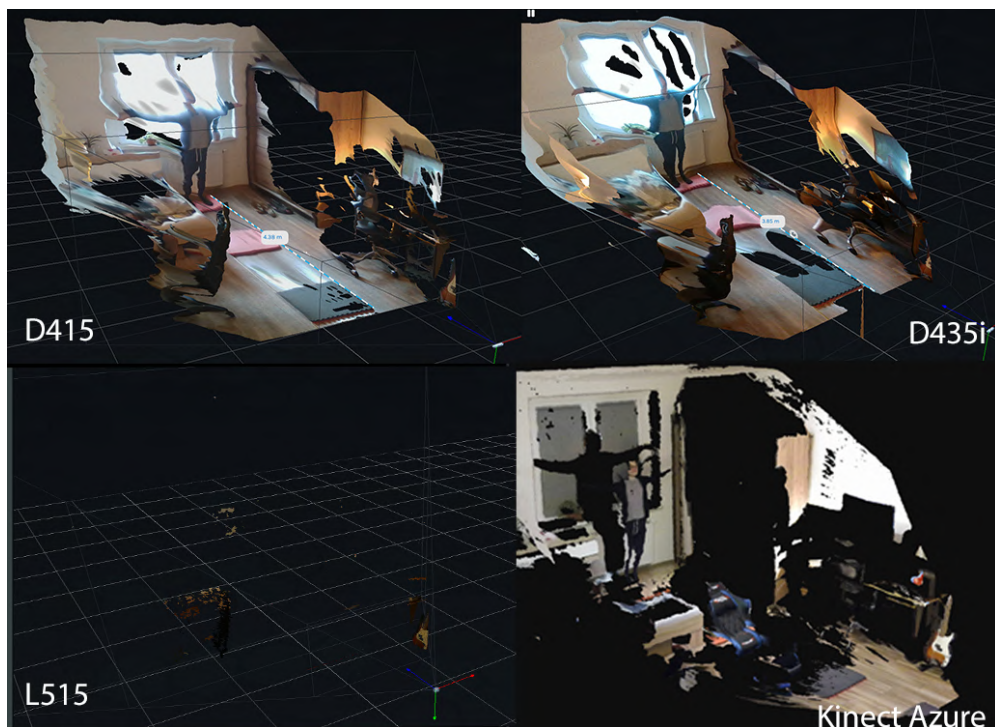
Figure 5.6. Pointcloud performance comparison.

■ 5.4.2 Software and SDKs Features

In the course of this study, it has been explored and tested the Nuitrack middleware, identifying it as the only available software solution on the market that utilizes data from RealSense for body tracking.

Nuitrack™ [63] is a 3D body tracking solution developed by 3DiVi Inc. It enables body motion analytics applications for virtually any widespread depth sensors - Orbbec Astra/Persee/Femto, Kinect v1/v2, Kinect Azure, Intel RealSense, etc.
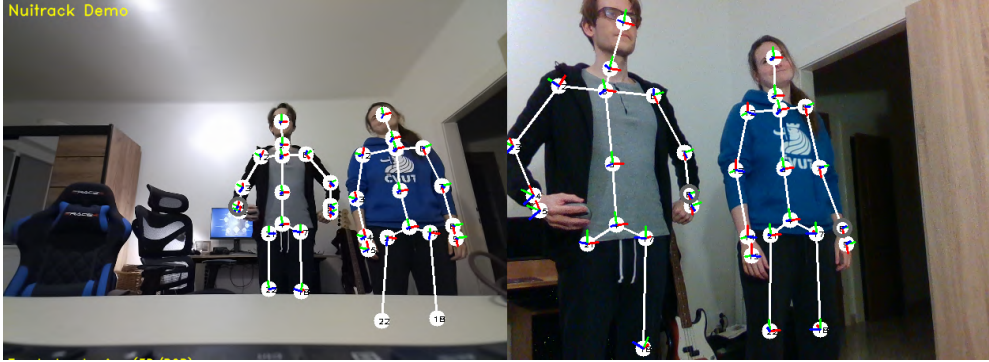
Figure 5.7. Separated input from Azure Kinect and Intel Realsense through one API. Note that Nuitrack allows up to 6 unique bodies to be tracked simultaneously.

Notably, Nuitrack can simultaneously handle multiple sensors within Unity, combining the multiple camera outputs or treating them as separate views, as seen in Figure 5.7 with only a small impact on the project performance. This feature is particularly beneficial for projects requiring complex sensor integration and real-time tracking capabilities. A small comparative test between Nuitrack [7] and the MIT-licensed solution for Azure Kinect, available from Microsoft's GitHub [8], has been made. The findings of this comparison are detailed in Table 5.2.

	Azure Kinect SDK	Nuitrack SDK
Sample scene, one avatar	29 FPS	59 FPS
Complex scene (10000+ tris) , one avatar	4 FPS	39 FPS
Number of joints tracked	32	19

Table 5.2. Comparison and Unity performance of Azure Kinect

It is important to note that the version of Nuitrack used for testing was the trial version, adjusted to a runtime of exactly three minutes. The Pro/AI version costs roughly 99$, and it also offers simple object tracking based on deep neural network learning, such as a ball or beanbag, required for exercises AC1 and AC2 from Section 4.2. However, should the final implementation phase not advance to LiDAR-based software solutions, Nuitrack represents a suitable candidate for body tracking applications, thanks to its unique features and compatibility with both RealSense sensors and Azure Kinect combined.

As the hardware selection phase neared completion towards the end of February, an unexpected opportunity arose with the loan of several LiDAR units, a detail further discussed in Section 5.5.2. While the trial version of Nuitrack offered functionality suited to the requirements for exercises AC1 and AC2, as mentioned in 4.2, its full potential is locked behind a Pro version.

However, with the unforeseen availability of additional LiDARs and the cost considerations of Nuitrack's Pro version, the approach to measuring was reconsidered and thus pivoted away from Nuitrack. The Kinect option was also set aside due to the author's development PC's specifications, which are marginally below the recommended requirements for optimal performance (as seen in Table 5.2). MediaPipe (see Section 3.3.1) has emerged as a particularly attractive option in this context. Despite not being

---

[7] https://nuitrack.com/

[8] https://github.com/microsoft/Azure-Kinect-Samples/tree/master

the most accurate, MediaPipe's intuitive operation and minimal system demands have proven to be well-suited for rapid development and testing on the available hardware. Its open-source nature and the absence of licensing fees further contribute to its appeal for the final implementation phase.

5.5   Software selection

This section evaluates various development platforms, complying with Functional and Non-Functional Requirements mentioned in Section 5.2.

5.5.1   Unity

Unity is a flexible and popular game engine known for its strong community support with a rich resource of knowledge, tutorials, and forums. It directly supports the sensors we're using, and since the author is already skilled in Unity, it's the quickest choice for this project. Unity's capable environment is great for quick development and integration, making it a solid tool for setting up the necessary tracking and visualization features.

In addition to Unity, Unreal Engine and TouchDesigner were also considered. Unreal Engine [9] is known for its powerful graphics capabilities and could be an alternative for projects that require high-quality visual output. However, for the specific needs of this project, the priority is Unity's sensor integration, namely the support of Pharus Entity tracking, as mentioned in Section 5.5.2.

TouchDesigner [10], on the other hand, is particularly appealing for creating visually stunning interactive environments and is commonly chosen for projection mapping projects. Despite this, its limited support for logging and motion-tracking data makes it less suitable for the detailed analytical requirements of this project, for example, Functional Requirement Number 5. (5.2.1).

5.5.2   Pharus

Pharus, a software developed by Otto Naderer, MSc., was introduced as part of his master's thesis at Johannes Kepler University Linz in June 2015. It presents a scalable, low-latency laser-ranging tracking system to track the detected position of entities in crowded environments. This system integrates data from multiple sensors to minimize occlusions, enabling robust tracking of individuals in crowded environments [64]. Pharus is currently integrated into the pipeline of Ars Electronica's DeepSpace [11] project, an interactive cave-like room, enabling virtual experience projected to 16 x 9 meter wall and floor combined space.

The need to utilize a 2-D LiDAR tracking system arises from its capability to output precise xy coordinates of tracked entity for floor projections - this is particularly advantageous compared to skeletal tracking, which struggles with exact spatial localization that will be needed for interactive projection. Additionally, the LiDAR system offers greater precision than a fixed depth sensor positioned above the subject, effectively reducing occlusions.

Pharus is compatible exclusively with LiDARs from SICK[12] and Hokuyo Automatic Co.[13]. CTU FEE provided two sensors for this setup: the Sick LMS151-10100 and the

---

[9] `https://www.unrealengine.com`

[10] `https://derivative.ca/`

[11] `https://ars.electronica.art/center/en/deepspace/`

[12] `https://www.sick.com/at/en/`

[13] `https://www.hokuyo-aut.jp/`

Hokuyo URG-04LX-UG01. Due to the lower performance of the Hokuyo sensor, it can be classified as inadequate, as stated in Table 5.3. Furthermore, Pharus requires a measuring area setting of either 190 degrees or 270 degrees for Hokuyo sensors, but this LiDAR only offers 240 degrees, as seen in Table 5.3.

	Scanning frequency	Aperture angle	Scanning Range
Sick LMS151-10100	50 Hz	270°	0.5m - 50m
Hokuyo URG-04LX-UG01	10Hz	240°	0.02m - 5.6m

Table 5.3. Analysis of Key Features for Selected 2D LiDAR Systems.

The inability to permanently adjust the measuring angle distorts the measuring area. When synchronizing these two sensors, individual positions are not confirmed; they are duplicated. Also, in practical deployment environments, the Hokuyo URG-04LX-UG01 sensor demonstrated considerable susceptibility to varying light conditions, significantly affecting its reliability. The sensor performed optimally in environments with minimal ambient light; however, it became nearly impractical to use in conditions that were not extremely dim. In contrast, the LMS151-10100 sensor exhibited constant performance across a broader range of lighting conditions, maintaining consistent accuracy and reliability. For these reasons, only the SICK LiDAR is used in subsequent implementations.

To facilitate wider usage and customization, this software is supported by an MIT-licensed Unity Deep Space Development Kit available on GitHub[14]. This kit includes laser tracking functionalities, operational control, and multi-projection setup that are pivotal for creating interactive installations within the Deep Space setting.

For this thesis, which is not directly conducted within the Deep Space environment, it was necessary to adapt the Unity 2018.x Development Kit to meet specific project requirements. This adaptation keeps only the laser tracking features of the Pharus system to collect data and interact in a new setting accurately.

Pharus offers two recording modes: recording entities based on the calculation of echoes (in this case, the position of feet in space) and their paths. Since DeepSpace SDK neither requires nor implements the direct positioning of feet and only utilizes entity tracking (which is understandable given the size of the projection area), it rarely uses just one sensor.

With contributions from Mr. Naderer, a special functionality was developed specifically for the requirements of this work, integrating the position of echoes in unconfirmed mode, meaning that multiple sensors' confirmation is not necessary for echo detection. This allows for the positioning of individual feet. However, it is highly prone to errors and the duplication effect, where one object may be detected as multiple entities and their echoes.

Data are transmitted using the User Datagram Protocol(UDP) through two protocols - TrackLink and TUIO. TUIO (Tangible User Interface Objects) is a protocol widely used in tangible user interfaces and interactive surfaces. It is designed to communicate the state of physical objects (tokens) and multi-touch events over a network [65]. In this work, Unity receives data through the TrackLink protocol as it parses the information about echo positions. Furthermore, the SDK functionality had to be modified, which does not account for echo implementation. While Pharus assigns permanent ID to entities within the scene, unlike skeletal detection, it cannot maintain

---

[14] `https://github.com/ArsElectronicaFuturelab/DeepSpaceDevKit`

these IDs if an object leaves the scene. During updates, however, it retains all echoes and associated properties such as orientation, velocity, and echoes, as seen in Table 5.4. Unfortunately, since the echoes do not have fixed IDs, they are overwritten with each update. Therefore, it is impossible to consistently assign and detect the right and the left foot based on these echoes.

Variable	Type	Description
trackID	public int	Unique identifier for each track
currentPos	public Vector2	Current position in meters
expectPos	public Vector2	Expected position in the next frame
relPos	public Vector2	Position in relative coordinates (TUIO style)
orientation	public Vector2	The track's current heading, normalized
speed	public float	Current speed in meters per second
state	public TrackState	Current state of the track
echoes	public List<Vector2>	Echoes associated with the track (TUIO style)

Table 5.4. Data transmitted between Pharus and DeepSpace LaserTracking SDK.

Chapter **6**
Prototypes

This part of the thesis was dedicated to designing various prototypes for selected methods from Chapter 2, which had the potential for further development. The prototype 6.1 was based on physiotherapy principles, although it was not further developed due to the lack of professional consultations. The next prototype focuses on balance exercises from the MABC-2 method, aiming to measure both quantitative observations and implement qualitative elements. This particular prototype laid the groundwork for the motion analysis in the final implementation, as described in detail in Chapter 7.

Both prototypes utilize MediaPipe pose detection from the TensorFlow library (described in Section 3.3.1), estimating poses in real-time by manually setting and calculating threshold values for distances and mutual angles between points to determine the pose.

Instructions for setting up both simple prototypes can be found in Appendix B.

6.1 Physiotherapy - DNS prototype

The prototype, intended more as a proof-of-concept, is designed to detect poor posture, inspired by Figure 2.1. The main idea of the posture analysis prototype lies in monitoring the vertical distances between shoulder and pelvic points, where a theoretically correct stance is expected to have these points aligned along one vertical axis (or at least slightly bent, as shown in Figure 6.1). The prototype also utilizes basic planimetric equations to measure the total angle difference between three points - horizontal shoulder line, shoulder and pelvis point, and horizontal pelvic point, pelvis, and shoulder.

Poor posture is visualized by a growing red horizontal line, as can be seen in Figure 6.2.
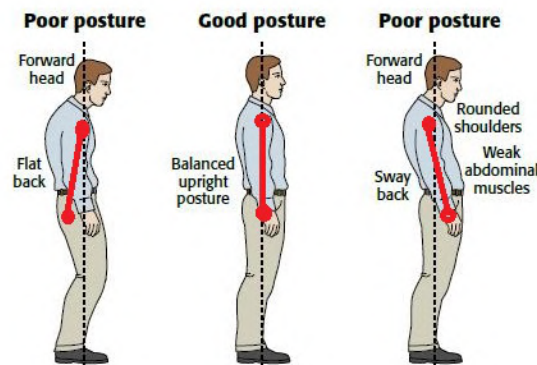
Figure 6.1. Postural Dysfunctions. One represents poor posture with a head leaning forward and flat back, the middle shows good, balanced upright posture, and the last demonstrates poor posture again, characterized by a forwardly leaning head, rounded shoulders, and weak abdominal muscles (image adjusted and taken from [66]).

a) good posture   b) bad posture
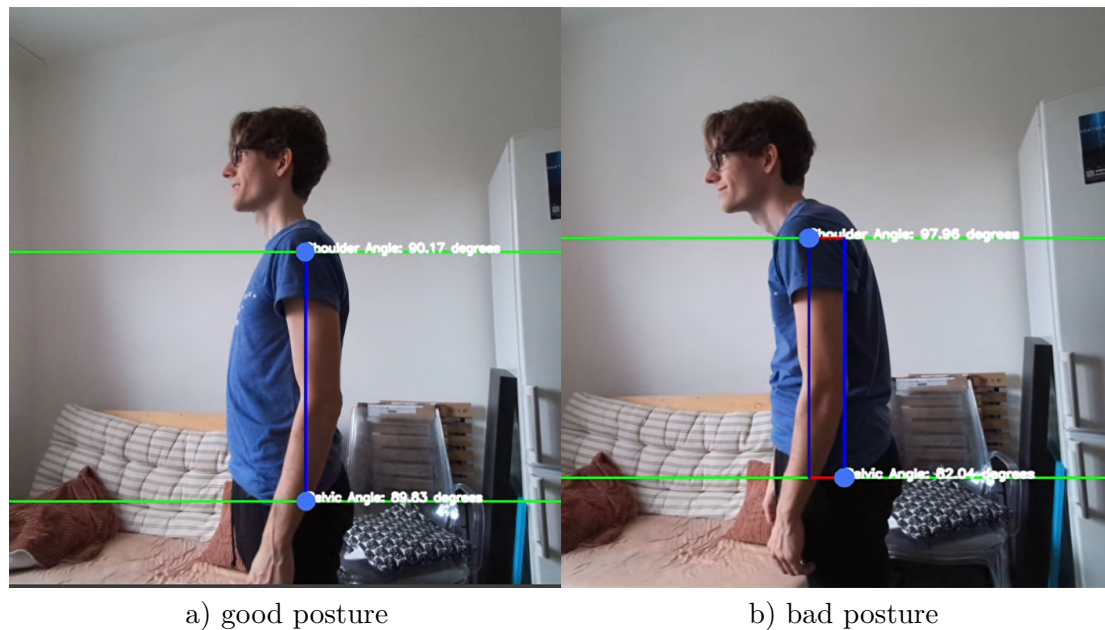
Figure 6.2. Postural Dysfunction, as detected by the prototype. a) shows correct posture stance, b) the person demonstrates typical posture that can be seen very often in desk-office jobs; head and body leaning forward, without abdominal engagement.

As previously noted, this prototype has not been validated due to the lack of professional staff for assessment—setting threshold values for angles and spinal incline would make sense for logging such an application. Additionally, the limitation of the Media Pipe library, which lacks additional points for the front of the pelvis and diaphragm, reduces the test's validity in tracking the diaphragm function within DNS.

6.2   Ergotherapy - MABC2 prototype

This basic prototype mirrors the physical test. Five sample scenes were created, containing `TrackingEntityManager` from the `DeepSpace.Lasertracking` part of SDK, discussed previously in Section 5.5.2, receiving and visualizing the spatial data from Pharus. The assessments keep the original dimensions and appearance, as required by FR4(5.2.1), and are projected on the ground, as seen in Figure 6.3. It has basic logging capabilities, where the Debug Console is redirected to a CSV file. The logging of the variables is identical to what was used in subsequent implementation, as will be referred to in Chapter 7 - correct throws, number of steps on the balance line, etc.
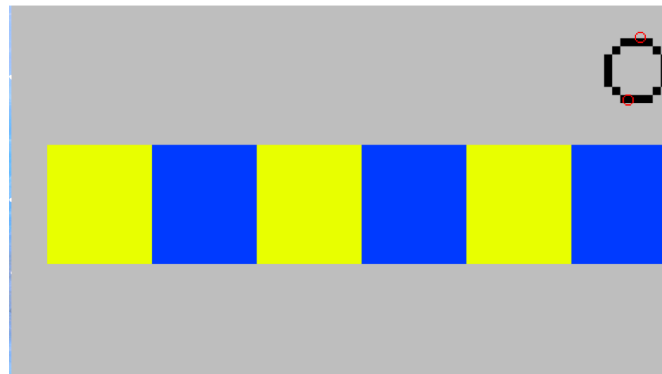
Figure 6.3. BAL3 scene from the prototype, with colored mats and visualization of standing TrackingEntity and its echoes (black circle with red dots)

■ 6.2.1 Skeletal tracking

This early low-fidelity prototype aimed to measure and detect the duration of standing on one leg, according to the MABC-2 BAL1 methodology discussed in Section 4.3, with qualitative observation considered. During the trial, the camera observes individual qualitative elements in real-time. It classifies them as true or false based on simple trigonometric reasoning. When the measured parameters are found in the camera, it waits for the user to raise their foot by indicating the START signal at the top corner of the display (see Figure 6.4).

Figure 6.4. Camera detects both ankles and knees of the participant.

The prototype measures the duration of leg standing, detects the leg chosen, and measures the qualitative criteria:

■ **Stiff-looking body** - determined by analyzing the shoulder width and the horizontal positions of the hands (wrists). If the difference in shoulder width is below a certain threshold and the horizontal positions of the hands are close to each other, the script considers the body as `stiff`. This condition suggests that the person has their arms close to their body and is not making exaggerated movements.

37

- **Flabby body** - Detected when the body is not stiff and shows large arm movements and swaying. Significant differences in the vertical positions of the elbows and notable deviation of the body's midpoint from the center are critical indicators of a `flabby` body.
- **Swaying violently** - Assessed by the midpoint movement between the ankles. If this point shifts markedly from the center, it suggests unstable movements, indicating `violent swaying` issues.
- **Exaggerated movements** - Identified by the considerable vertical distance between the shoulders and elbows. Large or dramatic arm movements qualify as `exaggerated.`. Both violent sway and exaggerated movements can be seen in Figure 6.5

a) exaggerated arm movements   b) swaying violently
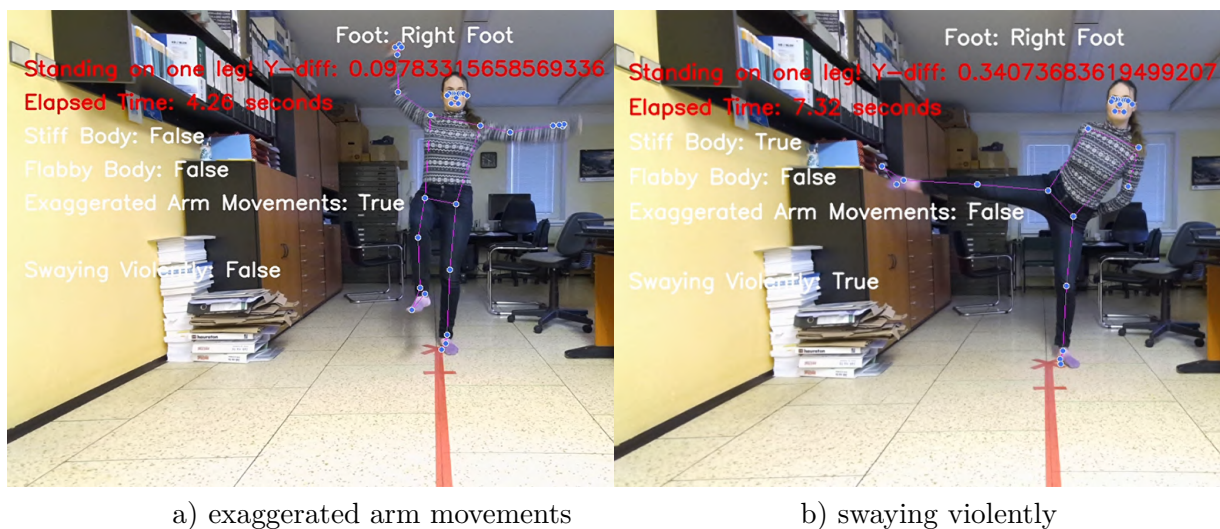
Figure 6.5. Markerless detection of one leg standing assessment - qualitative parts.

This prototype has one big downside - it measures the standing on the foot only if both ankle landmarks are visible - however, the correct performance of the exercise permits one leg to bend back, making it impossible for this simple prototype to handle. The subsequent implementation of detection was thus reinforced with a machine learning model, which will be elaborated later on in Section 7.1

Based on discussions with domain experts and considering the prototype's limitations observed during testing, it was decided that future development would not include detecting qualitative elements.

■ 6.2.2 Changes and variations

During the creation of the prototype, a few key changes must have been made. The first concerns the beanbag from the Aim and Catch circle. For purposes of this thesis, a few adjustments were made:

1. The author of this thesis does not own the original beanbag but measured the exact parameters - 13x6x5 cm.
2. Given the size of the beanbag, when the object landed in the measurement area, it was seldom registered as a single echo or not registered at all.
3. To suit the detection parameters, a tote-bag prototype has been created, as seen in Figure 6.6. The recreation of the original bag deviates by 7 centimeters horizontally,9 diagonally, and 15cm vertically.

Figure 6.6. A homemade prototype of Hogrefe's MABC2 Beanbag - a soft tote bag filled with socks, roughly scaled up to pass TrackingEntity identification.

The second part concerns the BAL2 assessment, where the child must walk on tiptoes on a 4.5-meter line. As will be discussed in the User Testing Chapter 8, none of the testing environments provided the possibility of suitable spatial and projection dimensions. The width of the line thus equals the width of the projection.

All remaining assessments are identical to the methodology.

Chapter 7
Implementation

This chapter describes the development of the system, focusing on the application of user-centric design principles mentioned earlier in Section 5.3 and meeting specific functional and non-functional requirements (5.2). It is divided into three sections: an overview of the system, implementation of individual tasks, and data gathering. Each section will detail the process, challenges, and solutions for implementing the interactive application and analysis system. The instructions on assembling and launching the application can be found in Appendix A.

7.1 System Overview

This section provides an overview of the integrated system architecture illustrated in Diagram 7.1, detailing how motion is captured and analyzed through interconnected modules.
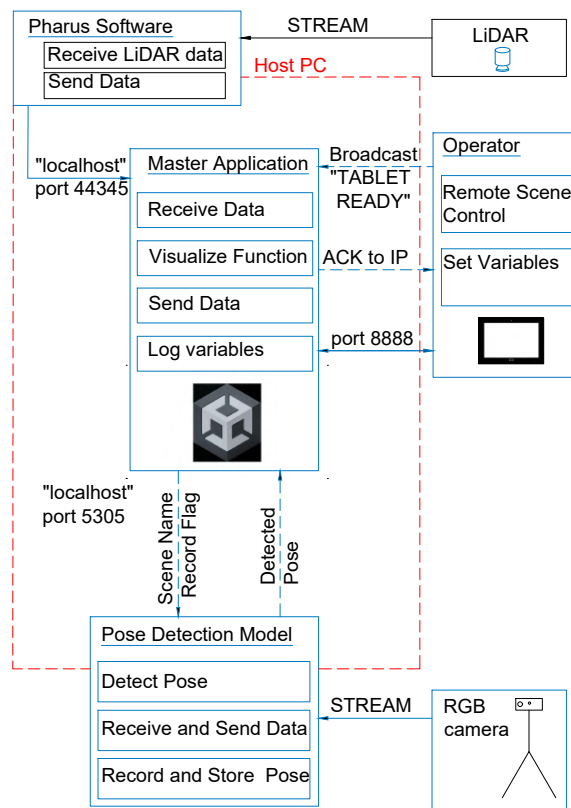
Figure 7.1. This diagram presents the system architecture integrating sensors, pose detection model, tablet controller, data flow and processing steps.

The system architecture in the diagram depicted in Figure 7.1 represents an integrated setup for capturing and analyzing motion through a series of interconnected modules. At the core, the Host PC acts as a central processing unit, running the Unity master application, which receives and manages LiDAR data via the Pharus Software, as described previously in Section 5.5.2, communicates between pose detection script and can be controlled over the network with Operator device, as will be elaborated further in Section 7.1.1.

Data flows between the components as follows: the LiDAR system collects spatial data and sends it to the Host PC, where the Pharus Software processes this data and forwards it to Unity. Unity's Master Application then utilizes this data to visualize the interactive application and to perform pose detection with the help of the RGB camera. The system is also equipped to log various variables (described in greater detail in Section 7.3) and manages scene-specific settings through an operator interface, which can set parameters remotely and send commands back to Unity for real-time adjustments.

■ 7.1.1 Operator Mode

As already mentioned in the initial assignment of the thesis, there is a requirement that the therapist himself can easily control the application. The application offers two modes - Operator and Host mode. Host mode runs on a computer that simultaneously runs projections and receives sensor data. The operator device communicates over the network with the guest computer. For the purposes of this implementation, the operator was an Android OS tablet, but any device with a compiled application could be the operator. Mobile devices, by default, cannot host a session but only serve as operators.
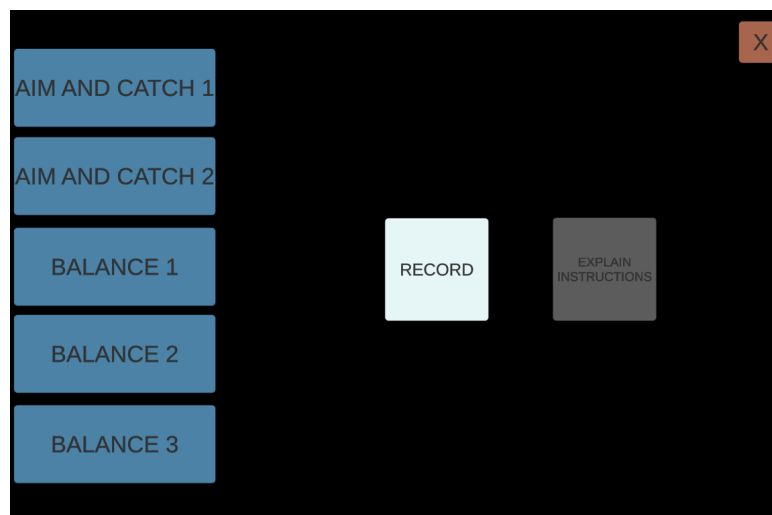
Figure 7.2. An overview of remote operator interface. The operator can set the scenes and toggle on/off the recording of the assessments and their oral explanation.

Figure 7.2 shows that the operator can switch between scenes and set the variables related to recording sessions and launching verbal explanations. As shown in Figure 7.1, the recording launches recording the coordinates in the skeletal tracking app, and verbal explanation enables spoken instructions, as will be mentioned further in Section 7.2.1. If the host computer does not register any connected device from which the main application is controlled, it can continue without the driver. This launches the first scene, AC1, and enables the UI Controller prefab, which behaves like a classic hamburger menu, as seen in Figure 7.3, where clicking in the upper left corner opens a

41

canvas with buttons identical to what the remote operator can control. Each assessment shares this prefab.
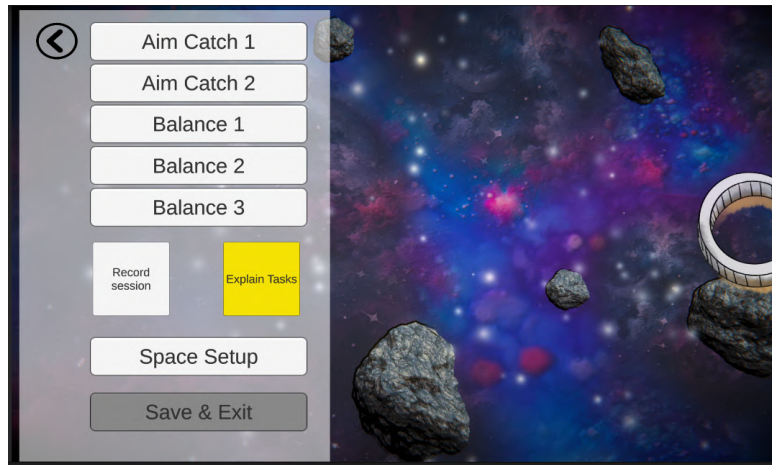
Figure 7.3. An overview of built-in user interface. The scenes and variables can be operated from the main application.

7.2 Assessments Implementation

Each of the MABC2 assessments shares the same fundamental architecture of the application. Since data from the Pharus is received in both relative (relative to Pharus map space, as seen in Figure 7.4) and world coordinates, all conversions are performed within these coordinates, specifically regarding the x and y coordinates of detection. In the Unity scene, the Tracking Entity Manager from Pharus SDK is placed at the origin. The positioning of the camera and the scene must mirror the calibrated space in Pharus, especially when dealing with a smaller scanned tracking area than the scanner ranging, as illustrated in Figure (7.4).

Each assessment features a UI Canvas projected at a location corresponding to Pharus's tracking space. Within this canvas, each component possessing a 'RectTransform' property has its own detector, identifying whether any entity is in that space. The basic detection process is based on determining whether a person is present on rectangles of sizes identical to the actual test mats.

Only the application's background is visible in the final implementation, but the actual detection occurs within the canvas. Initially, there was no need to adjust the z-axis placement (perspective axis), and for simulating MABC2, it was sufficient to project only the canvas. The Tracking Entity Manager thus provides data at positions x, y, 0. When creating visuals, to achieve an isometric view of the camera that suggests depth and eliminates the flat appearance of orthographic projection, neither the camera nor the canvas was rotated; instead, the background scene was adjusted. It should be noted that this solution is not ideal and could be improved further.
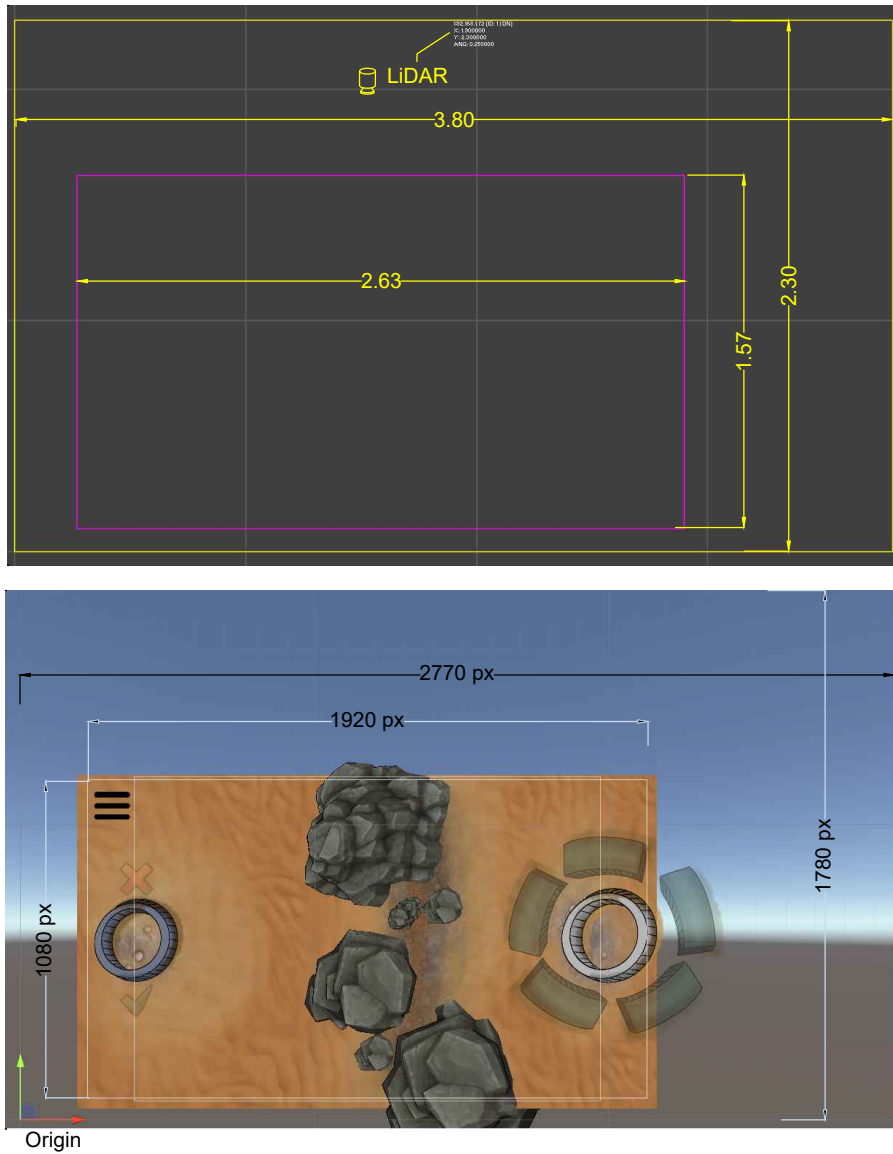
Figure 7.4. An annotated image showing the set space within the Pharus Space Setup. The yellow rectangle represents the mapping space with the LiDAR positioned. The purple rectangle is the actual detection space, with an identical size to the projection space. The space is deliberately reduced in size, as the map space is blocked in many places, which would lead to additional unwanted detections. The bottom image is then a cutout from the Unity application, where the `TrackingEntityManager` is placed at the origin of the coordinate scheme, and the application has to copy the setup from the top image - a simple conversion to pixels from meters is applied to make the canvas match the 1920x1080 resolution.

43

- **AC1** This task involves the therapist and the child. The child is shown an animation of a *swarming* circle, drawing the attention to step inside, as seen in Figure 7.5. The circle keeps the entity's ID (a child) during runtime, and if a child steps out, a warning will prompt their return to the ring.
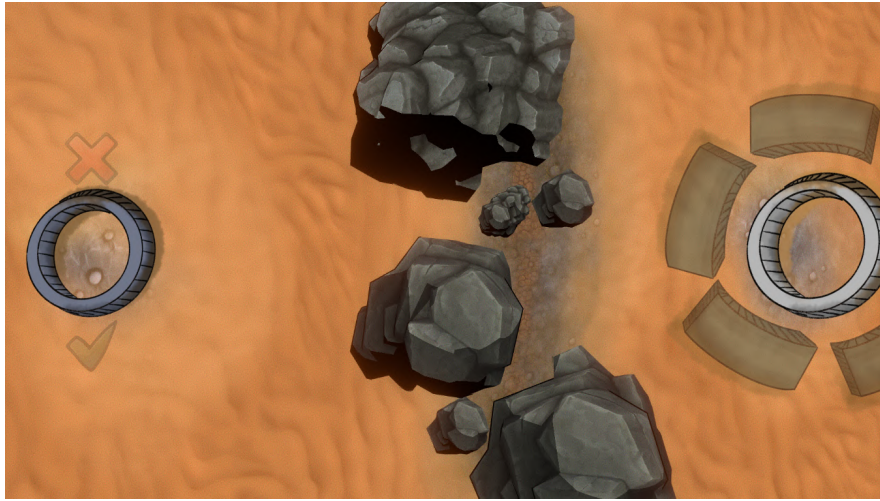
Figure 7.5. AC1 Scene - Desert with rocks that separates the visual space onto left and right halves.

The therapist throws a ball to the child. When the therapist steps into their designated circle, two faint symbols appear - a green flag on the right and a red cross on the left. After catching the ball, the therapist, based on the outcome of the throw, retrieves the ball and steps to either the left or the right, logging the result in the Logger instance. The child must stay within the circle for the entire measurement duration. The symbols are deliberately made barely visible (opacity set at 0.2) to keep the child unaware of the throw's success or failure.

- **AC2** This task, previously described in 4.2.5, relies solely on Pharus readings. The task is inspired by a stay in a space territory, where the child is required to shoot at a rotating planet.

At the beginning, a tutorial is played for the child, instructing them to stay in place and throw a ball.

Detection is divided into four sections, each triggering a boolean flag based on where the ball is detected, as shown in Figure 7.6:

- **Ball shooting area:** A blue rectangle with specified dimensions that detects the correct throw. If the ball lands in this area, it raises the flag and deactivates itself.
- **Incorrect throw area:** A red rectangle that indicates incorrect landing. If the ball is detected here, it deactivates the blue (correct) area to prevent sliding the ball across the area (like in curling). If the ball lands here, it raises the flag and deactivates itself.
- **Ball detection area:** An area with identical dimensions and locations activated after a throw. It can handle both cases: if the ball is not detected but was detected previously (meaning the ball moved away after the throw), the entity can step away from the throw site (child step zone). All variables are reset after the child steps back. If the ball remains in the zone, the child is expected to return and pick up the ball (since we have fixed Entity IDs, we can identify the ball).
- **Child step zone:** A white ring. If the child steps out of the ring without a ball being detected (stepped away during the throw), a warning sequence (animation) instructs the child to return to the ring.

44

a) Actual Scene   b) Detected Areas
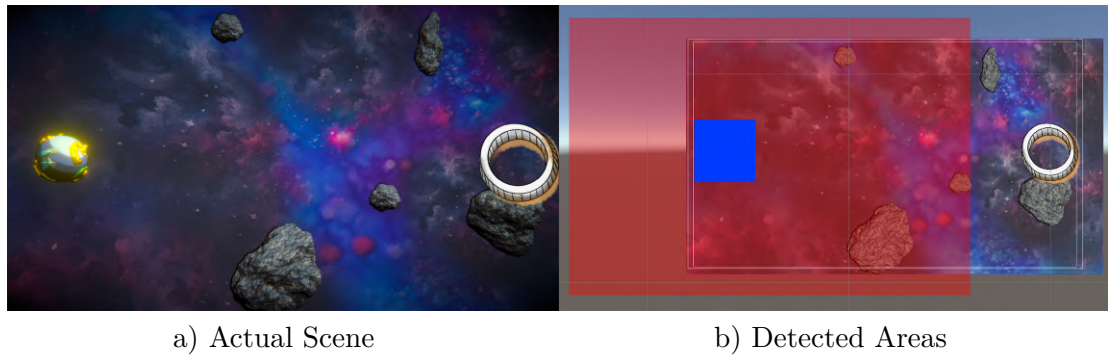
Figure 7.6. A space scene for AC2. a) is the projected scene, and b) are the underlying detection zones.

■ BAL1

In the BAL1 assessment, children are instructed to balance on either their left or right foot. At the beginning of the task, a pulsing circle guides the child into the measurement area, as previously used in AC1. Upon entering the measurement zone, a tutorial demonstrates correct and incorrect one-legged postures, as seen in Figure 7.7. During the assessment, which lasts for 30 seconds, the child sees a well filling with water, as discussed in the Interaction Design Section (see 5.3.1), and a magical forest that appears around them, as seen in Figure 7.8. A sound signals successful completion, and if the child stops balancing prematurely, the tutorial replays to encourage trying again on the other foot.

Figure 7.7. Animated tutorial for BAL1 Assessment. A correct balancing posture (left) and its incorrect form (right).

The assessment only measures two trials per foot, recording the best result for each, as detailed in Section 7.3.1.

This scene includes a UDP listener that receives pose data. The pose detector can signal three states: START, Standing on Left Foot, and Standing on Right Foot.

Based on these signals, the application adjusts visual elements:

- **START** - The PoseDetector detects both ankles aligned horizontally. Upon this detection, the application flags this measurement as valid and initiates the trial duration timer.
- **During the trial** - Interpolate the z-coordinate of the forest and the water level in the well based on the child's balance. This adjustment occurs only if the child maintains the same standing position on one foot, as the PoseDetector indicates.
- **No Detection Event** - If no new messages are received from the PoseDetector for 2 seconds, or if a `START` signal is detected again, the system logs the current trial result. It switches the tutorial animation to instruct standing on the opposite foot.

Figure 7.8. Finished BAL1 assessment.

The standing detection is secured in the pose detector program, a MediaPipe real-time detection taken from an earlier prototype (see Section 6.1), combined with a machine learning model. The need to employ a trained model arises from how the prototype estimation works. The real-time detection uses only the relative difference between the y-coordinate of the left and right foot (ankle) - if the visibility of both ankles is higher than the set threshold value for identifying the standing. However, the assessment permits the non-standing leg to be extended backward - and in this case, the pose detector loses one of the ankle points.

Recordings of various adult persons from the front and varying angles have been utilized, with balancing on one foot the correct and incorrect way. MediaPipe can detect poses from the recorded videos - a simple script for storing the annotated pose landmarks has been made to extract pose and hand landmarks using the holistic model from these recordings and store them in a CSV file. This storage format is consistent with the method used during application runtime to capture landmarks, as detailed in Section 7.3.2. Due to time limitations, the recorded dataset spans only five minutes and predominantly features correct poses. This data collection process is visualized in Table 7.1, illustrating that each frame read from the video is stored as a single row (registration) in the dataset.

For the creation of a machine learning model, a `Scikit-learn`[1] library for Python and `Pandas`[2] for data utilization was used. During the model training phase, a significant decision was made to exclude hand landmark data from the analysis.

---

[1] https://scikit-learn.org/stable/
[2] https://pandas.pydata.org/

Classified Pose	Number of registrations in dataset
LegStandCorrectLeft	2267
LegStandCorrectRight	1993
LegStandIncorrectRight	1397
LegStandIncorrectLeft	1142

Table 7.1. Distribution of various poses in the dataset.

This adjustment was crucial as the initial dataset inadvertently introduced a 'Clever Hans' [67] effect: incorrect postures often had hands hidden behind the torso, as seen in Figure 7.9, and the posture was labeled as correct when hands were visible, even without seeing the full torso. Focusing only on lower body landmarks and omitting hand and upper body landmark data diminished this effect completely.

Figure 7.9. An example of typical dataset entry. The image on the left represents an incorrect posture. The image on the right induces correct standing, with hand landmarks (denoted as black circles) detected.

The training process incorporated a 70/30 split between training and test data to evaluate the model's accuracy. In developing the pose detection model for the assessment, high accuracy rates *close to one* were achieved across different machine learning pipelines. Even with employing 10-fold cross-validation, the results were similar. Such near-perfect accuracies often signal a potential issue with the model, such as overfitting, where the model learns the training data too well, including its noise and anomalies, thus failing to generalize well to unseen data. One of the many reasons for this is class imbalances introduced in the dataset.

The Random Forest Classifier model was chosen from the machine learning pipeline for its robustness against overfitting and its effectiveness in handling high-dimensional datasets since multiple features must be considered in this project. A Random Forest Classifier [68] is a combined learning method for classification and regression that constructs multiple decision trees during training and outputs the class, which is the mode of the classes predicted by individual trees, thereby improving accuracy and controlling overfitting by averaging the results.

	Precision	Recall	F1-score
LegStandCorrectLeft	0.00	0.00	0.00
LegStandCorrectRight	0.30	0.79	0.43
LegStandIncorrectRight	0.75	1.00	0.86
LegStandIncorrectLeft	0.21	0.40	0.27
Overall Accuracy			**0.36**
Macro Average			**0.39**
Weighted Average			**0.24**

Table 7.2. Imbalance classes model evaluation on an unseen dataset. Precision is the ratio of correctly predicted positive observations to the total predicted positives, Recall is the ratio of correctly predicted positive observations to all actual positives, and F1 Score is the weighted average of Precision and Recall.

A standalone dataset with similar recorded poses was created to determine whether the model can generalize on unseen data, as seen in Table 7.2.

Performance Overview

- **LegStandCorrectLeft:** The model failed to correctly predict any instances of this class, as shown by the precision, recall, and F1-score of 0.00. This indicates a significant issue with the model's ability to generalize to this class [69].
- **LegStandCorrectRight:** Moderate recall (0.79) suggests that the model can identify this posture when it occurs. Still, low precision (0.30) indicates high false positives, implying many other postures are mistakenly classified as LegStandCorrectRight [69].
- **LegStandIncorrectLeft:** This class shows weak performance across precision (0.21) and recall (0.40), suggesting some ability to recognize this posture but with considerable inaccuracies [69].
- **LegStandIncorrectRight:** This class has high precision (0.75) and perfect recall (1.00), indicating the model performs well in identifying this specific posture when it appears and generally classifies it correctly [69].

The overall accuracy (0.36) is quite low, indicating that the model is correct only about a third of the time across all predictions. High performance on training data and poor generalization suggest that the model is, as was expected, overfitting to the training set, capturing noise and specific patterns not applicable to general cases.

Given the poor performance, likely due to the small dataset size and uneven class distribution, the decision was made to augment SMOTE (Synthetic Minority Oversampling Technique) on training data. SMOTE [70] is a technique that generates synthetic samples from the minority class to address the class imbalance, thereby enhancing the training process and potentially improving model accuracy, as seen in Table 7.3.

Performance Overview

- **LegStandCorrectLeft and LegStandCorrectRight:** both show very high recall rates (1.00) and decent precision, resulting in high F1 scores [71].
- **LegStandIncorrectLeft:** This class shows no correct predictions (precision and recall both at 0.00). The model fails to recognize this class entirely, possibly due to a lack of distinguishing features or insufficient training data that accurately represents this class [71].
- **LegStandIncorrectRight:** While achieving perfect precision (1.00), it has a recall of only 0.54. This means while the model reliably predicts this class, when it

	Precision	Recall	F1-score
LegStandCorrectLeft	0.79	1.00	0.88
LegStandCorrectRight	0.77	1.00	0.87
LegStandIncorrectRight	1.00	0.54	0.70
LegStandIncorrectLeft	0.00	0.00	0.00
Overall Accuracy			**0.8**
Macro Average			**0.61**
Weighted Average			**0.74**

Table 7.3. Scaled feature model evaluation on an unseen dataset.

decides to label an instance as LegStandIncorrectRight, it fails to identify more than half of the actual instances, resulting in many false negatives [71].

The model's overall accuracy is 0.80, indicating that 80% of all predictions are correct. However, this number is somewhat misleading due to the unbalanced performance across different classes. The macro average F1-score (0.61) gives a better picture of performance across classes, as it does not account for support (the number of samples for each class). The weighted average F1-score (0.74), which accounts for class imbalance, is higher, indicating that the model performs better on more frequently represented classes [71].

This artificially sampled model was used in the final implementation.

▪ BAL2

The Balance 2 assessment, as detailed in Section 4.2.7, introduces a visual metaphor by depicting a scenario where the participant walks on a rope suspended over a shark-infested reef, as seen in Figure 7.10.

An animated tutorial is displayed during the assessment when a child steps onto a platform that induces stability over the sea. This tutorial demonstrates the correct tiptoe walk and then the incorrect postures, providing clear visual guidance. After the animation ends, the sounds of a rope creaking begin, and the rope image is shown.

Figure 7.10. Virtual scene with a rope bridge suspended over a turquoise sea.

The balance line in the assessment is adjusted to fit the projected area, ensuring it aligns with the projection specification as discussed in Section 7.1. Since the width of the area is only 3.8 meters, and the required length of assessment is 4.5 meters, this problem can be solved by expanding the detected area alongside the whole width (4.5 meters), physically drawing a line on the floor at the projections end boundary and

removing the knot from finish tile. Still, for this thesis, this task has been reduced to the projection area, i.e., the width of 2.63 meters.

The task ends when the child reaches the second platform or manages to take 15 steps. Since Pharus does not recognize footsteps, individual echoes cannot be tracked behind ticks, and the task ends after reaching 15 steps; a simple algorithm for counting steps has been made:

The algorithm works on the histogram principle:

1. The line is divided into n-segments, where n is set to value 128 - this value was set based on repeated observations, where on the measured area, this number represented the best variant for sufficient separation of individual echos/steps.

2. The algorithm has a fixed value at which it declares a step. This value is quite low, hovering around 20-30 captures per bin throughout the testing (reffered in Chapter 8). This is mainly because the child can ignore the task and run across the line.

3. Then, it temporarily disables nearby bins after a footstep is registered to avoid repetitive counts from the same echo cluster.

4. Since walking on tiptoes at the same pace is generally quite difficult compared to a normal stride, a step surpassing a specific threshold value is classified as a heel raise.

5. Results are then logged to CSV files, providing a structured dataset for post-assessment analysis. The logged data includes echoes per bin, step counts, heel raise incidents, and timestamps.
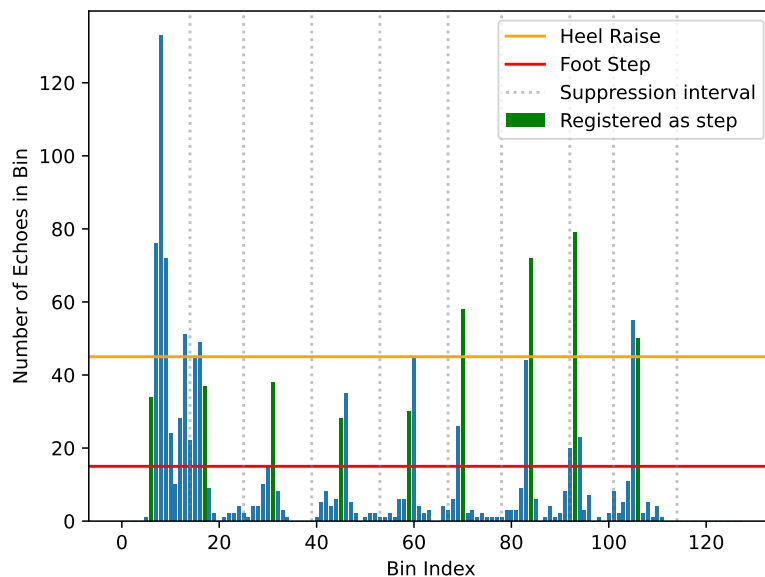
Figure 7.11. Footstep histogram. An example taken from User Testing No.2(See Section 8.2). This histogram shows the distribution of echoes across bin indices from 0 to 127. Green bars indicate the bins that registered the step. Horizontal lines at y=45 and y=15 represent the thresholds of the 'Heel Raise' and 'Foot Step'. Gray dotted lines indicate suppression intervals, extending ±10 bins from each central bin marked in green to prevent multiple bins in a cluster from being registered as multiple steps. For visual clarity, only the positive range is shown. Heel raises are counted per detected step, not per echo, regardless of which bin in a cluster is registered as a heel raise.

The steps are detected with very high accuracy. Still, since we cannot obtain blob size from Pharus (meaning the size of an echo), the detection of heel raise (tiptoe walk) is an extension of step registration, with mixed results (See Figure 7.11, where two of the steps were omitted from 'Heel Raise' detection, even though the test subject walked the line perfectly.).

The values of bins were set up based on the feedback from user testing, as will be mentioned in Chapter 8. Obtaining the blob size could refine the detection, as the area covered by the front part of the foot (standing on tiptoes) will naturally be smaller than that of the entire foot.
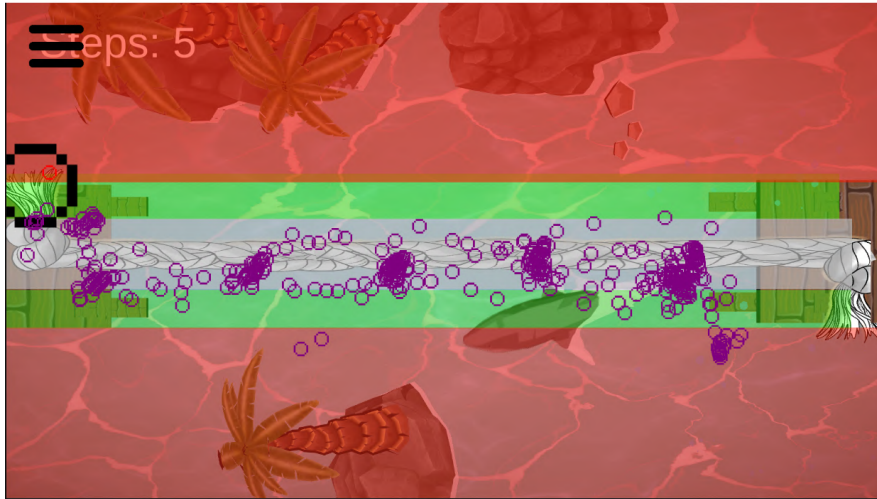
Figure 7.12. Detection zones for BAL2, with echoes recorded (purple). The red zones launch warning animations, grey is the correct line, and green bars detect overrun from the correct line. Black circle represents walking entity.

The original task involves counting the correct steps in a sequence and resetting the counter if a step is incorrect. Since determining the standing on tiptoes is difficult, this implementation considers a step correct if the echoes fall on the rectangle marked as the correct line. As can be seen from Figure 7.12, it is a gray line overlapping the drawn rope. The step is counted as correct if all echoes are within this area. The counter resets if an echo is detected in the adjacent green zones. Unfortunately, as mentioned in the Section 5.5.2, with the single-sensor setup, movements parallel to the sensor can obscure the projection area, rough simulations of the covered area and overly offset-placed echoes occur, particularly in incorrect zones, as is clearly shown in the aforementioned figure. Although the functionality to reset the counter in this solution exists, it is not being used during user testing, and it only tallies steps without resetting. For full functionality, at least two more precise sensors are necessary to capture the area completely.

51

▪ BAL3

This assessment features six tiles, using the metaphor of tiles floating on the icy seas. The child must safely hop from these unstable *ice tiles* to the secure *solid ground* tile. Figure 7.13 shows an example of this environment.
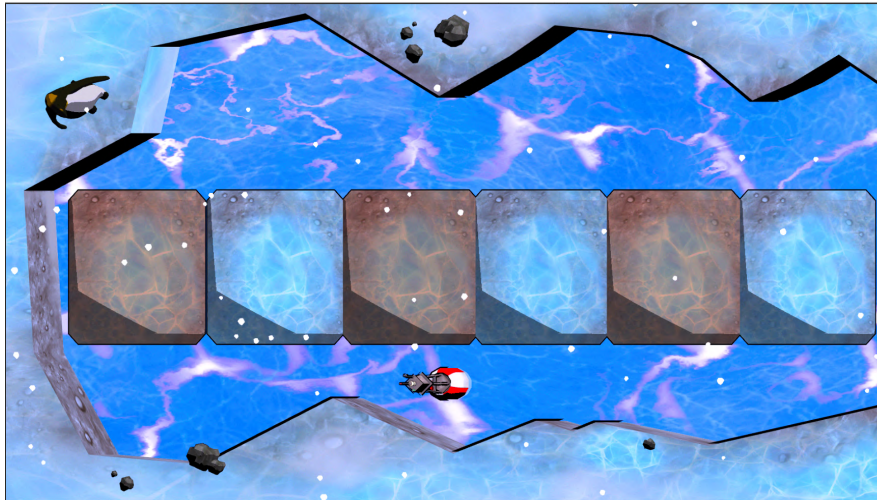
Figure 7.13. Virtual scene with six hexagonal tiles as ice floes in a sea, with a hopping task simulation to reach solid ground (last tile without animation).

The scene comprises six rectangular tiles, each equipped with a tracker that logs the time when the first contact is detected on the tile. At the beginning of the assessment, a float array is initialized to track the number of correctly executed hops. When a detection flag is triggered on a tile, the system verifies that the timestamp of the previous tile is not null, ensuring that the child does not skip any tiles.

Each tile only registers a correct hop if the child remains on it for at least one second, preventing them from merely walking over the tiles.

Upon completion of the activity, when the child steps off a tile, an animation is triggered, showing an ice floe collapsing into the sea. The task is completed if and only if the child reaches the last tile (ground).

▪ 7.2.1 Verbal instructions

The therapist can launch `Explain Assessment` through the controller app in the therapy setup. Each assessment includes brief verbal narration detailing the expected actions of the child, and it can also comment on whether the child's actions are correct or not, strictly adhering to the prescribed guidelines. These verbal instructions and feedback were generated using speechgen.io[3], a text-to-speech (TTS) generator.

A text-to-speech generator converts written text into spoken words. This technology utilizes synthesized voices to read text aloud and record it.

---

[3] `https://speechgen.io/`

7.3 Data utilization and gathering

This section utilizes both application runtime logging and the skeletal tracking of the application users.

7.3.1 Logging the assessments

The `LogManager` class is crucial for logging and exporting data during assessments within the Unity application. It is designed as a singleton to ensure it persists across different scenes without duplication. Each assessment control script has the right to write to variables in LogManager corresponding to its scene name. Also, part of Debug Console Output is redirected to LogManager for a better understanding of trial runtime.

It provides several functionalities:

■ **Debug Console Redirection**

The method `ExportLogsToCSV` captures log entries and data related to the trials. Upon the conclusion of a session, it exports these logs to CSV files for further analysis. Each row includes:

■ **Scene Name:** The Unity scene where the log was captured.
■ **Timestamp:** The exact time when the event was logged.
■ **Message:** Specific details or events captured during the session from the runtime, histogram measuring, and variable setting.

■ **Histogram Data Export**

For events like a completed BAL2 trial, the method `LogHistogram` records additional histogram data, as discussed previously in Balance 2 Assessment Implementation (See ??), which contains in each row the following items:

■ **Bin Index:** The index number for histogram bins.
■ **Echo Histogram Value:** The value of the histogram echo at each bin.
■ **Registered As Footstep:** Indicates whether the echo was registered as a footstep.
■ **Registered As Heel Raise:** Indicates whether the echo was registered as a heel raise.

When the `GameManager` signals that a child has finished walking, potentially multiple times during a session, the data is immediately stored in a uniquely named CSV file tagged with the histogram name and the timestamp of the measurement.

■ **Exporting MABC2 Form-Related Data**

The `ExportTestingDataToCSV` method within the LogManager class plays a pivotal role in quantitatively assessing the performance and calculating the Total Test Score for the MABC2 percentile classification. This method is triggered upon the conclusion of an assessment or manually by the therapist upon closure to generate CSV files containing detailed scoring data. Each entry in the CSV includes:

■ **Exercise:** Identifies the specific assessment activity, such as AC1, AC2, or balance-related measures.
■ **Value:** The numerical score or count derived from the assessment activity. Allows therapists to calculate the percentile of performance based on these recorded metrics.

Upon application closure, LogManager exports all logs and specific assessment data to separated CSV files, tagged with a timestamp, stored in the device's persistent data path, and then opens the folder location of saved data. This automated process ensures data is available for subsequent analysis and review.

■ 7.3.2 Skeletal tracking

As discussed previously in the application overview (7.1), the application communicates with the skeletal tracking script `PoseDetectorMABC2.py` via UDP to facilitate real-time data transfer. When the therapist initiates a session recording, the main application raises a recording flag, triggering the sending of data to a separate Python thread dedicated to handling this information.

The Python detector captures each data point, recording them into a CSV file using the holistic pose detector from MediaPipe, detailed in 3.3.1. The data includes the timestamp and the name of the current scene, followed by the detected pose. Each row in the CSV file comprises the following columns:

- **Timestamp:** The exact time the data was recorded.
- **Scene:** The current scene within the application.
- **Pose Detected:** The specific pose that was identified.

Following these, the file contains 33 groups of 4 columns, representing the positions and visibility of up to 33 pose points:

- **x, y, z:** The world coordinates (in meters) of each pose point, with the origin at the geometric center of the hips. The z-coordinate specifically represents the depth measurement in real-world 3D space, as described in Section 3.3.1
- **v:** Each point's visibility indicates whether the point is likely to be accurately estimated and visible within the frame.

The data is saved in world coordinates, allowing each row to be retrieved and visualized using MediaPipe's function `plot landmarks`, as seen in Figure 7.14.
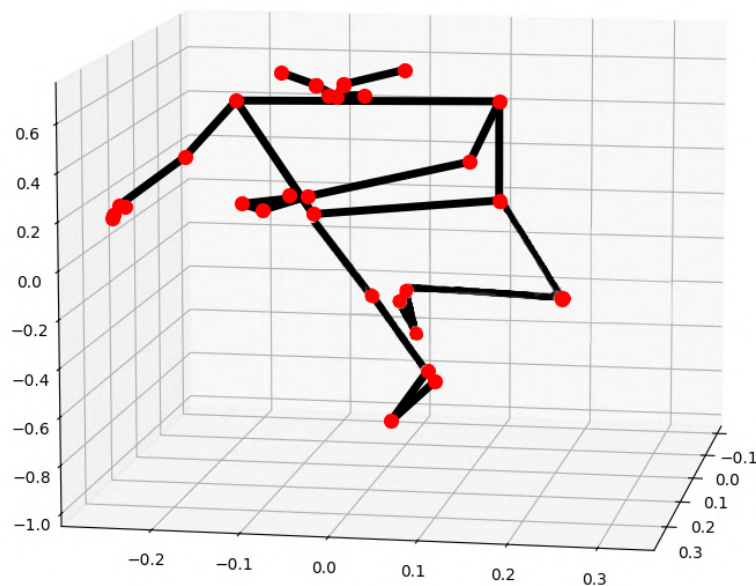
Figure 7.14. A plotted landmark model showing the PoseDetected flag for StandingIncorrectRight, indicating the child's improper posture by standing on the right foot with the left foot rested on the right knee, as detected by the pose estimation algorithm.

Chapter 8
User Testing

This chapter describes the user testing phase conducted in two sessions to evaluate different aspects of the application. The first session, held on April 15th and 16th, 2024, assessed the application according to the standard MABC-2 protocol. The second session, which took place on May 16th, 2024, tested the interactive features of the application. Subsequent sections detail the methodologies, results, and the iterative process of refining the application based on feedback from these sessions.

8.1 First Test Overview

The primary objective of the first test was to evaluate whether the basic measuring components — projection, floor measurement with Pharus, and skeletal tracking — could independently replicate a standard MABC-2 test environment. This test was crucial for validating the integration of these technologies as standalone assessment tools in the application. Testing follows the session guideline from Appendix C. Before the test began, a brief familiarization with the child was conducted, and the child's age and laterality were written down. All legal representatives have signed a consent form (see Appendix E) in which they agreed to participate in testing their child and recording the data. It was explained to each child that several small games would be played. The test was conducted like the traditional MABC-2 tests, with a few key differences, based on Prototype from Section 6.2.:

- Instead of using the physical mats from the MABC-2 kit, the test utilized projected scenes that corresponded with the necessary test aspects, thereby fulfilling Functional Requirement 4 (FR4) as referenced in Section 5.2.1.
- In the Aim and Catch task (AC2), children who successfully threw the ball were immediately provided with visual feedback in the form of a `good job` image flag projected onto the ground.
- The line for the Balance task (BAL1) was shortened to fit within the projected area.
- The visualization of tracked entities and their corresponding echoes on the ground was kept, mainly for the observer to interfere, should the incorrect evaluation by the system arise.

Visually, the setup mirrored the standard MABC-2 testing environment.

8.1.1 Test setup

- Testing environment:
  An empty 2.30 x 3.80 area, as seen in diagram depicted in Figure 8.1. The testing room had very dim lighting. The prototype was projected on a carpet, that absorbed most of the reflected light.
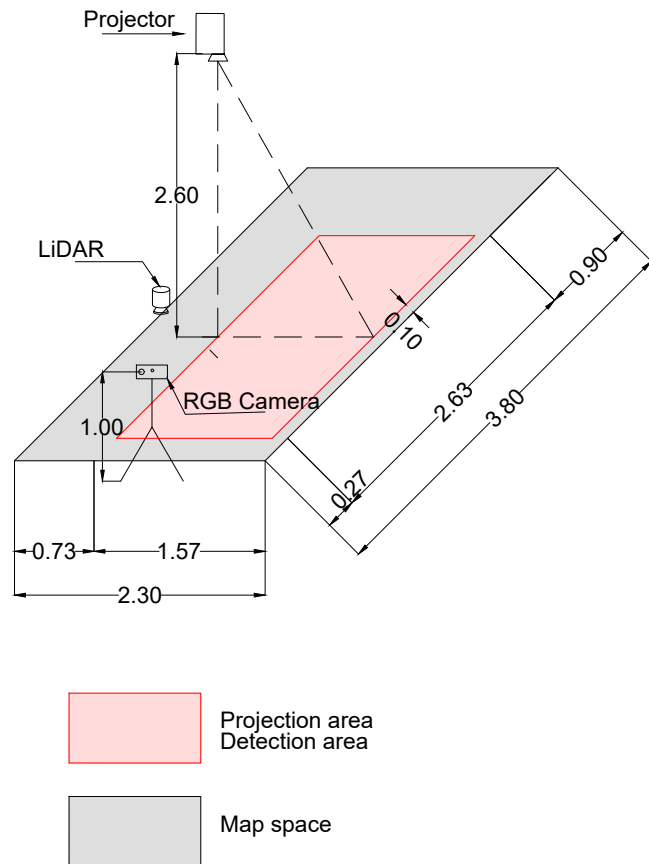
Figure 8.1. This diagram illustrates the initial testing setup, featuring a projector, an RGB camera, and LiDAR components.

- Number of participants: 3 (labelled as P1-P3)
- Used PC:
  - Model: MSI GP63 Leopard 8RE
  - CPU: Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz
  - GPU: GeForce® GTX 1060 with 6GB GDDR5, 1,280 CUDA Cores
  - OS: Windows 10 Home
  - RAM: 32 GB DDR4-2400
- Used RGB camera:
  - Model: Microsoft Azure Kinect
  - Field of view: 90° x 59°
  - Capturing mode: 1080p RGB at 30 FPS
- Used Projector:
  - Model: BenQ TK700STi
  - Technology: DLP
  - Native resolution: 3840 x 2160
  - Used resolution: 1920 x 1080
  - Luminosity: 3000 ANSI lm

8.1.2 Qualitative observations and Feedback:

User Feedback:

1. The third participant (P3) would have preferred a more interactive environment.

Observer Feedback:

1. Instructing a child while monitoring their performance and recording data in specific categories proved demanding. This could be due to the observer's inexperience in occupational therapy.
2. It is particularly challenging to evaluate the performance's qualitative and quantitative aspects and write them down on Testing sheets during the testing phase.
3. Children responded very positively when they could see their figures and footsteps (visualized trackingEntity and echoes) projected onto the ground.
4. Mediapipe strongly prefers adult detection. Although the child was standing in the initial T-pose, in front of the camera, if an adult came and sat even at the edge of the camera's detection, the model immediately retargeted to the adult, as seen in Figure 8.2, ignoring the detected child.

Figure 8.2. Pose detection, retargeted to sitting adult.

8.1.3 Main Outtakes

The results of the testing were discussed with Mgr. Štolová, who confirmed the observer's impressions and appreciated several aspects of the testing setup:

- Eliminating the need to repeatedly measure and set up tests is especially beneficial for patients with attention disorders, as it saves valuable time. This setup allows testing to proceed without delays, enabling a continuous focus on the tasks.
- Even without the use of advanced scanning methods, the mere simulation of the actual test using projection is usable.
- The fact that the testing sessions were recorded was highly valued for two main reasons:
  1. The possibility of reviewing the recordings confirms that managing and evaluating the test simultaneously is challenging, even for the therapist.
  2. Recordings provide a valuable resource for consultations with colleagues in cases of uncertainties about a child's performance during the tests.
- A good idea is to build an application that communicates with users in some manner. In therapeutic practice, therapists often encounter patients with autism spectrum disorders who may refuse to speak directly to therapists but are capable of communicating through tablet interfaces.

57

▪ This test confirmed that it is possible to simulate a selected part of the MABC-2 methodology without the need for physical test objects.

■ 8.1.4 Post-First Test Adjustments

▪ The measured histogram values were used in a subsequent implementation for step and heel raise detection.

▪ Due to the sensor performance, the detection areas were enlarged - in some cases, there was no correct detection due to detection offsets during the throw.

▪ The detection logic of the original prototype was used for the second test.

▪ Based on recommendations, it was decided to incorporate a voice interface for further testing, as mentioned in Section 7.2.1.

■ 8.2 Second Test Overview

The second round of testing tested the implemented interactive application from Chapter 7. The main objectives of the testing were:

1. To confirm or refute that the interactive methodology benefits real occupational therapy practice.

2. To confirm or refute that the tests can be fully automated.

The testing took place on the 17th of May in one of Prague's kindergartens. Twenty participants were expected, whose legal guardians were required to sign an Informed Consent Form (found in Appendix E) agreeing that their child would participate in the research. Unfortunately, for unknown reasons, only 7 people signed this Consent.

The test shares the same spatial changes in BAL1 and AC assessments, as in the First test (8.1)

■ 8.2.1 Test setup

▪ Testing environment: An empty 1.6 x 2.4 area, as seen in diagram depicted in Figure 8.3. The whole room was made of glossy linoleum with a thick pattern, as seen in Figure 8.4. The projection area had to be covered with a white cotton sheet to prevent light reflection into the sensors. The room could not even be partially darkened as windows were running the full length of the room from both directions. This was particularly reflected in the poor detection of BAL1, where daylight from the opposite window was reflected back into the RGB camera through this linoleum.
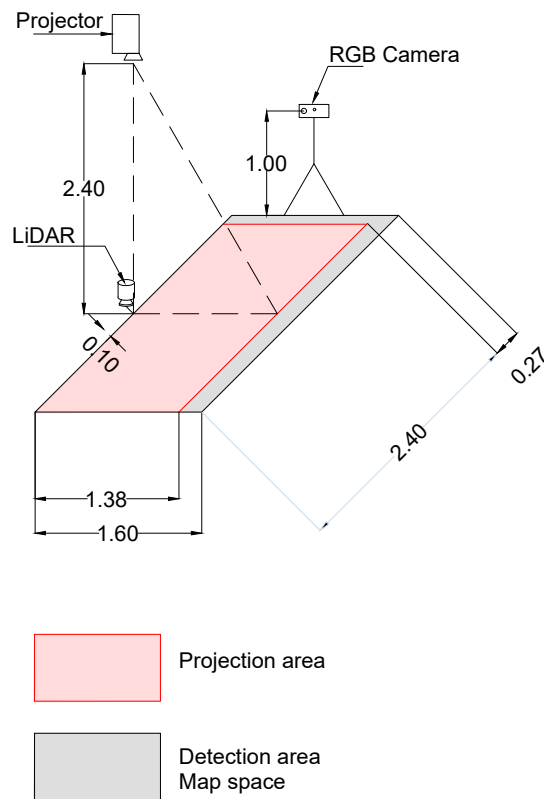
Figure 8.3. This diagram illustrates the final testing setup. The LiDAR is situated close to the projection area, resulting in more accurate detections.

Figure 8.4. The real testing environment.

59

- Number of participants: 7 (labeled as P4-P10)
- Used PC:
  - Model: MSI GP63 Leopard 8RE
  - CPU: Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz
  - GPU: GeForce® GTX 1060 with 6GB GDDR5, 1,280 CUDA Cores
  - OS: Windows 10 Home
  - RAM: 32 GB DDR4-2400
- Used RGB camera:
  - Model: Microsoft Azure Kinect
  - Field of view: 90° x 59°
  - Capturing mode: 1080p RGB at 30 FPS
- Used Projector:
  - Model: BenQ TK700STi
  - Technology: DLP
  - Native resolution: 3840 x 2160
  - Used resolution: 1920 x 1080
  - Luminosity: 3000 ANSI lm

▪ 8.2.2 Session Guideline

Testing follows the session guideline from Appendix D. Before the test began, a brief familiarization with the child was conducted, and the child's age and laterality were written down. It was explained to each child that several small games would be played. The participants were divided into two groups, as seen in Table 8.1.:

- Group A followed the standard methodology of the MABC2 test, i.e., the explanation of the exercises, the practice session, and the actual testing.
- Group B had only the testing part, without prior explanation or practice trials. This group, however, has been launched the verbal explanation (mentioned in Section 7.2.1) at the start of the tasks.

Participant	Group
P4	A
P5	B
P6	A
P7	B
P8	A
P9	B
P10	B

Table 8.1. Participants and category in which they were listed.

This was followed by a debriefing of the participants, gathering participants' insights, and asking a few questions from the questionnaire part of the Session Guide.

■ 8.2.3 Feedback

User Feedback:

1. The most popular interaction among most children was the shooting down of the AC2 planet, highlighted by participants P4, P5, P6, P8, and P10.
2. According to P7, one of the most enjoyable components was the tightrope walking in BAL2. This implementation is perceived as hilarious, mainly because of the tension perceived by the child and the visual aspect. Participants P8 and P10 confirmed this.
3. AC1 was confusing for most participants. All the participants from Group B noted this.

Observer Feedback:

1. The individual personality and developmental differences in the same age group were evident.
2. Boys generally had a more active and enthusiastic approach than girls. Some of them tried to test the limits of the application (stepping out of measurement zones in BAL2, expecting what could happen, for example).
3. Children reacted more positively to recorded instructions. This approach reduced the overall test duration but required more time for them to understand and complete the trial successfully. However, most felt a bigger sense of accomplishment when they did.
4. The task always begins on one side; however, in BAL2 and BAL3, almost all participants tried replicating the assessment from the end to the beginning by turning back after the first completion. After discovering that it does not react to them, they have adjusted to the testing logic, that the task always begins on one side.
5. The children from group B were hesitant to step inside the pulsing circle from AC1, one of them even refusing to do so without vigorous verbal encouragement. This might have been caused by initial distrust in stepping into the projected area.
6. Everyone in Group B performed their first attempt at the mats in BAL3 incorrectly-jumping over multiple tiles or not jumping at all. The system detected this. It was only the second jump that was correctly performed.
7. One of the participants (P6) wore glittery clothing and shoes, significantly interfering with the testing process. Due to the LiDAR's susceptibility to light, the intense lighting from the DLP projector, and the RGB image segmentation's sensitivity to lighting conditions, the glitters reflected light rays randomly. This led to numerous detection problems, such as a low detection rate and ghosting effects, where the system inaccurately indicated the child's position as nearby when they were distant.
8. During the testing, the part of detection using a pre-trained model seemed to fail most of the time of detecting a standing child, even though the accuracy during setup ranged around 60-70% (and was tested independently on three adult persons that were present during testing) and was least contributing to the BAL1 standing detection.

Personal Introspective:

1. Shortly before the beginning of testing, I encountered several unexpected technical complications and malfunctions that required the entire setup to be reconfigured at the last minute. This situation led to significant stress, as I had to rush to prepare the equipment and environment anew. The urgency and pressure affected me personally, causing a feeling of unease that persisted throughout the session.

2. Amid these technical issues and the resulting stress, I struggled with operational aspects that are typically routine. For instance, I repeatedly forgot the exact locations of the detected areas within the test setup, namely AC2, stepping in during testing and being identified a few times as an incorrectly thrown ball. Additionally, my ability to operate the recording camera effectively was compromised. These lapses in operational efficiency could have affected the quality of data collected, particularly in terms of visual documentation.

3. The testing schedule was tightly bound to the strict timetable of the kindergarten. Despite a significant delay in starting the tests, we were fortunate enough that the principal allowed us to extend and double the testing duration, accommodating a slight modification in the children's daily schedule. This flexibility enabled us to move forward as planned, though we still had to rush some parts of the testing process to fit within the new timeframe.

4. At the end of the test, the operator's tablet malfunctioned, and I was forced to use scene switching and recording via the UI, which was implemented as a rescue option—this slightly prolonged the last two participants' measurements.

■ 8.2.4 Qualitative Observations

The testing recordings were reviewed by Mgr. Radka Beranová, DiS., assistant of the Department of Occupational Therapy, Faculty of Health Studies, at Jan Evangelista Purkyně University, who uses the MABC-2 on a daily basis;

■ The therapist was surprised that all the children were timid and hesitant at the beginning of the test - in such young, healthy individuals, she would have expected a more positive attitude towards interactive, technology-based games.

 This may have been due to communication from the kindergarten staff, although the exact cause is unknown. The researchers did not know in advance what the children had been told before; in any case, they were selected individually, and the rest had to wait in the next room; against the researchers' wishes, the teacher still supervised them. For all participants except the P9, this mistrust subsided during the testing.

■ **P4** Throughout the testing, it was evident that this participant was an athletic individual and had excellent coordination. During the test, he was very active, intentionally stepping out of the measurement area to see how the application would respond, among other things, as seen in Figure 8.5.

 The participant is right-handed and very dominantly so. This was particularly noticeable during the AC circuit and especially when jumping on mats, where he did not use the pendulum motion of his arms, thus leading to overloading and an early landing on his right foot, as seen in Figure 8.6. However, the jump attempts can still be classified as successful. Although the results do not indicate motor difficulties, based on the recordings, the therapist would recommend compensation on the left side.

Figure 8.5. Participant 4, trying out the application's limits. The application warns him to return to the measurement area.

Figure 8.6. Participant 4, landing on his right foot first.

- **P5** During the evaluation of this patient, the therapist noted the importance of understanding the social background of each child — how often they are physically active or if they participate in any sports, for instance, according to Mgr.Beranová, this child exhibits signs of motor dyspraxia, even though she communicated and interacted well with the projection. The first difficulties were noticeable when standing on one leg. The child had problems managing a straight posture; at the same time, it could be caused by a change of focus point, as seen in Figure 8.7.
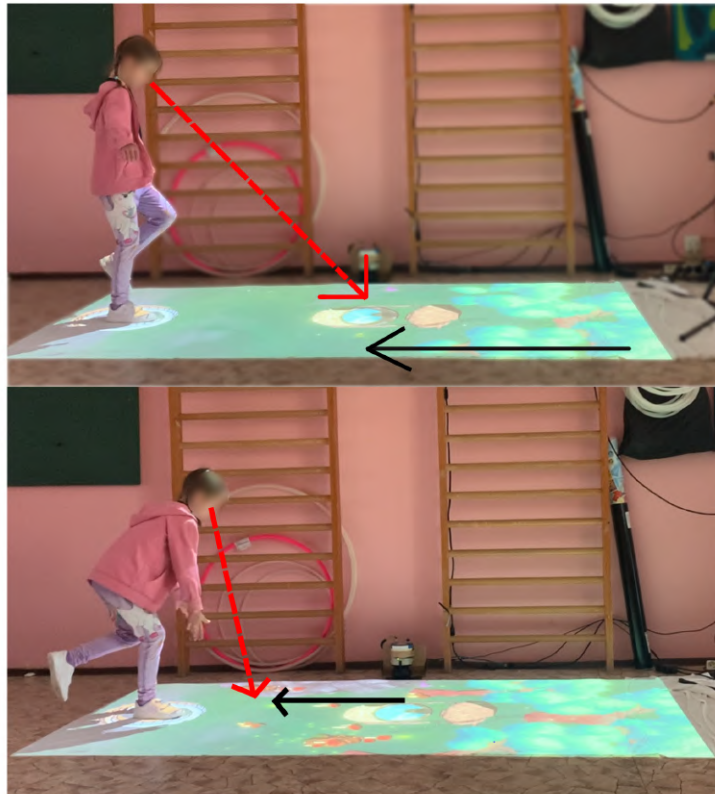
Figure 8.7. Participant 5, fixating its sight on the filling well (red dashed line) whilst the forest keeps growing towards the participant (horizontal black line). The child switches its focus point to the appearing forest, resulting in losing balance and falling over (bottom image).

The child's locomotor system showed signs of dyspraxia. Although an early landing on the right foot was detected in P4, a similar issue occurred with P5 during a two-footed jump. However, given the participant's previous performances in other assessments, these attempts would be considered incorrect. They are also rated incorrect because the movements completely omit upper body motion during the jump, as seen in Figure 8.8.

Figure 8.8. Participant 5, landing on his right foot first. Note that this landing was preceded by stiffness in the upper body, engaging only the legs during the maneuver.

- **P6** This participant does not exhibit significant signs of dyspraxia and, therefore, should be classified within the range of individuals without motor difficulties.
- **P7** Although the participant passed the entire test correctly, and the resulting test score would place him as an individual without motor difficulty range, the recordings show poor postural control.
- **P8** The footage shows that to collect the ball in AC2, the participant repeatedly goes around the entire projection area, trying not to step inside, as seen in Figure 8.9. This may have occurred, given that he stepped out of the circle several times while making catches in AC1 and receiving a warning message.

Figure 8.9. P8, avoiding contact with the projection surface.

This participant does not exhibit significant signs of dyspraxia and, therefore, should be classified within the range of individuals without motor difficulties.

- **P9** At the beginning of testing, this participant's behavior is timid, even frightened. As P9 belongs to group B, i.e., without prior explanation, she is invited to inspect the area - the first recorded verbal instruction is heard only after stepping into the circle in AC1.

While inspecting the AC1 area, she repeatedly circled most of the projection surface, carefully avoiding the circle. Her movement was highly stereotypical, repeating this pattern for over a minute and a half. Consequently, she had to be verbally prompted to enter the circle to continue.

Figure 8.10. P9, stereotypically circling around the area.

This participant completely ignores the computer-generated voice command and is fixated on the test performer during the whole duration of the testing. Also, she has no intrinsic sense of reward from the attempts made. E.g., when throwing in AC2, it is rewarded with an explosion, and the comment *great* (a well-received interaction component by other children) when hit correctly, the only affirmation had to come

from the tester; otherwise, the child would not continue with the test, as seen in Figure 8.11.

Figure 8.11. P9, expecting the verbal confirmation after each jump.

Although the overall performance places Participant P9 in the non-motor difficulty range, the therapist suggests some form of intervention based on her posture, hands clasped throughout the test, as seen in both Figures 8.10 and 8.11, reactions to instructions, and complete lack of involvement in the game, which is not normal for the given age group. She was described as bradypsychic and bradykinetic, terms that refer to slowed mental processing and reduced speed of physical movements, respectively.

▪ **P10.** Unlike the previous participant, P10 exhibits great openness in non-verbal communication and posture, as shown in Figure 8.12. Similarly to P4, P10 tests the limits of the system- namely the short registration time for a correct BAL1 attempt, where P10 repeatedly lifts her foot every second. The system evaluates this action as correct since it does not account for such rapid, repetitive interactions. This participant does not exhibit significant signs of dyspraxia and, therefore, should be classified within the range of individuals without motor difficulties.

Figure 8.12. Open body language with positive reaction to projection.

▪ 8.2.5 Main Outtakes

The results of the testing were discussed with Mgr. Radka Beranová, DiS., assistant of the Department of Occupational Therapy, Faculty of Health Studies, at Jan Evangelista Purkyně University.

- It is recognized that the AC1 is evaluated by a therapist.
- Using arrows to explain AC2 is confusing.
- In BAL1, the visual element of the growing forest is distracting, as seen in Figure 8.7. Although there was no change in the focus point (the spot the child concentrates on) for all participants except for P5, this represents a significant issue if it is meant to be an interventional method. A change in the position of the head and eyes leads to changes in the vestibular system, and balance exercises typically require focusing in a horizontal line of sight.
- A possible reason why all children from Group B always skipped over multiple ice tiles in their first attempt at BAL3 may lie in the coloring of the blocks, with one of the colors being red, which is associated with a forbidden zone or danger.
- The BAL3 assessment should be able to detect an early step with one foot, which is crucial in evaluating the correctness of the trial. However, given the hardware's performance, since this event is usually measured in milliseconds, it is currently impossible.
- Verbal guidance is an excellent asset - these commands are the same for everyone and are very strict. During testing, the therapist may introduce their own bias into the patient's guidance, losing the momentum of accurate explanation.
- Although it is common practice to start each attempt from a specific place, it is acceptable for balance exercises to have bilateral measurements.
- The training model is likely to perform very poorly. The model is learned on specific coordinates and trained on adults. The so-called *anthropometry*, i.e., the length of the bone distance measurements, works in children. Although the neural network can retarget to any skeleton, the body-to-bone ratio in children differs and must be considered during model training.
- Like Mgr. Štolová, Mgr. Beranová also confirmed that the assessment of children is done retrospectively from recordings. This approach is partly to prevent children from feeling that they are being evaluated and partly because it is time-consuming to do so otherwise.
- **A significant benefit is the motivation and gameplay.** During conventional testing, which lacks a game element, it is common for children to sabotage the testing because it is tedious and unengaging. Additionally, this motivation allows for observing other character traits in children. For instance, it helped identify several problematic personality traits in participant P9, which, although possibly noticeable in conventional testing, would not have been as evident.

8.2.6 Results

The recorded logged data, described in Section 7.3, can be found in the designated folder of the submission - either from the GitLab repository or in the submission package, as described further in Appendix G

In this section, interest was focused on the duration of the pure testing time per participant, excluding the briefing and debriefing periods.

The results shown in Table 8.2 and the graph shown in Figure 8.13 suggest that participants in Group B generally required less time to complete the tests than Group A participants. However, given the small sample size of only seven participants, it is impossible to conclude that Group B performed better than Group A. Instead, these results should indicate potential tendencies within the data rather than conclusive findings [72].

Participant	Group	Duration (minutes)
P4	A	11:55
P5	B	8:12
P6	A	11:37
P7	B	6:26
P8	A	10:24
P9	B	10:00
P10	B	7:04

Table 8.2. Total testing time required for each participant.
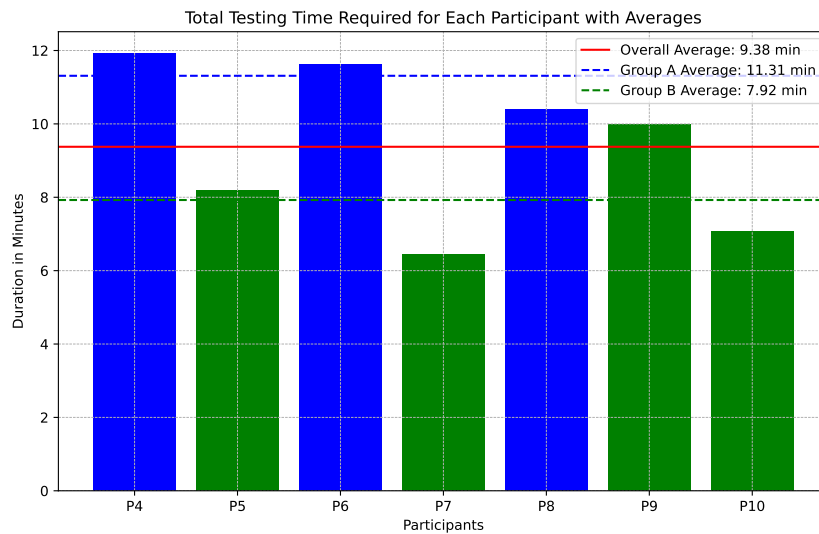
Figure 8.13. The bar graph illustrates the total testing times for each participant, segmented by groups. The dashed lines represent the average times for Groups A and B, highlighting the quicker completion tendency of Group B.

.

8.3 Summary

■ The greatest benefit identified was not the automation of the therapist's work but the integration of a gaming element into the testing process. The product cannot fully replace a therapist but is a valuable tool to enhance therapeutic practices, as referred in 8.2.

■ Early results show that participants new to the system might complete tasks quicker because of the streamlined and interactive setup, though more data is needed to draw firm conclusions.

■ Detection technology and user interface updates are needed to better handle different behaviors and environmental issues.

■ Responses to the interactive and automated features varied; some children tried to test the system's limits, while others were hesitant or confused without direct verbal instructions.

■ Children generally liked the visual feedback and interactive elements, which helped keep them engaged.

68

■ Using recorded instructions made the tests faster and smoother, though some children needed extra encouragement.

Chapter 9
Conclusion

The thesis addresses the integration of motion capture technologies in pediatric physiotherapy and ergotherapy, focusing on the application of markerless methods and attempting to showcase the fundamental ideas and methodologies through the development of simple prototypes, following the UCD principles, as previously mentioned in Chapter 8.

9.1 Chapter Summary

In Chapter 2, Method Research, common diagnostic methods in physiotherapy and ergotherapy were analyzed, focusing on their applicability and limitations in pediatric assessments. Chapter 3, Movement Capture Methods, explored a range of motion capture technologies, examining marker-based and markerless methods. In Chapter 4, the assessments from MABC-2 were analyzed in detail. It can be stated that some of the methods can be significantly abstracted, and their core ideas adapted into somewhat different concepts, namely the balance and aiming parts. However, due to the limitations mentioned in Section 4.3, it is not feasible to fully convert the complete MABC-2 methodology into a precise solution that could significantly expedite the therapist's work; implementing truly accurate systems would be way beyond this thesis's scope.

The application was developed, implementing AC and BAL assessments for the youngest user testing group using spatial LiDAR tracking, skeletal tracking with a machine learning model, and projection mapping to create a gamified test version in an interactive environment. In Chapter 8, two tests were conducted to determine whether using such technologies is beneficial and whether the methodology can be automated to the extent that a therapist is unnecessary.

Although this was not proven, on the contrary, testing showed that it is an excellent supplement for the therapist, which saves a significant amount of time. The average test duration, including setup and evaluation, is one hour; this solution takes about twenty minutes per test.

9.1.1 Requirements Fulfillment

The application successfully fulfilled several key functional requirements:

- **FR1 -** The application implemented part the MABC-2 method.
- **FR2 -** The application exclusively supports Aim-Catch Circle.
- **FR3 -** Projection technology has been utilized to display the testing method visually.
- **FR4 -** Ensured that the dimensions of projected objects in assessments were maintained accurately per the physical MABC-2 toolkit to adhere to ensure clinical validity.
- **FR5 -** The application records, manages, and stores motion data (Section 7.3).

70

- **NFR1 -** Immediate feedback and rewards are provided. This could possibly be elaborated further.
- **NFR2 -** The app design supported ongoing consultations with domain experts throughout development.
- **NFR3 -**It can be stated that the interface is visually appealing and engaging to maintain the attention of the user group.
- **NFR4 -** Implemented precise measurement capabilities to so that the distances and dimensions within the application reflect those of the physical objects used in traditional MABC-2 assessments.

9.2 Results

The greatest contribution of this work is that it brings up other important aspects and interventional cues into this part of the methodology that the therapist can look for, such as responses to game stimuli and the child's inner motivation to continue in the testing.

Furthermore, making the application interactive opens up possibilities for better communication with patients on the autism spectrum, who might otherwise have difficulty enduring the full length of testing or might not communicate at all. The same applies to patients with attention disorders.

Chapter 10
Follow-up work

This section discusses future ideas for improvements and potential continuation of the project. Since this thesis aimed to develop a simple application or prototype, this chapter looks at options for additional implementation and research.

10.1 Technical improvements

This section discusses how the existing solution could be technically enhanced in terms of the technologies used and the integration of new ideas.

10.1.1 Projection technology

In future iterations of this work, it would be advisable to move away from using a single-projector setup for a logical reason—orthographic projection leads to occlusion, resulting in the shadowing and covering of parts of the projection area, as seen in Figure 10.1. Additionally, due to the spatial limitations of most occupational therapy testing centers, creating a single-projector orthographic projection from a distance sufficient to project at least 4.5 meters, as required by BAL2, onto the ground is not feasible.

Figure 10.1. Occlusion of the projection area

This issue could be resolved using multiple synchronized projectors or by implementing projection onto a so-called LED floor. An LED floor, a surface that can be walked on and is covered with light-emitting diodes, can display any video content and likely mitigate the problem of occlusion and limited projection distance.

■ 10.1.2 Detection technology

Using just one LiDAR is unsufficient. For adequate scanning, at least two LiDARs are necessary to ensure continuous detection in the x and y axes without the risk of covering the detection area, similar to a projector setup. Naturally, LiDARs with a higher scanning frequency than 50 Hz would be more suitable for detection.

As evident from both user tests in Chapter 8, the trained models did not detect children. For future work, training the neural network on appropriate data, specifically from pediatric patients, would be highly beneficial. Additionally, there is a significant need to expand the dataset, as mentioned in the implementation section 7.1; the size of the testing dataset used in this study is inadequate. Detection could also be expanded beyond single-leg standing, such as the BAL1 exercise, to include activities like correctly catching a bag in AC1, where detecting object interaction with the body could be possible.

■ 10.1.3 Testing outtakes integration

Integrating insights from the second round of testing, as previously mentioned in Section 8.2.5, would be beneficial. Future updates should focus on improving how users interact with the application and how automated features work to make them quicker and more precise. These changes will help the application meet therapeutic needs more effectively. Also, adding advanced feedback mechanisms, as mentioned in the upcoming Section 10.2, would allow for adjustments during assessments depending on how participants behave. More development could include options for therapists to customize tests and more thorough testing with larger groups to gather more data and improve the application.

■ 10.2 Automation

This section discusses potential advancements in automating processes to speed up assessment procedures further.

10.2.1 Automated Form Fill

Figure 10.2. Automated Form Fill Template.

For this thesis, an output record form for the AB1 age group was created using Optical Character Recognition (OCR) and manual input with Adobe Acrobat Pro[1] from a scanned record sheet, as seen in Figure 10.2.1.

OCR technology converts various document types into editable and searchable formats, including scanned papers, PDF files, and digital images. The total test score for the MABC-2 test is calculated based on percentile tables, which vary every six months of a child's age, resulting in six different tables for the 3-6 year age group AB1. These tables have also been digitized and are available from a GitLab repository of this project[2]. Although there was an initial plan to automate the logging of variables from measurements directly into the test sheet using Python form filling [73], this was not implemented due to time constraints. However, this would be a straightforward interpretation that could significantly automate the testing process.

10.2.2 Better replay options

Another area for improvement would be creating a skeletal data playback system. The existing function provided by MediaPipe for plotting these points is inadequate. Alongside this player, and in response to therapists' requests, a player for standard camera recordings would be developed. This would allow therapists to retrospectively review the footage with annotated data, such as the recording date and segmentation into different scenes. If certain actions like a throw are detected, these would be displayed in the console. Additionally, the system would enable therapists to make notes directly on the video feed, similar to the functionality found in software like Kinovea 3.4.2.

---

[1] https://www.adobe.com/cz/acrobat/acrobat-pro.html
[2] https://gitlab.fel.cvut.cz/koudebar/dp_koudelkova_barbora

10.3   Possible Applications

This section discusses potential applications of integrating new methodologies in occupational therapy assessments and therapeutic environments, or transforming the ideas of them.

10.3.1   Real Usage

During ongoing consultations with the Department of Occupational Therapy at Jan Evangelista Purkyně University in Ústí nad Labem, a potential collaboration emerged that could involve integrating new methodologies as alternatives to parts of the standard test. However, one limitation is noted: the analysis only covered BAL and AC components, omitting the MD component for one age group - AB1. The testing includes three age groups, where the older groups would not only interact with projections but also use physical objects like balance mats and wall throws. Therefore, it is necessary to explore these options further.

10.3.2   Sensory Rooms

The potential application I envision is in transforming these into simple motion-based games with projections and basic angle/movement tracking to familiarize children with the tasks or as integration into so-called sensory rooms for dyspraxic and autistic patients. to familiarize children with the tasks. This idea builds on the work of the artistic collective Initi [3], which, in one of its projects, also utilizes LiDAR-based tracking to create simple interactive applications for children, such as throwing activities at interactive walls in museums and similar environments.

Sensory rooms are unique therapeutic spaces designed to develop and engage the sensory skills of individuals, particularly those with sensory processing disorders, autism, developmental disabilities, or dementia (as depicted in Figure 10.3). These rooms are carefully designed to provide a controlled environment that can either calm or stimulate the senses through light, color, sound, and tactile objects. The key features of sensory rooms include all kinds of different stimuli, such as various sounds, textures, lights, light tubes, and sensory projections.
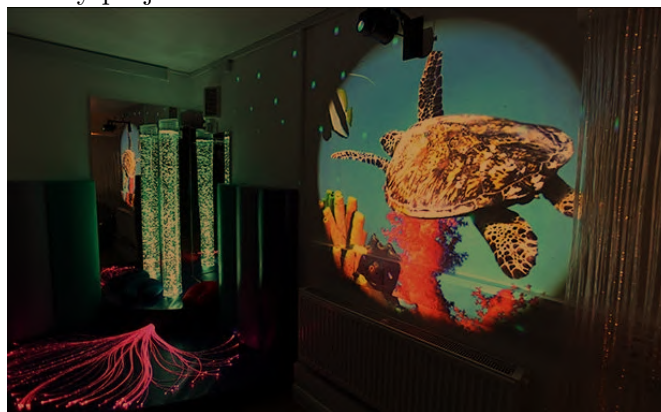
Figure 10.3. A sensory room at Brunel University. Image taken from [74]

Implementing the critical ideas of the MABC-2 and incorporating them into sensory interplay could be an exciting challenge.

---

[3] https://www.initi.org/

Chapter 11
Glossary

AC	■	Aim and Catch Circle
ADL	■	Activities of Daily Living
BAL	■	Balance Circle
BBT	■	The Box and Block Test
BVR	■	Biplanar Videoradiography
CSV	■	Comma-separated Value
DCD	■	Developmental Coordination Disorder
DNS	■	Dynamic Neuromuscular Stabilization
FIM	■	The Functional Independence Measure
FOV	■	Field of View
FPS	■	Frames Per Second
GDPR	■	General Data Protection Regulation
ISSS	■	Integrated Spinal Stabilizing System
LED	■	Light-Emitting Diode
LiDAR	■	Light Detection and Ranging
MABC-2	■	Movement Assessment Battery for Children, Second Edition
MD	■	Manual Dexterity Circle
PPT	■	Purdue Pegboard Test
RMSE	■	Root Mean Square Error
SDK	■	Software Development Kit
SMOTE	■	Synthetic Minority Over-sampling Technique
STA	■	Soft Tissue Artifacts
TTS	■	Text-To-Speech
UCD	■	User-centered design
UDP	■	User Datagram Protocol

References

[1] *Top 19 Applications Of Deep Learning and Computer Vision In Healthcare* [Online]. Available from `https://viso.ai/applications/computer-vision-in-healthcare/`. Accessed: 2024-05-12.

[2] NOROUZI-GHEIDARI, Nahid, Mindy F LEVIN, Joyce FUNG, and Philippe ARCHAMBAULT. Interactive virtual reality game-based rehabilitation for stroke patients. In: *2013 International Conference on Virtual Rehabilitation (ICVR)*. 2013. pp. 220–221.

[3] ADLAKHA, Shagun, Deepak CHHABRA, and Pratyoosh SHUKLA. Effectiveness of gamification for the rehabilitation of neurodegenerative disorders. *Chaos, Solitons & Fractals*. Elsevier, 2020, Vol. 140, pp. 110192.

[4] SPOLEČNOST, Česká Vojtova. *Vojtův princip – diagnostika a terapie*. Available from `https://vojtovametoda.com/vojtova-metoda/`. Accessed: 2023-12-21.

[5] *Rehabilitační léčba některých druhů funkční ženské sterility metodou Ludmily Mojžíšové* [Online]. Available from `https://www.levitas.cz/2018/02/rehabilitacni-lecba-nekterych-druhu-funkcni-zenske-sterility-metodou-ludmily-mojzisove/`. Accessed: 2023-12-21.

[6] *Dynamická neuromuskulární stabilizace (DNS)* [Online]. Available from `https://fyzioklinika.cz/poradna/clanky-o-zdravi/111-dynamicka-neuromuskularni-stabilizace-dns`. Accessed: 2023-12-21.

[7] JANDA, Z., and A. KOBESOVÁ. Tajemnství pohybové medicíny. *Téma* [Online]. 2011, pp. 18–20. Available from `https://www.rehabps.com/download/dns_pdf_tema.pdf`. Accessed: 2023-01-19.

[8] FRANK, Clare, Alena KOBESOVA, and Pavel KOLAR. Dynamic neuromuscular stabilization & sports rehabilitation. *International journal of sports physical therapy*. North American Sports Medicine Institute, 2013, Vol. 8, No. 1, pp. 62.

[9] *ergoterapie* [Online]. Available from `https://www.nzip.cz/rejstrikovy-pojem/1958`. Accessed: 2023-12-17.

[10] *Dyspraxia* [Online]. Available from `https://www.healthdirect.gov.au/dyspraxia`. Accessed: 2023-12-17.

[11] KOLAR, P, J SMRZOVA, and A KOBESOVA. Developmental coordination disorder-Developmental dyspraxia. *Ceska a Slovenska Neurologie a Neurochirurgie*. CZECH MEDICAL SOCIOLOGY, 2011, Vol. 74, No. 5, pp. 533–538.

[12] *How to Administer the Purdue Pegboard Test- Standardized Assessment* [Online]. Available from `https://youtu.be/CR33qxOHBck?si=jEp6iTlEzvZHls-I`.

[13] MARVIN, Katie. *Purdue Pegboard Test (PPT)* [Online]. Available from `https://strokengine.ca/en/assessments/purdue-pegboard-test-ppt/`. Accessed: 2023-12-17.

[14] RYBÁŘOVÁ, K., J. SÝKOROVÁ, L. MARKOVCOVÁ, and COL. *Česká rozšířená verze manuálu pro Box and Block Test (BBT)* [Online]. Available from `https://rehab ilitace.lf1.cuni.cz/file/372/bbt-albertov.pdf`. Accessed: 2023-12-17.

[15] HENDERSON, SE, DA SUGDEN, and AL BARNETT. edition 2. *Movement assessment battery for children. London, UK. The Psychological Corporation.* 2007.

[16] *MABC-2 - Test motoriky pro děti*. Available from `https://hogrefe.cz/mabc-2`. Accessed: 2023-12-17.

[17] VANASKOVA, Eva, and OTHERS. Testovanı v neurorehabilitaci. *Neurologie pro praxi.* 2005, Vol. 6, No. 6, pp. 311–314.

[18] KRIVOŠÍKOVÁ, M., and A. KRULOVÁ. *Funkční míra nezávislosti – měřící nástroj v posouzení soběstačnosti* [Online]. Available from `https://www.uzis.cz/res/ file/akce/20171107-klasifikon/2017-11-07-15_krulova.pdf`. Accessed: 2023-12-18.

[19] GRANGER, Carl V, Byron B HAMILTON, John M LINACRE, Allen W HEINEMANN, and Benjamin D WRIGHT. Performance profiles of the functional independence measure. *American journal of physical medicine & rehabilitation.* LWW, 1993, Vol. 72, No. 2, pp. 84–89.

[20] CORRIGAN, John D, Kip SMITH-KNAPP, and Carl V GRANGER. Validity of the functional independence measure for persons with traumatic brain injury. *Archives of physical medicine and rehabilitation.* Elsevier, 1997, Vol. 78, No. 8, pp. 828–834.

[21] STINEMAN, Margaret G, Richard N ROSS, Roger FIEDLER, Carl V GRANGER, and Greg MAISLIN. Functional independence staging: conceptual foundation, face validity, and empirical derivation. *Archives of physical medicine and rehabilitation.* Elsevier, 2003, Vol. 84, No. 1, pp. 29–37.

[22] TIMBECK, Rebecca J., and Sandi J. SPAULDING. Ability of the Functional Independence Measure™ to Predict Rehabilitation Outcomes After Stroke: A Review of the Literature. *Physical & Occupational Therapy In Geriatrics.* Taylor Francis, 2004, Vol. 22, No. 1, pp. 63-76. Available from DOI 10.1080/J148v22n01_04. Available from `https://doi.org/10.1080/J148v22n01_04` .

[23] SHU, Sherney. *Knee muscle interaction during gait* [Online]. Available from `http s://medium.com/@ruby09040502/knee-muscle-interaction-during-gait- d3438366502f`.

[24] PAYNE, Andrew. *Sagittal, frontal and transverse body planes: Exercises and Movements* [Online]. Available from `https://blog.nasm.org/exercise-programmi ng/sagittal-frontal-traverse-planes-explained-with-exercises`. Accessed: 2023-12-21.

[25] KUTCHA, Leslie. *Planes of motion: Sagittal, Frontal, transverse: Blog* [Online]. Available from `https://physiquedevelopment.com/planes-of-motion-sagitt al-frontal-transverse/`.

[26] ANDRIACCHI, Thomas P, Eugene J ALEXANDER, MK TONEY, C DYRBY, and Jeffrey a SUM. A point cluster method for in vivo motion analysis: applied to a study of knee kinematics. 1998.

[27] METCALF, CD, C PHILLIPS, A FORRESTER, J GLODOWSKI, K SIMPSON, C EVERITT, A DAREKAR, L KING, D WARWICK, and AS DICKINSON. Quantifying soft tissue artefacts and imaging variability in motion capture of the fingers. *Annals of biomedical engineering.* Springer, 2020, Vol. 48, pp. 1551–1561.

[28] MIRANDA, Daniel L, Michael J RAINBOW, Joseph J CRISCO, and Braden C FLEMING. Kinematic differences between optical motion capture and biplanar videoradiography during a jump–cut maneuver. *Journal of biomechanics*. Elsevier, 2013, Vol. 46, No. 3, pp. 567–573.

[29] KESSLER, Sarah E, Michael J RAINBOW, Glen A LICHTWARK, Andrew G CRESSWELL, Susan E D'ANDREA, Nicolai KONOW, and Luke A KELLY. A direct comparison of biplanar videoradiography and optical motion capture for foot and ankle kinematics. *Frontiers in bioengineering and biotechnology*. Frontiers Media SA, 2019, Vol. 7, pp. 199.

[30] MORO, Matteo, Giorgia MARCHESI, Filip HESSE, Francesca ODONE, and Maura CASADIO. Markerless vs. marker-based gait analysis: A proof of concept study. *Sensors*. MDPI, 2022, Vol. 22, No. 5, pp. 2011.

[31] *Vicon Documentation* [Online]. Available from `https://docs.vicon.com/pages/viewpage.action?pageId=107483982`. Accessed: 2023-12-21.

[32] CARSE, Bruce, Barry MEADOWS, Roy BOWERS, and Philip ROWE. Affordable clinical gait analysis: An assessment of the marker tracking accuracy of a new low-cost optical 3D motion analysis system. *Physiotherapy*. Elsevier, 2013, Vol. 99, No. 4, pp. 347–351.

[33] BELLUSCI, Giovanni, Daniel ROETENBERG, Fred DIJKSTRA, Henk LUINGE, and Per SLYCKE. Xsens MVN MotionGrid: Drift-free human motion tracking using tightly coupled ultra-wideband and miniature inertial sensors. *Xsens Technologies White Paper*. 2011, pp. 1–10.

[34] MOESLUND, T.B., A. HILTON, V. KRUGER, and L. SIGAL. *Visual Analysis of Humans: Looking at People*. Springer London, 2011. ISBN 9780857299963. Available from `https://books.google.cz/books?id=LFUOkgEACAAJ`.

[35] STRAKA, Matthias, Stefan HAUSWIESNER, Matthias RUTHER, and Horst BISCHOF. Simultaneous shape and pose adaption of articulated models using linear optimization. In: *European Conference on Computer Vision*. 2012. pp. 724–737.

[36] *How Mediapipe calculates the Z World Coordinate* [Online]. Available from `https://github.com/google/mediapipe/issues/5261`. Accessed: 2024-05-12.

[37] *Holistic landmarks detection task guide* [Online]. Available from `https://developers.google.com/mediapipe/solutions/vision/holistic_landmarker`. Accessed: 2024-05-12.

[38] SIVA, Loges. *An easy guide for pose estimation with Google's MediaPipe* [Online]. Available from `https://medium.com/mlearning-ai/an-easy-guide-for-pose-estimation-with-googles-mediapipe-a7962de0e944`.

[39] GUPTA, Amit, Pratisthit Lal SHRESTHA, Bhola THAPA, Raman SILWAL, and Raju SHRESTHA. Knee Flexion/Extension Angle Measurement for Gait Analysis Using Machine Learning Solution "MediaPipe Pose" and Its Comparison with Kinovea. In: *IOP Conference Series: Materials Science and Engineering*. 2023. pp. 012004.

[40] HII, Chang Soon Tony, Kok Beng GAN, Nasharuddin ZAINAL, Norlinah MOHAMED IBRAHIM, Shahrul AZMIN, Siti Hajar MAT DESA, Bart van de WARRENBURG, and Huay Woon YOU. Automated Gait Analysis Based on a Marker-Free Pose Estimation Model. *Sensors*. MDPI, 2023, Vol. 23, No. 14, pp. 6489.

[41] MENYCHTAS, Dimitrios, Nikolaos PETROU, Ioannis KANSIZOGLOU, Erasmia GIANNAKOU, Athanasios GREKIDIS, Antonios GASTERATOS, Vassilios GOURGOULIS,

Eleni Douda, Ilias Smilios, Maria Michalopoulou, and others. Gait analysis comparison between manual marking, 2D pose estimation algorithms, and 3D marker-based system. *Frontiers in Rehabilitation Sciences*. Frontiers Media SA, 2023, Vol. 4.

[42] Chen, Jason. *Theia 3D 2023.1 Documentation* [Online]. Available from `https://gizmodo.com/microsoft-xbox-360-kinect-launches-november-4-5563148`. Accessed: 2023-12-28.

[43] *Azure Kinect DK* [Online]. Available from `https://azure.microsoft.com/en-us/products/kinect-dk/`. Accessed: 2023-12-28.

[44] Cippitelli, Enea, Samuele Gasparrini, Susanna Spinsante, and Ennio Gambi. Kinect as a tool for gait analysis: validation of a real-time joint extraction algorithm working in side view. *Sensors*. MDPI, 2015, Vol. 15, No. 1, pp. 1417–1434.

[45] Dolatabadi, Elham, Babak Taati, and Alex Mihailidis. Concurrent validity of the Microsoft Kinect for Windows v2 for measuring spatiotemporal gait parameters. *Medical Engineering Physics*. 2016, Vol. 38, No. 9, pp. 952-958. ISSN 1350-4533. Available from DOI https://doi.org/10.1016/j.medengphy.2016.06.015. Available from `https://www.sciencedirect.com/science/article/pii/S1350453316301291`.

[46] Yeung, Ling-Fung, Zhenqun Yang, Kenneth Chik-Chi Cheng, Dan Du, and Raymond Kai-Yu Tong. Effects of camera viewing angles on tracking kinematic gait patterns using Azure Kinect, Kinect v2 and Orbbec Astra Pro v2. *Gait & posture*. Elsevier, 2021, Vol. 87, pp. 19–26.

[47] Gálai, Bence, and Csaba Benedek. Feature selection for Lidar-based gait recognition. In: *2015 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM)*. 2015. pp. 1-5.

[48] LeapMotion. *How Does the Leap Motion Controller Work?* Available from `https://medium.com/@LeapMotion/how-does-the-leap-motion-controller-work-9503124bfa04`. Accessed: 2023-11-15.

[49] Taylor, Travis S. *Introduction to laser science and engineering*. CRC Press, 2019.

[50] Harrap, Rob, and Matt Lato. An overview of LIDAR: collection to application. *NGI publication*. 2010, Vol. 2, pp. 1–9.

[51] Zheng, Li, and Tao Song. Gait analysis algorithm for lower limb rehabilitation robot applications. *Mechanical Sciences*. Copernicus Publications Gottingen, Germany, 2023, Vol. 14, No. 2, pp. 315–331.

[52] *Kinovea taekwondo* [Online]. Available from `https://youtu.be/CR33qxOHBck?si=jEp6iTlEzvZHls-I`.

[53] *Theia 3D 2023.1 Documentation* [Online]. Available from `https://www.theiamarkerless.ca/faqs/faqs.html##how-does-theia3d-work`. Accessed: 2023-12-21.

[54] Ito, Naoaki, Haraldur B Sigursson, Kayla D Seymore, Elanna K Arhos, Thomas S Buchanan, Lynn Snyder-Mackler, and Karin Gravare Silbernagel. Markerless motion capture: What clinician-scientists need to know right now. *JSAMS plus*. Elsevier, 2022, Vol. 1, pp. 100001.

[55] Wren, Tishya AL, Pavel Isakov, and Susan A Rethlefsen. Comparison of Kinematics Between Theia Markerless and Conventional Marker-Based Gait Analysis in Clinical Patients. *Gait & Posture*. Elsevier, 2023.

[56] Sahir, Sofiane. *Canny Edge Detection : Step by Step in Python* [Online]. [Accessed 12-05-2024.

[57] HADWIN, Kathryn J, Greg WOOD, Sally PAYNE, Christopher MACKINTOSH, and Johnny VV PARR. Strengths and weaknesses of the MABC-2 as a diagnostic tool for developmental coordination disorder: An online survey of occupational therapists and physiotherapists. *PLoS One*. Public Library of Science San Francisco, CA USA, 2023, Vol. 18, No. 6, pp. e0286751.

[58] MACINTYRE, Christine, and Kim MCVITTY. Movement and learning in the early years: supporting dyspraxia (DCD) and other difficulties. Sage, 2004.

[59] LAVIOLA JR, Joseph J, Ernst KRUIJFF, Ryan P MCMAHAN, Doug BOWMAN, and Ivan P POUPYREV. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[60] *Principy práce s dětmi v MŠ* [Online]. Available from `https://inkluzivniskola.cz/principy-prace-s-detmi-s-omj-v-ms`. Accessed 22-05-2024.

[61] ANGEL, Ed, and Dave SHREINER. Teaching a shader-based introduction to computer graphics. *IEEE Computer Graphics and Applications*. IEEE, 2011, Vol. 31, No. 2, pp. 9–13.

[62] OPENAI. *Assistance with Depth sensor output Image Description* [OpenAI Chatbot Assistance]. Available from `https://www.openai.com/chatgpt`. Conversation on [02.03.2024.

[63] INC., 3DiVi. *Nuitrack SDK* [Online]. Available from `https://github.com/3DiVi/nuitrack-sdk/tree/master`. Accessed: 2024-05-23.

[64] NADERER, Otto. Crowd tracking and movement pattern recognition/submitted by: Otto Naderer. 2015.

[65] KALTENBRUNNER, Martin, Till BOVERMANN, Ross BENCINA, Enrico COSTANZA, and OTHERS. TUIO: A protocol for table-top tangible user interfaces. In: *Proc. of the The 6th Int'l Workshop on Gesture in Human-Computer Interaction and Simulation*. 2005. pp. 1–5.

[66] CLARKE, David. *Symptoms, causes and treatment of bad posture* [Online]. Available from `https://www.thephysiocompany.com/blog/stop-slouching-postural-dysfunction-symptoms-causes-and-treatment-of-bad-posture`.

[67] LINDWURM, Eugen. *Deep learning, meet Clever Hans* [Online]. Available from `https://towardsdatascience.com/deep-learning-meet-clever-hans-3576144dc5a9`. Accessed: 2024-05-21.

[68] *Understand random forest algorithms with examples* [Online]. Available from `https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/` . Accessed: 2024-05-12.

[69] OPENAI. *Description of Model Performance on Unseen Data* [OpenAI Chatbot Assistance]. Provided guidance on evaluating machine learning model performance on unseen data.

[70] NAIR, Aashish. *Create Artificial Data With SMOTE* [Online]. Available from `https://towardsdatascience.com/create-artificial-data-with-smote-2a31ee855904`. Accessed 21-05-2024.

[71] OPENAI. *Description of Model Performance on Unseen Data Using SMOTE* [OpenAI Chatbot Assistance]. Provided guidance on evaluating machine learning model performance using SMOTE on unseen data.

[72] OPENAI. *Description of Bargraph* [OpenAI Chatbot Assistance]. Provided guidance on describing plotted values from bargraph.

[73] LIO. *Never Fill Out Another Boring Online Form. Use Python to make it automatic!* [Online]. Available from `https://softwaresennin.medium.com/never-fill-out-another-boring-online-form-use-python-to-make-it-automatic-e84afa6b066a`. [Accessed 22-05-2024.

[74] BUCHANNUN, Joe. *Experience five minutes of peace in Brunel's new Sensory Room* [Online]. Available from `https://www.brunel.ac.uk/news-and-events/news/articles/Experience-five-minutes-of-peace-in-Brunels-new-Sensory-Room`. Accessed: 2023-05-24.

Appendix A
Readme

This document provides instructions on how to set up and run the Unity application, Pharus software, and posture analysis Python scripts for the diploma thesis.

A.1 Unity Application Setup

- Ensure you have Unity version 2023.2.2f1 installed on your system.
- Clone the project repository from GitHub:
    git clone `https://gitlab.fel.cvut.cz/koudebar/dp_koudelkova_barbora.git`
- Navigate to the folder *MABC2*.
- Open the cloned project in Unity by selecting 'Open' from the Unity Hub and navigating to the project directory.
- To run the application within Unity, click the 'Play' button in the Unity editor.
- To run the application standalone:
    - Navigate to the folder *MABC2/Build*.
    - Run *mabc2.exe*

A.2 Pharus Software Configuration

- Install the Pharus software if not already installed.
- Configure and set up the sensors and space OR
- Navigate to the folder *Miscellaneous/PharusRecordings*, open terminal in folder, where Pharus.exe is, and type:
    ./pharus.exe `NAMEOFSCENE.`*rec*
- Configure the Pharus software to connect with the Unity application:
    - Please ensure the software is set and enabled to send data to the correct IP address and port expected by the DeepSpace application.
    - Do not close Pharus or terminal during Unity runtime

A.3 Installing and Running the Python Scripts

- Prerequisites:
    - Python 3.6 or higher
    - Libraries: MediaPipe, OpenCV, NumPy, scikit-learn, pickle
    - Install the libraries using pip: *pip install mediapipe opencv-python numpy scikit-learn*

83

- Clone the Python scripts from the repository:
  git clone `https://gitlab.fel.cvut.cz/koudebar/dp_koudelkova_barbora.git`
- Navigate to the script directory: *cd MediaPipescripts/UnityRuntimeDetection*
- Run the desired script:
  - *python PoseDetectorMABC2.py* - to start real-time pose detection and send data to Unity. This MUST be launched during runtime.
  - *python PoseClassificatorMABC2.py* - to collect data for model training.
  - *python TrainPoseClassificator.py* - to train or retrain the model.

A.4   Setup

- To calibrate space with the application correctly, in Unity Editor, set the prefab `TrackingReceiveHandler` to dimensions identical to the Pharus space setup and the desired resolution that will be projected.
- If part of the tracking area in Pharus is set to obstructed and the projection area is partial, change the `Canvas` prefab location to the corresponding coordinates of the projection area.

A.5   Troubleshooting

- Ensure all components (Unity, Pharus, Python) are correctly installed and configured.
- Check for software updates regularly to ensure compatibility and performance.
- For connectivity issues with Pharus software, verify network settings and server logs.
- If Python scripts fail to run, ensure all dependencies are installed, and the Python version is correct.

Appendix B
How to run prototypes

This document provides instructions on how to set up and run the posture analysis Python script.

B.1 Prerequisites

- Python 3.6 or higher
- MediaPipe
- OpenCV
- NumPy

B.2 Installation

1. Ensure Python is installed on your system.
2. Install the required libraries using pip:

    pip install mediapipe opencv-python numpy

B.3 Running the Script

1. Download the posture analysis scripts from the provided GitLab repository:
    git clone `https://gitlab.fel.cvut.cz/koudebar/dp_koudelkova_barbora.git`
2. Navigate to the directory where the script is located: *MediaPipeScripts/Prototypes*
3. Run the scripts using the following command:

    python DiafragmXPelvicFloor.py for running posture prototype 6.1 *python AnkleDiff.py* for running posture prototype 6.2

B.4 Using the Application

For the Protype Number One:

- Stand in front of the camera at a suitable distance to ensure your full upper body is visible.
- The script will display a window showing your live camera feed with posture analysis.
- Stand on your left side facing the camera, perform the posture as required by the assessment.
- The angles at the shoulder and pelvic regions will be calculated and displayed.
- To exit the application, press 'q' on your keyboard.

For the Protype Number Two:

- Stand in front of the camera at a suitable distance to ensure your full upper body is visible.
- The script will display a window showing your live camera feed with posture analysis and indicating START.
- Perform the posture as required by the MABC-2 BAL1 4.2.6 assessment.
- The BAL1 criteria will be calculated and displayed.
- To exit the application, press 'q' on your keyboard.

B.5   Troubleshooting

If you encounter any issues, ensure that:

- The camera is properly connected and configured on your computer.
- All the required Python libraries are installed.
- You have the latest version of the script from the GitLab repository.

Appendix **C**
Session Guide for 15.04.2024

Ověření použitelnosti aplikace implementující ergoterapeutické vyšetření MABC-2

Participant id ....   Datum...................

Věk.....................

Děkujeme Vám, že se zúčastníte studie o ověření použitelnosti aplikace implementující ergoterapeutické vyšetření MABC-2, studii zaměřené konkrétně na hodnocení hrubé motoriky a postřehu pomocí tohoto testu. Tato studie také testuje efektivitu bezmarkerového snímání a projekce v reálném čase. Cílem je zjistit, jak dobře děti zvládnou různé motorické úkoly a jak přesně můžeme tyto úkoly zaznamenávat pomocí našeho systému.

Struktura Testování:

1. Zahřívací Část:
   - Nejprve provedeme zahřívací část, kde vám ukážeme a vysvětlíme jednotlivé cviky, které budete během testu provádět.
   - Zaměříme se na základní pohyby jako jsou skoky, chůze po čáře a balanční cvičení.
2. Testovací Část:
   - Následuje testovací část, kde budete cvičení opakovat samostatně. Prosíme, abyste se snažili cvičení provádět co nejlépe podle instrukcí a věnovali pozornost přesnosti provedení.
   - Při provádění budeme sledovat vaše pohyby pomocí bezkontaktních metod, které zaznamenávají pouze informace o významných bodech (klouby) vašeho těla.
3. Sběr a Analýza Dat:
   - Vaše výkony budou zaznamenávány na papír, tak, jak je prováděno standardní měření s ergoterapeutem.
   - Současně bude systém automaticky zaznamenávat a analyzovat vaše pohyby pro porovnání s manuálním hodnocením.

Děkujeme za vaši účast a spolupráci v této studii!

Appendix **D**
Session Guide for 17.05.2024

Ověření použitelnosti aplikace implementující ergoterapeutické vyšetření MABC-2

Participant id ....   Datum...................

Věk....................   Lateralita.................

Skupina.............

Děkujeme Vám, že se zúčastníte studie o ověření použitelnosti aplikace implementující ergoterapeutické vyšetření MABC-2, studii zaměřené konkrétně na hodnocení hrubé motoriky a postřehu pomocí tohoto testu. Tato studie také testuje efektivitu bezmarkerového snímání a projekce v reálném čase. Cílem je zjistit, jak dobře děti zvládnou různé motorické úkoly a jak přesně můžeme tyto úkoly zaznamenávat pomocí našeho systému.

Struktura Testování:

V rámci tohoto uživatelského testování budeme zkoumat, jak děti reagují na instrukce a jak se dokážou samy orientovat v prostředí bez předchozího vysvětlení. Děti budou rozděleny do dvou skupin:

Skupina A: Této skupině budou nejprve vysvětleny a ukázány jednotlivé cviky během zahřívací části. Následně budou cviky opakovat podle předchozích instrukcí v testovací části.

Skupina B: Děti v této skupině přeskočí zahřívací část a bez předchozího vysvětlení budou provádět cviky pouze na základě projekce, která je bude vizuálně navádět.

1. **Zahřívací Část:**
   - Nejprve provedeme zahřívací část, kde vám ukážeme a vysvětlíme jednotlivé cviky, které budete během testu provádět.
   - Zaměříme se na základní pohyby jako jsou skoky, chůze po čáře a balanční cvičení.

2. **Testovací Část:**
   - Následuje testovací část, kde budete cvičení opakovat samostatně. Prosíme, abyste se snažili cvičení provádět co nejlépe podle instrukcí a věnovali pozornost přesnosti provedení.
   - Při provádění budeme sledovat vaše pohyby pomocí bezkontaktních metod, které zaznamenávají pouze informace o významných bodech (klouby) vašeho těla.

3. **Sběr a Analýza Dat:**
   - Vaše výkony budou zaznamenávány na papír, tak, jak je prováděno standardní měření s ergoterapeutem.
   - Současně bude systém automaticky zaznamenávat a analyzovat vaše pohyby pro porovnání s manuálním hodnocením.
   - Pro účely této studie budeme také provádět neveřejný kamerový záznam, který nám umožní přesněji analyzovat pohybové vzory
   - Poprosíme Vás o následné zodpovězení krátkých otázek

Děkujeme za vaši účast a spolupráci v této studii!

Post test

Prosíme Vás o vyplnění následujícího dotazníku. Jednotlivé body ve vztahují k Vaší celkové zkušenosti s testovaným systémem.

Jaká část testu se Vám nejvíce líbila? Proč?
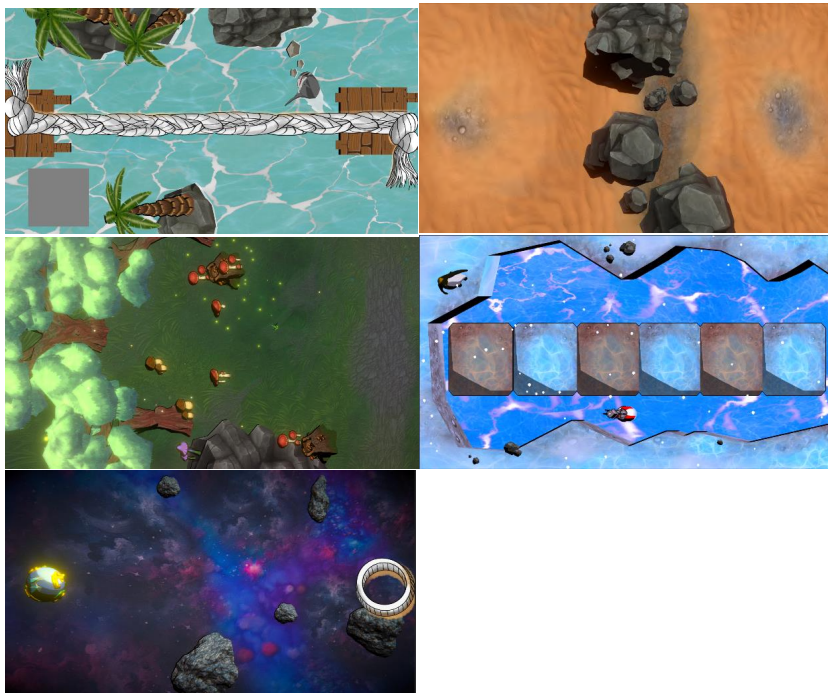
Je něco, co byste chtěl/a změnit na prostředí, ve kterém se testovalo, bylo pro Vás dostatečně interaktivní?

Měl/a jste pocit, že Vám nějaké cvičení nešlo? Dokázal/a byste ho specifikovat? Proč?

Měl/a jste pocit, že na Vás projekce reagovala? Jinými slovy – když jste dělal/a nějaké pohyby, tak se na podlaze něco stalo?

Ověření použitelnosti aplikace implementující MABC-2   13. – 17.5. 2024

Vyjádřete svou míru souhlasu s výroky týkajících se projekce.

	Naprosto souhlasím	Spíše souhlasím	Nevím	Spíše nesouhlasím	Naprosto nesouhlasím
a) Projekce je interaktivní – pořád se v ní něco děje, je na co se dívat.	1	2	3	4	5
b) Chápu přesně, co se po mě chce.	1	2	3	4	5
c) Nedokáži se soustředit na úkol, musím se stále dívat na projekci	1	2	3	4	5
d) Chtěl/a bych, aby mi někdo slovně vysvětlil, co mám dělat.	1	2	3	4	5

Post test - Hodnocení testujícího

	Naprosto souhlasím	Spíše souhlasím	Nevím	Spíše nesouhlasím	Naprosto nesouhlasím
a) Dítě se nedokáže soustředit na úkol, je rozptýlené	1	2	3	4	5
b) Dítě nechápe, co je po něm chtěno	1	2	3	4	5
c) Dítě vkládá do pokusu energii a jistou míru entuziasmu	1	2	3	4	5
d) Dítě se rychle adaptuje na změnu pokynů	1	2	3	4	5
e) Dítě dodržuje pokyny po celou dobu měření	1	2	3	4	5
f) Dítě má dobrou flexibilitu a rozsah pohybu	1	2	3	4	5
g) Dítě má dobrou koordinaci pohybů	1	2	3	4	5

Appendix E
Testing Consent Form

INFORMOVANÝ SOUHLAS S ÚČASTÍ VE VÝZKUMU

Název studie: **Ověření použitelnosti aplikace implementující ergoterapeutické vyšetření MABC-2**

Vedoucí výzkumník: **Ing. David Sedláček, Ph.D.**

Instituce: **ČVUT v Praze**

Jméno účastníka: ....

1. ÚČEL STUDIE

Cílem projektu je vyvinout a implementovat aplikaci s interaktivními herními prvky dle metodiky MABC-2, používanou v ergoterapii jako intervenční metoda pro určení možné dyspraxie dítěte. Hlavním cílem her bude ověřit přesnost a možnost využití moderních technologií projekce a bezmarkerového snímání pohybu pro částečnou automatizaci této metodiky, konkrétně její motorické části, čímž se zjednoduší a zrychlí příprava, kterou nyní ergoterapeuti vyžadují pro svou práci. Tento přístup poskytuje jedinečnou možnost učinit metodiku pro děti dostupnější a atraktivnější. Navržené řešení nejprve potřebujeme otestovat s dětmi bez významných omezení.

Žádáme Vás tímto, abyste jako zákonný zástupce dobrovolně umožnil(a) zapojení Vašeho dítěte do výzkumné studie. Tento formulář obsahuje informace o této studii. Rádi bychom se ujistili, že jste obeznámen(a) s účelem této studie a s tím, co pro Vás účast v této studii bude znamenat. Prosíme, nepodepisujte tento formulář, pokud nerozumíte některé části tohoto textu. V takovém případě se nás neváhejte na cokoliv zeptat.

Díky Vám a Vašim názorům budeme schopni upravit produkt tak, aby jeho použití bylo efektivní a snadné.

2. ÚČAST VE STUDII

Očekáváme, že se této studie zúčastní 10 až 15 osob. Kritérii pro výběr účastníků je věk účastníka, který musí být v okamžiku studie v rozsahu více než 4 až 7 let.

Účast v této studii je dobrovolná. Očekáváme, že se zúčastníte jednoho sezení, které se bude skládat z krátkého rozhovoru následovaného uživatelským testem, který není nikterak náročný, ani nevyžaduje jakoukoli přípravu ze strany uživatele. Celková doba sezení bude přibližně 20 minut.

3. ÚHRADA NÁKLADŮ

Veškeré výdaje spojené s Vaší účastí ve studii budou hrazeny z prostředků ČVUT. Malou pozornost od nás, jako výraz našich díků za Váš čas, dostanete i v případě, že se rozhodnete předčasně ukončit vaši účast ve studii.

4. UTAJENÍ ZÍSKANÝCH INFORMACÍ

Vaše odpovědi zaznamenáme do našich poznámek, které budou vedeny anonymně, a nijak z nich nebude možné zpětně identifikovat Vaši osobu. Informace budou zpracovány během analýzy získaných dat a objeví se ve výsledcích projektu, ale opět jen tak, že nebude možné dohledat, od koho jsme je získali.

V průběhu experimentu můžeme dělat zvukový nebo obrazový záznam pro potřeby pozdějšího vyhodnocení. Takto získaná data nebudou přímo publikována, ale poslouží jen pro potřeby členů výzkumného týmu související s tímto experimentem.

5. NÁHRADA MOŽNÝCH ŠKOD

Podpisem tohoto dokumentu se nezříkáte svých zákonných práv žádat odškodnění v případě úrazu, který byste utrpěl(a) v přímé souvislosti s touto studií.

6. DOBROVOLNÁ ÚČAST A AUTORIZACE

Vaše rozhodnutí zúčastnit se této studie je zcela dobrovolné.

7. ODSTOUPENÍ OD STUDIE ČI STAŽENÍ AUTORIZACE

Budete-li si to kdykoliv přát, můžete odstoupit od této studie bez jakýchkoliv negativních důsledků. Pokud se rozhodnete od studie odstoupit později (tj. nebudete chtít, abychom s Vašimi údaji dále pracovali), kontaktujte prosím písemně vedoucího výzkumníka:

- Ing. David Sedláček, Ph.D.
  Katedra počítačové grafiky a interakce, ČVUT FEL
  Karlovo náměstí 13, 121 35 Praha 2
  Tel.: +420-22435-7589

Od okamžiku Vašeho odstoupení nebudou již Vaše data nově zpracovávána v žádných dalších fázích výzkumného projektu. Nebude však možné měnit již existující publikované dokumenty či hotové a odevzdané výstupy z projektu.

8. KONTAKTNÍ OSOBY

Pro další informace o Vašich právech coby účastníka experimentu, či pokud nebudete spokojen(a) se způsobem, jak je tato studie provozována, můžete kontaktovat výzkumníka odpovídajícího za testování uživatelských rozhraní. Pokud máte libovolnou otázku či pokud utrpíte jakoukoliv újmu během studie, prosím kontaktujte vedoucího projektu:

- Ing. David Sedláček, Ph.D.
  Katedra počítačové grafiky a interakce, ČVUT FEL
  Karlovo náměstí 13, 121 35 Praha 2
  Tel.: +420-22435-7589

9. POTVRZENÍ

Zaškrtněte jednu z těchto variant

□ Přečetl(a) jsem si informace v tomto formuláři.
□ Informace v tomto formuláři mi přečetl: ....

Na veškeré otázky, které jsem měl, mi odpověděl: ....

Dobrovolně souhlasím s účastí na tomto výzkumu provedeném pracovníky ČVUT FEL. Obdržím podepsanou kopii tohoto formuláře.

Jméno a příjmení účastníka: ....

Jméno a příjmení zákonného zástupce:.......................................

Datum: ....

Podpis zákonného zástupce: ....

Jméno a příjmení osoby
přebírající tento souhlas: ....

Datum: ....

Podpis této osoby: ....

Appendix F
List of Assets Taken From Unity Asset Store

List of Asset Store Assets used in thesis:

- Simple Medievil Well - VIS GAMES
- Stylized Planet-Pack Full - ONE POTATO KINGDOM STUDIO
- Asteroids Pack - MARK DION
- Fish - PolyPack - ALASTRA INFINITE
- Imperial Penguin - FELIS CHAUS
- Simple Buoy - CROSS WALK STUDIOS
- Tree, Cartoon style - TEXTURE ME!
- Polygonal's Low-Poly Partition Effects - POLYGONAL STUFF
- Simple FX - Cartoon Particles - SYNTNY STUDIOS
- Hand Painted Textures - ALEX LUSTH
- Unlit Stylized Water Pack - ARTUR KOKOTOV
- Stylized Low Poly Rocks - JUSTCREATE
- Simple toon - DMITRYI CHALOVSKIY
- Deep Space 8K Development Kit - ARS ELECTRONICA FUTURELAB
- Grass And Flowers Pack 1 - VLADISLAV POCHEZHERTSEV
- Free Hut Pack - STORMBRINGER STUDIOS

Appendix **G**
Attachments

```
├─README.TXT..................................................................submission
├─src...............................................................source code directory
│       ├─ MABC2................................................................Unity project
│           ├─ Assets...........................................................Unity packages
│               ├─DeepSpace................Ars Electronica scripts and packages
│               ├─Scripts......................................Own scripts for application
│           ├─ Packages.............................................................Unity packages
│           ├─ Project settings................................................Build application
│           └─ Build................................................................Build application
│       ├─MediaPipePrototypes.......................................................Prototypes
│       └─ Runtime detection.....Scripts using for training and launching model and detection
├─ media............................................................................media directory
└─ text................................................................all textmedia directory
        ├─Forms.............................................................various forms, guidelines
        ├─Data..................................................................recorded testing data
        └─thesis.pdf....................................................................................thesis
```

95