

MAZE GAME ADJUSTMENT FOR BALANCE REHABILITATION USING
WIRELESS PRESSURE SENSORS INSOLES AND REINFORCEMENT
LEARNING



A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Telecommunication and Computer Engineering
Suranaree University of Technology
Academic Year 2024

การเรียนรู้แบบรีอินฟอร์สเมนต์เพื่อปรับระดับความยากของเกมในการฟื้นฟู



นายชัยเดช อังสุวรรณ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมโทรคมนาคมและคอมพิวเตอร์

มหาวิทยาลัยเทคโนโลยีสุรนารี

ปีการศึกษา 2567

MAZE GAME ADJUSTMENT FOR BALANCE REHABILITATION USING
WIRELESS PRESSURE SENSORS INSOLES AND REINFORCEMENT LEARNING

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's Degree.

Thesis Examining Committee



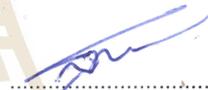
(Dr. Pattra Wattanapan, M.D.)

Chairperson



(Asst. Prof. Dr. Wipawee Usaha)

Member (Thesis Advisor)



(Assoc. Prof. Dr. Soodkhet Pojprapai)

Member (Thesis Co-advisor)



(Assoc. Prof. Dr. Piyaporn Mesawad)

Member



(Asst. Prof. Dr. Chitapong Wechtaisong)

Member



(Assoc. Prof. Dr. Yupaporn Ruksakulpiwat)

Vice Rector for Academic Affairs and
Quality Assurance



(Assoc. Prof. Dr. Pornsiri Jongkol)

Dean of Institute of Engineering

ชัชเดช อังสุวรรณ : การเรียนรู้แบบรีอินฟอร์สแมนท์เพื่อปรับระดับความยากของเกมในการ
พื้นฟู (MAZE GAME ADJUSTMENT FOR BALANCE REHABILITATION USING
WIRELESS PRESSURE SENSORS INSOLES AND REINFORCEMENT LEARNING)

อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.วิภาวี อูสาหะ , 60 หน้า.

คำสำคัญ : การฟื้นฟูสมรรถภาพ/จุดศูนย์กลางแรงกด/การเรียนรู้แบบรีอินฟอร์สแมนท์/ซีเรียสเกม/
แผ่นรองเท้าอัจฉริยะ

การหกล้มเป็นปัญหาสำคัญในผู้สูงอายุและผู้ที่มีข้อจำกัดด้านการเคลื่อนไหว ซึ่งมักนำไปสู่
อาการบาดเจ็บ ความหวาดกลัว และคุณภาพชีวิตที่ลดลง ภาวะสุขภาพบำบัดเป็นวิธีที่มีประสิทธิภาพใน
การช่วยเพิ่มความสมดุลและฟื้นฟูการเคลื่อนไหว อย่างไรก็ตาม วิธีการฟื้นฟูแบบดั้งเดิมอาจมีความ
ซ้ำซากและน่าเบื่อ ทำให้ผู้ป่วยขาดแรงจูงใจและไม่สามารถมีส่วนร่วมในกระบวนการบำบัดในระยะ
ยาวได้อย่างเต็มที่

เพื่อแก้ปัญหานี้ งานวิจัยนี้จึงพัฒนา เกมฟื้นฟูสมดุร่างกาย SuraSole maze ที่สามารถ
ปรับระดับความยากได้โดยอัตโนมัติตามความสามารถของผู้เล่น ด้วยการประยุกต์ใช้เทคนิค
Reinforcement learning (RL) ด้วยวิธี Q-learning ซึ่งเป็น Machine learning วิธีหนึ่งเกม
SuraSole maze นี้ถูกออกแบบมาเพื่อส่งเสริมการพัฒนาสมดุลของผู้เล่นผ่านการฝึกถ่ายน้ำหนักใน
รูปแบบของเกมเขาวงกต โดยใช้ข้อมูล COP (Center of pressure) จากแผ่นรองเท้าติดเซ็นเซอร์
ไร้สาย SuraSole ทั้งนี้ RL จะปรับระดับความยากของเกมให้เหมาะสมกับสถานะการทรงตัวของผู้
เล่น เกมนี้ถูกเลือกเพราะเกมประเภทเขาวงกตสามารถเพิ่มแรงจูงใจในการฟื้นฟูสมดุผ่านความสนุก
ความท้าทาย และความร่วมมือของผู้เล่น โดยแต่ละด่านมีระดับความยากที่ปรับเปลี่ยนได้เพื่อ
ส่งเสริมพัฒนาการของผู้เล่น

SuraSole maze นับว่าเป็น Serious games (SG) หนึ่งที่มีศักยภาพในการฟื้นฟูสมดุ
เนื่องจากสามารถเก็บข้อมูลเชิงกายภาพของผู้ป่วยในระหว่างการเล่นเกม โดยเฉพาะการใช้ เซ็นเซอร์
ไร้สายแบบสวมใส่ (Wearable sensors) เช่น SuraSole ซึ่งสามารถเก็บข้อมูลเกี่ยวกับการเคลื่อนไหว
และการทรงตัวแบบไดนามิก ข้อมูลจาก COP นี้ช่วยให้นักบำบัดสามารถวิเคราะห์สถานะของผู้ป่วย
และออกแบบโปรแกรมการฝึกที่เหมาะสมกับความต้องการเฉพาะบุคคล

การทดลอง ดำเนินการในสองระยะ ระยะแรกเป็นการฝึกโมเดล RL ด้วยข้อมูลจำลอง
(Simulated Data) เพื่อสร้างแบบจำลองผ่านการรันการจำลองด้วยภาษา Python โดยมีวัตถุประสงค์
เพื่อพัฒนานโยบายการเลือกระดับของเกมส์ที่เหมาะสมตามค่า COP ในเบื้องต้นก่อนการนำไปใช้

ทดสอบกับอาสาสมัครต่อไป ทั้งนี้โมเดลถูกฝึกให้เรียนรู้การปรับระดับความยากของด่านในสถานการณ์ต่าง ๆ เพื่อลด training time ของ RL ในการตัดสินใจของระบบ ระยะที่สองเป็นการทดลองกับอาสาสมัครสุขภาพดีจำนวน 15 คน อายุระหว่าง 21-44 ปี โดยให้ผู้เข้าร่วมเล่นด่านที่ออกแบบสำหรับผู้ที่สมรรถนะร่างกายไม่ดีเพื่อทดสอบความปลอดภัยของเกม และด่านที่ใช้ทดสอบการปรับระดับความยากด้วย RL ตัวชี้วัดของการที่บันทึกค่าพารามิเตอร์ที่เกี่ยวข้องได้แก่ RMS amplitude, RMS velocity, และ COP velocity เพื่อตรวจสอบการพัฒนาสมรรถนะก่อนและหลังการเล่นเกม รวมถึงการสำรวจความคิดเห็นเกี่ยวกับความพึงพอใจในการเล่น และความปลอดภัยของเกม

ผลการทดลอง แสดงให้เห็นว่าเกม SuraSole maze สามารถเพิ่มคะแนนสมดุล (Balance score) โดยรวมจากค่าเฉลี่ยก่อนเล่น 89.6 ± 9.23 เป็น 95.5 ± 1.6 หลังการเล่น จากคะแนนเต็ม 100 ซึ่งมีความแตกต่างอย่างมีนัยสำคัญ ($p = 0.0161$) และยังลดค่า RMS amplitude ในทิศทาง Medio-lateral และ Anterior-posterior จาก 0.12 ± 0.07 m เป็น 0.08 ± 0.06 m และจาก 0.070 ± 0.036 m เป็น 0.068 ± 0.036 m ตามลำดับ รวมถึงลดค่า RMS velocity ในทิศทาง Medio-lateral และ Anterior-posterior จาก 0.043 ± 0.046 m/s เป็น 0.020 ± 0.010 m/s ตามลำดับ และค่า COP velocity โดยรวมลดลงจาก 0.0013 ± 0.0011 m/s เป็น 0.0007 ± 0.0003 m/s ซึ่งสะท้อนถึงการพัฒนากายในการทรงตัวแบบองค์รวม แบบสอบถามหลังการทดลองระบุว่า 75% ของผู้เข้าร่วมพึงพอใจกับหลักการของเกม และ 96% เห็นด้วยว่าเกมนี้มีความปลอดภัย

ผลลัพธ์เหล่านี้ชี้ให้เห็นว่าเกม SuraSole maze มีศักยภาพในการเป็นเครื่องมือบำบัดความสมดุล โดยช่วยส่งเสริมการฟื้นฟูและอาจจะช่วยลดความเสี่ยงในการล้มของผู้สูงอายุหรือผู้ที่มีปัญหาการทรงตัว ทั้งนี้ในอนาคตควรขยายการทดลองกับกลุ่มตัวอย่างที่มีปัญหาด้านสมดุลที่หลากหลาย เพื่อเพิ่มความครอบคลุมและพัฒนาระบบให้เหมาะสมยิ่งขึ้น

สาขาวิชา วิศวกรรมโทรคมนาคม

ปีการศึกษา 2567

ลายมือชื่อนักศึกษา โจนาห์ อังควาน

ลายมือชื่ออาจารย์ที่ปรึกษา [ลายมือ]

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม [ลายมือ]

CHAIDECH ANGSUWAN: MAZE GAME ADJUSTMENT FOR BALANCE
REHABILITATION USING WIRELESS PRESSURE SENSORS INSOLES AND
REINFORCEMENT LEARNING.

THESIS ADVISOR: ASST. PROF. DR. WIPAWEE USAHA, 60 PP.

Keyword: Rehabilitation/Center of pressure/Reinforcement learning/Serious
games/Smart insoles

Falls are a major concern among the elderly and those with mobility limitations, often leading to injuries, fear, and reduced quality of life. Physical therapy is effective for improving balance and restoring mobility. However, conventional rehabilitation methods can be repetitive and demotivating, reducing patient engagement in long-term therapy.

To address this issue, the SuraSole maze game was developed as a balance rehabilitation tool capable of automatically adjusting its difficulty levels based on the player's abilities using Reinforcement learning (RL) with Q-learning. The SuraSole maze game was designed to promote balance improvement through weight-shifting exercises in a maze format, utilizing Center of pressure (COP) data collected from SuraSole wireless sensor insoles. The RL system dynamically adjusts the game's difficulty level to match the player's balance status. This maze-based approach was chosen because such games can enhance motivation through fun, challenge, and engagement, with each level being adaptable to encourage player progression

SuraSole maze is considered to be Serious games (SG) for balance rehabilitation, which enables the collection of patient data during gameplay. Wearable sensors like SuraSole are particularly effective in capturing dynamic movement and balance data. The COP data allows therapists to analyze the patient's condition and design personalized training programs tailored to individual needs.

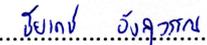
The experiment was conducted in two phases. The first phase involved training the RL model using simulated data with Python to develop an initial policy for selecting appropriate game levels based on COP values. This step aimed to optimize the model's ability to adjust the difficulty levels of game stages under different scenarios while reducing the training time of the RL algorithm. In the second phase, the experiment was conducted with 15 healthy volunteers aged 21–44 years. Participants played levels designed for individuals with poor balance to test the game's safety and the RL-based adaptive difficulty system. COP data, including RMS amplitude, RMS velocity, and COP velocity, were recorded before and during gameplay to assess balance improvements. Participant feedback on satisfaction and safety was also collected.

The experimental results showed that the SuraSole maze game improved the Balance score from a pre-play average of 89.6 ± 9.23 to 95.5 ± 1.6 post-play from a scale of 100, with a statistically significant difference ($p = 0.0161$). The RMS amplitude decreased in the medio-lateral and anterior-posterior directions from 0.12 ± 0.07 m to 0.08 ± 0.06 m and 0.070 ± 0.036 m to 0.068 ± 0.036 m, respectively. The RMS velocity decreased in the medio-lateral and anterior-posterior directions from 0.043 ± 0.046 m/s to 0.020 ± 0.010 m/s. The COP velocity also decreased from 0.0013 ± 0.0011 m/s to 0.0007 ± 0.0003 m/s, reflecting an overall improvement in balance control. Post-experiment surveys revealed that 75% of participants were satisfied with the game principles, and 96% agreed that the game was safe for use.

These findings indicate that the SuraSole maze game has the potential to serve as an effective balance rehabilitation tool, promoting recovery and possibly reducing fall risks among the elderly or individuals with balance impairments. Future studies should involve a more diverse sample of participants with varying balance issues to enhance the system's applicability and further refine its effectiveness.

School of Telecommunication Engineering

Academic Year 2024

Student's Signature 

Advisor's Signature 

Co-advisor's Signature 

ACKNOWLEDGEMENT

This thesis has been successfully completed through the profound grace of my parents. I would like to express my deepest gratitude to them for giving me life, for their unconditional love, care, and warm support throughout my journey. This success is a reflection of their unwavering pride and encouragement.

I am sincerely grateful to the individuals and groups who have greatly contributed to the academic and research aspects of this work. My heartfelt thanks go to Asst. Prof. Dr. Wipawee Usaha, for her invaluable knowledge, continuous guidance, life advice, and constant encouragement throughout the research process. I would also like to express my deep appreciation to Asst. Prof. Dr. Soodkhet Pojprapai, for his support in problem-solving and thoughtful research guidance, as well as his mentorship in academic and personal growth.

I extend my thanks to the faculty members of the Telecommunication Engineering Program, School of Engineering, Suranaree University of Technology, for their valuable teaching and advice throughout my academic years. I am also thankful to the staff of the Scientific and Technological Instruments Center, Suranaree University of Technology, for their kind assistance and technical support.

Special thanks go to Mr. Adisak Rattananamlom, Dr. Thita Sonklin, Dr. Mongkol Kongtungmon, all my siblings, friends, and fellow companions, for their advice, motivation, and lasting friendship throughout this journey.

Lastly, I am deeply thankful to all the teachers, past and present, who have imparted knowledge across many fields, and to my beloved parents and family members who have supported me and nurtured me with great care. Their belief in me helped me overcome all difficulties. I dedicate all the merits and achievements of this thesis to my parents, teachers, and all beloved relatives who have been my inspiration and strength.

Chaidech Angsuwan

TABLE OF CONTENTS

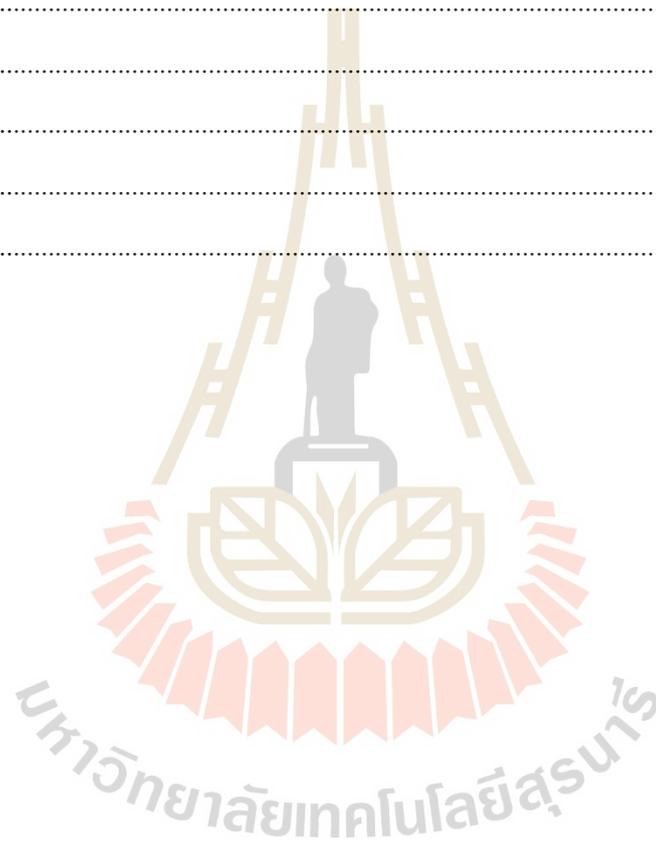
	Page
ABSTRACT (THAI)	I
ABSTRACT (ENGLISH)	III
ACKNOWLEDGEMENT	V
TABLE OF CONTENTS	VI
LIST OF TABLES	IX
LIST OF FIGURES	X
LIST OF ABBREVIATIONS	XII
CHAPTER	
I INTRODUCTION	1
1.1 Introduction	1
1.2 Objectives	6
1.3 Contributions	6
1.4 Scope of research	6
1.5 Hypotheses	6
1.6 Expected Benefits	7
II LITERATURE REVIEW	8
2.1 Introduction	8
2.2 SuraSole	9
2.3 SuraSole maze game	10
2.4 Reinforcement learning algorithm	12
2.5 Literature of RL for difficulty adjustment in rehabilitation games	12
2.6 Balance ability assessment	15

TABLE OF CONTENTS (Continued)

	Page
2.7 Summary	19
III RESEARCH METHODOLOGY	20
3.1 Introduction	20
3.2 Research design	21
3.3 Participants	21
3.4 Research tools	21
3.5 Experimental procedures	21
3.5.1 Phase 1: Development of RL-based game adjustment ..	21
3.5.2 Phase 1: Game concept development.....	22
3.5.3 Phase 2: Experimental evaluation with participants	23
3.6 Data analysis.....	26
3.7 Summary	26
IV RESULTS	27
4.1 Introduction	27
4.2 Individual balance metrics	28
4.3 Balance score	30
4.4 Average reward	33
4.5 Player Satisfaction	34
4.5.1 Patient survey	35
4.5.2 Game principles	35
4.5.3 Game safety	35
4.6 Summary	36
V CONCLUSIONS	37

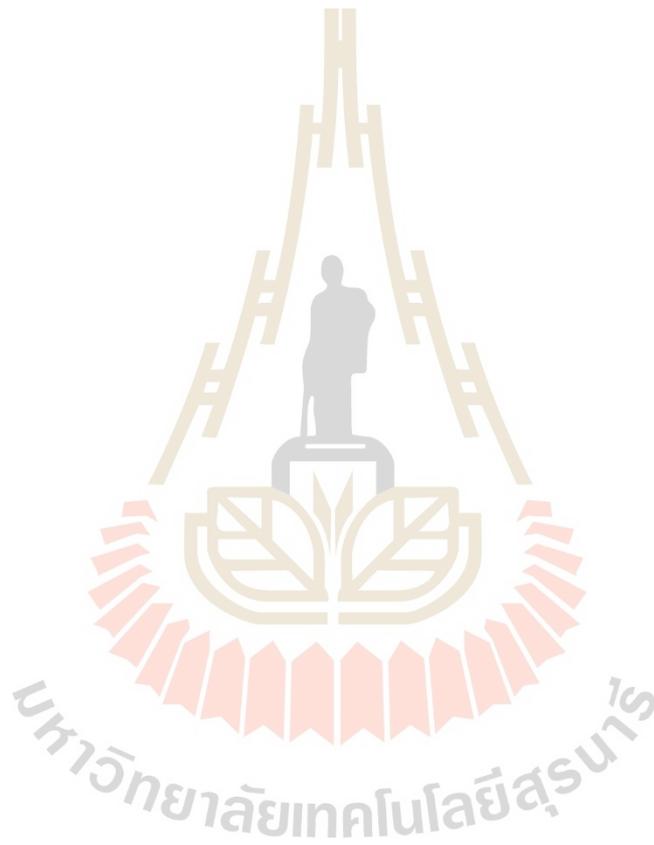
TABLE OF CONTENTS (Continued)

	Page
5.1 Conclusion	37
5.2 Future work	38
REFERENCES	40
APPENDIX A	45
APPENDIX B	47
APPENDIX C	58
BIOGRAPHY	60



LIST OF TABLES

Table	Page
1.1.1 Comparison of core aspects in this research	5



LIST OF FIGURES

Figure		Page
1.1.1	(a) Illustration of the data acquisition of center of pressure using pressure-sensing insoles as the measurement device (b) Sample of COP while standing.....	3
1.1.2	Nintendo Wii Balance Board (WBB).....	4
2.2.1	SuraSole smart insole.....	9
2.2.2	Sensor points distributed across specific foot zones.....	10
2.3.1	SuraSole maze game.....	11
3.5.1	The 9 maze levels of COP in Quadrant 3.....	22
3.5.2	Training maze game.....	23
3.5.3	RL process for game level selection based on COP data.....	25
3.5.4	Game testing with participants wearing SuraSole insoles.....	25
4.2.1	Average RMS velocity in the X and Y axes before and after gameplay.....	29
4.2.2	Average RMS amplitude in the X and Y axes before and after gameplay.....	29
4.2.3	Average COP velocity before and after gameplay.....	30
4.3.1	Balance score of the participants.....	31
4.4.1	Relationship between the number of gameplay rounds and the average reward of participants.....	33
4.5.1	Analysis of feedback from game assessment in each category.....	34
B.1.1	Division of quadrants in various states.....	47
B.1.2	COP states and RL-selected levels for Quadrant 3, (a) SuraSole maze game interface showing COP in Quadrant 3 before the game (b) SuraSole maze game level for a sample COP in Quadrant 3.....	48

LIST OF FIGURES (Continued)

Figure	Page
B.1.3 COP States and RL-selected Levels for Quadrant 5, (a) SuraSole maze game interface showing COP in Quadrant 5 before the game, (b) SuraSole maze game level for a sample COP in Quadrant 5.....	49
B.1.4 SuraSole maze game level for a sample COP in Quadrant 1	49
B.1.5 SuraSole maze game level for a sample COP in Quadrant 2	50
B.1.6 SuraSole maze game level for a sample COP in Quadrant 4.....	50
B.2.1 Block diagram of the user interaction with the game.....	51
B.3.1 State transitions of Quadrant 1 to various states	52
B.3.2 State transitions of Quadrant 5 to various states	53
B.4.1 State transitions of maze level F in Quadrant 1 to various states	53
B.4.2 Number of action selections for each state.....	54
B.4.3 Sample of average reward of the RL model depicting consistent improvement	55
B.4.4 Q-values for state 1 or Quadrant 1	55
B.4.5 Q-values for state 2 or Quadrant 2	56
B.4.6 Q-values for state 3 or Quadrant 3	56
B.4.7 Q-values for state 4 or Quadrant 4	57
B.4.8 Q-values for state 5 or Quadrant 5	57

LIST OF ABBREVIATIONS

SG	=	Serious games
RL	=	Reinforcement learning
COP	=	Center of pressure
VR	=	Virtual reality
WBB	=	Wii Balance Board
ADC	=	Analog-to-digital converter
FSR	=	Force sensitive resistor
ASD	=	Autism spectrum disorder
ML	=	Medio-lateral direction
AP	=	Anterior-posterior direction
<i>Mah</i>	=	Mahalanobis distance
<i>LST</i>	=	Lower scoring threshold
α	=	Learning rate
γ	=	Discount factor
MDP	=	Markov decision process
<i>s</i>	=	State
<i>s'</i>	=	Next state
<i>a</i>	=	Action
<i>a'</i>	=	Next action
<i>r</i>	=	Reward
<i>Q</i>	=	Q-value

CHAPTER I

INTRODUCTION

1.1 Introduction

Falls are a common issue among the elderly, individuals with decreased independence, and those with physical limitations (Danielsen et al., 2016). The global population aged 65 and older is projected to exceed 1.5 billion by 2050, making the risks and complications associated with falls a significant concern (United Nations, 2020). Falls can result in severe injuries, increased activity limitations, heightened fear of falling, depression, and a reduced quality of life. Additionally, medical costs related to falls are substantial, estimated at around \$50 billion in 2015 (Florence et al., 2018).

Studies have shown that physical therapy can improve balance, prevent falls, and enhance independence among the elderly (Thomas et al., 2010). Physical therapy is, therefore, an effective intervention to help older adults regain lost abilities and return to independent living. However, traditional rehabilitation activities are often repetitive, leading to boredom and a lack of motivation among patients to engage in physical therapy (Teruel et al., 2019).

Previous research has found that patients who enjoy rehabilitation activities are more likely to spend time on them (Hocine and Gouaich, 2011). This suggests that designing rehabilitation activities in the form of games has the potential to enhance the rehabilitation process by creating an engaging environment that motivates patients (Sekhavat, 2017). Serious games (SG) have been specifically designed for rehabilitation purposes and are widely used due to their ability to attract players and effectively support the rehabilitation process (Maheu-Cadotte et al., 2021).

The difficulty level of rehabilitation games can usually be typically determined by therapists based on the initial statistical profile of the player, or by the scores derived from the in-game achievements at the end of the game. However, data collected from the player's physical condition during the game play may provide

insights to the balance impairments of the player. Such data input for SG can be obtained from sensor devices, such as force pressure plates (Riedmann, 2022; Baranyi, 2013), wearable sensor devices (Agrawal et al., 2023) or virtual reality (VR) (Mariselvam et al., 2023) hardware systems. While VR systems provide immersive experiences, the costs and system effort-demanding set up may pose challenges to implement. On the other hand, force pressure plates may be less costly, however, the static set up may limit the dynamics of players' movements. On the other hand, wearable sensors, such as wireless pressure sensors embedded in insoles (Agrawal et al., 2023), can offer a more natural degree of freedom for movements and enable insights of dynamics of the player's balance impairment (Agrawal et al., 2023).

Furthermore, pressure sensors from force plates or wireless insoles can provide insights to the Center of pressure (COP) of the player. The COP is the point where all forces act on a surface, resulting from the combination of forces exerted by the body on the ground. The COP plays a critical role in maintaining balance and preventing falls during standing or walking. Furthermore, COP serves as an indicator of pressure distribution on the surface and is commonly used in the analysis of balance and movement. Measuring COP helps enhance walking efficiency, reduce the risk of fall-related injuries, and promote greater stability and safety in walking (Physiopedia, 2023). The measurement of COP is conducted while the participants stand upright on a pressure sensor. The device records the pressure exerted on the surface in all directions. This data is used to analyze balance and body stability.

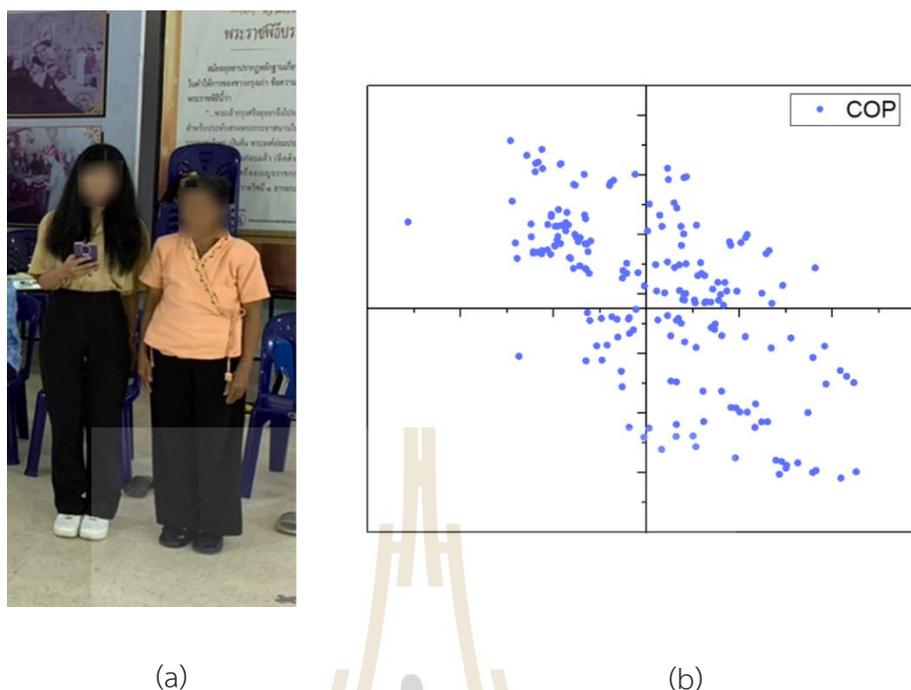


Figure 1.1.1 (a) Illustration of the data acquisition of center of pressure using pressure-sensing insoles as the measurement device (b) Sample of COP while standing

In (Riedmann, 2022), the Nintendo Wii Balance Board (WBB) is a widely used low-cost force plate to measure the COP of the player to feed as input to a SG for balance impairment treatment. In particular, the CoP data was used to adjust the difficulty according to the player's balance ability and enhance the player's experience in a personalized manner. To evaluate the player's balance ability, a heuristic method was used with baseline data from healthy players. The difficulty of each game level has been based on the physiotherapists requirements. However, the game dynamically adjusts the game level difficulty based on the user performance. Since level generation is parameterized by patient balance scores, it relies on the patient's balance score obtained from the heuristic algorithm to quantify the patient's balance ability, to yield meaningful results. The heuristic model is a problem-solving approach that emphasizes speed by using predefined rules or principles to find a "good enough" solution in situations where achieving a precise answer would require excessive resources. Although it does not guarantee the best possible outcome, heuristics can provide suitable solutions within a limited timeframe (Beheshti and Shamsuddin, 2013).



Figure 1.1.2 Nintendo Wii Balance Board (WBB) (Nintendo., 2025)

In (Riedmann, 2022), the proposed SG called Walk in the Park is an extension of a previous rehabilitation SG called RehabLabyrinth (Baranyi et al., 2013). While RehabLabyrinth utilized static game levels that require manual selection by therapists or physicians, Walk in the Park introduced automatic adaptive game levels. These adaptive levels are designed to boost patient motivation and facilitate intensive and repetitive exercises, making them more effective compared to the original SG.

Recently, the concept of Reinforcement learning (RL), a branch of artificial intelligence, has been introduced to promote rehabilitation and engagement. RL learns through trial-and-error interactions with its environment, receiving feedback to refine decision-making strategies over time. This ability sets RL apart from traditional Machine learning, which focuses on processing pre-existing datasets. RL, on the other hand, can adapt and evolve continuously based on real-world scenarios (Jayaraman et al., 2024).

RL has been applied in various fields, including healthcare, due to its ability to adapt and learn from complex environments. Its capacity to learn continuously from data can significantly improve clinical trial outcomes and healthcare strategies (Jayaraman et al., 2024). For instance, research on personalized chemotherapy dosing (Zhao et al., 2009; Ahn and Park, 2011; Ebrahimi Zade et al., 2020) utilizes RL to optimize the amount of chemotherapy tailored to individual patients. Another example is stroke rehabilitation, where RL is used with robotic arms equipped with two-degrees-of-freedom (DoFs) to facilitate arm rehabilitation. These robots monitor exercise methods to enhance therapeutic efficiency and ensure patient safety during exercise. This approach not only increases patient engagement but also accelerates the rehabilitation process effectively (Choutri et al., 2023).

In the context of balance rehabilitation, RL has been proposed to adjust game difficulty based on patients' skills and abilities. RL can significantly enhance the effectiveness of SG for physical rehabilitation for patients by providing personalized and adaptive training experiences (Hornak et al., 2019). This ensures maximum efficiency in the rehabilitation process through dynamic difficulty adjustment. Specifically, RL algorithms can dynamically modify the difficulty level of tasks based on the patient's performance. If the patient performs well, the game increases the challenge to promote progress conversely, if the patient struggles, the game reduces the difficulty to lower barriers. RL can also generate personalized feedback for each patient, focusing on areas requiring improvement. This capability further enhances the efficiency of the rehabilitation process (de Oliveira Andrade et al., 2014).

In this thesis, we therefore develop a game level difficulty adjustment method customized for each user based on RL using the user's COP directly obtained during the game. The SG is based on the SuraSole maze game (Suratec, 2024), which is designed by physiotherapists for balance impairment treatment. The SuraSole maze game is based on the patient's COP obtained from SuraSole wearable wireless pressure insoles to gather real-time patient movements, instead of a static balance board as in (Riedmann, 2022; Baranyi, 2013). The patient's sensor COP data is fed into the RL algorithm, which then updates the game environment and tasks based on this information. In particular, the RL technique called Q-learning (see Appendix B), is used to learn policies that determine the best actions (game level) to maximize patient progress. Table 1.1.1 compares the features of the SGs referred to in this research work. The objectives and contributions of this thesis proposal are as follows.

	RehaLabyrinth [10]	Walk in the park [9]	Surasole Maze	This_Thesis
Gamification elements	✓	✓	✓	✓
COP Calibration	✓	✓	✓	✓
Balance evaluation	X	✓	X	✓
Performance feedback	✓	✓	✓	✓
Difficulty adaptation	X	✓	X	✓
Method of difficulty adaptation	X	Heuristics	X	Reinforcement learning
Equipment				

Table 1.1.1 Comparison of core aspects in this research

1.2 Objectives

1.2.1 To develop a serious game based on reinforcement learning that can adjust to the participant's Center of pressure (COP) to suggest balance training game levels for each player.

1.2.2 To test the proposed game on healthy subjects to assess their COP, gait parameters, engagement, and safety of the serious game.

1.3 Contributions

1.3.1 Development of a prototype SG for balance impairment rehabilitation, designed as a future framework that utilizes adaptive difficulty adjustment within the game based on patients' abilities, leveraging reinforcement learning to support effective balance recovery tailored to individual needs.

1.3.2 Use of COP collected in during the serious game from the wireless pressure sensor insoles.

1.4 Scope of research

1.4.1 Develop a game difficulty adjustment system for balance rehabilitation using Reinforcement learning with simulation and data collected from wireless pressure sensor insoles (SuraSole).

1.4.2 Test the game on healthy participants to evaluate balance, safety, and engagement in gameplay.

1.4.3 Develop computational processes and analysis methods for Center of Pressure (COP) data to adjust the game levels to suit individual players.

1.5 Hypotheses

1.5.1 Using Reinforcement learning for game difficulty adjustment can improve players' balance.

1.5.2 COP data collected from wireless pressure sensor insoles can be effectively processed to adjust game difficulty based on players' abilities.

1.5.3 Players will be more engaged and find the game enjoyable when difficulty levels are dynamically adjusted according to their skills.

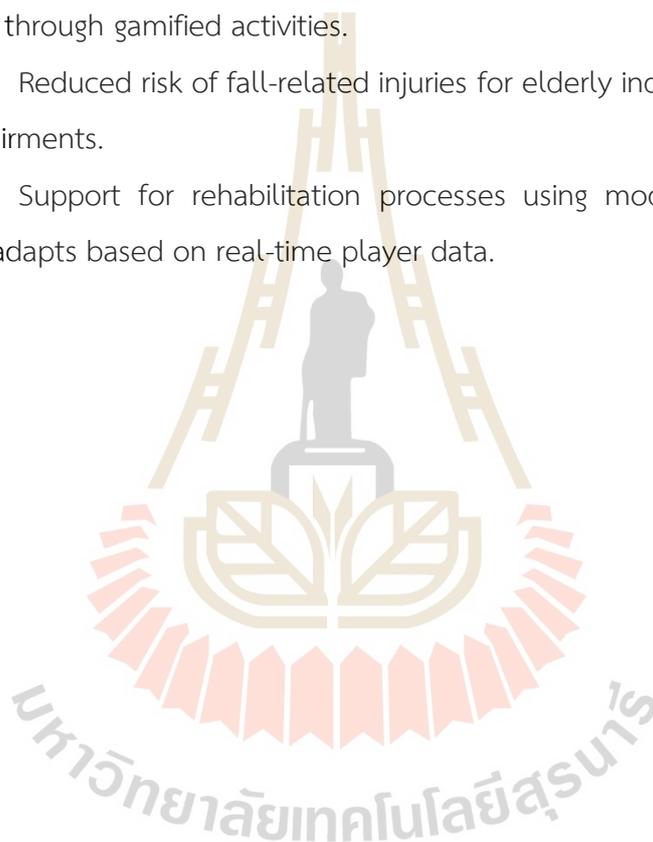
1.6 Expected Benefits

1.6.1 A new tool for balance rehabilitation that adapts game difficulty based on players' COP.

1.6.2 Increased motivation and enjoyment for individuals undergoing balance rehabilitation through gamified activities.

1.6.3 Reduced risk of fall-related injuries for elderly individuals or those with balance impairments.

1.6.4 Support for rehabilitation processes using modern technology that dynamically adapts based on real-time player data.



CHAPTER II

LITERATURE REVIEW

2.1 Introduction

The development of games for balance rehabilitation has gained increasing attention in recent years due to their potential to enhance the effectiveness of therapy and motivate patients. Previous research has developed various approaches to integrate gaming technology into rehabilitation processes to improve balance and movement. One of the widely recognized techniques is the application of Reinforcement learning (RL), which can enhance the effectiveness of Serious games (SG) for physical rehabilitation (Hornak et al., 2019). RL's ability to automatically adjust difficulty levels ensures that each training session is tailored to the patient's ability (Sekhavat, 2012; Mariselvam et al., 2023; Khabbaz et al., 2023).

In the context of balance rehabilitation, related studies emphasize the importance of gaming technology that can dynamically adapt to the unique characteristics of individual players. This study adopts these concepts to develop the SuraSole maze game, which uses RL to adjust game difficulty levels based on Center of pressure (COP) data measured from SuraSole insoles. The goal is to support effective balance rehabilitation while evaluating the game's impact on motivation and balance control capabilities among diverse target groups.

This chapter is divided into three main sections:

- 1) Technology of SuraSole Insoles and SuraSole maze game: Section 2.2-2.3 focus on the structure and functionality of SuraSole and the mechanisms of the SuraSole maze game.

- 2) Concept of RL: Section 2.4 explains the principles of RL and how it is applied in rehabilitation games.
- 3) Calculation of balance parameters: Section 2.5 describes methods for measuring and analyzing various parameters related to balance assessment.

2.2 SuraSole

SuraSole is a smart insole based on wearable sensor technology and the Internet of Things. The SuraSole sole is designed to help doctors, physical therapists and patients by providing home-based rehabilitation walking and balance impairment physical therapy, which can monitor the development of physical therapy without having to come to the hospital. This helps save the cost of imported equipment and travel and face-to-face treatment time for physiotherapy at the hospital. SuraSole enables balance training information by transmitting information via mobile phones or computer to the hospital or doctors and physiotherapists.

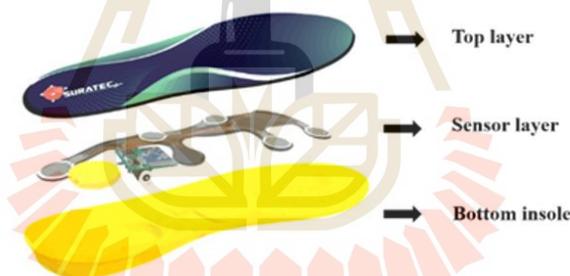


Figure 2.2.1 SuraSole smart insole

The insole system utilized in this study was embedded with multiple Force Sensitive Resistor (FSR) 8 sensors, each measuring 18 mm in diameter, integrated within both insoles. These sensors were strategically positioned to monitor variations in pressure. The collected data were segmented into five distinct zones: the hallux (HA1), fourth toe (HA2), medial forefoot (M1), central forefoot (M3), lateral forefoot (M5), midfoot (MF), lateral foot (LF) and heel (HF). The FSR sensors were connected to a microcontroller through a voltage divider configuration, with the output processed by a 10-bit Analog-to-Digital Converter (ADC). The system was calibrated to measure forces within a range of 0 to 196.2 N (0 to 20 kg), ensuring a response time of less than 10 microseconds. The sensors operated at a sampling frequency of 20 Hz, transmitting

real-time data via Bluetooth to a smartphone, where the information was subsequently uploaded to a database server for analysis.

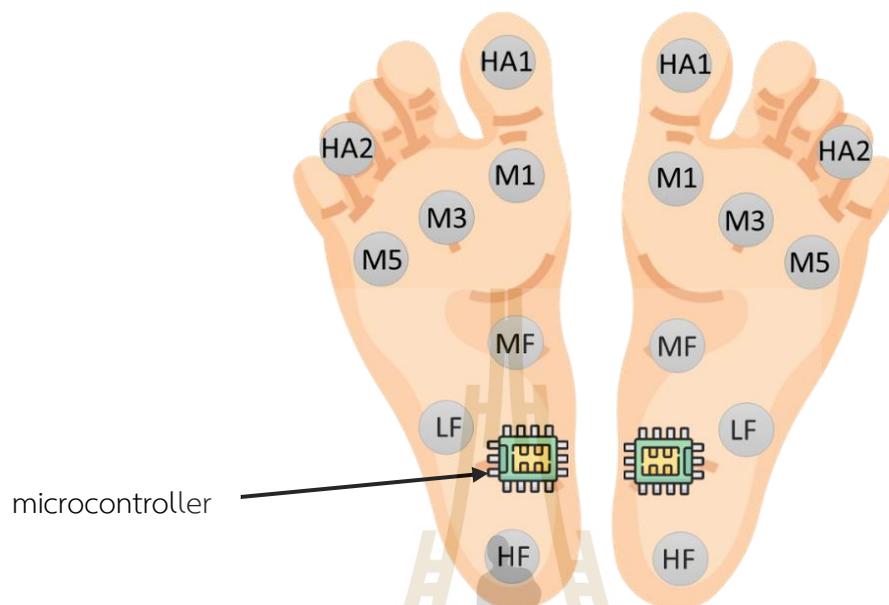


Figure 2.2.2 Sensor points distributed across specific foot zones

2.3 SuraSole maze game

The SuraSole maze game is a rehabilitation and physical therapy SG gaming application that helps patients improve their balance and weight transfer abilities (Suratec, 2024). Developed by physiotherapists, the game shown in figure 2.3.1 (a)–(g) offers an engaging experience and tracks progress through scoring, fostering motivation and user satisfaction. The objective of the game is to navigate a ball through a maze, shown in Figure 2.3.1 (f), created from obstacles to reach the finish line. Additionally, players must collect bonus points to score. Each level is timed, with the clock starting when the player moves and stopping when they reach the finish line. Once the player reaches the finish line, the game concludes.

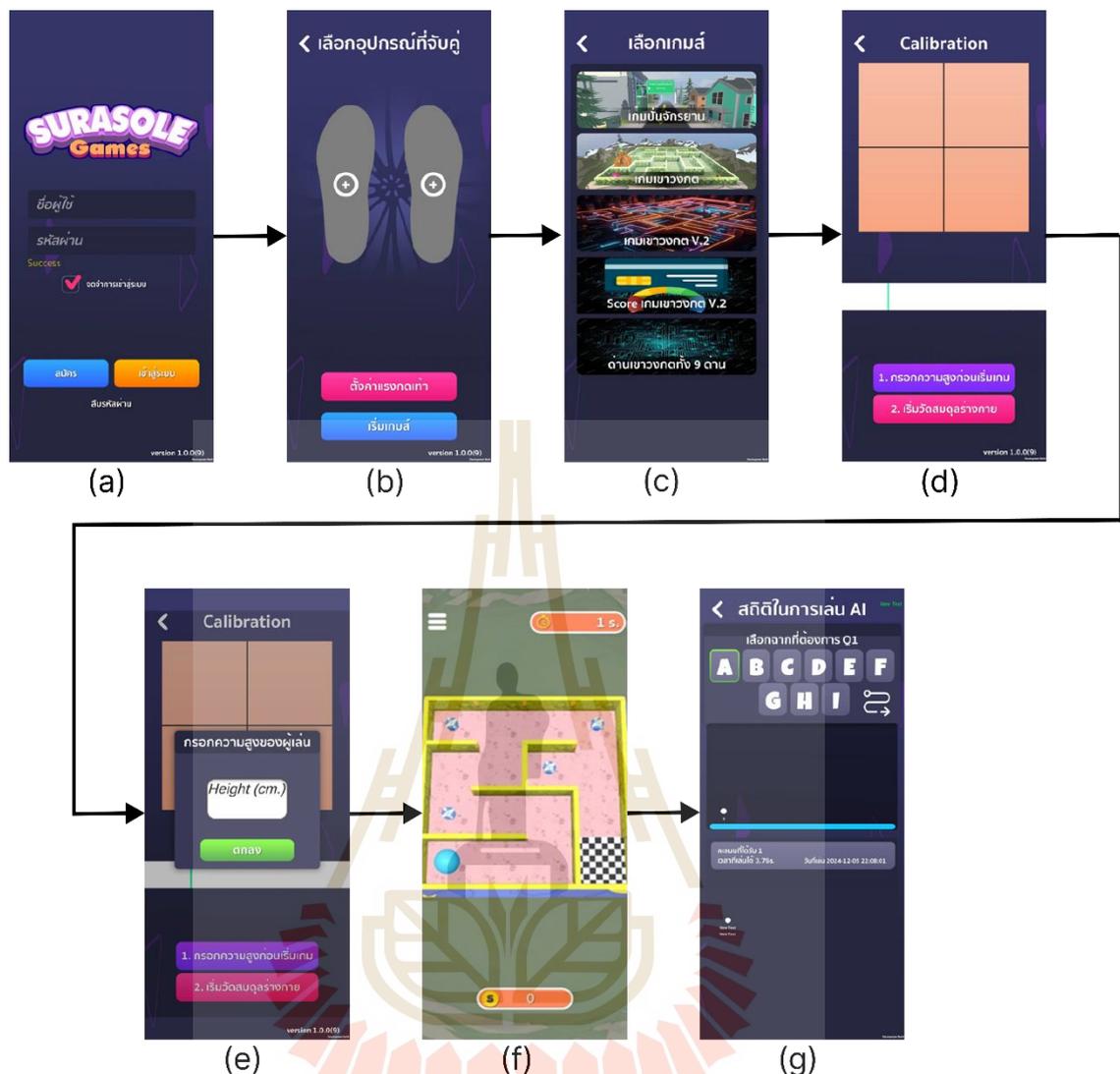


Figure 2.3.1 SuraSole maze game

In the process of balance rehabilitation, patients often attempt to compensate for their weakest points of COP by shifting more weight to the stronger side that is unaffected. Such actions can negatively impact the outcomes of rehabilitation. Physical therapists can observe these compensatory efforts and provide guidance for improvement. Game level design should require patients to shift their weight towards their weaker side, promoting more effective rehabilitation outcomes (Riedmann, 2022).

This thesis therefore develops a RL model that adjusts the game level difficulty in which the path of the ball is chosen to reach the game's goal, by focusing on path selection that can shift weight towards the player's weakest direction of the COP to help the ball reach the finish line. The primary outcome is to achieve improvement in

balance in terms of COP. The secondary outcome is to stimulate internal motivation and encourage full participation in the health rehabilitation and physical therapy process, aiming to improve rehabilitation outcomes. For more details on mapping the SuraSole maze game to the RL framework, see Chapter 3.

2.4 Reinforcement learning algorithm

Reinforcement learning (RL) represents a distinct category within machine learning algorithms, primarily operating on datasets characterized by their uncertainty and complexity. In this framework, RL agents are rewarded for making correct decisions and penalized for incorrect ones. The core objective is to maximize the cumulative reward, effectively treating each problem as if it were a game. This unique approach allows the model to adapt and learn new skills over time, making RL particularly suited for applications such as AI gaming, robotic navigation, and any other application requiring the acquisition of new and complex abilities. Through this process, RL aims to continually improve its decision-making capabilities, thereby enhancing its performance in the designated tasks. For more details on RL, see Appendix A.

2.5 Literature of RL for difficulty adjustment in rehabilitation games

RL plays a crucial role in healthcare and physical therapy, particularly in the development of SG that aim at dynamically adjusting game difficulty levels to optimally challenge each player (Jayaraman et al., 2025; Seyderhelm and Blackmore, 2021). Moreover, RL enhances player immersion in therapy by making the game more engaging as difficulty is adjusted according to their skills, thereby motivating and increasing participation in the physical rehabilitation process.

A significant difference between RL and traditional control systems is that, while control systems depend on theoretical models, RL is driven by outcomes. This enhances the efficiency of treatment methods to suit individual needs by evaluating and comparing different approaches within a specific personal context. For instance, RL facilitates the selection of the most effective treatment strategy for an individual by weighing all possible options for that person, thereby personalizing the

rehabilitation or healthcare process with precision that adapts to the changing ability levels of the player.

RL also reduces the burden on medical professionals in monitoring and adjusting treatment programs by utilizing automated systems that can adapt treatments appropriately in real-time. Moreover, RL enhances the user experience through various rehabilitation games or programs, making the process more engaging and motivating for patients to actively participate in their rehabilitation. RL not only adjusts the difficulty level of rehabilitation to suit individual needs but also increases the likelihood of long-term success in the rehabilitation process (Jayaraman et al., 2024).

In (Sekhavat, 2017), a rehabilitation system was introduced that automatically adjusts the difficulty level of games based on patients' skills using the RL technique. This system was specifically designed to enhance motor recovery, particularly in stroke patients. The game utilized in this study involved players controlling a character through a Kinect motion sensor, which could capture body positions and movements. Players were required to raise their arms to strike illuminated balls along a curved line. The system dynamically adjusted parameters such as the character's speed, ball size, and the distance between curves to match the players' abilities. In the RL process, the system used metrics like the difference between wins and losses and scores from each round as indicators of the state. Actions included adjusting the character's speed, ball size, and the distance between curves to tailor the game's difficulty to the players' skills. Rewards were designed to encourage continuous gameplay and skill development, such as reducing score differences and increasing opportunities for progression in each round. The results of this system demonstrated that MPRL was effective in both enhancing player satisfaction and promoting long-term motor control recovery.

In (Khabbaz et al., 2023), the SmartBird game was developed to enhance communication skills in children with Autism Spectrum Disorder (ASD). The game integrates RL with Fuzzy logic to dynamically adjust the game difficulty according to the player's skills. Players control a bird in the game to avoid obstacles, escape eagles, and dodge hunters' projectiles. The objective of the game is to improve the gameplay

ability and communication skills of children with ASD. In the RL process, the game's state is defined by the player's skill level, measured using a factor called the Skill Factor, which reflects the ability to avoid obstacles and enemies in the game. Rewards are determined by the reduction in the character's in-game health, and the system utilizes the RL to learn how to select actions, such as increasing or decreasing the number of obstacles or enemies, to match the player's skill level. The goal is to balance the challenge and enjoyment of the game while promoting long-term communication skill development. Experimental results from a sample group of 15 children with ASD showed that the game significantly improved their communication skills and gameplay abilities. The game's adaptive design particularly contributed to maintaining the children's engagement and supporting their continuous skill development during gameplay.

In (Mariselvam et al., 2023), developed a virtual reality (VR) rehabilitation game utilizing RL to support children with Down syndrome, particularly those reliant on wheelchairs. The game was designed to enhance both physical and cognitive skills, such as muscle control, hand-eye coordination, and social interaction. Players used a VR headset and controllers to play a game focused on controlling a ball on a board to score points. In the RL process, the system analyzed the player's abilities and used states (e.g., hand movements, responses to actions, and ball positions) as key indicators. Rewards were given when players successfully moved the ball to the desired position, while actions (e.g., adjusting the board's tilt or the ball's movement) were determined to help players develop missing skills. The primary goal of the system was to build confidence and physical control in children while providing an enjoyable and skill-appropriate experience. Experimental results demonstrated that RL significantly improved players' performance and scores. Additionally, the game showed potential for long-term development in children with Down Syndrome. However, further system training is required to substitute human caregivers in certain scenarios.

From the review of related research, it was found that studies applying SG for balance rehabilitation using RL remain limited, reflecting opportunities for further development in this area. In particular, there is potential to advance the dynamic

adjustment of game difficulty to suit a diverse range of target patients. It is anticipated that this effort will lead to the creation of new approaches that can be effectively applied in rehabilitation processes.

2.6 Balance ability assessment

The Balance score is a measure used to assess an individual's balance ability. In the context of balance rehabilitation, the Balance score serves as a crucial parameter for evaluating the effectiveness of rehabilitation processes by comparing scores before and after the intervention (Riedmann, 2022). In this study, the Balance score is utilized to analyze the impact of the developed SuraSole maze game. It involves calculating various parameters to evaluate improvements in the participants' balance ability, both quantitatively and qualitatively. Additionally, it acts as a key indicator for refining the game to effectively meet the needs of diverse target groups.

We used a Balance score, ranging from 0 to 100, to assess balance ability during weight-shifting exercises. This score allows for the measurement of efficiency and progress in balance rehabilitation and provides important information about the participants' balance deficits. The relevant parameters include COP velocity, RMS amplitude in the anterior-posterior and lateral directions, and RMS velocity in the anterior-posterior and lateral directions.

The calculation of the COP from SuraSole sensor data involves two main directions COP in the mediolateral direction (ML or COP_x) and COP in the anterior-posterior direction (AP or COP_y). The calculation can be performed as follows,

$$COP_x = \frac{\sum_{i=1}^8(F_{Li} \times X_{Li}) + \sum_{i=1}^8(F_{Ri} \times X_{Ri})}{\sum_{i=1}^8(F_{Li} + F_{Ri})} \quad (2.1)$$

$$COP_y = \frac{\sum_{i=1}^8(F_{Li} \times Y_{Li}) + \sum_{i=1}^8(F_{Ri} \times Y_{Ri})}{\sum_{i=1}^8(F_{Li} + F_{Ri})} \quad (2.2)$$

where

F_{Li} and F_{Ri} are the force measured by the i -th sensor on the left foot and right foot, X_{Li} and X_{Ri} are the position in the X-axis of the i -th sensor on the left foot and right foot,

Y_{Li} and Y_{Ri} are the position in the Y-axis of the i -th sensor on the left foot and right foot.

Sway path refers to the area traversed by the COP over a specific period, serving as an indicator of postural control capability. An increase in the Sway path area may reflect a decrease in balance-maintaining ability. However, the value can also be high in cases where the body exhibits good balance or low in cases of unstable posture, which might not always correspond to the actual ability to control balance (Palmieri et al., 2002).

The Sway path can be analyzed separately in the mediolateral direction (ML_sway), representing lateral movements, and in the anterior-posterior direction (AP_sway), representing forward and backward movements. These can be calculated using equations 2.3 and 2.4, respectively,

$$ML_sway = \sqrt{\frac{1}{N} * \sum_{i=1}^N (\Delta COP_x[i])^2} \quad (2.3)$$

$$AP_sway = \sqrt{\frac{1}{N} * \sum_{i=1}^N (\Delta COP_y[i])^2} \quad (2.4)$$

where

N is the number of COP measurements in the experiment,

$\Delta COP_x[i]$ is change in COP values along the X-axis between the i -th data point and the previous point $i - 1$ calculated as $COP_x[i] - COP_x[i - 1]$,

$\Delta COP_y[i]$ is change in COP values along the Y-axis between the i -th data point and the previous point $i - 1$ calculated as $COP_y[i] - COP_y[i - 1]$.

COP velocity refers to the rate of change of the position of the center of pressure (COP) over time. An increase in COP velocity indicates a decline in postural control, while a decrease in COP velocity suggests a reduction in the ability to maintain posture (Palmieri et al., 2002). The COP velocity can be calculated from,

$$COP\ velocity = \frac{\frac{1}{N} \sum_{i=1}^N \sqrt{(\Delta COP_x[i])^2 + (\Delta COP_y[i])^2}}{Duration} \quad (2.5)$$

where

N is the number of COP measurements in the experiment,

$\Delta COP_x[i]$ is change in COP values along the X-axis between the i-th data point and the previous point i-1,

$\Delta COP_y[i]$ is change in COP values along the Y-axis between the i-th data point and the previous point i-1,

Duration is total time span of the COP measurement.

RMS amplitude measures the average absolute displacement around the mean COP. A decrease in RMS amplitude is associated with an improvement in the ability to maintain an upright posture, whereas an increase in RMS amplitude indicates a decline in postural control (Palmieri et al., 2002). The parameter can be determined by

$$RMS_{amp_x} = \sqrt{\frac{1}{N} * \sum_{i=1}^N (x_i - \mu_x)^2} \quad (2.6)$$

$$RMS_{amp_y} = \sqrt{\frac{1}{N} * \sum_{i=1}^N (y_i - \mu_y)^2} \quad (2.7)$$

where

N is the number of COP measurements in the experiment,

x_i and y_i are the data points of COP_x and COP_y respectively,

μ_x and μ_y are the mean values of COP_x and COP_y respectively.

RMS velocity refers to the change in velocity over a certain period. It is a measure that reflects the fluctuations or variations in velocity. 'An increase in RMS velocity corresponds to an increase in COP displacement, which indicates hurried or uncontrolled weight distribution during the experiment' (Palmieri et al., 2002). The RMS velocity in x and y directions are defined as

$$RMS_{vel_x} = \sqrt{\frac{1}{N-1} * \sum_{i=1}^N (\Delta x_i - \mu_{\Delta x})^2} \quad (2.8)$$

$$RMS_{vel_y} = \sqrt{\frac{1}{N-1} * \sum_{i=1}^N (\Delta y_i - \mu_{\Delta y})^2} \quad (2.9)$$

where

N is the number of COP measurements in the experiment,

Δx_i and Δy_i are the changes at each point of COP_x and COP_y respectively,

$\mu\Delta x$ and $\mu\Delta y$ are the mean values of the changes in COP_x and COP_y respectively.

Typically, the Balance scores are compared before and after gameplay among the participants. This parameter highlights the effectiveness of the game in improving balance. Additionally, this comparison can be used to further analyze factors influencing rehabilitation outcomes.

After obtaining the individual balance parameters, we calculated the overall Balance score as described in (Riedmann, 2022), which reflects the relationship between various attributes to assess balance ability during gameplay. This score has a maximum value of 100, with the score decreasing as the deviation from the baseline distribution increases. The greater the deviation from the baseline distribution, the lower the balance ability score. The Balance score can be determined from

$$BS = \max \left(0, 1 - \frac{Mah(x)}{LST} \right) * 100 \quad (2.11)$$

where $Mah(x)$ represents the Mahalanobis distance of the new observation x which can be calculated from

$$Mah(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)} \quad (2.12)$$

where

x is the data vector of the participant (e.g., COP velocity, RMS amplitude, and RMS velocity),

μ is the mean data vector of normal individuals (e.g., COP velocity, RMS amplitude, and RMS velocity),

Σ^{-1} is the covariance matrix,

T denotes the transpose of the vector.

The LST is the Lower Scoring Threshold for the Balance score, allowing participants with deviations from the typical distribution but with notable changes to increase their overall score, which can be calculated from

$$LST = \mu_{Mah} + 5 * \sigma_{Mah} \quad (2.13)$$

where

μ_{Mah} is the mean of the Mahalanobis distance for the entire sample group,

σ_{Mah} is the standard deviation of the Mahalanobis distance for the entire sample group.

2.7 Summary

Chapter 2 presents the background on technology, theoretical principles, and methodologies related to the development of the SuraSole maze game. It emphasized the use of the SuraSole smart insoles for collecting balance data, the application of RL for dynamic game difficulty adjustment, and the analysis of balance parameters crucial for rehabilitation. All this information provides a foundation for the development and evaluation of the balance rehabilitation game.



CHAPTER III

RESEARCH METHODOLOGY

3.1 Introduction

To achieve the objectives of this research, the development of a system capable of automatically adjusting the difficulty levels of a balance rehabilitation game and evaluating its effectiveness requires a clear and appropriate research methodology. This study divides the process into three key phases: the development of the Reinforcement learning (RL) model, the design and development of the SuraSole maze game, and the evaluation of the system through experiments with participants.

In the first phase, the research focuses on developing the RL model to adjust the game's difficulty levels based on balance metrics, such as COP (Center of Pressure), measured using SuraSole insoles. This phase includes the training and testing of the RL model to enable the system to select levels that align with the player's state.

In the second phase, the design of the SuraSole maze game emphasizes creating levels that correspond to the player's COP state in each quadrant. Input from experts, such as physiotherapists, is incorporated into the game design to ensure it suits players with specific characteristics.

In the final phase, experiments with participants are conducted to evaluate the game's effectiveness both quantitatively and qualitatively. Data is collected on balance metrics, and participant feedback is surveyed regarding their satisfaction, safety, and motivation to engage in the game.

Thus, the Research Methodology chapter outlines the detailed process, from RL model development to game design and evaluation.

3.2 Research design

This study is an experimental research with the objective of developing a balance rehabilitation game that automatically adjusts difficulty levels based on players' abilities using RL. The study also aims to analyze the impact of the game on participants' balance and assess its effectiveness in improving balance and enhancing motivation for participating in balance rehabilitation.

3.3 Participants

The study involved 15 participants aged between 21-44 years (mean age 26.73 ± 6.65 years 12 males and 3 females). Participants were voluntarily recruited in the laboratory and were required to have no physical balance limitations or injuries that could affect their movement.

3.4 Research tools

- a) **SuraSole maze game:** A game developed using RL that can adjust difficulty levels based on players' COP data. The game consists of levels designed to assess and improve players' balance.
- b) **SuraSole insoles:** Wireless sensor-equipped insoles used to collect COP data.
- c) **Post-Game evaluation questionnaire:** A survey designed to measure participants' satisfaction, including their opinions on the game's safety, difficulty, and engagement.

3.5 Experimental procedures

The research methodology was conducted in two phases: Phase 1 involved developing the RL-based game adjustment system and focused on designing the conceptual framework of the game, Phase 2 evaluated the game with participants.

3.5.1 Phase 1: Development of RL-based game adjustment

In this phase, a RL model was developed to adjust the game's difficulty levels using simulated data. The RL model was designed to learn decision-making for

selecting game levels based on COP parameters. Details of the RL model training process and results are provided in Appendix B.

3.5.2 Phase 1: Game concept development

The game concept was inspired by (Riedmann, 2022; Baranyi, 2013), where players guide a ball through a maze by practicing weight shifting on SuraSole insoles. The game levels were designed to cover the COP states of players in five quadrants based on their balance capabilities.

Quadrants 1-4: Each quadrant contains nine game levels designed to improve players' balance at different states. From Figure 3.5.1 the difficulty of each level is ordered alphabetically (A-I), where "A" represents the easiest level and "I" represents the most difficult. The alphabet after the quadrant indicator (e.g., "A-I") specifies which quadrant the level belongs to.

Quadrant 5: This quadrant contains four game levels, as players in this state already have good balance. The focus here is on maintaining balance and evaluating the RL model's ability to guide players to this optimal state.

The level designs for Quadrants 1-4 were adapted from (Baranyi et al., 2013) and refined based on feedback from physiotherapists to ensure they meet players' needs and effectively improve balance. The RL framework for the SuraSole maze game are presented and the experiment results detailed in Appendix B.

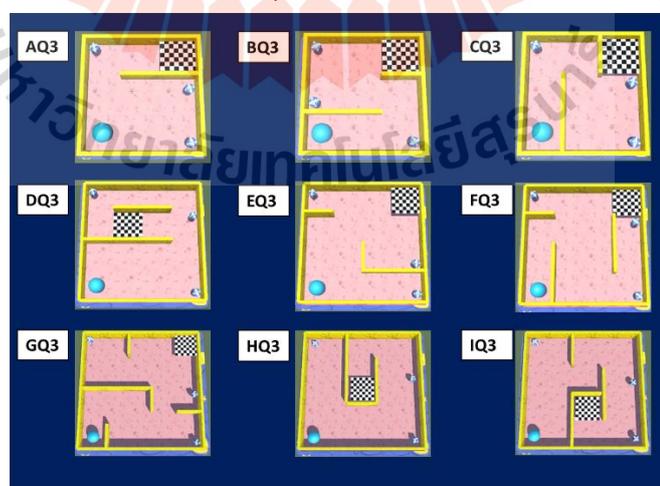


Figure 3.5.1 The 9 maze levels of COP in Quadrant 3

3.5.3 Phase 2: Experimental evaluation with participants

Once the RL model and the game were developed, the model was integrated into the SuraSole maze game for experimental evaluation with participants in a laboratory setting. The details of the procedure are as follows.

a) Preparation of participants

Participants underwent a training session to familiarize with the maze game developed to improve balance abilities. During the game, participants wore researcher-provided shoes equipped with SuraSole insoles, which were customized to fit their foot size. The game was connected via Bluetooth to the SuraSole insoles in each shoe. Participants were first trained to play the game to familiarize themselves with the game mechanics, including controlling the ball's movement and learning how to navigate through the levels. The training game is shown in Figure 3.5.2.



Figure 3.5.2 Training maze game

b) Assessment of game level for balance impairments

All participants played all levels, each designed for individuals with balance impairments, one round per level. These levels were evaluated for safety and to ensure they did not pose risks to players. Participants with normal balance were used to assess whether these levels were appropriate for individuals who may have balance limitations. In this part of the experiment, the RL model was not applied, as the participants already had normal balance. A post-game evaluation was then carried out using filled-in questionnaires. Results can be found in section 4.5.

c) Assessment of game level for normal balance

Participants were tested on levels designed for individuals with normal balance. They played five rounds in total. For each round, the Q-learning model selected the most suitable level for the player based on their COP data, measured before starting the game and again before starting a new level. This part of the experiment aimed to evaluate whether the model could accurately process players' COP status and select levels that provided appropriate challenges for each individual. A pre-and post-game COP-related measurements were evaluated. The results are presented in section 4.2-4.4.

Figure 3.5.3 illustrates the learning process of the RL model during gameplay. The process begins with measuring the player's COP to determine the current state. The measured COP data is then fed into the RL model, which evaluates and selects the most appropriate maze level for the player at that moment (Action).

After the player completes the selected level, the COP is measured again to assess how well the player can control the COP to stay near the center point. This COP data is used to calculate the reward based on the equation (B.1.1) presented in APPENDIX B. The closer the COP is to the center point, the higher the reward the player receives.

The obtained reward is then used to update the Q-value for the specific state and action using the Q-update equation (A.1.1) in APPENDIX A, enabling the model to learn how effective that particular action is for the given state.

Subsequently, a new state S' is generated from the COP measurement after gameplay, and the RL model selects a new level accordingly. This cycle repeats continuously, allowing the model to learn and improve its decision-making efficiency over time, ensuring that the level selection remains appropriate to the player's ability and progress at each stage.

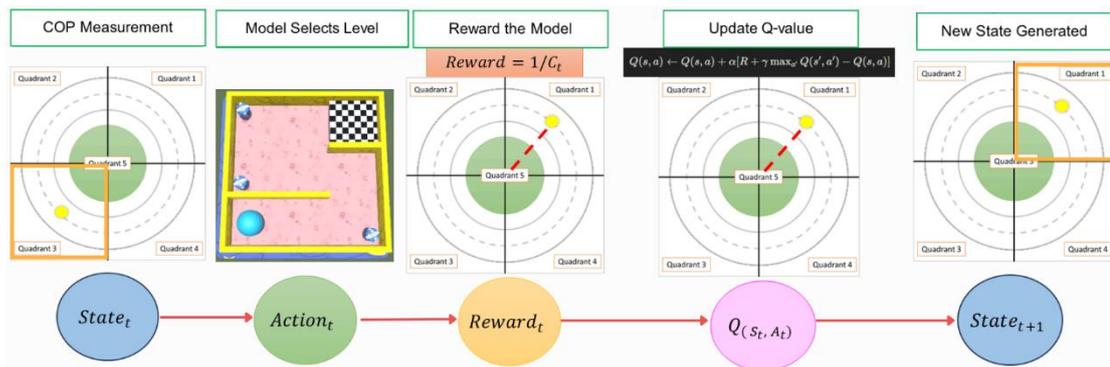


Figure 3.5.3 RL process for game level selection based on COP data

d) Evaluation of motivation for rehabilitation

After completing the designated levels, participants completed a questionnaire to assess whether these levels could enhance their motivation to engage in rehabilitation. This evaluation provided insights into whether the game could positively influence patients' behaviors and encourage participation in physical therapy. Evaluation results are given in section 4.5.



Figure 3.5.4 Game testing with participants wearing SuraSole insoles

3.6 Data analysis

a) Analysis of balance metrics

a.1 Compare the Balance score in equation (2.11) before and after gameplay to evaluate the effectiveness of the game in improving participants' balance.

a.2 Analyze RMS amplitude, RMS velocity, and COP velocity parameters in equation (2.5-2.9) to assess changes indicating improvements in balance in each direction (medio-lateral and anterior-posterior).

b) Analysis of game level adjustment through RL

b.1 Examine the Average reward in equation (A.4.1) obtained by participants in each round to evaluate whether the RL system could adjust game difficulty appropriately to match players' abilities.

c) Analysis of participants' feedback

c.1 Use a questionnaire to measure participants' levels of Satisfaction and Safety with the game.

c.2 Analyze responses to questionnaire items related to Motivation for Rehabilitation, such as enjoyment, challenge, and willingness to play again.

3.7 Summary

Chapter 3 presented the research methodology for the development and evaluation of the SuraSole maze game. This began with the development of a RL model to dynamically adjust the game's difficulty based on the COP state of the players and the design of maze levels based on recommendations from physical therapists to ensure suitability for the participants. Additionally, quantitative data collection and participant feedback were conducted to evaluate the game's effectiveness in terms of safety, improving balance ability, and motivating engagement in the rehabilitation process.

CHAPTER IV

RESULTS

4.1 Introduction

The analysis in this research focuses on evaluating the effectiveness of the SuraSole maze game, developed to improve balance through the application of Reinforcement learning (RL). The aim is to assess changes in balance-related parameters, such as RMS amplitude, RMS velocity, and COP velocity, in both the medio-lateral and anterior-posterior directions. These parameters reflect the participants' ability to control COP movement and overall stability.

The results include a comparison of the Balance score before and after gameplay to evaluate whether the game levels and difficulty adjustments, based on players' COP states, effectively enhance physical balance. Additionally, the Average reward data was utilized to measure the effectiveness of the RL system in adapting the game difficulty to match the players' abilities.

The analysis of participants' feedback was designed to assess levels of satisfaction, safety, and motivation for balance rehabilitation. This feedback provides insights into the game's ability to promote motivation and encourage patients to engage in rehabilitation activities consistently.

Therefore, the experiments conducted in this study not only emphasizes quantitative evaluations through physical parameters but also incorporates qualitative assessments based on participants' feedback and experiences. This combined approach provides crucial information for further development and improvement of the game to suit target groups with diverse characteristics.

4.2 Individual balance metrics

Based on the analysis of the parameters used to evaluate balance, it was found that after playing the game, the COP velocity, RMS amplitude, and RMS velocity in both directions showed certain trends of improvement. These trends indicate an enhancement in the ability to control movement and maintain balance of the COP, with the following details.

1. RMS amplitude (rms_amp_x): Decreased from 0.12 ± 0.07 to 0.08 ± 0.06 m, indicating a reduction in COP displacement in the medio-lateral direction, reflecting improved control of lateral balance, with $p = 0.0698$ (not statistically significant) from the paired t-test.

2. RMS Amplitude (rms_amp_y): Slightly decreased from 0.070 ± 0.036 to 0.068 ± 0.036 m, indicating increased stability in the anterior-posterior direction. Although the change was not pronounced, it still suggested a slight improvement trend, with $p = 0.4190$ (not statistically significant) from the paired t-test.

3. RMS velocity (rms_vel_x): Decreased from 0.043 ± 0.046 to 0.020 ± 0.010 m/s, indicating a reduction in COP velocity in the medio-lateral direction, suggesting improved control over lateral movement, with $p = 0.0252$ (statistically significant) from the paired t-test.

4. RMS velocity (rms_vel_y): Decreased from 0.013 ± 0.010 to 0.010 ± 0.003 m/s, indicating a reduction in COP velocity in the anterior-posterior direction, which reflects better control over forward-backward balance, with $p = 0.1191$ (not statistically significant) from the paired t-test.

5. COP velocity (velocity): Decreased from 0.0013 ± 0.0011 to 0.0007 ± 0.0003 m/s, indicating a reduction in the overall variability of COP, which is a sign of improved balance stability, with $p = 0.0471$ (statistically significant) from the paired t-test.

It should be noted that the experiment so far has been carried out on healthy volunteers with no balance impairment. Additional experiments with participants who have balance limitations are recommended to obtain more comprehensive data and to demonstrate the game's potential in effectively enhancing balance control skills.

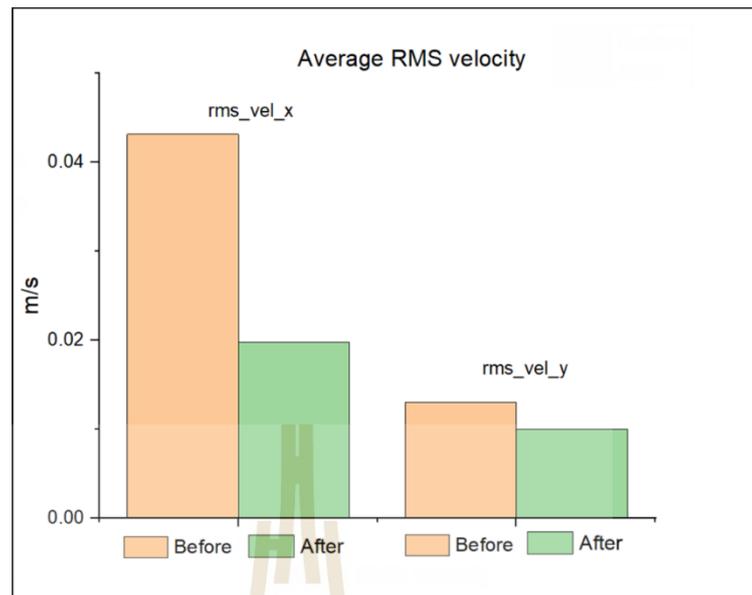


Figure 4.2.1 Average RMS velocity in the X and Y axes before and after gameplay

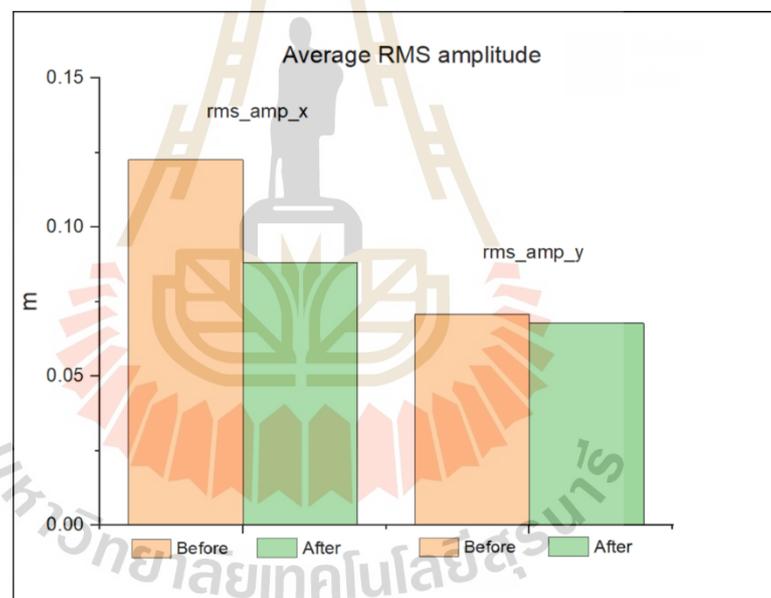


Figure 4.2.2 Average RMS amplitude in the X and Y axes before and after gameplay

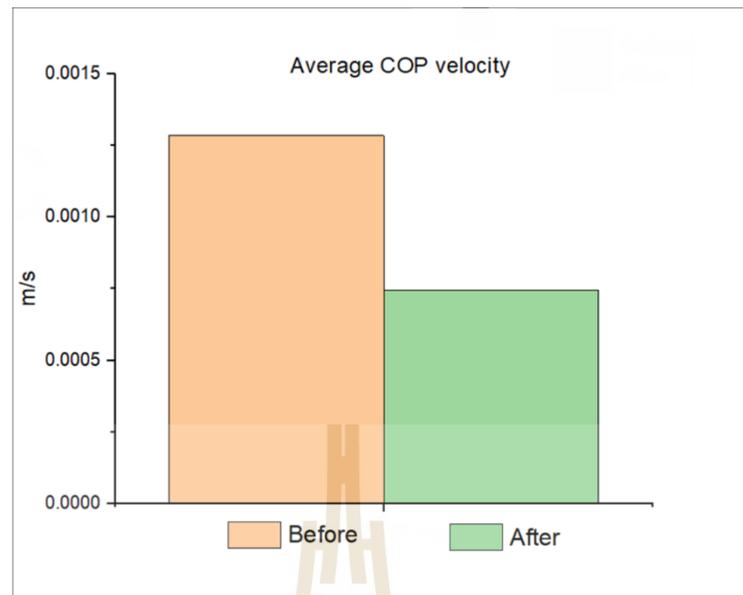


Figure 4.2.3 Average COP velocity before and after gameplay

4.3 Balance score

This section describes the details of the assessment conducted with the participants, including a summary of the findings from the balance ability evaluation to reflect the suitability of the game used for testing. It also analyzes whether the scores obtained from the experiment adequately reflect balance improvement. Additionally, a comparison of Balance scores in equation (2.11) between participants before and after gameplay is included to verify whether the designed game can help improve balance.

To validate the comparison results, a Paired t-test was performed to calculate the significance level (p-value), which indicates whether the difference between pre- and post-experiment scores is statistically significant (Datatab, 2024). A significance level of $p < 0.05$ was set as the criterion. If the p-value obtained is less than or equal to 0.05, the difference between the pre- and post-experiment scores is considered significant. This suggests that the designed game can effectively improve balance ability with statistical significance.

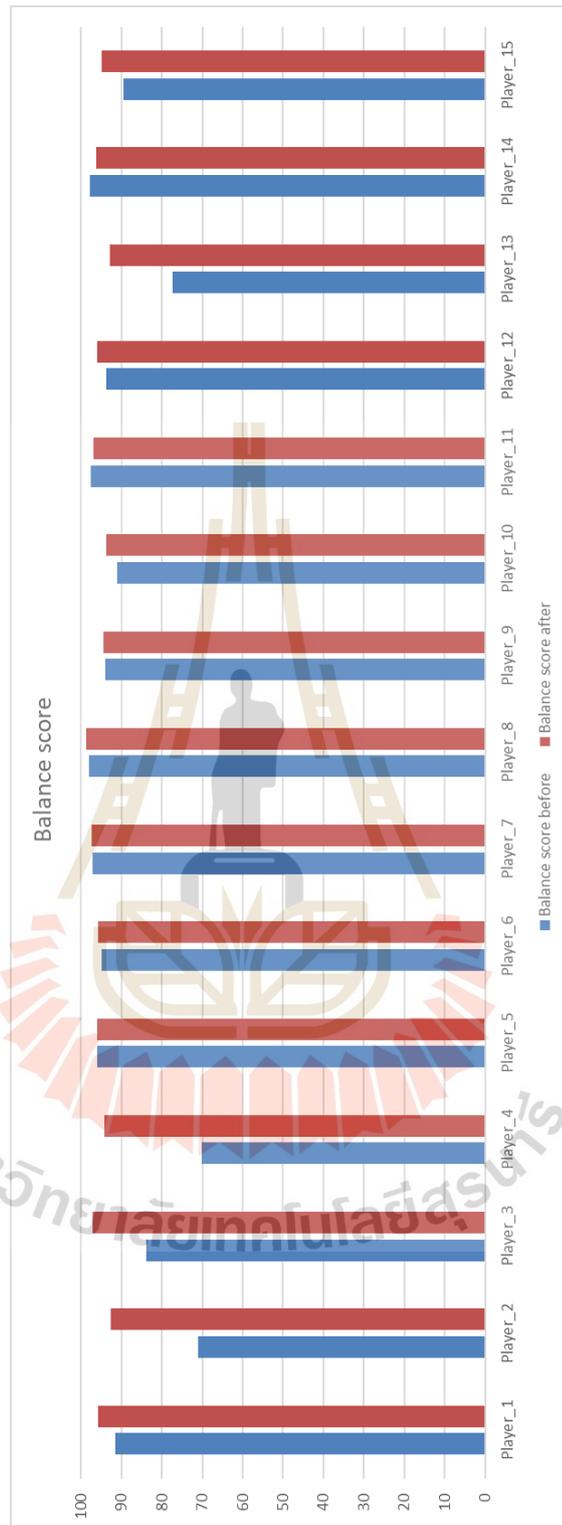


Figure 4.3.1 Balance score of the participants

The graph shows that most participants exhibited some improvement in their Balance score after playing the game compared to before the experiment (blue bars).

This reflects the game's potential to effectively enhance participants' balance ability. Notably, Player 3 and Player 13 demonstrated a significant improvement in their scores, increasing from 83.86 to 97.13 and 77.34 to 92.87, respectively. Additionally, Player 2 and Player 4, who had low initial scores, showed more progress, improving from 70.94 to 92.67 and 70.14 to 94.14, respectively. This indicates that the game may help develop balance control skills even in individuals with initially lower balance abilities.

Meanwhile, some participants, such as Player 7, Player 8, and Player 11, showed no significant difference in their scores before and after the experiment. This could be attributed to their already stable balance abilities or other factors, such as the game's challenge level not being appropriate for these participants. External factors, such as familiarity with the equipment or a lack of challenge during gameplay, may have also influenced their performance.

Overall, the average Balance score for all participants improved from 89.6 ± 9.23 before the experiment to 95.5 ± 1.6 after playing the game. Using the paired t-test using the significance level of 5% resulted in a p-value of 0.0161, there is a significant in the improvement of the average Balance score of the playing the game.

It is important to note that the participants in this study were healthy individuals with normal balance abilities, which might have limited the observed improvements compared to individuals with balance impairments. Therefore, further studies should be conducted with participants who have balance limitations to comprehensively evaluate the game's effectiveness. Additionally, adjusting the game's difficulty level to better match the capabilities of different groups of players, along with considering factors such as familiarity with the equipment and player engagement during the experiment, could further enhance the game's design. These improvements would ensure that the game effectively meets the needs of its target groups and maximizes its potential for use in balance rehabilitation.

4.4 Average reward

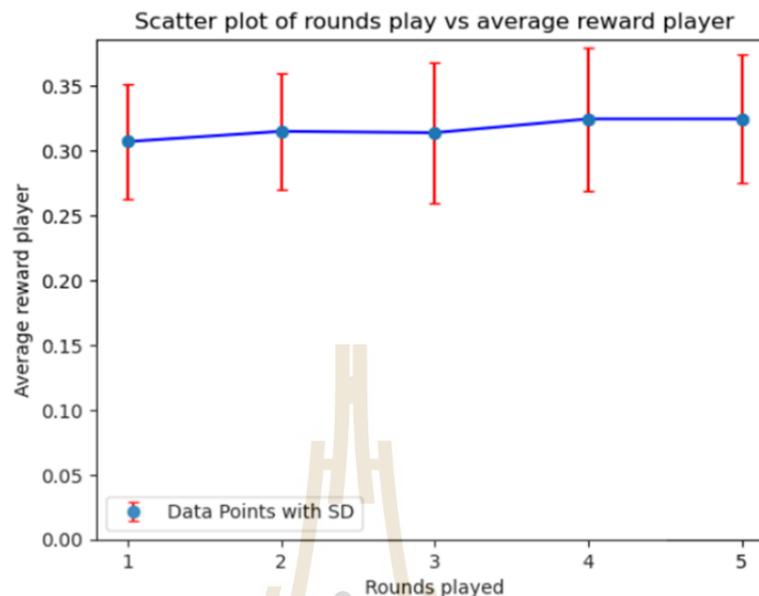


Figure 4.4.1 Relationship between the number of game play rounds and the average reward of participants

From Figure 4.4.1, it is evident that the Average reward tends to increase as the number of gameplay rounds increases. This indicates that the RL model helps players move the COP closer to the center over time, reducing the COP distance from the center and thereby increasing the Average reward, where by the immediate reward is given by equation (A.4.1). This trend reflects the Q-learning model's ability to learn and adapt to improve player performance as the number of gameplay rounds increases.

Using the Average reward from all participants provides an overall perspective on the RL model's effectiveness in helping players improve COP control. However, the variability in the reward values suggests differences in individual player abilities. Adjusting the model's parameters, such as learning rate or exploration rate, could further enhance performance at both the individual and overall levels.

Additionally, it should be noted that the participants in this experiment were general individuals with normal COP, who already had a good foundation in balance. This may have contributed to the model's consistent learning and performance improvements. However, familiarity with the game mechanics over multiple rounds may also have been a factor in players' improved COP control, resulting in reduced

COP distance and increased rewards. Therefore, separating the effects of RL from the familiarity with the game should be considered in future experiments.

The results of this experiment also indicate that while the RL model can improve COP control overall, reducing the variability in reward values may require fine-tuning parameters such as the learning rate or exploration rate. This would allow the model to learn more consistently and respond more effectively to players with varying abilities.

4.5 Player Satisfaction

After playing the game, participants completed a satisfaction assessment questionnaire to collect information about their experiences and opinions regarding the developed game. The purpose of this assessment was to measure the level of satisfaction, enjoyment, game safety, and the game's ability to promote balance rehabilitation. Analyzing the data from this assessment helps to understand the participants' responses to the game format and difficulty, as well as to evaluate whether the game design can create appropriate motivation and challenge for training. The questionnaire can be found in Appendix C.

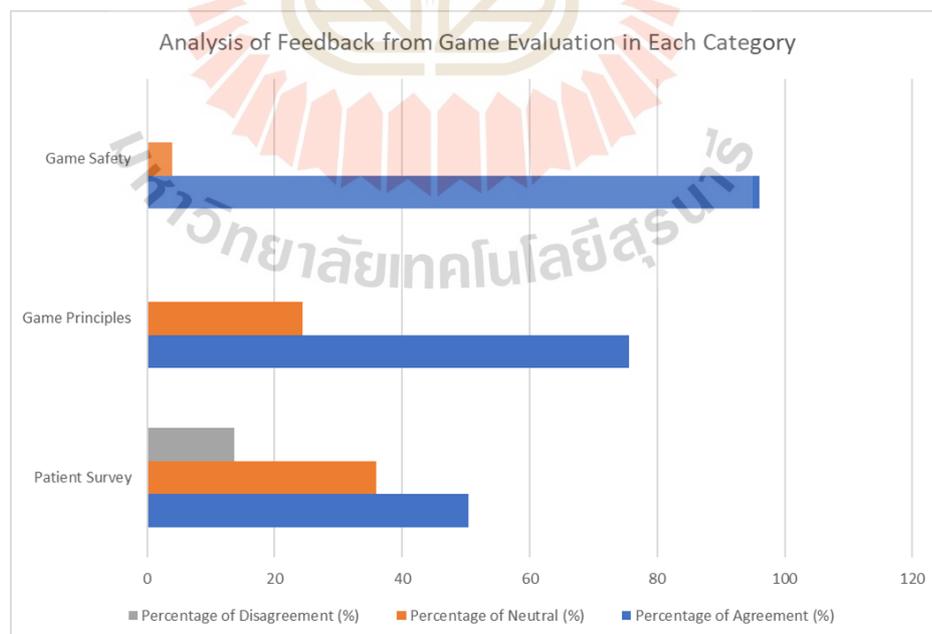


Figure 4.5.1 Analysis of feedback from game assessment in each category

From Figure 4.5.1, the evaluation is categorized into three aspects, patient survey, game safety, and game principles. The analysis of the results is as follows.

4.5.1 Patient survey

This category shows that 50.43% of participants expressed "agreement," indicating a generally positive reception from more than half of the participants. However, 35.9% responded "neutral," suggesting that a considerable number of participants did not perceive a significant difference in outcomes. Additionally, 13.68% responded with "disagreement," highlighting areas that may require improvement to enhance patient satisfaction in the future.

4.5.2 Game principles

The feedback for this category was positive, with 75.56% of participants expressing agreement. This result highlights that the maze levels were appropriately designed to match the patients' skill levels and demonstrates the game's effectiveness in terms of usability and technical principles. However, 24.44% of participants responded neutral, suggesting areas where the design may not have fully met expectations. There were no disagreement responses, indicating strong alignment between the maze design and user needs overall.

4.5.3 Game safety

The feedback for the Game Safety category was strongly positive, with 96% of participants expressing agreement. This indicates that participants perceived the game as safe for use, particularly in reducing risks related to physical fatigue, technical errors, or other safety-related concerns. The high level of agreement reflects the robustness of the game's design in prioritizing user safety, which is a critical factor for therapeutic and rehabilitative applications.

On the other hand, 4% of participants responded neutrally, which may stem from uncertainty or a lack of clarity regarding certain aspects of safety. Nevertheless, the overall positive feedback reinforces the game's reliability and its ability to provide a safe environment for participants.

4.6 Summary

Chapter 4 summarizes the analysis related to the potential of the SuraSole maze game, which was developed to promote improvements in balance ability. The focus was on measuring key parameters such as RMS amplitude, RMS velocity, and COP velocity, which indicate trends toward improvement in the overall balance score of participants. The results showed a significant improvement in the Balance score, with $p = 0.0161$. which may reflect the potential of the game in supporting the development of participants' overall movement control skills. Furthermore, the increase in Average reward per round indicates the trend of the RL model in adjusting the game's difficulty levels to match the COP status of the players, potentially supporting the continuous improvement of participants' abilities.

However, it should be noted that this experiment was conducted with a group of volunteers who had normal balance and good baseline stability, which may mean the results do not fully reflect the game's effectiveness for individuals with balance impairments. Additionally, all experiments were conducted in a controlled laboratory environment, which may not fully align with real-world conditions where various environmental factors are present.

Feedback from participants also indicated positive satisfaction regarding the game's safety, usability design, and its ability to motivate rehabilitation engagement. The data from both quantitative and qualitative evaluations suggest that the SuraSole maze game holds potential for use in rehabilitation processes. Nevertheless, further studies are needed with sample groups experiencing balance issues and experiments conducted in environments that more closely resemble real-world settings to comprehensively assess the game's potential.

CHAPTER V

CONCLUSIONS

5.1 Conclusions

This study demonstrates the potential of the developed maze game in improving balance ability and promoting engagement in the rehabilitation process. The experimental results showed that participants' Balance scores increased after playing the game from 89.6 ± 9.23 to 95.5 ± 1.6 , reflecting the game's effectiveness in enhancing balance control. Additionally, changes in parameters such as RMS amplitude, RMS velocity, and COP velocity further suggested the game's potential to enhance stability and movement control.

Moreover, the increasing Average reward per round highlights the capability of the RL model to learn and adapt, helping participants better control their COP. This improvement led to a reduction in COP distance and a continuous increase in Reward. The use of Average reward across all participants reflects the overall effectiveness of the model. Additionally, the system's ability to adjust game difficulty automatically based on the participants' COP status further supported the continuous improvement of participants' performance. However, the variability in Average reward observed suggests differences in individual players' abilities. The participants in this experiment were individuals with normal COP and good baseline balance, which contributed to the consistent effectiveness of the RL model. Furthermore, familiarity with the gameplay may have also contributed to reducing COP distance and increasing Average reward. Future experiments should aim to separate the effects of RL adaptation, difficulty adjustments, and gameplay familiarity to ensure more accurate assessments.

Feedback from participants reflected positive satisfaction, especially in the "Game Safety" category, which received the highest level of agreement, emphasizing the game's

appropriateness and safety for use. The "Game Principles" category also highlighted the suitability of the technical design. However, responses from the "Patient Survey" indicated the need for further improvements on challenge physical ability patients.

It is important to note that this experiment was conducted with participants who had normal balance abilities. This may have limited the observable changes compared to patients with balance impairments. Therefore, future studies should be conducted with patients who have balance limitations to comprehensively evaluate the game's effectiveness. Additionally, the game's difficulty levels should be adjusted to better match the capabilities of different groups of users.

In conclusion, the developed maze game demonstrates its potential as tool for balance rehabilitation while also motivating patients to engage in physical therapy. Nevertheless, further improvements are necessary to ensure the game can accommodate a more diverse range of users. Long-term studies are also recommended to validate the game's effectiveness in broader contexts such as, for the elderly and patients with Parkinson's disease, stroke, sarcopenia, and balance impairments.

5.2 Future work

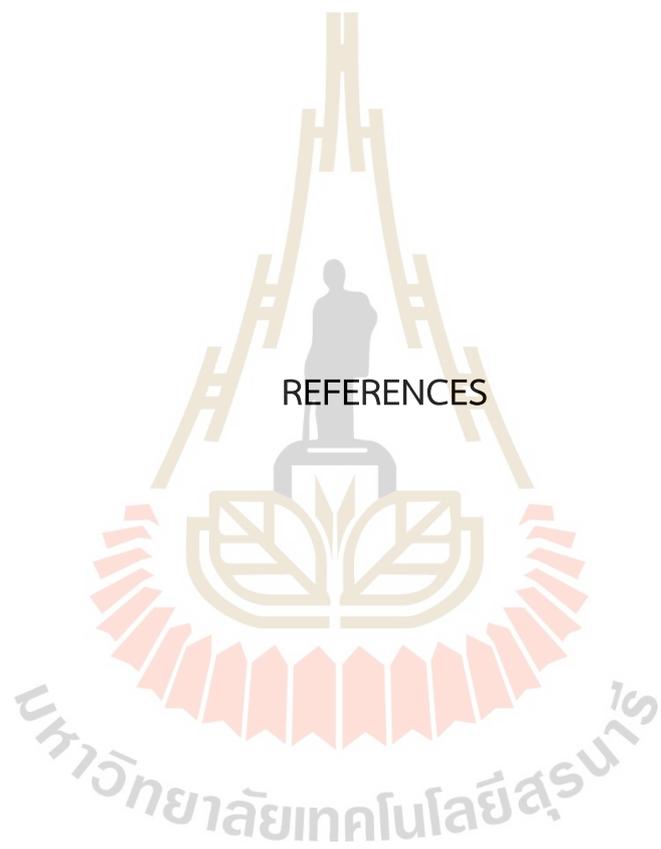
Future research will focus on trials involving elderly individuals with balance impairments to evaluate the game's effectiveness in restoring balance within the target group. Efforts will also be made to enhance game engagement and optimize maze levels to match the abilities and limitations of older adults and incorporate supportive features such as user-friendly instructions and feedback mechanisms that enhance motivation. Additionally, long-term follow-ups will assess the sustainability of the game's outcomes and its potential to reduce the risk of falls among elderly users.

For future development, consideration will be given to adding an additional state that reflects not only the position of the COP within a specific quadrant but also includes information about the distance from the center point. In the current system, only the COP quadrant is considered; for example, if the COP is at a first position close to the center and at a second position farther away, although both positions are in the

same quadrant, the system still selects the same maze level. Incorporating distance-based states would allow the model to distinguish more precisely and select actions that are better matched to the player's actual balance ability.

Additionally, there are plans to enhance the game's ability to detect improper playing behaviors, such as applying weight incorrectly, which may induce ball movement, but from incorrect posture. Mechanisms for monitoring and alerting will be added to ensure that players apply weight properly, promoting effective balance training and preventing ineffective or incorrect practice.





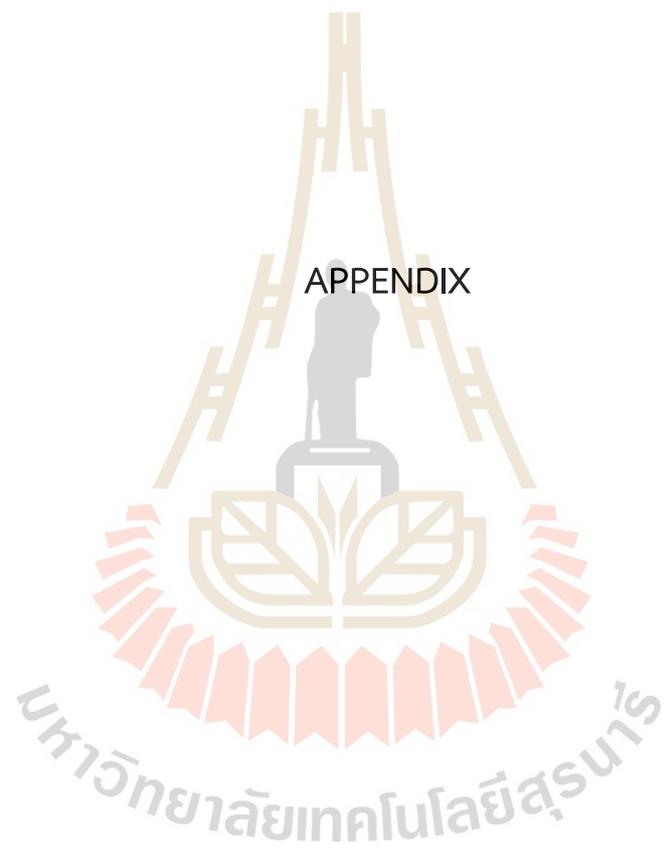
REFERENCES

- Agrawal, D. K., Usaha, W., Pojprapai, S., & Wattanapan, P. (2023). Fall risk prediction using wireless sensor insoles with machine learning. **IEEE Access, 11**, 23119–23126. <https://doi.org/10.1109/ACCESS.2023.3252886>
- Ahn, I., & Park, J. (2011). Drug scheduling of cancer chemotherapy based on natural actor-critic approach. **Biosystems, 106**, 121–129.
- Baranyi, R., Willinger, R., Lederer, N., Grechenig, T., & Schramm, W. (2013). Chances for serious games in rehabilitation of stroke patients on the example of utilizing the Wii Fit Balance Board. In **2013 IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH)** (pp. 1–7). IEEE. <https://doi.org/10.1109/SeGAH.2013.6665319>
- Beheshti, Z., & Shamsuddin, S. M. (2013). A review of population-based meta-heuristic algorithm. **International Journal of Advances in Soft Computing and Its Applications**, 5(1), 1–35.
- Choutri, K., Fareh, R., Rahmati, M., Bettayeb, M., Fadloun, S., & Lagha, M. (2023). Reinforcement learning fractional order PID controller for upper limb rehabilitation robot. In **2023 International Conference on Fractional Differentiation and its Applications (ICFDA)** (pp. 1–6).
- Danielsen, A., Olofsen, H., & Bremdal, B. A. (2016). Increasing fall risk awareness using wearables: A fall risk awareness protocol. **Journal of Biomedical Informatics, 63**, 184–194.
- Datatab. (2024). Paired t-test tutorial. Retrieved December 23, 2024, from <https://datatab.net/tutorial/paired-t-test>
- de Oliveira Andrade, K., Fernandes, G., de Paula Caurin, G. A., Siqueira, A. A. G., Romero, R. A. F., & de L. Pereira, R. (2014). Dynamic player modelling in serious games applied to rehabilitation robotics. In **2014 Brazilian Symposium on Robotics and Latin American Robotics Symposium (SBR-LARS)** (pp. 211–216). <https://doi.org/10.1109/SBR.LARS.ROBOCONTROL.2014.41>

- Ebrahimi Zade, A., Shahabi Haghighi, S., & Soltani, M. (2020). Reinforcement learning for optimal scheduling of glioblastoma treatment with temozolomide. **Computer Methods and Programs in Biomedicine, 193**, 105443.
- Florence, C. S., Bergen, G., Atherly, A., Burns, E., Stevens, J., & Drake, C. (2018). Medical costs of fatal and nonfatal falls in older adults. **Journal of the American Geriatrics Society, 66**(4), 693–698.
- Hocine, N., & Gouaich, A. (2011, July). Therapeutic games' difficulty adaptation: An approach based on player's ability and motivation. In **2011 16th International Conference on Computer Games (CGAMES)** (pp. 257–261). IEEE.
- Hornak, D., Jascur, M., Ferencik, N., & Bundzel, M. (2019). Proof of concept: Using reinforcement learning agent as an adversary in serious games. In **2019 IEEE International Work Conference on Bioinspired Intelligence (IWOB)** (pp. 111–116). <https://doi.org/10.1109/IWOB147054.2019.9114431>
- Jayaraman, P., Desman, J., Sabounchi, M., et al. (2024). A primer on reinforcement learning in medicine for clinicians. **NPJ Digital Medicine, 7**, 337. <https://doi.org/10.1038/s41746-024-01316-0>
- Khabbaz, A., Fateh, M., Pouyan, A., & Rezvani, M. (2023). Designing a serious game for children with autism using reinforcement learning and fuzzy logic. **Journal of AI and Data Mining, 11**(3), 375–390.
- Maheu-Cadotte, M.-A., Cossette, S. R., Dubé, V., Fontaine, G., Lavallée, A., Lavoie, P., Mailhot, T., & Deschênes, M.-F. (2021). Efficacy of serious games in healthcare professions education: A systematic review and meta-analysis. **Simulation in Healthcare, 16**(3), 199–212.
- Mariselvam, J., Rajendran, S., & Alotaibi, Y. (2023). Reinforcement learning-based AI assistant and VR play therapy game for children with Down syndrome bound to wheelchairs. **Mathematics, 8**, 16989–17011.
- Nintendo. (n.d.). **Wii Fit Plus**. Retrieved May 14, 2025, from <https://www.nintendo.com/en-gb/Games/Wii/Wii-Fit-Plus-283905.html>

- Palmieri, R. M., Ingersoll, C. D., Stone, M. B., & Krause, B. A. (2002). Center-of-pressure parameters used in the assessment of postural control. *Journal of Sport Rehabilitation*, 11*(1), 51–66.
- Physiopedia. (2023, April 25). Center of pressure (COP). In *Physiopedia**. Retrieved February 25, 2024.
- Riedmann, J. (2022). Analysis, design and prototypical implementation of a serious game for balance rehabilitation providing personalized level generation based on machine learning techniques [Diploma thesis, Technische Universität Wien]. *repositUM*. <https://doi.org/10.34726/hss.2022.105020>
- Sekhavat, Y. (2017). MPRL: Multiple-periodic reinforcement learning for difficulty adjustment in rehabilitation games. In *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)**. <https://doi.org/10.1109/SeGAH.2017.7939260>
- Seyderhelm, A., & Blackmore, K. (2021). Systematic review of dynamic difficulty adaption for serious games: The importance of diverse approaches. *SSRN Electronic Journal**. <https://doi.org/10.2139/ssrn.3982971>
- Suratec. (2024). Suratec official website. Retrieved November 27, 2024, from <https://www.suratec.co.th/th/>
- Teruel, M. A., López-Jaquero, V., Sánchez-Cifo, M. A., Navarro, E., & González, P. (2019). Improving motivation in wrist rehabilitation therapies. In *Human-Computer Interaction – INTERACT 2019** (pp. 199–206). https://doi.org/10.1007/978-3-030-24097-4_24
- Thomas, S., Mackintosh, S., & Halbert, J. (2010). Does the 'Otago exercise programme' reduce mortality and falls in older adults? A systematic review and meta-analysis. *Age and Ageing*, 39*(6), 681–687.
- United Nations, Department of Economic and Social Affairs. (2020). *World population ageing**. New York, NY: United Nations.
- Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, 8*, 279–292. <https://doi.org/10.1007/BF00992698>
- Zhao, Y., Kosorok, M. R., & Zeng, D. (2009). Reinforcement learning design for cancer clinical trials. *Statistics in Medicine*, 28*, 3294–3315.

APPENDIX



APPENDIX A

Q-learning

Q-learning is a popular Reinforcement learning technique used to find the optimal action-selection policy for a given finite Markov decision process (MDP) (Jayaraman et al., 2024). It aims to maximize the cumulative reward an agent can achieve by learning the best actions to take from each state. In the section on the RL Framework, we will discuss it in Appendix B.

1. **Agent and Environment:** The agent interacts with the environment, which is represented by states (\mathbf{s}) and actions (\mathbf{a}).

2. **Q-values:** The agent maintains a table of Q-values, $Q(\mathbf{s}, \mathbf{a})$, which estimates the expected utility (or cumulative reward) of taking action \mathbf{a} in state \mathbf{s} and following the optimal policy thereafter.

3. **Initialization:** Initially, Q-values are typically set to arbitrary values.

4. **Policy:** The agent selects actions based on its current Q-values, often using an exploration-exploitation strategy (e.g., ϵ -greedy policy) to balance between exploring new actions and exploiting known ones.

5. **Learning:**

5.1 **Experience:** The agent takes an action \mathbf{a} in state \mathbf{s} , receives a reward r , and transitions to a new state \mathbf{s}' .

5.2 **Update Rule:** The Q-value for the state-action pair (\mathbf{s}, \mathbf{a}) is updated using the formula:

$$Q(\mathbf{s}, \mathbf{a}) \leftarrow Q(\mathbf{s}, \mathbf{a}) + \alpha [r + \gamma \max_{\mathbf{a}'} Q(\mathbf{s}', \mathbf{a}') - Q(\mathbf{s}, \mathbf{a})] \quad (\text{A.1.1})$$

where

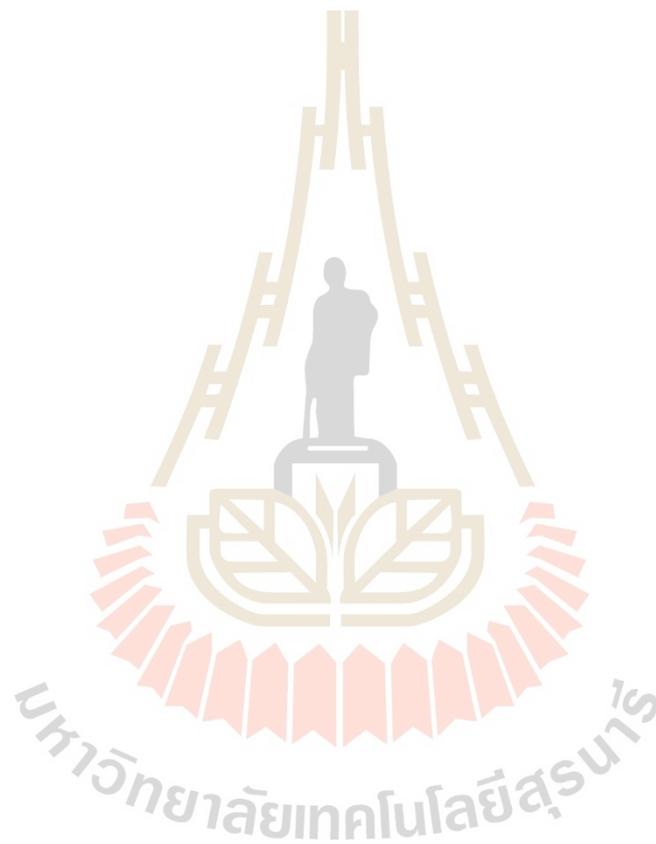
α is the learning rate (how much new information overrides old information),

γ is the discount factor (how much future rewards are valued compared to immediate rewards),

$\max_{\mathbf{a}'} Q(\mathbf{s}', \mathbf{a}')$ is the maximum estimated Q-value for the next state \mathbf{s}' .

6. **Convergence:** Over time, as the agent explores the environment and updates its Q-values, the Q-values converge to the true Q-values, allowing the agent to learn the optimal policy for selecting actions.

Q-learning is model-free, meaning it doesn't require knowledge of the environment's dynamics, making it versatile and widely applicable in various Reinforcement learning problems.



APPENDIX B

RL-based SuraSole Maze Game Level Difficulty Adjustment

The Q-learning method is one of the popular and widely used Reinforcement learning (RL) algorithms. Its goal is to enable an agent to learn the best decisions in various situations through trial and error. It uses the concept of Q-values, which measure the value of taking an action in a given state. These Q-values help the agent determine which actions will yield the highest long-term rewards (Watkins and Dayan, 1992).

B.1 Defining the states, actions and rewards for SuraSole maze game

State: The state refers to the condition of the patient, which in this framework is the COP of the patient or the current situation. In this work, we define the states where the coordinates of the COP are located, i.e., as Quadrant 1, Quadrant 2, Quadrant 3, Quadrant 4, and Quadrant 5. Each quadrant represents a different patient condition based on the COP value as shown in Figure B.1.1.

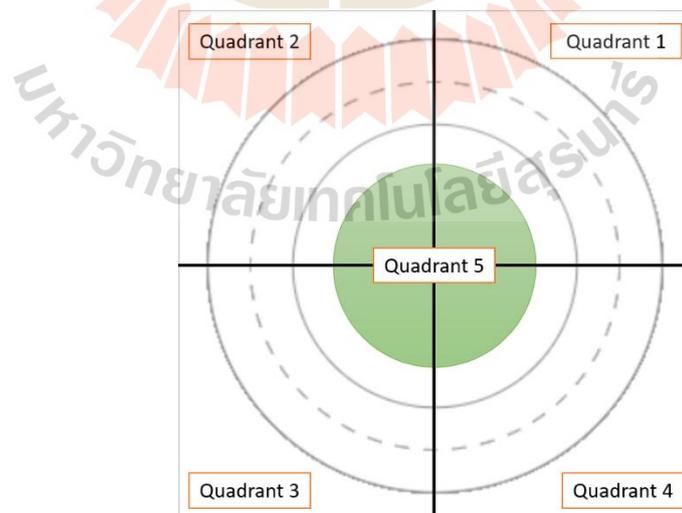


Figure B.1.1 Division of quadrants in various states.

Action: The action refers to the decisions made by the agent to improve the current state of the player. In this framework, actions are defined as the 9 maze levels, each of which is designed based on the (Baranyi et al., 2013) template and corresponds to the COP state. For example, if a participant's COP is located in Quadrant 3, the RL system will select one of the 9 levels specifically designed for Quadrant 1 to match the player's state. For Quadrant 5, which is designed for individuals with normal COP conditions, the RL system will select from only 4 levels, as illustrated in Figures B.1.2 and B.1.3.

Additionally, to illustrate the maze designs of other Quadrants, we have included Figures B.1.4, B.1.5, and B.1.6, which present the maze layouts for Quadrants 1, 2, and 4, respectively. Each maze is designed to provide challenges that correspond to the COP positions of players within those specific Quadrants, enabling the RL system to select levels that are well-matched to the player's abilities and progress with precision.

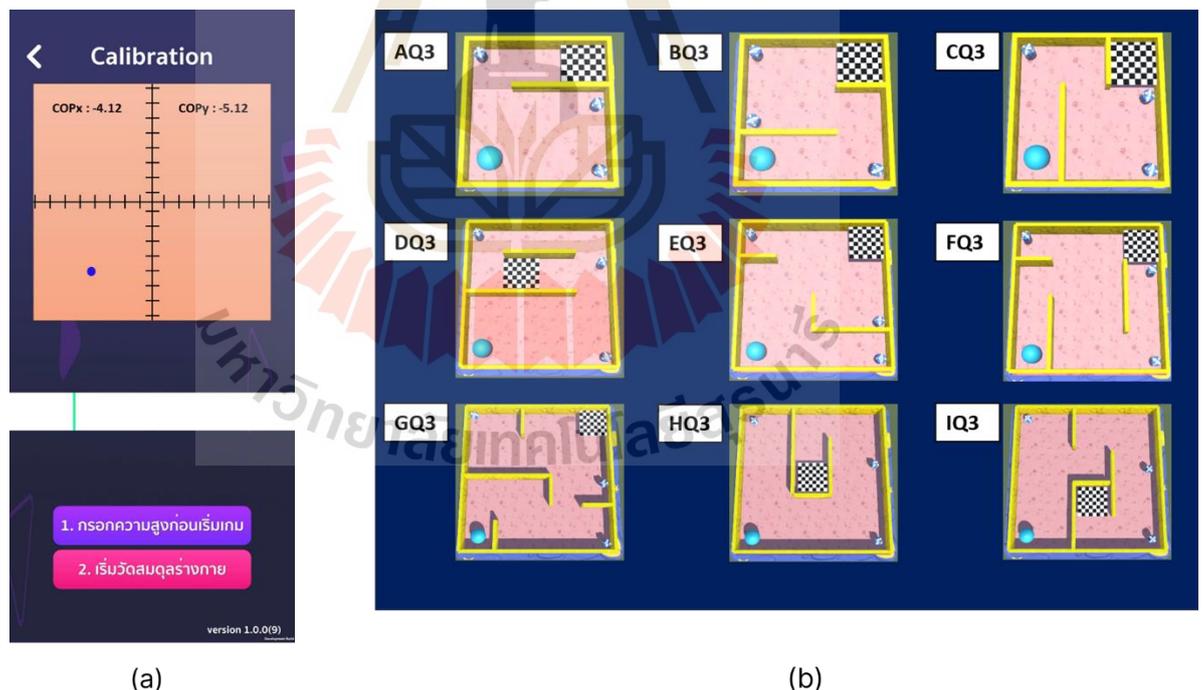


Figure B.1.2 COP states and RL-selected levels for Quadrant 3, (a) SuraSole maze game interface showing COP in Quadrant 3 before the game (b) SuraSole maze game level for a sample COP in Quadrant 3

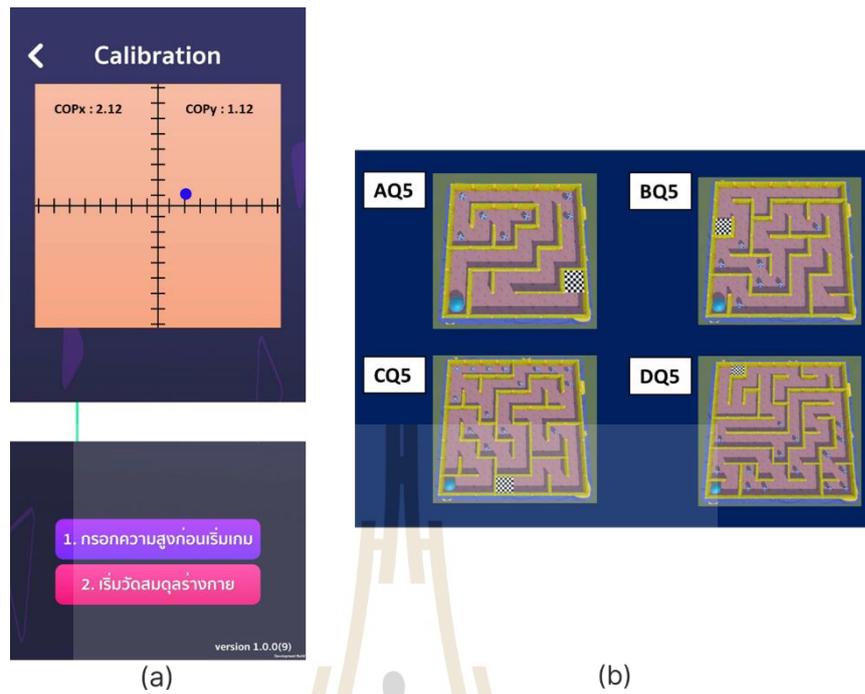


Figure B.1.3 COP states and RL-selected levels for Quadrant 5, (a) SuraSole maze game interface showing COP in Quadrant 5 before the game, (b) SuraSole maze game level for a sample COP in Quadrant 5



Figure B.1.4 SuraSole maze game level for a sample COP in Quadrant 1

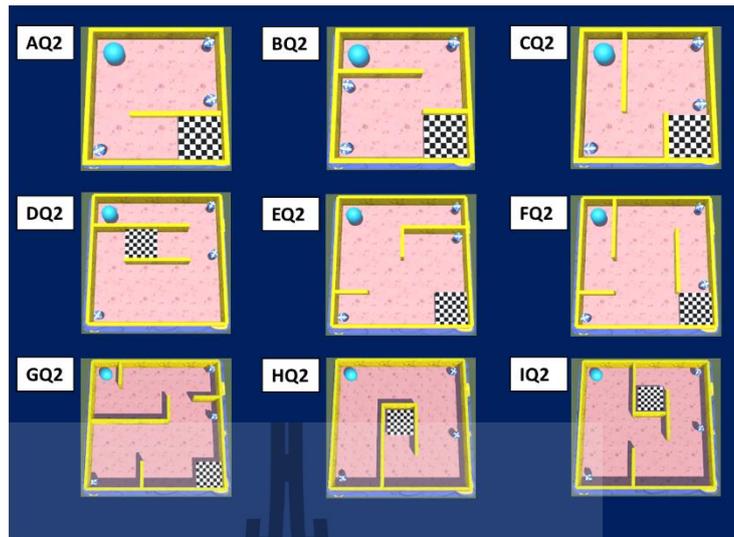


Figure B.1.5 SuraSole maze game level for a sample COP in Quadrant 2

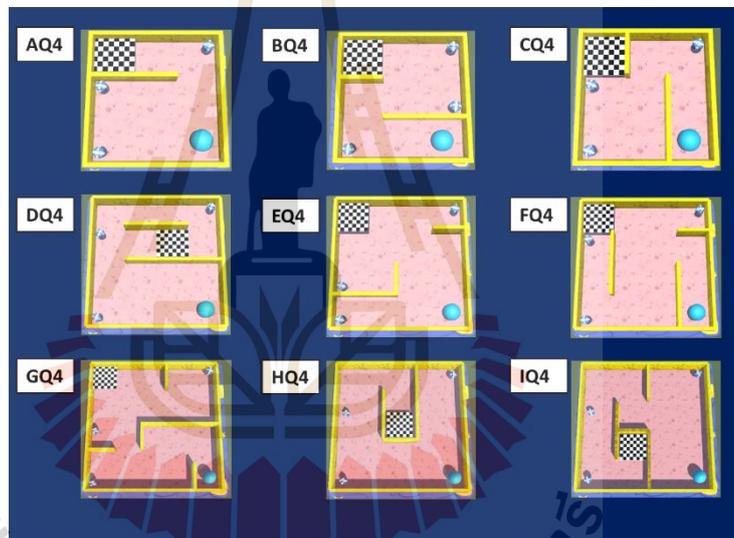


Figure B.1.6 SuraSole maze game level for a sample COP in Quadrant 4

Reward: In this framework, the reward is given as $\text{Reward} = \frac{1}{C}$, where C is the distance between the COP coordinate (X, Y) and the origin, calculated using the formula

$$C = \sqrt{X^2 + Y^2} \quad (\text{B.1.1})$$

where, X is the COP value on the x-axis and Y is the COP value on the y-axis.

Note that the closer (X, Y) is to the origin, the better the balance and the larger the reward which depicts the desired outcome of improved balance.

B.2 Q-learning framework for SuraSole maze game

Q-table: The Q-table stores Q-values for every state-action pair. In this case, we have a total of 5 states corresponding to the 5 defined quadrant. Quadrant 1 to 4 has 9 possible actions (levels) whereas Quadrant 5 has 4 possible actions (levels). The learning Rate (α) is set to $\alpha = 0.001$, The exploration rate is set to 0.8 during training, The discount factor (γ) is set to $\gamma = 0.5$.

Exploration and Exploitation: In the early stages of training, the agent will use exploration to try out different actions. With an exploration rate of 0.8, the agent will randomly choose actions 80% of the time to find the best way to proceed, and 20% of the time, the agent uses exploitation, selecting actions with the highest known Q-values at that particular state.

Data Collection and Processing: Actions and their outcomes are recorded and continuously used to update the Q-table. Training data is processed to improve the accuracy of Q-values, which are then used to select appropriate maze levels for each patient. Block diagram as shown in Figure B.2.1.

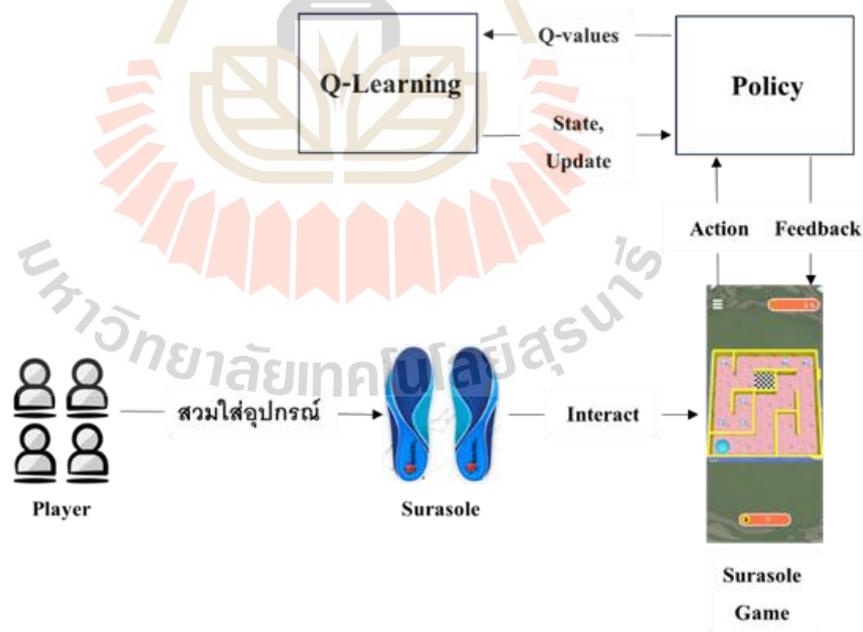


Figure B.2.1 Block diagram of the user interaction with the game

B.3 Model training

Since we initially did not have any patient data to train the model, we simulated data using the state transition diagram technique to simulate the COP transition of patient states.

State Transition Diagram: This diagram is a state-action probability transition model that shows state transitions, consisting of nodes representing different states (Wikipedia, 2024). Here, each node represents a patient's state, and transitions are represented by arrows from one state to another, with each arrow having a probability associated with it, ensuring that the total probability equals 1.

Example of state transitions

When the player is in State 1 or Quadrant 1, state transitions occur as shown in Figure B.3.1.

The state transitions for Quadrants 2, 3, and 4 are similar to those of Quadrant 1.

For Quadrant 5, or the optimal state, state transitions occur as shown in Figure B.3.2.

The goal is to create a game that can select appropriate maze levels for each patient's COP.

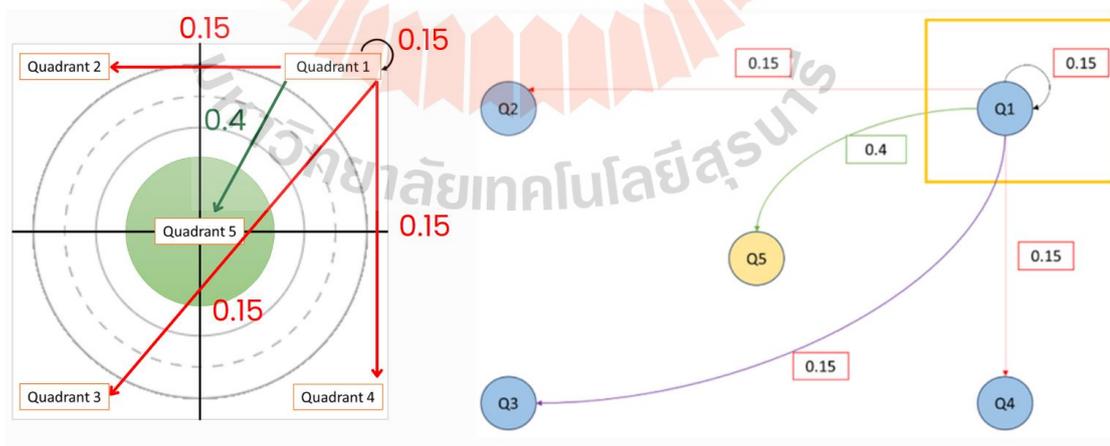


Figure B.3.1 State transitions of Quadrant 1 to various states.

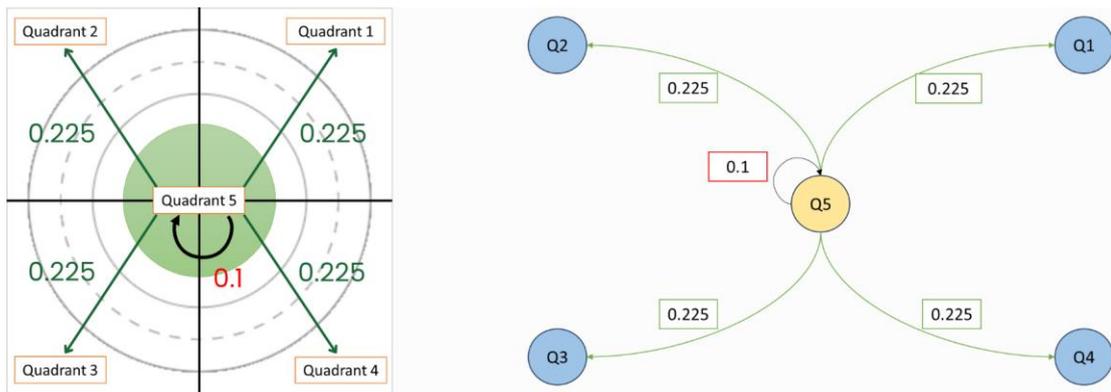


Figure B.3.2 State transitions of Quadrant 5 to various states.

B.4 Model validation

We validated the model's performance by simulating patient COP performance that could potentially help patients reach a favorable state. In this experiment, we assumed that the maze level F is the most suitable level (optimum action) for COPs in Q1, Q2, Q3, and Q4. In particular, level F uses a state transition diagram with a probability of 0.6 for transitioning to a favorable state Q5 and 0.1 for transitioning to other states (Q1, Q2, Q3, and Q4). An example of the state transition diagram for maze level F in the state of Quadrant 1 is shown in Figure B.4.1. For other levels other than level F, the state transitions remain consistent with the original pattern, as shown in Figure B.3.1.

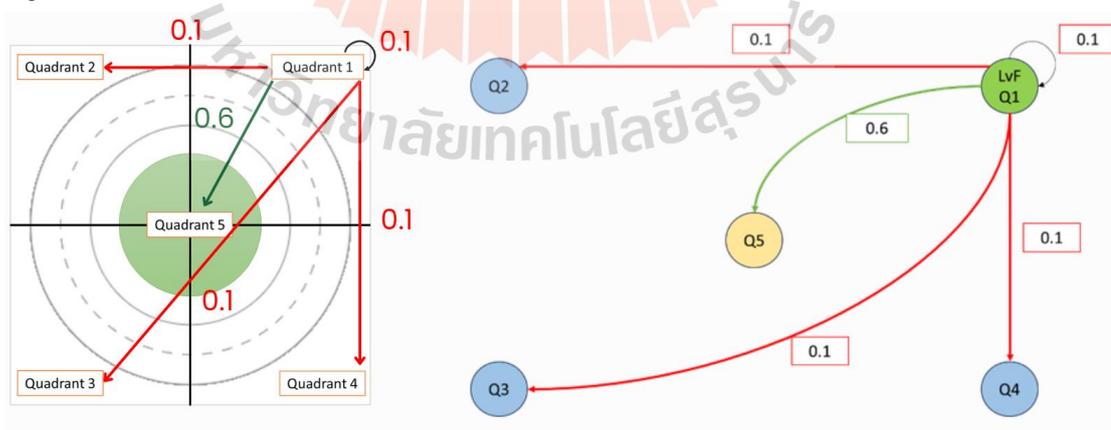


Figure B.4.1 State transitions of maze level F in Quadrant 1 to various states.

To further evaluate the overall performance of the model, we calculated the Average reward per round using Equation (B.4.1), which serves as an indicator of the

model's performance consistency. This Average reward reflects the cumulative success of the model in guiding participants toward favorable states across multiple rounds of testing. The results from the experiment demonstrated that the model could consistently select actions that lead to favorable states i.e. level F, as shown in Figure B.4.2.

$$Avg_Reward = \frac{\sum_{i=1}^n C_i^{-1}}{n} \quad (B.4.1)$$

where C_i^{-1} is the reciprocal of the distance from the center of the COP circle in test round i , which serves as a measure of postural stability. A smaller distance indicates greater stability, n is the total number of test rounds.

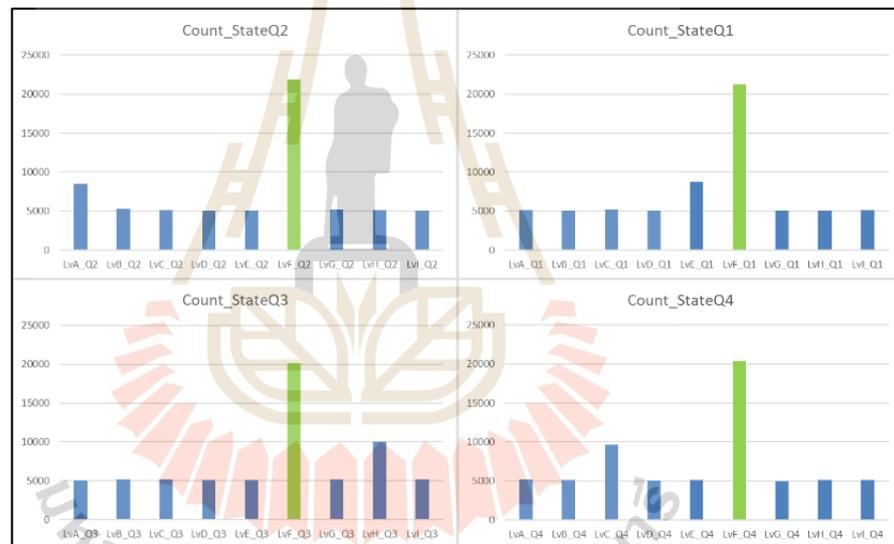


Figure B.4.2 Number of action selections for each state

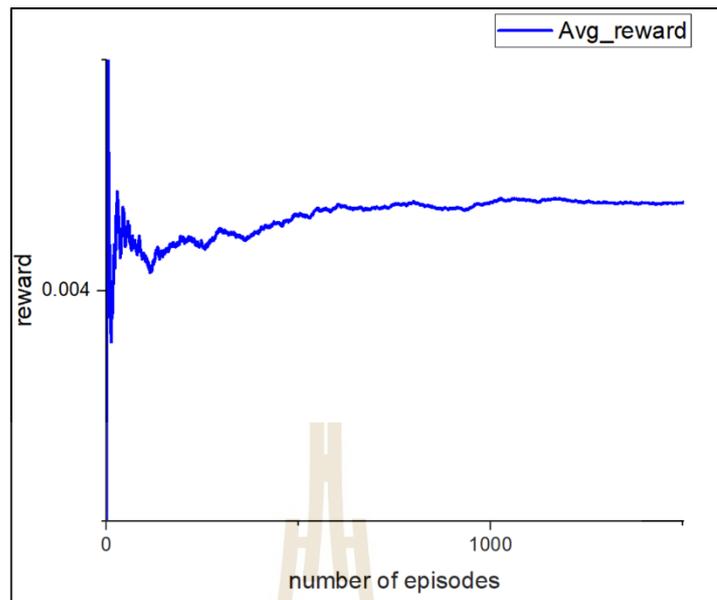


Figure B.4.3 Sample of average reward of the RL model depicting consistent improvement

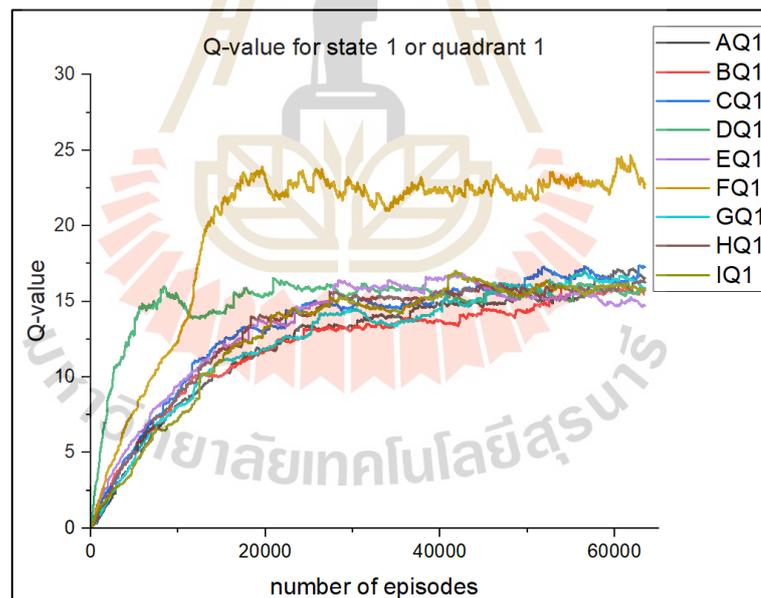


Figure B.4.4 Q-values for state 1 or Quadrant 1

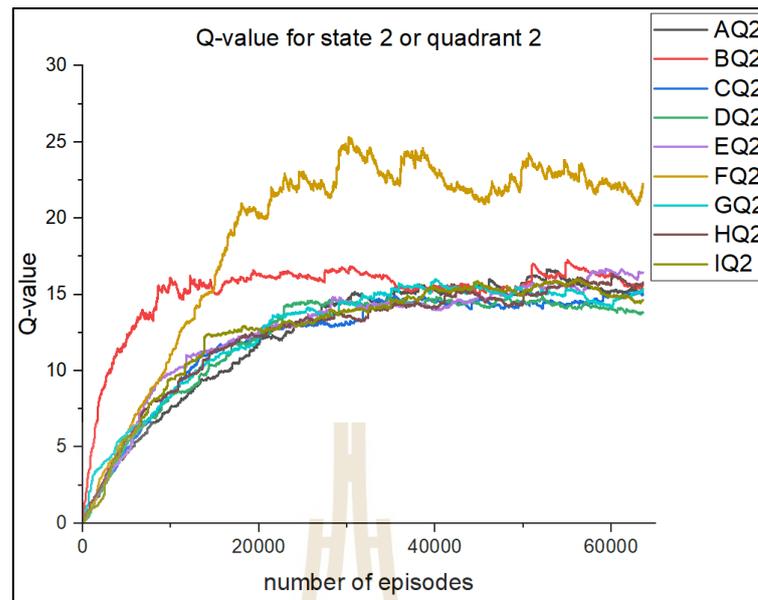


Figure B.4.5 Q-values for state 2 or Quadrant 2

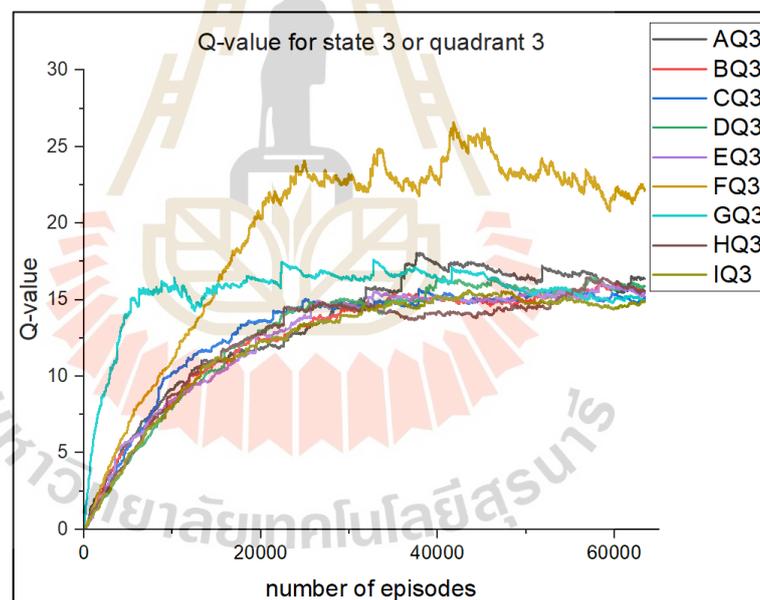


Figure B.4.6 Q-values for state 3 or Quadrant 3

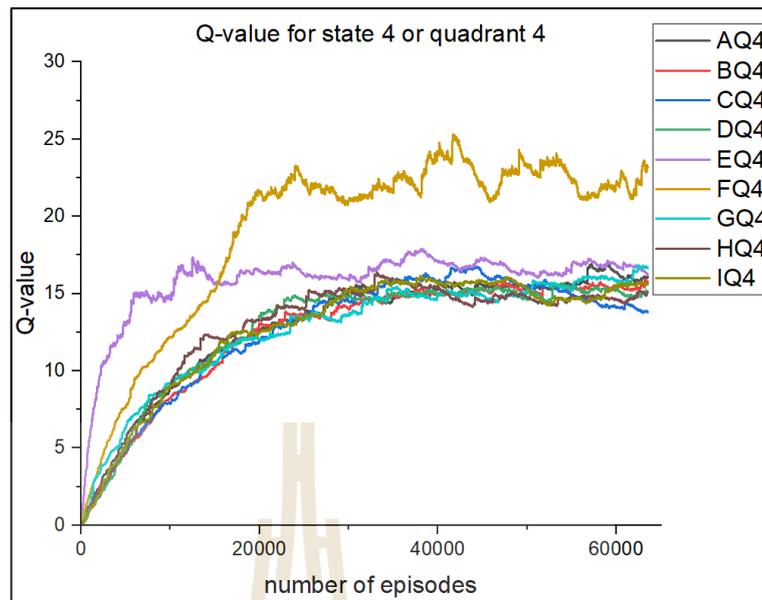


Figure B.4.7 Q-values for state 4 or Quadrant 4

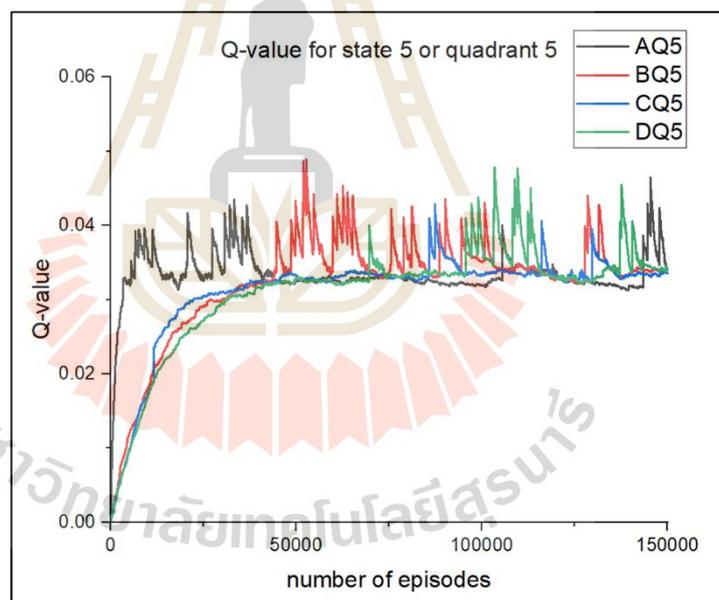


Figure B.4.8 Q-values for state 5 or Quadrant 5

APPENDIX C

Assessment questionnaire

This section compiles information from questionnaires designed to evaluate the opinions and satisfaction of participants across various aspects, including Game safety, Game principles, and the Patient survey. Each section contains questions that address key topics such as the safety of the equipment, the appropriateness of game difficulty levels, the comfort during use, and the motivation to incorporate the game as part of the rehabilitation process. The data gathered from these questionnaires plays a crucial role in analyzing the suitability and effectiveness of the SuraSole maze game, contributing to its further development for more comprehensive and practical applications in the future.

C.1 Patient survey

Patient Survey			
Statement	Agree	Neutral	Disagree
P1) The game was too tiring and did not adjust its difficulty to suit me.			
P2) The exercise (weight shifting) was too easy.			
P3) I would be happy if this game were used more often in my therapy sessions.			
P4) I have no motivation to continue playing.			
P5) The game frustrated me.			
P6) I would play this game at home.			
P7) It was difficult to follow the principles of the game.			
P8) The game was tiring, but manageable.			
P9) I was gradually pushed to my performance limits.			
Remark :			

C.2 Game principles

Game Principles			
Level Creation	Good	Neutral	Needs Improvement
L1) The movement behavior evaluation calculated by the application accurately reflects the patient's balance issues.			
L2) The levels created by the application match the patient's weak directions (not only focusing on the weakest).			
L3) The application adjusts the difficulty level according to the patient's skill progression sufficiently.			
L4) The application allows the physiotherapist to progressively customize the patient's level.			
Game Principles	Good	Neutral	Needs Improvement
S1) The difficulty curve of the game does not exceed the patient's capabilities.			
S2) The patient's progression is clearly visible during the game.			
S3) There is no noticeable delay between weight shifts on the SuraSole and the movements in the game.			
S4) Weight-shifting training with SuraSole is an effective practice for balance rehabilitation.			
S5) I can imagine this game being an additional tool for balance rehabilitation in the future.			
Remark :			

C.3 Game safety

Game Safety
Please rate each statement from 1 (Strongly Disagree) to 3 (Strongly Agree).

- The player feels confident that playing this game is safe and does not pose any harm to the user.
1 2 3
- During gameplay, the player does not feel at risk of accidents or falling.
1 2 3
- The difficulty level of the game does not create any risk of harm to the player.
1 2 3
- The player feels that using the equipment (Surasole Insoles) is safe, does not pose any harm, and does not cause discomfort.
1 2 3
- The player thinks this game is suitable for individuals with balance or mobility issues.
1 2 3
- The player feels that the game does not cause any pain or discomfort during gameplay.
1 2 3
- The difficulty level of the game is not overly pressuring to the point of causing injury to the player.
1 2 3
- The sensor system and equipment connection function safely without causing any harm to the player.
1 2 3
- I feel that the game can adjust its difficulty without negatively impacting the player's physical condition.
1 2 3
- Overall, the player feels that this game is safe to use, does not pose any harm, and is suitable for rehabilitation purposes.
1 2 3

BIOGRAPHY

Mr. Chaidech Angsuwan was born on February 22, 1998, in Mueang District, Nakhon Phanom Province. He completed his high school education at Marywitthaya School in Mueang District, Nakhon Ratchasima Province. He earned his Bachelor's degree in Engineering (Telecommunications Engineering) from Suranaree University of Technology, Nakhon Ratchasima Province, in 2020. In 2022, he pursued a Master's degree in Telecommunications and Computer Engineering at the School of Engineering, Suranaree University of Technology.

