Proposal for a Sensory Integration Self-system based on an Artificial Intelligence Speaker for Children with Developmental Disabilities: Pilot Study

YeJin Wee[1], and OnSeok Lee[1,2,*]
[1] Department of Software Convergence, Graduate School, Soonchunhyang University
22, Soonchunhyang-ro, Asan, 31538, Republic of Korea
[2] Department of Medical IT Engineering, College of Medical Sciences, Soonchunhyang University
22, Soonchunhyang-ro, Asan, 31538, Republic of Korea
[*]Corresponding author: Onseok Lee
[e-mail: leeos@sch.ac.kr]

Received July 5, 2022; revised March 25, 2023; accepted April 1, 2023;
published April 30, 2023

Abstract

Conventional occupational therapy (OT) is conducted under the observation of an occupational therapist, and there are limitations in measuring and analyzing details such as degree of hand tremor and movement tendency, so this important information may be lost. It is therefore difficult to identify quantitative performance indicators, and the presence of observers during performance sometimes makes the subjects feel that they have to achieve good results. In this study, by using the Unity3D and artificial intelligence (AI) speaker, we propose a system that allows the subjects to steadily use it by themselves and helps the occupational therapist objectively evaluate through quantitative data. This system is based on the OT of the sensory integration approach. And the purpose of this system is to improve children's activities of daily living by providing various feedback to induce sensory integration, which allows them to develop the ability to effectively use their bodies. A dynamic OT cognitive assessment tool for children used in clinical practice was implemented in Unity3D to create an OT environment of virtual space. The Leap Motion Controller allows users to track and record hand motion data in real time. Occupational therapists can control the user's performance environment remotely by connecting Unity3D and AI speaker. The experiment with the conventional OT tool and the system we proposed was conducted. As a result, it was found that when the system was performed without an observer, users can perform spontaneously and several times feeling ease and active mind.

Keywords: Remote Therapy, AI speaker, Virtual Reality, Human-Computer Integration, Occupational Therapy.

A preliminary version of this paper appeared in KIPS Fall Conference 2019, November 1-2, Jeju, Republic of Korea. This version includes a concrete analysis and supporting implementation results on proposed sensory integration self-system. This study was supported by the Soonchunhyang University Research Fund, the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) [No. 2022R1A2C1010170] and the BK21 FOUR (Fostering Outstanding Universities for Research) [No. 5199990914048].

http://doi.org/10.3837/tiis.2023.04.010 **ISSN : 1976-7277**

1. Introduction

Developmental disability refers to a condition in which communication and cognitive development are delayed and abnormal, and development is not carried out in accordance with age. Globally in 2016, it is estimated that approximately 52 million children under the age of 5 had developmental disorders including epilepsy, intellectual disability, vision loss, hearing loss, autism spectrum disorder (ASD), and attention deficit hyperactivity disorder [1]. In Korea, as of 2019, the proportion of children with intellectual and ASD was about 0.087% for those under 5 years old, and about 0.78% for those aged 5-14 [2]. In particular, the importance of early diagnosis and treatment of children with developmental disabilities is emphasized because delayed or abnormal motor development at an early age can lead to developmental language and learning problems, attention problems, and poor academic and social skills [3, 4]. Motor development and cognitive development are interrelated; therefore, children with motor disabilities tend to experience cognitive delay and learning disabilities, and children with cognitive disabilities tend to experience motor problems [5-7]. Currently, research on early measures for children with developmental disabilities and for the improvement of their activities of daily living (ADL) continues through motor and cognitive therapy [3, 8-11], among which sensory integration approach (SI) is a that enables them to recognize and organize the senses gained through various physical activities and to acquire strategies for effective use of the body [12]. Children acquire developmental skills and form a sense of satisfaction, mastery, and self-direction [12]. The occupational therapy (OT) based SI stimulates senses such as tactile, vestibular senses, proprioceptive senses, sight, and auditory senses, thereby stimulating the brain nerve to induce normal nerve development [12]. Specifically, the efficient use of hands is a way to stimulate the brain, which promotes brain activity in a person with impaired cognitive function, leading to improvements in function [13]. This is based on neuroplasticity—a theory by which the neural pathways in the brain are reconstructed by external stimuli, experiences, and learning. In addition, OT programs based SI that apply hand motor training can not only improve hand motor function but also ADL [14].

However, it is difficult to identify and collect quantitative performance indicators through OT currently performed in clinics because they are diagnosed with an observer of an occupational therapist. And they are less efficient in terms of time and cost as patients must visit the hospital directly. If patients cannot visit the hospital, it has negative effects such as a worsening of their condition, an increase in the workload of the therapist and burdens of family, so it is required to overcome spatial limitations [15-17]. An alternative diagnostic and treatment method that obtains objective data through Human-Computer Interaction (HCI) devices and virtual reality (VR) is proposed [18-22]. In a study conducted with Leap Motion Controller (LMC, Leap Motion, Inc., San Francisco, CA, USA) by implementing the box and block test (BBT), an OT tool that evaluates hand motor ability in VR, it was confirmed that the conventional BBT and system BBT in a virtual space showed similar performance results, and the objective analysis was possible based on the data collected by tracking the real-time movement [19]. VR helps the user focus on virtual space by eliminating distractions in their surroundings and leaving only the performance environment [23]. This removes spatial constraints so that users can perform in any space and increase their immersion, and good results can be expected in terms of performance. Besides, it is also reasonable in terms of time costs because it can be performed at home without having to visit a hospital. Repetitive training is very important for OT [24, 25]. VR does not run out of resources such as paper and pencils, and preparation is not needed to set up the performance environment; thus, it can be conveniently set up multiple times. In addition, users can see the flow of performance in VR—for example, by checking their movements and interacting with objects—so that they feel a sense of achievement and are motivated to perform repeatedly. However, if a complex OT tool is implemented in a virtual space, the occurrence of many errors due to LMC during OT makes user rather be reluctant to perform and easily lose interest [26]. Unlike reality, in virtual space, user cannot feel the sense of physical contact, so it is difficult for users to interact with objects, and objects that do not properly reflect the physical laws of reality make the user more confused [26]. Therefore, it is important to adopt an OT tool that can be implemented in the virtual space and to modify it to suit the environment.

Moreover, conventional OT is performed by a therapist that assists and observes patients during the treatment. As patients are sometimes afraid of making a negative impression, they tend to only disclose information that will be seen positively [27, 28]. This leads the therapist to have a less accurate assessment of the patients' performance, which can lead to severe results. On the other hand, in the treatment with artificial intelligence (AI) without personality, patients respond more comfortably and honestly [27, 28]. Through treatment with humans, patients may perceive that they are being evaluated and try to make a good impression; therefore, treatment with artificial intelligence can lead to a more honest performance by patients [27, 28]. This could help occupational therapists more accurately diagnose and treat patients.

It is critical to overcome the preceding limitations because OT provides motivation for training as patients check the performance of the treatment and feel a sense of achievement as they perform it repeatedly in OT. Therefore, the purpose of this study was to improve children's ADL by inducing them to stimulate the brain nerve using OT based SI, which allows them to develop the ability to use their bodies effectively using various senses. Also, data on the performance results of users can be accurately and continuously recorded for objective comparison and analysis, allowing for more accurate treatment such as detecting abnormalities of users and performing them intensively. In addition, immediate visual and auditory feedback on performance allows users to perform on their own without being disturbed by external factors.

The main contributions in this paper are summarized as follows:

- It helps the user perform more accurate and active performance by reducing the psychological burden on the observer using AI speaker.
- Users felt more interest and concentration through visual and auditory feedback using the AI speaker and virtual space.
- A quantitative analysis is possible based on objective data obtained using LMC.
- Using AI speaker to observe the user's performance from a distance and change the patient's performance environment to present more free remote medical possibilities.

The rest of this paper is organized as follows. Section 2 presents how to implement OT tool in virtual space and interwork it with AI speaker using a web server. After that, in section 3, the overall performance flow of the proposed system is described, and the various feedback the user will receive and how the user and occupational therapist uses the AI speaker is explained. Also, an experiment using this system and conventional OT tool is conducted, and the results are analyzed in section 3. And, Section 4 summarizes and discusses this paper, and this paper is concluded in section 5.

2. Materials and Methods

This study has a system design, as shown in **Fig. 1**. Users perform OT using an LMC and receive visual feedback in virtual space. In addition, the web server can interwork with the virtual space and the AI speaker to exchange data. This allows users to communicate visual and auditory feedback of performance.

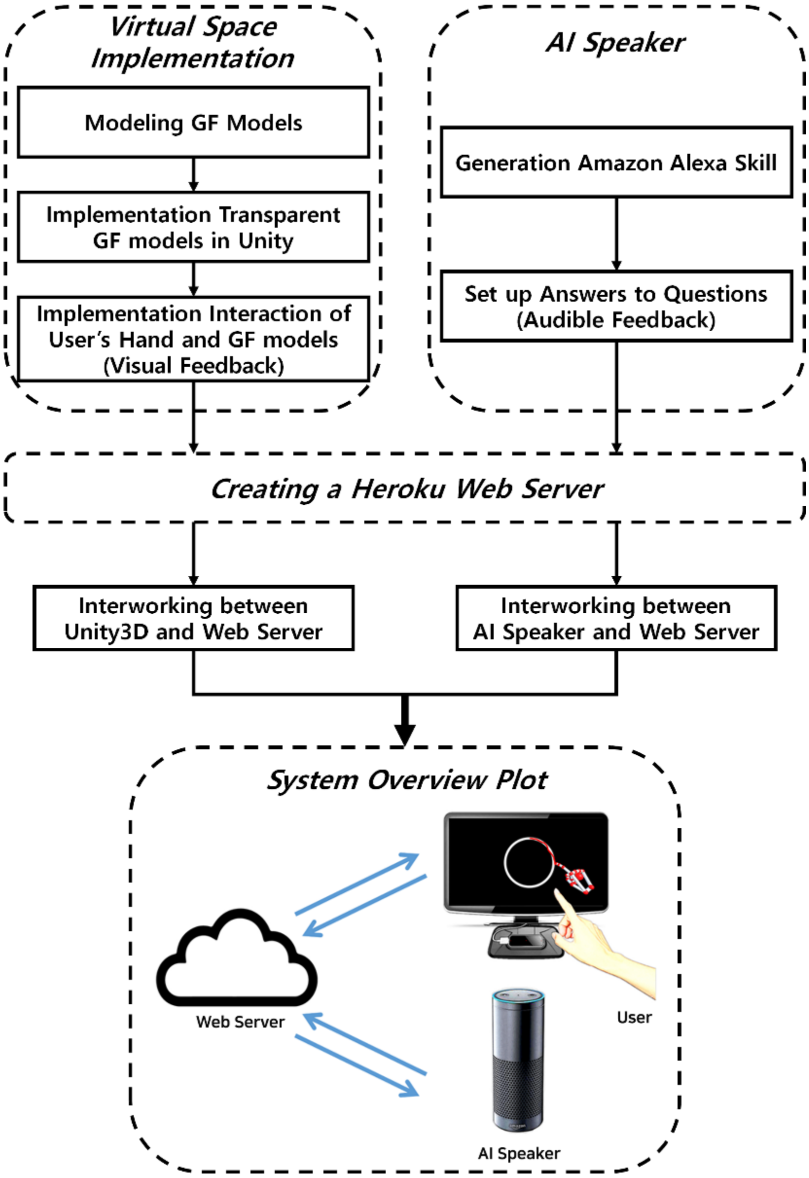


Fig. 1. Design of a sensory integration self-system based on an artificial intelligence speaker for children with developmental disabilities.

2.1 System Implement Methods

2.1.1 OT Tool: Dynamic Occupational Therapy Cognitive Assessment for Children (DOTCA-ch)

In this study, the tool of dynamic occupational therapy cognitive assessment for children (DOTCA-ch), which is currently used in clinics, is implemented and performed in virtual space. DOTCA-ch is a tool that is applied to children based on the Lowenstein Occupational Therapy Cognitive Assessment (LOTCA), a cognitive evaluation tool for adults with neuropsychological disabilities [29].

Traditional standardized cognitive assessments are inherently static, in that they only judge performance outcomes for diagnosing and quantifying cognitive impairment; however, this static assessment does not meet the goal of cognitive assessment that assessment of intelligence should be an assessment of learning ability. On the other hand, DOTCA-ch is a dynamic assessment in which an intervention process is performed to assess not only performance outcomes but also the children's learning ability in the performance process. Dynamic assessment refers to the difference between what children can do independently and what they can do with help and guidance from others, and the intervention process is a catalyst for learning that helps children's inner awakening as well as a process in which the therapist directs specified tasks for improvement of cognitive ability. In this process, children understand the types of information essential to complete the task and can achieve learning simultaneously within the assessment process [29].

DOTCA-ch consists of 22 subtests according to five areas of cognition: orientation, spatial perception, praxis, visuomotor construction, thinking operations. In this study, we selected copy geometric forms (GF), one of the subtests in the visuomotor construction area, which is the ability to organize and compose movements through visual stimulation. Five are the models of GF: circle, triangle, rhombus, complex, and cube (**Fig. 2**). According to the developmental milestones, it is specified that children can copy in the order of circle, square, triangle, rhombus, and more complex figures in the domain of the fine motor among the development domains[30, 31]. From this, we can see that children tend to be more difficult to copy diagonal lines rather than horizontal or vertical lines, and also they feel difficult to copy shapes made of lines or curves rather than simple lines. Also, **Table 1** represents the specified criteria for each model [29]. The size of the model is not considered, and no eraser is provided. If the patient wants to modify a model, it must be redrawn from the beginning.

The assessment of the visuomotor construction includes the process of before mediation, immediate memory, mediation, after mediation, delayed memory, and the scores and time it takes are recorded accordingly. In this study, assessment using GF was designed and applied so that it does not lose the characteristics of OT in clinics and can be properly performed with a computer.

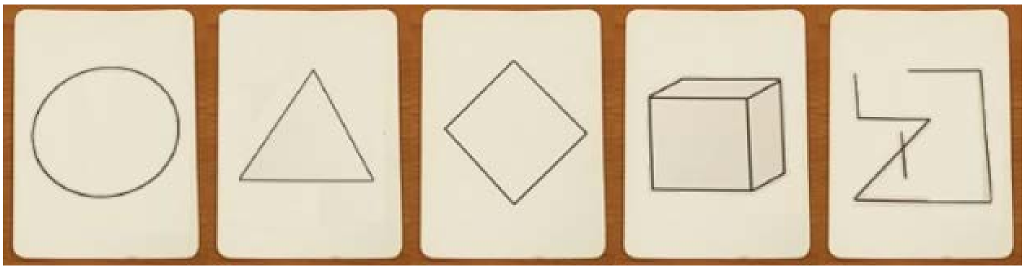


Fig. 2. Five models of copy geometric forms (GF Models) in DOTCA-ch: circle, triangle, rhombus, cube, and complex.

Table 1. Criteria for GF Models specified by DOTCA-ch.

[1]GF models	Criteria
Circle	The ratio of height and width in each quadrant should not exceed 2:1.
Triangle	Three distinct sides. One corner is higher than the others.
Rhombus	Four distinct corners. A difference of less than $\frac{1}{16}$ of an inch is permitted. The four corners have an angle of 90°.
Complex	All angles are equal to the sample **Fig. 2**. The diagonal is split by a short vertical line.
Cube	All sides are connected at the correct point, including three diagonal lines.

[1]GF: Copy geometric forms of DOTCA-ch

2.1.2 3D Modeling of GF Models and Implementation of Virtual Space

Models of GF were modeled using Solidworks 2014 (Dassault Systems SolidWorks Corp., MA, USA). Because the original GF model is a two-dimensional figure, it has been properly modified and configured to be implemented in a three-dimensional virtual space. When modeling these, we referred to the standard of models presented in DOTCA-ch and ensured that the model did not deviate from the geometrical form. In addition, it was modeled considering the thickness of each side of the model and the thickness of the front and rear (z-axis), so that the user does not incur problems during the performance.

Modeled files were converted to .obj files that were available in Unity 2018.2.4. (Unity3D, Unity Technologies, San Francisco, CA, USA) and implemented in Unity3D virtual space (**Fig. 3**). Because the size of the model is not taken into account by the assessment criteria, each model is displayed solely on the screen, so that sufficient hand motion can be performed to the extent that it does not exceed the normal operating range of the LMC. Afterward, experiments were carried out to modify these models to the appropriate size and position. Additionally, the models were modeled transparently so that users could check whether their hands were inside or outside the model and the traces of the movement path.

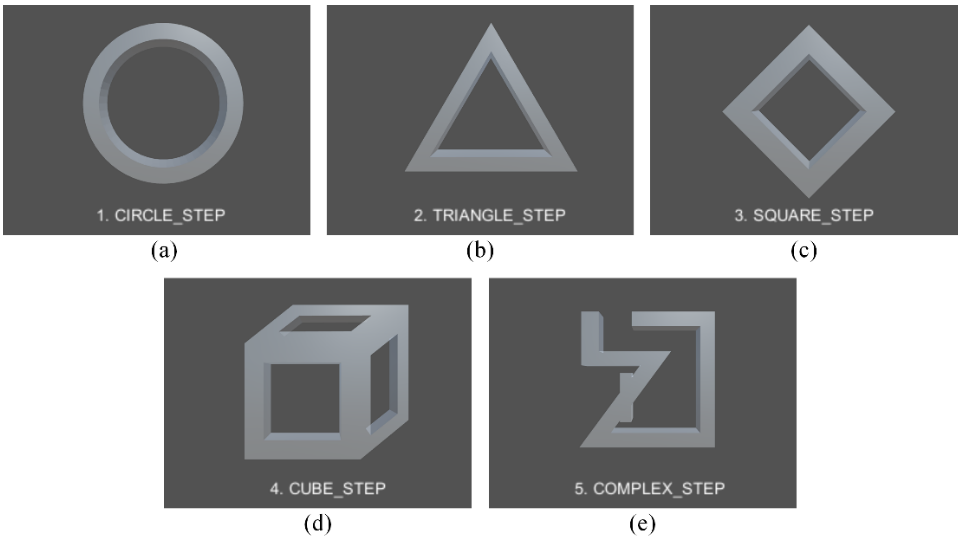


Fig. 3. GF Models implemented transparently in Unity3D to determine the user's hand position and path of movement: (a) is a circle step; (b) is a triangle step; (c) is a rhombus step; (d) is a cube step; (e) is a complex step expressing a three-dimensional cube in a two-dimensional picture in GF.

We used Unity 2018.2.4. (Unity Technologies, San Francisco, CA, USA) and LMC Orion 4.0.0 as tools for the performance.

Unity3D is a game engine that provides an intuitive graphical user interface (GUI) and a development environment of 2D and 3D [23]. Currently, it is actively used in various fields such as 3D animation, VR, and augmented reality (AR). It supports various platforms, so the scope of development environments such as mobile, web, and console is very wide and uses C# and JavaScript as a development language. In this study, it was used to visually represent the GF models of DOTCA-ch and the user's hand motion data tracked by the LMC on the screen. It also provides real-time information about the performance, such as the time of performance, the name of the GF models, and the number of times a user's hand is out of the model.

The LMC is a device that tracks the trajectory and movements of the hands very accurately and is distinguished from Kinect, a well-known full-body motion tracking device that precisely tracks hands in a small space [32, 33]. The LMC consists of two RGB cameras and three IR cameras, and image data obtained using LMC devices is performed image processing in the Leap Motion Service software, and consequently is recon structed to 3D construction and the hand's trajectory are extracted [19, 20]. The device can detect movement of 0.01mm, enabling recognition of finger gestures and motion in virtual space [34]. Users can identify their hands in virtual space with the LMC, and all data during the performance were tracked and recorded in real time.

In this system, we display the user's movements and visually represent the real-time path movement in the Unity3D virtual space using its API that supports the LMC [35, 36]. Based on the data from LMC and Unity3D, visual feedback is provided to users. In addition, to make the performance smoother, we made it possible to move on to the next/previous step or initialize the movement's path by taking a specified gesture. First, when the thumb, index, and little finger were extended at the same time if the gestured hand was the right hand, the next step was carried out, else if the left hand, the previous step was carried out. At this time, the Extended Finger Detector was used to check whether the corresponding finger was extended, and the SceneChange script was implemented so that it could cross the steps. When users want to perform the step again, they can initialize their movement's path by facing both hands upward. At this time, the direction of the palm can be checked using the Palm Direction Detector, and the step is reset using the Detector Logic Gate.

2.1.3 Communication between AI speaker and Unity3D

This system enables them to exchange data with each other and give correct feedback to the users accordingly as they perform by creating a web server for interworking between the Unity3D and AI speaker.

We used the Heroku service (salesforce.com, Inc., San Francisco, CA, USA) to build a web server for interworking between the Unity3D and AI speaker. Heroku is a platform-as-a-service (PaaS) cloud service that supports a variety of programming languages and is easy to install and deploy, making it easy to host a website. We used Node.js to upload and read commands. To access the web server created earlier and import data to it, Unity3D repeatedly runs the coroutine that accesses the web server. According to commands uploaded on the web server, Unity3D adjusts the performance environment accordingly. This command is uploaded by the AI speaker and is executed in real time because it is checked in Unity3D every frame.

We used Amazon's Echo Plus (Amazon.com, Inc., Seattle, WA, USA), which is the most proficient in natural language processing and can be developed as an open source. AI speaker is a device that applies artificial intelligence, which allows the user to naturally communicate with it by learning the intention and patterns of conversation. AI speakers made by Amazon define it as a single skill to create the intended dialogue environment. It consists of a function specifying examples of a sentence that a user might question and the terms that receive the data to be saved and providing a corresponding answer. These functions are the settings of the Alexa Developer Console. The automation function must be set in the AWS Lambda of Amazon Web Services for the AI speaker to run automatically as a setting in the Alexa Developer Console. AWS Lambda is a serverless computing service that allows code to run automatically without a server [37]. The function in AWS Lambda is connected as a trigger to the skill of the AI speaker, and the performance is coded accordingly. In addition, commands are delivered from the AI speaker to the web server.

2.2 System Conducting Methods

2.2.1 System Flow

In conventional DOTCA-ch, patients look at the five GF models, copy them identically, and then receive interventions for incorrect performance by an occupational therapist. However, in this study, we replaced the intervention process conducted by a therapist with that provided visual and auditory feedback using Unity3D and AI speaker so that the user can correct the wrong performance on their own (**Fig. 4**). Because performance and mediation are fulfilled simultaneously, they can immediately correct the wrong performance. The five GF models were drawn one by one in order. To initiate the performance, the user has to maintain a pinch gesture by which the thumb and index finger meet, and the remaining fingers stretch out; when the gesture is released, the saving of coordinate data and of the hand's path is stopped. Only thumb coordinate data with a pinch gesture are tracked and saved.

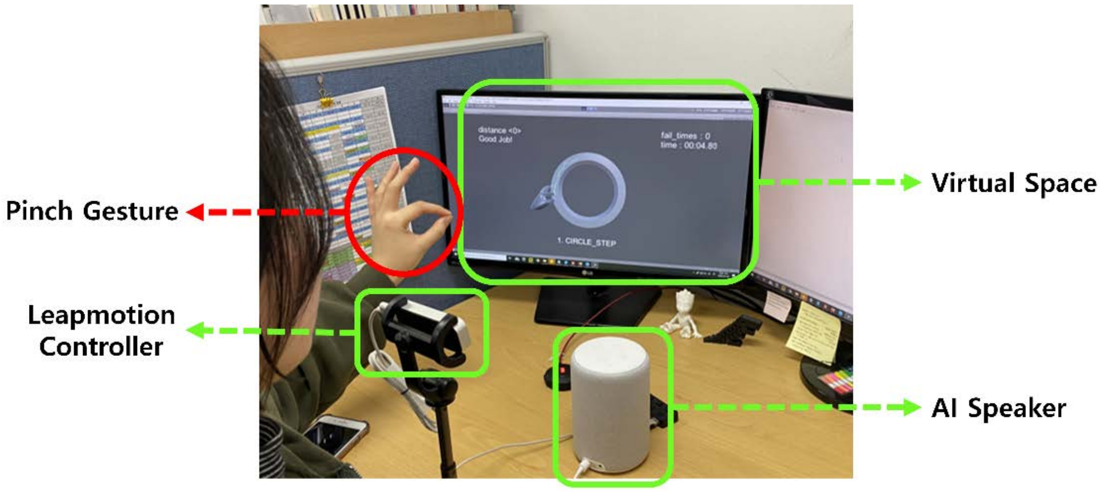


Fig. 4. Main components to operate this system.

Three are the main types of feedback obtained during the performance. First, the distance from the GF model and the number of times the user's hand is out of it are displayed in text on the screen. The second is that the GF model has been set up transparently, so it is possible to check whether the user's hand is the back of the model or at the front. Third, the movement path of the user's hand in the Unity3D virtual space is displayed in real time. If the user's hand is inside the GF model, it is displayed in green; if it is outside, it is displayed in red to immediately notify the wrong performance.

This system flow can be summarized in three steps as follows:

1. Before starting, the user first asks the AI speaker how to perform the system and comprehends it.
2. The user keeps the pinch gesture and draws along without leaving the GF model. At this time, visual feedback on incorrect performance is provided for the intervention. If the user makes a mistake, he can make the specified gesture or ask the AI speaker to initialize the performance's path and redraw it.
3. After completion of drawing all five GF models, the user requests the AI speaker to show the upper, lower, left, and right sides of the model to check the performance result.

2.2.2 Communication with Virtual Space

This system provides various feedback to assist users' performance, enabling them to check their own performance path in virtual space and obtain information on correct performance through color feedback and transparent characteristics of the model. In addition, the AI speaker provides intervention for difficult performance as well as the role of the occupational therapist who directs the performance, so that dynamic assessment can take place. This allows users to perform themselves autonomously and independently.

When users perform, path markers are displayed to indicate the movement's path that they draw along the GF model with the pinch gesture. Only when users make a pinch gesture, the path markers are displayed, and the tracked trajectory is saved. To intuitively indicate whether the user is drawing along the GF model properly, the path marker is colored green when the user's hand is inside the model and red when it is outside to help the user correct the position of the hand (**Fig. 5**). We used the Collider.ClosestPointOnBounds() method provided by Unity3D to measure the distance between the GF model and the hand object. This method measures the distance between the specified object and the nearest object that meets the condition. The Physics.OverlapSphere() method was used to determine if the hand object was inside the GF model. It detects objects within a certain range around a particular point, which we set where the thumb meets the index finger. We programmed a user's communication in virtual space using the two methods mentioned earlier in this system.

(a) (b)

Fig. 5. Color feedback indicating which side of the GF model the hand is on: (a) shows positive visual feedback with green color, indicating when the hand is inside the GF model; (b) expresses negative visual feedback with red color, indicating when the hand is outside the GF model.

Users can view performance results from the top, bottom, and right sides of the GF model to check how far their hand movements deviate from the model (**Fig. 6**). It could help the user and the therapist determine where the user's performance is skewed. The command read from the web server is divided into words to determine its purpose. In the case of a command that checks the performance in a different direction of the GF model, the camera object in Uniy3D is positioned in a certain direction and turns toward it. After assigning the location value for the Transform.position of the camera object, let it turn toward the model using the Transform.LookAt() method. The coordinates indicating how the camera object moves and looks at are different depending on the direction that the user requests; thus, they should be specified accordingly.

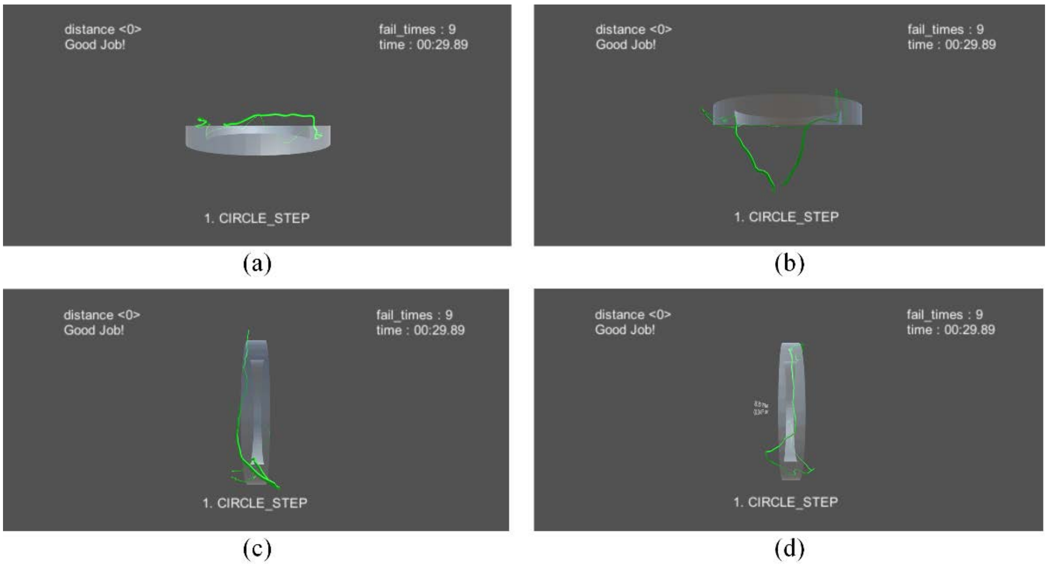


Fig. 6. Views of the GF model from different directions for user to check their own performance results: (a) shows the upper side of the model; (b) shows the lower side of the model; (c) shows the left side of the model; (d) shows the right side of the model.

2.2.3 Interaction with AI speaker

Also, this system answers users' questions through the AI speaker to help them perform smoothly and receives requests for performance. The user can ask various questions to the AI speaker and receive answers about how to perform overall and how to adjust the environment of the virtual space through an AI speaker. The user should request the AI speaker to check performance results from the top, bottom, right, and left aspects. In order to verify how the path of hand movement is skewed, the user should ask the AI speaker to show performance results on different sides of the GF model. When users request feedback from the AI speaker, it uploads their requests to the web server, and Unity3D reads the requests uploaded to the web server to provide the user with visual feedback.

2.3 Method of the Comparison of the System and Conventional OT tool

2.3.1 Experimental Setup

Eight participants (1 man, 7 women; age: 25 ± 2.35 years) with no problems to upper body and cognitive ability conducted the experiments, and they are familiar with an observer. Before experiment, they practiced enough to get used to system for about 10 minutes. First of all, the participants performed the existing OT tool in the presence and absence of an observer, respectively, and then the system proposed in this paper was also conducted twice in the same way. In each situation with and without an observer, participants performed the conventional OT tool and the system proposed in this paper, and conducted a questionnaire to ask questions about what they felt. The questions include whether the system operated stably for users to perform, whether the user was able to perform with ease and an active mind depending on the presence of an observer. And it also includes a questionnaire asking which method has a higher interest and concentration level when comparing the conventional OT tool with the system. This study was conducted after receiving approval from the Research Ethics Committee (1040875-202104-SB-029), and the research consent form that subjects signed.

2.3.2 Statistic Analysis Method

First, the stability of the system was measured based on the responses that users felt about how stable they were able to perform the system. Stability is questioned by dividing into how reliably the system, where the role of LMC plays a large role, recognizes the user's hand, and whether the system, where the role of the AI speaker plays a large role, properly recognizes and performs user commands. Also, a chi-square test was performed to find out whether the user's ease and active mind in performance vary according to the presence or absence of an observer. And, to analyze the interest and concentration level that users feel comparing the conventional OT tool and system, we compared the frequency % of user's responses to each method. At this time, SPSS version 21.0 (IBM, New York, NY, USA) was used as a tool for analysis of results.

3. Results

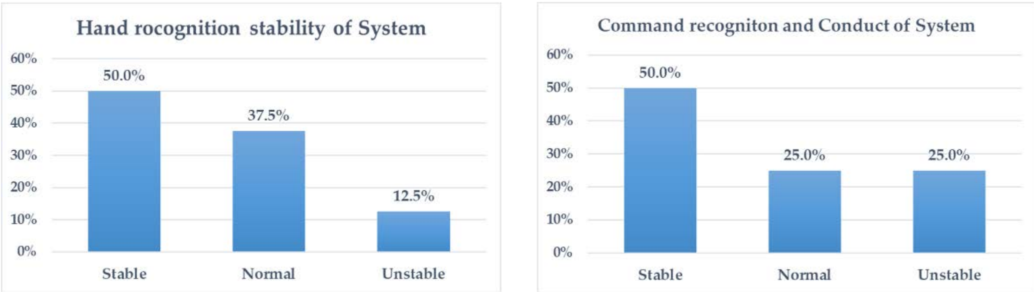
3.1 The evaluation and comparison of the System

3.1.1 The stability of the System

Fig. 7 shows the percentage of answers to how stable the performance of the system feels after users it.

In **Fig. 7a**, it indicates the degree to which the user feels whether the system reliably recognizes the user's hand, and the users who responded as 'stable or normal' accounted for 87.5% of the total, and the system reliably recognizes the user's hand. Thus, it was found that it helped the user to perform OT smoothly. For 12.5% of the respondents who said they were 'unstable' in the survey, they said it was difficult to perform while maintaining the pinch gesture, which is a gesture for performance, and they also said it was difficult to recognize the spatial sense of Unity.

In **Fig. 7b**, it shows that to the extent that users feel that the system reliably recognizes and performs their commands, 75.0% of the users responding as 'stable or normal' felt that the system reliably recognized the user's request and performed what they wanted. However, users who responded as 'unstable' accounted for 25.0% of the respondents, indicating that quite a few users felt unstable about the user's demand recognition and performance stability. Users who responded that it was unstable said that AI speaker did not properly recognize the user's pronunciation and that it did not process the request at once and only conducted the user's command with requesting several times. This indicates that there is a need for improvement in the recognition problem of AI speakers themselves, and the interworking between the web server and Unity3D.



(a) (b)

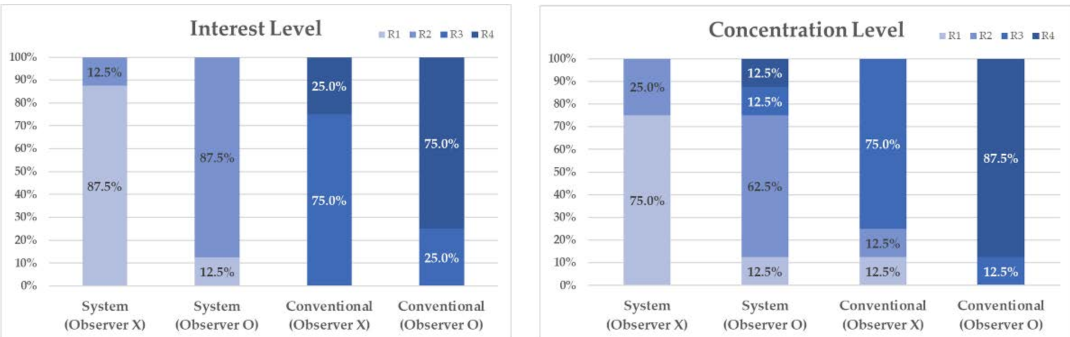
Fig. 7. User's responses to the stability of the system: (a) shows the percentage of responses to how well the system recognizes the user's hand; (b) shows the percentage of responses to how well the system recognizes the user's request and performs it properly.

3.1.2 The Comparison of the Interest and Concentration

To determine which of the four methods users felt most interested in and were able to perform with concentration, frequency analysis was performed on the results of users' responses to them.

First, looking at the interest level in performance (**Fig. 8a**), overwhelmingly many users answered that it was most interesting to perform the system without an observer, and there was a tendency to answer that the system was more interesting than the conventional OT tool. As for the reason that they responded that they performed more interestingly when performing the system without an observer, they answered that it was burdensome to perform multiple times satisfactorily because the observer was next to them and that it was embarrassing for the observer to watch from the side. Also, they answered that the reason for giving low ranking to the conventional OT tool was that it was a simple performance, so it was less interesting and that they participated passively.

Regarding the concentration level on performance (**Fig. 8b**), the response that performing the system without an observer was able to perform with more concentration than other methods was overwhelmingly high, and it was performed more intensively when performing the system than the conventional OT tool. However, some users ranked 1st and 2nd in performing the conventional OT tool without an observer, and for this reason, the method was easy, so they answered that they were more focused.



(a) (b)

Fig. 8. Ranking frequency graphs of interest and concentration that users felt in the four experiments: R1, R2, R3, and R4 in the legend are preference rankings given by users and represent the 1st, 2nd, 3rd, and 4th rank, respectively; (a) shows the ranking frequency of the interests users felt while experimenting; (b) shows the ranking frequency of the concentration users felt while experimenting.

3.2 The Comparison of the Ease and Active Mind with the Conventional Method

Fig. 9 shows the rank frequency graphs of the ease and active mind felt by users in four experiments. The proportion of the 1st is the same as the system and conventional without an observer, but if checking the proportion of the 2nd, it can be seen that the system without an observer takes up a higher proportion. Therefore, it was found that users were able to actively carry out the performance with a more ease mind when performing with the system without an observer. Also, comparing the system and conventional OT tool with an observer, it can be seen that users prefer the system more because the proportion of the system is higher when the proportion of the 1st and 2nd combined is considered.

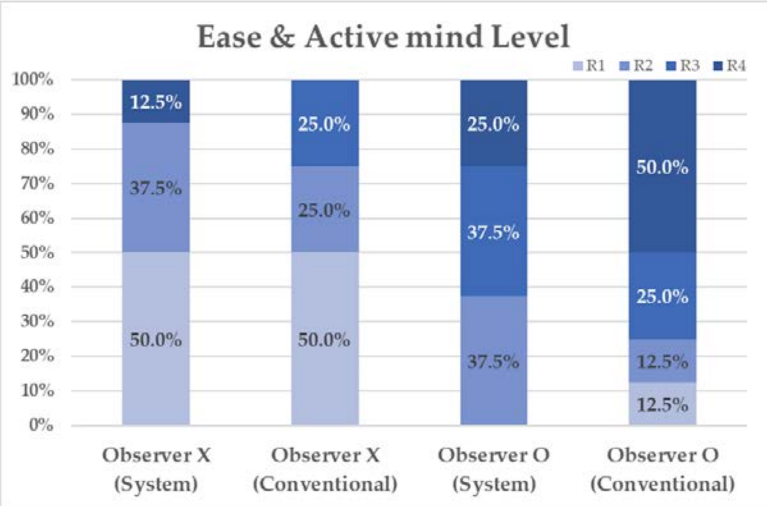


Fig. 9. Ranking frequency graphs of the ease and active mind that users felt in the four experiments: R1, R2, R3, and R4 in the legend are preference rankings given by users and represent the 1st, 2nd, 3rd, and 4th places, respectively.

A chi-square test was performed to analyze the relationship between the presence/absence of an observer and the user’s ease and active mind when performing OT. The null hypothesis was set as ‘The user's ease and active mind will be the same regardless of the occupational therapist’s residence when performing OT. As a result, there is a statistically significant difference in the level of ease and active mind the user feels in the performance between the conventional OT tool with an observer and the system without an observer, and between the system with an observer and conventional OT tool without an observer.

In more detail, it is concluded that regardless of the performance of the conventional OT tool and the system, the null hypothesis was adopted ($p>.05$) when divided only by the observer’s presence or absence, ’The user’s ease and an active mind are the same regardless of the occupational therapist's residence’. However, considering the performance of the conventional OT tool and system, the null hypothesis was rejected for performing between the conventional OT tool with an observer and the system without an observer ($p<.05$). Also, the null hypothesis was rejected for performing between the system with an observer and conventional OT tool without an observer ($p<.05$).

4. Discussion

In this study, we proposed a system that enables children with developmental disabilities to perform repetitive treatments on their own by using Unity3D and LMC, performing treatments in virtual space that overcome space and time limitations. This system implemented the GF model of DOTCA-ch—a dynamic OT tool for children with developmental disabilities used in clinics—in the virtual space and allows users to draw along it. We aim to enhance the ADL of children with developmental disabilities by providing hand movement and visual and auditory feedback of performance with OT based SI, which stimulates the brain nerve through the various senses inducing normal nerve development. Furthermore, an AI speaker was used to reduce the user's performance burden according to the observation when performing the treatment, and the occupational therapist as well as the user can control the user's performance environment on Unity3D through interworking with the web server.

In conventional clinical practice, not all patients can be provided with exactly the same explanation and performance environment. However, this system can provide objective assessment because it obtains and evaluates the user's motion data in virtual space in real time and provides a consistent performance environment for the user by interworking an AI speaker with Unity3D. And, in dynamic assessment, the mediation process that which the therapist points out and corrects wrong performance is very important, and this system provides various feedback that enables users to immediately identify and correct the wrong performance at the same time. In addition, continuous repetition is important for OT, and patients may not feel the progress of their treatment and become tired of it. Nevertheless, this system makes patients repeat their performance as they can feel a sense of achievement and interest in seeing their progress directly on the screen. This was the same result in the survey conducted in this paper. Users answered that performing with VR was more interesting than the conventional OT tool of drawing on paper, and this is because it checked the drawing performances in real time, so they were able to perform more actively and harder.

This system has the potential for data errors due to mechanical limitations. The LMC is a device specialized in finger gestures and motion that allows finger tracking very accurately, but with one LMC, there is an unrecognizable part because it only observes the user's movements in one direction, which causes occlusion. In this study, however, since the user maintains the pinch gesture with the palm of one hand facing the screen, the occlusion problem is acceptable rather than tasks that need both hands are used and the trajectory of the hands is complex. Moreover, the accuracy of the pinch gesture performance can be used to assess users' performance. In addition, in this study, to minimize errors caused by LMC, an OT tool that does not require to implement the physical laws of reality was applied so that there is no problem performing it in virtual space, and feedback with objects was implemented to make it easier for the user to interact with OT tool objects. And, Haptic feedback, one of the challenges of virtual space, is an important part that determines the degree of user immersion, so we will study how to address it. And we provide directional feedback on incorrect performance as text on the monitor rather than the voice. Because of this, users are required to check the texts on the screen while paying attention to the movement of the hand at the same time. So, their concentration can be distracted, which can lead to deviation of motion data. In the future, it is thought that it can lead to more correct performance by allowing the user to focus only on drawing along the model if we improve this and add a function that give feedback on the directions of movement to users by voice and encourage them.

Another limitation is that the GF subtest of DOTCA-ch was originally performed for users to look at and copy the models; however, in this study, users draw along the models, and the procedures for dynamic assessment in the order of performance-mediation-performance were summarized. Thus, an experiment is needed to verify whether this system is effective in rehabilitation despite these transformations, though this study did not conduct one. The main purpose of this paper is to propose a system that eliminates the performance burden of having to make a good impression on users and to allow occupational therapists to realize remote rehabilitation by interworking an AI speaker with virtual space. Therefore, in this paper, we emphasized the contents of this and conducted experiments based on this, and as a result, users answered that it was the most convenient and active way to use this system in the absence of an observer. For this reason, users answered that they were able to try various attempts they wanted to try because there was no observer and that the psychological burden of performing multiple times was reduced. However, some users asked they can't know whether they were doing well because there was no one to help them. It was found that the role of supporting and encouraging the patient, like an occupational therapist, is an important factor in leading the correct performance, and an AI speaker should be able to perform this role as well. In addition, some users replied that it is more comfortable and active to perform the conventional OT tool in the presence of an observer, and for that reason, they answered that drawing on paper is easier because the system needs to be drawn not only the shape of the GF model but also the 3D spatial sense. This is because the rejection of unfamiliar methods was felt more than the presence/absence of an observer, and it is considered to be the result of feeling a burden of a 3D spatial sense. Rather, this is considered to be a part that suggests the possibility of expanding it to treatment related to 3D spatial cognition.

In this study, it was applied to occupational therapy with a relatively simple sequence. However, it is expected that the positive effect on introverted users will be further maximized if they perform tasks that require higher concentration and active activity. Later, An experiment to verify the therapeutic effect of the system will be conducted, and a method to quantitatively assess the results will be devised in subsequent studies. In addition, an objective assessment method with numerical performance data obtained through this system will be established, and the progress of treatment will be analyzed with the data accumulated through repeated performance. And, it will be possible that the occupational therapist predicts other users' further progress or use it to establish the treatment direction if data from multiple users are gained, since the user's hand motion coordinate data during a performance is recorded and saved in real time. Also, this study will be applicable to other rehabilitation tools related to the upper extremities and cognition as it gathers and analyzes the three-dimensional movement of the user's hand, which is drawn along the instruction model. Depth data may be analyzed later to enable the diagnosis and treatment of the three-dimensional space as well as the recognition of the two-dimensional model.

5. Conclusion

In this study, an AI speaker and virtual space were used to create a system that allows users to steadily perform on their own to enhance their ADL. We aimed to overcome the problems of conventional OT such as much information loss, inconsistent treatment environment, and time and space limitations. And we focused on users' ability to perform the treatment honestly because they feel burdened by the presence of observers. As a result, users answered that using the system in the absence of an observer is suitable for more active and repetitive performance when performing, and it was confirmed that it occupies the highest ranking in terms of interest and concentration.

This system is intended to be used as an aid to assist occupational therapists and presented to compensate for the shortcomings of existing methods. By using an AI speaker and virtual environments, patients can perform treatment regardless of space, and occupational therapists are able to control their occupational environment through an AI speaker. Through this, the patient can take an assertive attitude toward treatment without the burden of observers, and the therapist can check the accumulated patient's data and establish future treatment directions. It may also lead to an extension of research on a patient's language ability through an AI speaker.

References

[1] B. O. Olusanya, A. C. Davis, D. Wertlieb, N.-Y. Boo, M. Nair, R. Halpern, H. Kuper, C. Breinbauer, P. J. De Vries, and M. Gladstone, "Developmental disabilities among children younger than 5 years in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016," *The Lancet Global Health*, 6(10), pp. e1100-e1121, 2018. Article (CrossRef Link)

[2] KOSIS, "Number of The Registered Disabled-by Age, Types of Disability and Gender (The Whole Country):2007-2019," *Statistics Korea*, 2020. [Online]. Available: https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT_11761_N003&conn_path=I2&language=en

[3] H. Kim, A. G. Carlson, T. W. Curby, and A. Winsler, "Relations among motor, social, and cognitive skills in pre-kindergarten children with developmental disabilities," *Research in developmental disabilities*, vol. 53-54, pp. 43-60, 2016. Article (CrossRef Link)

[4] M. MacDonald, and L. L. McIntyre, "The relationship of age, early motor skills and observable child behaviors in young children with developmental delays," *Research in developmental disabilities*, vol. 93, pp. 103445, 2019. Article (CrossRef Link)

[5] A. Diamond, "Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex," *Child development*, 71(1), pp. 44-56, 2000. Article (CrossRef Link)

[6] E. Hartman, S. Houwen, E. Scherder, and C. Visscher, "On the relationship between motor performance and executive functioning in children with intellectual disabilities," *Journal of Intellectual Disability Research*, 54(5), pp. 468-477, 2010. Article (CrossRef Link)

[7] S. Houwen, L. Visser, A. van der Putten, and C. J. R. i. D. D. Vlaskamp, "The interrelationships between motor, cognitive, and language development in children with and without intellectual and developmental disabilities," *Research in Developmental Disabilities*, vol. 53-54, pp. 19-31, 2016. Article (CrossRef Link)

[8] E. I. Blanche, M. C. Chang, J. Gutiérrez, and J. S. Gunter, "Effectiveness of a sensory-enriched early intervention group program for children with developmental disabilities," *The American Journal of Occupational Therapy*, 70(5), pp. 7005220010p1-7005220010p8, 2016. Article (CrossRef Link)

[9] C. R. S. Araújo, A. A. Cardoso, and L. d. C. Magalhães, "Efficacy of the cognitive orientation to daily occupational performance with Brazilian children with developmental coordination disorder," *Scandinavian Journal of Occupational Therapy*, 26(1), pp. 46-54, 2019. Article (CrossRef Link)

[10] L. Anderson, J. Wilson, and G. Williams, "Cognitive Orientation to daily Occupational Performance (CO-OP) as group therapy for children living with motor coordination difficulties: An integrated literature review," *Australian occupational therapy journal*, 64(2), pp. 170-184, 2017. Article (CrossRef Link)

[11] I. Y. Sung, J. Y. Jeon, K. J. Yun, J. S. Yuk, E. M. Byun, H. W. Yoo, T. S. Ko, H. W. Kim, and D. H. Jang, "Development of tablet personal computer-based cognitive training programs for children with developmental disabilities whose cognitive age is less than 4 years," *Medicine*, 99(2), pp. e18674, 2020. Article (CrossRef Link)

[12] B. L. Jeong, "The Effects of Improving Social Skills with Structured Sensory Integration Program on Children with Developmental Disabilities," *Korean Journal of Converging Humanities*, 7(1), pp. 79-99, 2019. Article (CrossRef Link)
[13] Y. J. Lee, H. J. Kim, H. I. Kim, K. J. Jung, and Y. A. Yang, "The correlation of board game performance according to cognitive function and hand dexterity," *The Journal of Occupational Therapy for the Aged and Dementia*, 11(2), pp. 19-28, 2017. Article (CrossRef Link)
[14] H. H. Kim, B. G. Hwang, and B. K. Yoo, "The Effects of the Hand-function Training Applied Sensory Integration Group Treatment Program on the Hand-function and ADL ability of Children with Developmental Disability," *Korean Journal of Physical and Multiple Disabilities*, 54(1), pp. 127-143, 2011. Article (CrossRef Link)
[15] H. J. Smaling, B. Tilburgs, W. P. Achterberg, and M. Visser, "The Impact of Social Distancing Due to the COVID-19 Pandemic on People with Dementia, Family Carers and Healthcare Professionals: A Qualitative Study," *International journal of environmental research and public health*, 19(1), pp. 519, 2022. Article (CrossRef Link)
[16] Y. Suarez-Balcazar, M. Mirza, V. L. Errisuriz, W. Zeng, J. P. Brown, S. Vanegas, N. Heydarian, D. Parra-Medina, P. Morales, and H. Torres, "Impact of COVID-19 on the Mental Health and Well-Being of Latinx Caregivers of Children with Intellectual and Developmental Disabilities," *International Journal of Environmental Research and Public Health*, 18(15), pp. 7971, 2021. Article (CrossRef Link)
[17] S. Zhang, Y. Hao, Y. Feng, and N. Y. Lee, "COVID-19 Pandemic Impacts on Children with Developmental Disabilities: Service Disruption, Transition to Telehealth, and Child Wellbeing," *International journal of environmental research and public health*, 19(6), pp. 3259, 2022. Article (CrossRef Link)
[18] Á. Aguilera-Rubio, A. Cuesta-Gómez, A. Mallo-López, A. Jardón-Huete, E. D. Oña-Simbaña, and I. M. Alguacil-Diego, "Feasibility and Efficacy of a Virtual Reality Game-Based Upper Extremity Motor Function Rehabilitation Therapy in Patients with Chronic Stroke: A Pilot Study," *International Journal of Environmental Research and Public Health*, 19(6), pp. 3381, 2022. Article (CrossRef Link)
[19] S. Kim, S. Park, and O. Lee, "Development of a Diagnosis and Evaluation System for Hemiplegic Patients Post-Stroke Based on Motion Recognition Tracking and Analysis of Wrist Joint Kinematics," *Sensors*, 20(16), pp. 4548, 2020. Article (CrossRef Link)
[20] E. B. Kim, E. Y. Kim, and O. Lee, "Meta-Analyzing the Writing Process of Structural Language to Develop New Writing Analysis Elements," *Applied Sciences*, 10(10), pp. 3479, 2020. Article (CrossRef Link)
[21] C. Li, L. Cheng, H. Yang, Y. Zou, and F. Huang, "An Automatic Rehabilitation Assessment System for Hand Function Based on Leap Motion and Ensemble Learning," *Cybernetics and Systems*, 52(1), pp. 3-25, 2020. Article (CrossRef Link)
[22] P. Fernández-González, M. Carratalá-Tejada, E. Monge-Pereira, S. Collado-Vázquez, P. S.-H. Baeza, A. Cuesta-Gómez, E. D. Oña-Simbaña, A. Jardón-Huete, F. Molina-Rueda, and C. B.-B. de Quirós, "Leap motion controlled video game-based therapy for upper limb rehabilitation in patients with Parkinson's disease: a feasibility study," *Journal of neuroengineering and rehabilitation*, 16(1), pp. 1-10, 2019. Article (CrossRef Link)
[23] K. N. Fong, Y. TANG, S. Karen, K. Andy, C. Cherry, and W. Yuko, "Effects of Task-Specific Virtual Reality Training Using a Leap Motion Controller on Hemiparetic Upper Extremity in Patients with Stroke," *Research Square*, 2020. Article (CrossRef Link)
[24] Y. Yu, M. Kodama, H. Matsuwaki, K. Taniguchi, S. Matsumoto, H. Yamanaka, I. Fukuda, M. Shimodozono, and K. Kawahira, "Research of Rehabilitation Device for Hemiplegic Knee Flexion Based on Repetitive Facilitation Exercise," in *Proc. of International Conference on Intelligent Robotics and Applications*, Springer, Cham, pp.158-167, 2016. Article (CrossRef Link)
[25] J. Kim, and J. Yim, "Effects of high-frequency repetitive transcranial magnetic stimulation combined with task-oriented mirror therapy training on hand rehabilitation of acute stroke patients," *Medical science monitor: international medical journal of experimental and clinical research*, vol. 24, pp. 743-750, 2018. Article (CrossRef Link)

[26] S. N. Gieser, C. Gentry, J. LePage, and F. Makedon, "Comparing objective and subjective metrics between physical and virtual tasks," in *Proc.of International Conference on Virtual, Augmented and Mixed Reality*, Springer, Cham, pp. 3-13, 2016. Article (CrossRef Link)
[27] G. M. Lucas, J. Gratch, A. King, and L.-P. Morency, "It's only a computer: Virtual humans increase willingness to disclose," *Computers in Human Behavior*, vol. 37, pp. 94-100, 2014. Article (CrossRef Link)
[28] J. Gratch, G. M. Lucas, A. A. King, and L.-P. Morency, "It's only a computer: the impact of human-agent interaction in clinical interviews," in *Proc. of the 2014 international conference on Autonomous agents and multi-agent systems*, pp.85-92, 2014. Article (CrossRef Link)
[29] N. Katz, S. Golstand, R. T. Bar-Ilan, and S. Parush, "The Dynamic Occupational Therapy Cognitive Assessment for Children (DOTCA–Ch): a new instrument for assessing learning potential," *The American Journal of Occupational Therapy*, 61(1), pp. 41-52, 2007. Article (CrossRef Link)
[30] R. J. Scharf, G. J. Scharf, and A. Stroustrup, "Developmental Milestones," *Pediatrics in review*, 37(1), pp. 25-38, 2016. Article (CrossRef Link)
[31] R. J. Gerber, T. Wilks, and C. Erdie-Lalena, "Developmental milestones: motor development," *Pediatrics in review*, 31(7), pp. 267-277, 2010. Article (CrossRef Link)
[32] L. Bo, C. Zhang, H. Cheng, and B. Baoxing, "Fingertip data fusion of Kinect v2 and leap motion in unity," *Ingenierie des Systemes d'Information*, 23(6), pp. 143-159, 2018. Article (CrossRef Link)
[33] M. Iosa, G. Morone, A. Fusco, M. Castagnoli, F. R. Fusco, L. Pratesi, and S. Paolucci, "Leap motion controlled videogame-based therapy for rehabilitation of elderly patients with subacute stroke: a feasibility pilot study," *Topics in stroke rehabilitation*, 22(4), pp. 306-316, 2015. Article (CrossRef Link)
[34] M. Mohandes, S. Aliyu, and M. Deriche, "Arabic sign language recognition using the leap motion controller," in *Proc.of 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, pp. 960-965, 2014. Article (CrossRef Link)
[35] W. Ding, X. Zhu, B. Xu, Y. Xu, K. Chen and Z. Wan, "The Interactive Modeling Method of Virtual City Scene Based on Building Codes," *KSII Transactions Internet and Information Systems*, 15(1), pp. 74-89, 2021. Article (CrossRef Link)
[36] S. Lee, M. Hong, S. Kim and S. J. Choi, "Effect Analysis of Virtual-reality Vestibular Rehabilitation based on Eye-tracking," *KSII Transactions on Internet and Information Systems*, 14(2), pp. 826-840, 2020. Article (CrossRef Link)
[37] H. Kim, H. Kim, K. Chun and Y. Chung, "Experience in Practical Implementation of Abstraction Interface for Integrated Cloud Resource Management on Multi-Clouds," *KSII Transactions on Internet and Information Systems*, 11(1), pp. 18-38, 2017. Article (CrossRef Link)

Yejin Wee received the B.S. degree and M.S. degree in Medical IT Engineering and Software Convergence from the Soonchunhyang University, Asan, Korea, in 2020 and 2022, respectively. Her current research interests include data science, medical image processing and computer vision.

OnSeok Lee received the B.S., M.S., and Ph.D. degrees from Korea University, in 2005, 2007, and 2011, respectively. Currently, he is a Professor in Medical IT Engingeering, Soonchunhyang University, Asan, Korea.