Collaborative School Mental Health System: Leveraging a Conversational Agent for Enhancing Children's Executive Function

Doeun Park, Ph.D.
HCI Lab, Yonsei University Seoul, South Korea
doeun14@naver.com

Myounglee Choo
HCI Lab, Yonsei University Seoul, South Korea
myounglee.choo@gmail.com

Minseo Cho
HCI Lab, Yonsei University Seoul, South Korea
pcat11@yonsei.ac.kr

Jinwoo Kim, Ph.D.
HCI Lab, Yonsei University Seoul, South Korea
jinwoo@yonsei.ac.kr

Yee-Jin Shin, M.D., Ph.D.
Department of Child and Adolescent Psychiatry, College of Medicine, Yonsei University Seoul, South Korea
yjshin@yuhs.ac

ABSTRACT

Attention deficit hyperactivity disorder (ADHD) is a common childhood psychiatric disorder. Schools can play a vital role in the early detection and treatment of mental health issues. However, stigma and fear regarding mental health often prevent schools from engaging in active interventions. ADHD is characterized by deficits in executive function, a critical contributor to children's self-directed behavior. We developed a conversational agent to assist children in planning and accomplishing daily tasks, with the aim of enhancing their executive function. We also designed supportive systems for both parents and teachers, proposing a collaborative school mental health system that incorporates various stakeholders. Through practical implementation with first-graders, this study confirmed the system's potential to improve structured living and symptoms among children with ADHD. Surveys involving parents and teachers confirmed that the application improved executive function and reduced inattention. Therefore, we suggest an enhanced mental health support system.

CCS CONCEPTS

• **Human-centered computing** → Human computer interaction (HCI); Empirical studies in HCI.

KEYWORDS

Attention Deficit Hyperactivity Disorder (ADHD), Executive Function, Conversational Agent, School Mental System

ACM Reference Format:
Doeun Park, Ph.D., Myounglee Choo, Minseo Cho, Jinwoo Kim, Ph.D., and Yee-Jin Shin, M.D., Ph.D.. 2024. Collaborative School Mental Health System: Leveraging a Conversational Agent for Enhancing Children's Executive Function. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3613904.3642593

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642593

1 INTRODUCTION

The prevalence of childhood attention deficit hyperactivity disorder (ADHD) is approximately 7.2%, and it is the most common mental illness experienced by children during the prepubertal period [43, 56]. ADHD is characterized by primary features of inattention, hyperactivity, and impulsivity. Most ADHD cases are diagnosed after children start attending elementary school [51]. In many cases, ADHD is associated with problematic behaviors that can be spotted in the patient's school life [6, 26]. During the transition from kindergarten to elementary school, children undergo drastic changes in their lives [18]. Specifically, children need to adapt to a somewhat controlled environment in which they are asked to remain seated during class or may be allowed to use the washroom only at certain times [15]. However, children diagnosed with ADHD, which is characterized by defects in executive functions, tend to find it difficult to adapt to such controlled environments [45, 55]. Children with ADHD not only have difficulty concentrating in class but also disturb other students by making unusual noises or even interrupting class by rolling on the floor.

A key classifier of ADHD is a lack of executive function [7, 56]. Executive function refers to the ability to appropriately maintain a problem-solving mechanism to achieve future goals [29]. Well-developed executive function, in turn, can have a positive influence in developing an ability to flexibly control one's mental representations when acting in goal-oriented ways [25]. In other words, children become able to plan, organize, control, and regulate their own actions, beginning to get better at managing their lives moving into adulthood. It is also important for children other than those diagnosed with ADHD to effectively control and cope with themselves in various areas of daily lives and at school [27].

Today, children are familiar with a range of online devices, such as smartphones and tablets [24, 32]. Moreover, since the outbreak of COVID-19, online devices have become more common in our lives and are widely used in many different virtual interactions, such as non-face-to-face classes and medical appointments. The prevalence of technology use increased by 15% during the COVID-19 pandemic [1]. Learning services through tablets have become popular, and children are now required to study autonomously in an online space where they have easy access to fun distractions with a single tab on the screen. Therefore, for them to become more proactive in their lives, it is important to cultivate their executive function.

Schools can offer a valuable platform to promote mental health, prevent mental health issues, and support children facing emotional difficulties [39]. Parents may not be aware of their child's ADHD symptoms and may think of them as a child's growth process; therefore, they may not suspect ADHD symptoms in their child. However, teachers can recognize children with ADHD in better than their normative peers [34]. Therefore, teachers may play an important role in the ADHD diagnosis process. Nevertheless, teachers may find it burdensome to recommend that parents have their children's mental health assessed. Because mental illnesses still have a negative perception, parents might not accept their children's illness and may be conscious of the social stigma attached to their children [52, 53]. Hence, there is a compelling need to develop a novel school-based child mental health support system to address these challenges.

The distribution of prototypes in schools can greatly enhance the accessibility for students and parents from the perspective of applying screening or intervention [9, 31]. This paper proposes a new strategy for enhancing students' executive function using their school as a platform. We are currently conducting a long-term iterative study on a conversational agent that can help improve executive function in children. This study has children as subjects who will carry out their daily tasks by communicating with an agent that assists them in setting up necessary daily tasks by themselves. This study aimed to determine whether the developed conversational agent can help improve children's executive function. In particular, we aimed to determine whether it improves their executive function or relieves existing ADHD symptoms. An additional goal was to present a child mental health support system that involves schools and listening to various stakeholder opinions. Through this process, we aimed to confirm the usability and feasibility of the proposed system.

Twenty schools participated in this study, and participants were recruited from among the first-grade elementary school students. The participants were provided with an application featuring a conversational agent that they could use for six weeks. Additionally, the parents of the children were provided access to a monitoring application. School teachers were also provided with a monitoring web page to oversee the prototype usage of students in their respective classes. The research was designed as a pre-post study with intervention and control groups, and a total of 132 children completed the study. In this study, we observed positive changes in the children's executive function and ADHD symptoms after using the prototype. It also enabled the discovery of a mental health system through collaboration between schools and families that was useful and feasible for a large group of children.

2 RELATED WORKS

2.1 Executive Function and Daily Tasks Performed by Children Themselves

Executive function is defined as self-directed behavior, including goal selection and prolonged self-control towards the achievement of goals, which is widely covered not only in the clinical field but also in the field of child development [33, 48]. Executive function enables children to maintain their level of concentration, retain pertinent objectives, resist impulses, endure frustration they might face, weigh the outcomes of their actions, and be prepared for the future [42]. This aligns with the discovery that children with strong executive function often exhibit enhanced academic performance [41]. However, beyond academic performance, executive function is generally considered a vital ability to possess for people to successfully achieve goals in various situations they may encounter in the course of their lives [50].

Studies using brain imaging methods have shown that executive function is deeply related to the frontal lobe and continues to develop and change until adolescence [30, 50]. Executive function is known to develop most rapidly from early childhood through the lower grades of elementary school, although the timing of development in each area can vary [8, 10].

To improve executive function, it is important to allow children to develop their own routines that are made out of their autonomous decisions and for them to get used to those routines. First, children should be able to control their daily lives willingly. A study of children aged 6–7 years showed that children had higher self-directed executive functioning when they performed activities that they chose to do compared to when instructed by their parents [23]. Therefore, children must experience the entire process of managing and performing activities based on their decisions. Second, children should become accustomed to it through repetition. High levels of practice, repetition, child engagement, and real-world generalization are considered some of the most important factors in interventions aimed at improving executive function [2].

Therefore, in this study, a prototype was developed to enable elementary school children to independently set their own goal behaviors, perform these behaviors, and evaluate the outcomes. The aim was to promote the enhancement of children's executive function through interaction with this prototype.

2.2 The Need to Introduce a New Management System Involving Schools for Children with ADHD

ADHD prevalence is known to be 5–7% worldwide [20]. However, in a study using Korea's National Health Insurance database, less than 1% of children were diagnosed or treated, confirming that less than 1/5 were receiving treatment [40].

Schools can play a valuable role in the screening and treatment of ADHD. First, teachers can suspect that a student is an ADHD patient more quickly than anyone else because ADHD symptoms can differ somewhat from those of typically developing children [44]. Teachers work with many children and experience various cases [4]. Therefore, teachers can play a significant role in screening children for ADHD, thereby leading to early treatment.

Second, schools can also play a crucial role in the treatment of children with ADHD. In the United States, medication and school support are identified as treatments applied to 2/3 children [38]. In this study, 66.9% answered that they were taking medication and 64.7% were receiving school support. In particular, in the United States public education system, Individualized Education Programs (IEPs), or 504 Plans, are provided to children with ADHD, and 51.6% of them receive such education [12]. Such programs have made significant contributions to ADHD treatment.

Third, the primary benefit of school mental health programs is that it increases student participants' accessibility and reduces the chance that they will feel attacked by negative stereotypes [9, 31]. Despite significant improvements in awareness of mental illness, there are still fears that other people are judgmental and biased towards certain mental health issues [5]. Having a relevant program as part of universal school practice is expected to reduce the level of concern.

Thus, school-based support systems could play a significant role in the diagnosis and treatment of ADHD. Furthermore, schools can have a significant impact on children with ADHD, as well as on those with broader mental health concerns, throughout the diagnostic and treatment process.

Therefore, this study applied a research design that considered children who may have ADHD but have not yet been diagnosed. Additionally, the intervention applied in this study, which targeted children with ADHD, also demonstrated benefits for typically developing children.

3 PROTOTYPING

In this study, we used three distinct software applications, each tailored to address the specific requirements of its respective user group. These applications include an Android-based program equipped with a conversational agent to aid children in their daily tasks, a mobile application designed for parental monitoring, and a webpage intended for teacher management (Figure 1). The overall structure involved providing the child with an application equipped with a conversational agent and allowing parents and teachers to monitor and support the child's application usage.

3.1 Conversational Agent for Children's Daily Tasks

This study is based on voicebot interaction research previously conducted as a tangible product targeting children with ADHD [13, 14]. The voicebot had a physical presence that increased its effectiveness in children with ADHD. However, its limited versatility made it challenging to conduct experiments with a large number of participants. In this study, we integrated a chatbot builder, NAVER CLOVA, into an Android application that is compatible with tablet PCs to develop a conversational agent. In other words, we redeveloped an application that allows voice interaction. This prototype is an application that helps children start and complete their daily tasks, which they set up with their parents (Figure 1 (a)). The daily task list was designed to address the overall daily tasks of children in lower elementary school grades. There are a total of 18 daily tasks, including waking up in the morning, brushing teeth after dinner, preparing bags, and reading books.

In previous studies, self-instruction and behavioral parenting training have been incorporated into the interaction process to assist children with ADHD from a cognitive-behavioral therapy perspective [13]. Cognitive-behavioral therapy enables children to develop strategies for monitoring and managing their own executive function [11]. In particular, self-instruction involves acquiring strategies for self-regulation, which has been found to be effective in regulating behavior in children's daily lives and learning domains [58]. Behavioral parenting training is widely recognized as one of the most effective interventions [54]. An exemplary case involves the use of tokens to reinforce positive behaviors and reduce negative behaviors in children [37].

The application developed for this research was designed using a "Goal, Plan, Do, and Check" process, primarily based on self-instructional steps [16, 21]. This process enables children to 1) set their own task goals, 2) receive reminders about these goals, 3) implement them, and 4) confirm their progress. Rewards were provided to children through a token economy system. When a child completed their daily tasks, they received a star sticker, and upon successfully completing all tasks for the day, the star sticker was replaced with a character sticker. In addition, a character card was also given (Figure 1 (b) and (c)).

When impaired executive function impairs the process of regulating behavior through internally represented forms of information, obtaining externalized forms of information is helpful [46, 47]. Notably, the physical representation of information must be externalized when performing a task [47]. Therefore, we determined that voice interaction might be necessary for children with ADHD. The prototype tablet PC application delivers both visual and auditory cues. This prototype uses cues to convey information about a child's situation.

A conversational agent, which has widely used functions such as ordering goods and weather guidance, is structured to appropriately answer its user's question [60, 64]. In our prototype, the typical relationship between the user and the chatbot is reversed. This means that ForME was designed to communicate with its users by prompting them to respond with several command words with limited intent, so that conversations can be conducted in a rule-based manner. The core interaction of this prototype enables child users to individually perform several actions one after the other.

Children would normally deliver a speech with four main intentions [14]. In some situations, dialogue may be determined by three or five different intents, depending on the context. First, there is the correct answer section, where the child has completed a certain task successfully and proceeds to the next task. When the current activity has not been performed completely or is performed incorrectly, the second section is defined as in-progress. In the third section, the unable-to-process section, ForME is unable to process the child's response because of a lack of clarity in the child's speech or accent. The final section, or the no-response section, occurs when there is no response from the child at all. The children's responses were organized such that the prototype could maintain communication with them based on their responses.

These tasks can be performed in two ways. First, the agent ForME provides a sound alarm at a preset time, saying "What should we be doing right now?", so that the child can fulfill their set goal. Second,

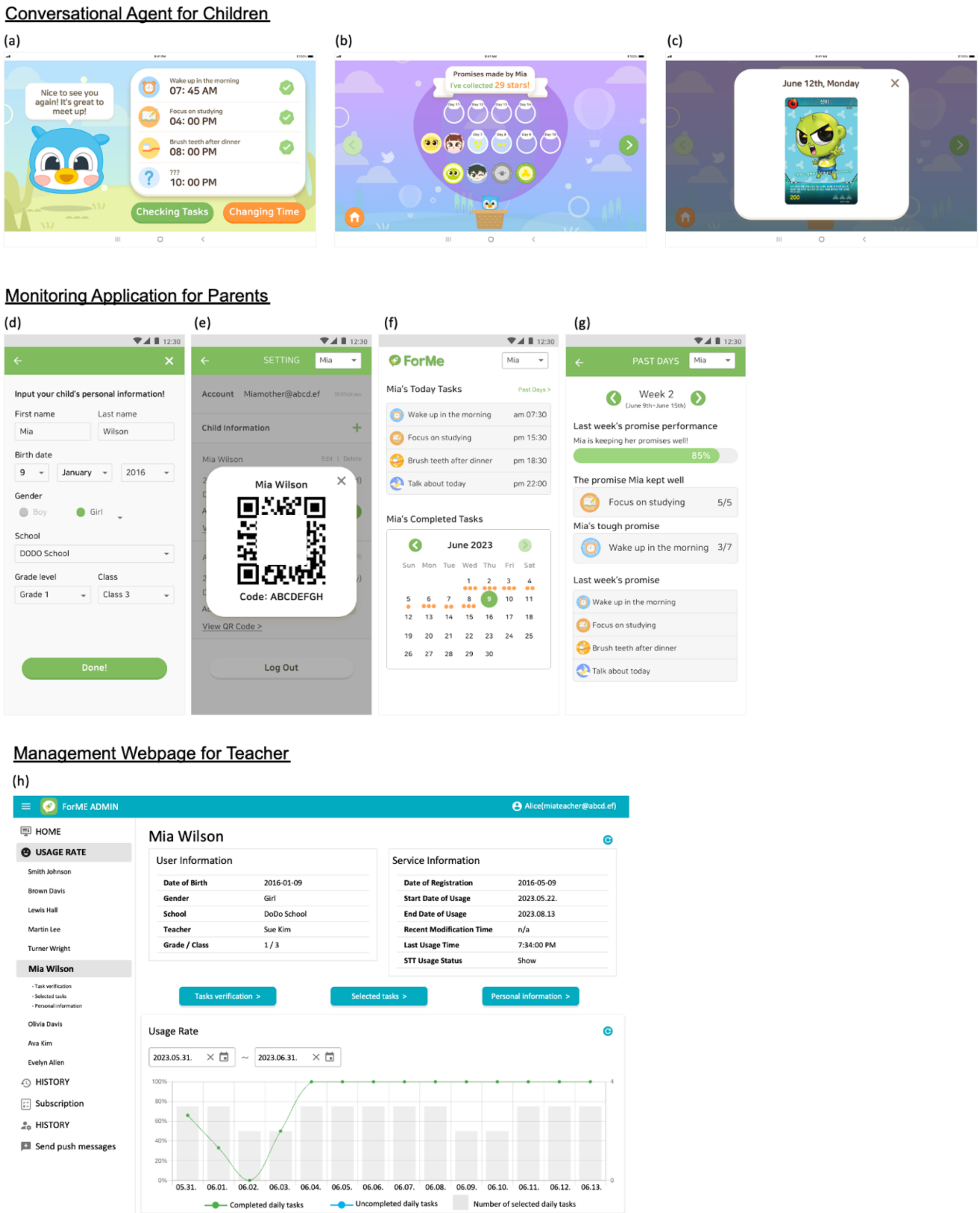

Figure 1: Screens of Applications for Children, Parents, and Teachers.

the child can start fulfilling the goal by tapping the ForME icon in the application.

The cue was designed with reference to smart speakers and displays. Figure 2 shows a child and the ForME agent working together to complete the task of "Brushing teeth after dinner" when the prototype alarm goes off at the pre-arranged time. The agent delivers a cue to the user through two channels: sounds and visuals. The sounds consist of two components: a beep that serves as a cue to help participants better understand the contents of the conversation and a dialog that contains the actual conversation. There are four different beep sounds: ringing, ding, tick tock, and ding-dang-dong. Ringing indicates the start and end of the entire task. Ding signals when users start and stop talking. Tick tock indicates the start of the timer, and ding-dang-dong indicates the end.

For each phase, multiple paraphrased versions of the dialogues are generated, totaling five or more alternatives for a given scenario. One of these alternatives is randomly chosen and presented to the user. Visual cues accompany each step to illustrate screen modifications. Adjustments in the steps and the agent's expressions reflect advancements in the task. Figure 2 illustrates children's responses and actions during their interactions with the ForME agent.

When the user does not mention the required keywords or does not speak at all during each step, the agent asks for a response so that the user can provide an answer that includes the keywords. For example, when the agent asks the user which task should be performed (see Figure 2) and the user provides a wrong answer, such as "wash my hands" instead of "brush my teeth," the agent produces an output that says "You know. . .there's something you have to do three times a day to avoid an ache! Tell me what that is!", which allows the user to derive the correct answer via the hint given. The keywords that the agent perceives as the correct answer at each step were developed using the Natural Language Understanding function provided by the chatbot engine, which was developed to analyze and process user intentions.

We configured the prepared scenarios using the NAVER CLOVA chatbot builder, which is equipped with natural language processing technology for Korean as well as a built-in machine learning algorithm. The Speech-to-Text functionality embedded in Android was prioritized. However, in cases where the analysis of a child's speech is ambiguous and intent recognition is challenging, we devised a system to verify using CLOVA Speech-to-Text. The server was divided into a Web App and API Server. The Web-based app is deployed to the client using the service through Google Firebase Hosting, and it processes service requests such as conversations and appointment schedule checks from the client and returns the results. Because the CLOVA Chatbot Builder does not provide built-in responses for waiting for appointment times or handling no-response intent, we implemented Chatbot Middleware to enable these functionalities.

3.2 Monitoring Application for Parents

We developed a Unity-based mobile application for parents to monitor their children's use of ForME. Thus, parents can install the parent application regardless of the operating system of their smartphone.

Parents and children can use the parent application to register their child's information, input school details, and enter the corresponding school code, as illustrated in Figure 1(d). Subsequently, as depicted in Figure 1(e), they can obtain a QR code, which the child can use to log in to the child application.

Parents can track their children's daily task performance and view the tasks they have completed through a dedicated web page for parents (Figure 1(f) and (g)). It is important to note that updates on task performance are not in real time but become accessible after a day. Real-time alerts and notifications were not incorporated to promote the children's autonomy and to avoid excessive intervention. Instead, parents were encouraged to agree with their children before using ForME, allowing the child to maintain a high degree of autonomy. However, if difficulties arose, the parents were informed that they could assist their children.

3.3 Management Webpage for Teacher

We developed a webpage that allows homeroom teachers at the school to manage each class's students who participated in the research. Homeroom teachers can issue codes for their respective classes and distribute them to link student accounts to their own.

As depicted in Figure 1(h), each homeroom teacher can access and review their students and monitor how they perform daily tasks. School administrators and counselors can also access information on students in each class and check their task performance rates.

As a result, each teacher is able to individually manage the daily lives of the participating children at their homes. Counselor teachers also have the ability to gain insight into how the children are living at home, making it possible to use this information for counseling purposes.

4 METHOD

4.1 Procedures

This study received institutional review board approval. This study was conducted on children in their first year of elementary school. The schools were recruited through a metropolitan city education office in Korea, and 20 schools participated in the study. Considering the number of students attending each school, two schools were paired, and using a random number table, each pair was randomly assigned to the intervention and control groups. As a result, 10 schools were assigned to the intervention group and the other 10 to the control group. The total number of students from the 10 schools assigned to the intervention group was 1,046 and the control group was 929.

Letters promoting participant recruitment were distributed by each school to each household. The application for study participation was collected with the consent from the students and their guardians. The researchers then had a call with each parent to explain the study in detail. Subsequently, the students and their parents provided consent to participate in the study through electronic signatures within mobile messages, and each student was given a screening number. The list of participants was finalized by the researchers after confirming whether the student had 1) mental diseases other than ADHD and 2) registered for other clinical studies.

The first questionnaire for parents was delivered to each household by mail. Once the parents completed the questionnaire, they were asked to take a picture of the document and return it to the researcher via phone. The researchers checked for missing items and collected the survey data. The missing survey items were requested from the parents again. Questionnaires for teachers about each student were distributed and collected through e-mail. The researchers once again checked for missing information and processed the survey data.

Participants in both the intervention and control groups completed their first questionnaires in May. After completing the survey,

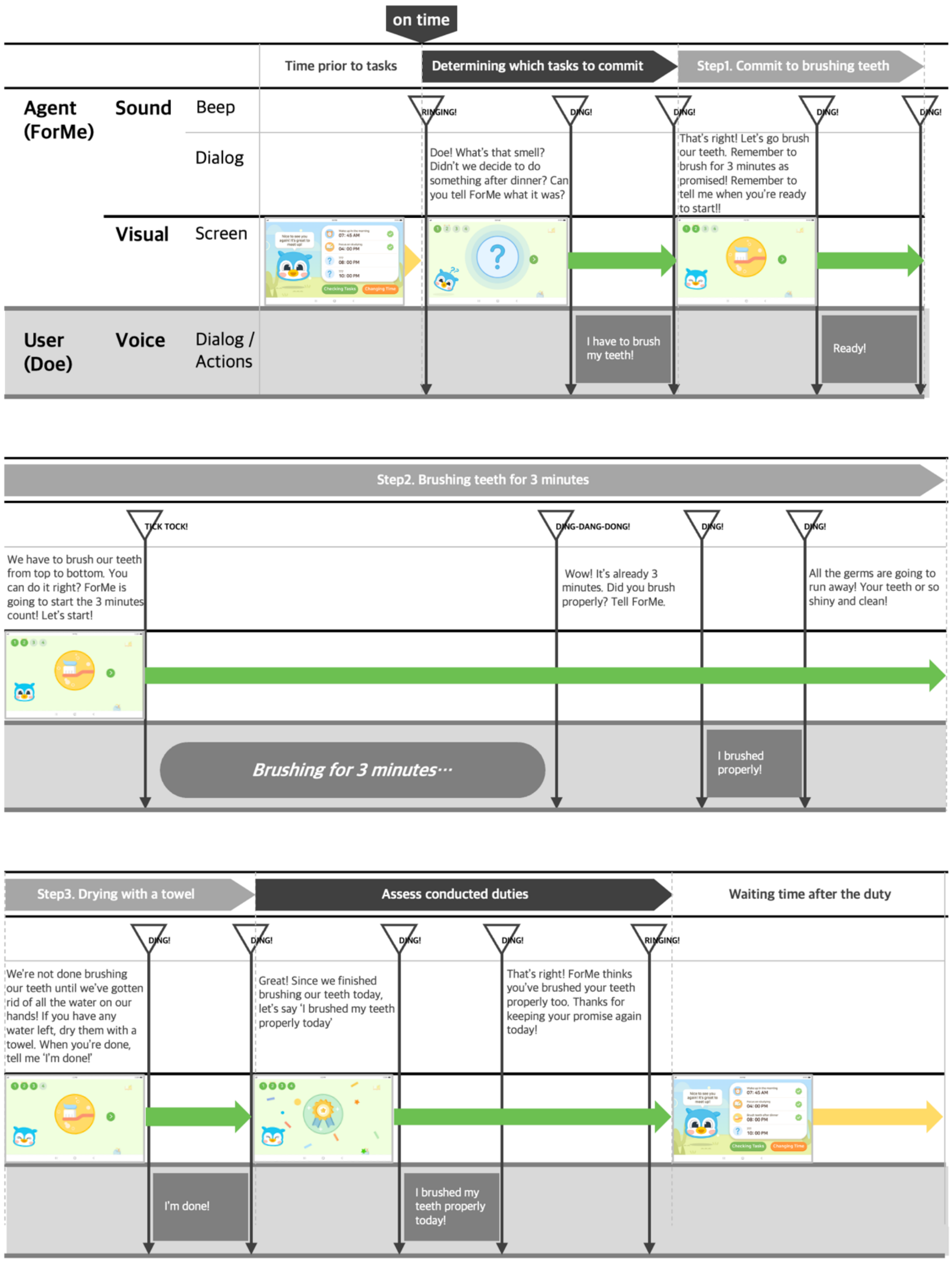

Figure 2: Task Completion Scenario between Agent and User

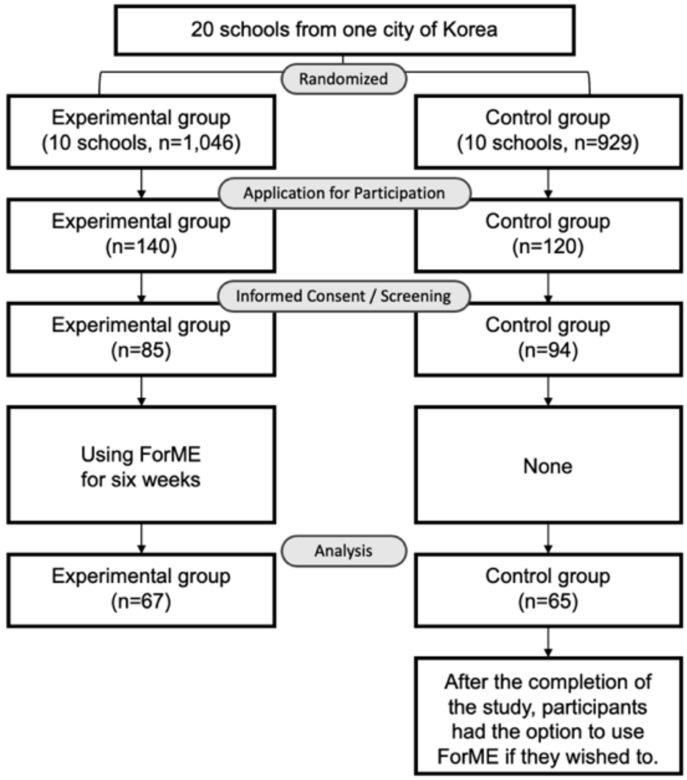

Figure 3: Flowchart Showing the Participant Enrollment and Assessment Process

the intervention group was instructed to use ForME for the next six weeks by entering a code that they were given. The second questionnaire was administered in June. After completion of the survey, the control group had the option to use ForME for six weeks if they wished.

After completing all phases, we collected feedback from parents and teachers about the program through an online survey.

4.2 Participants

A total of 260 participants applied for the study through the recruitment letter sent out from each school. The researchers contacted the children's parents to confirm the necessary information, resulting in a final count of eligible participants who registered as subjects. There were several reasons why some of them had to eventually give up registering, which included not having enough time to conduct two surveys, because it was not a mandatory program assigned by the school, and not having an Android device at home. Ultimately, 179 participants completed the first survey, among which, 132 completed both surveys. Figure 3 presents more detailed information regarding this.

4.3 Outcome Measures

4.3.1 Executive Function. Children's executive function was assessed using the Behavior Rating Inventory of Executive Function (BRIEF). BRIEF is a parental questionnaire designed to assess executive function in children aged 5 to 18 in their everyday lives [19]. This survey consists of 86 statements describing different behaviors. Parents were asked to assess each statement using a rating scale ranging from 1 (never) to 3 (often). The BRIEF questionnaire includes eight clinical scales, which are further categorized into two index scores. The Behavioral Regulation Index (BRI) comprises three scales: inhibition, shifting, and emotional control, while the Metacognition Index (MI) includes five scales: initiation, working memory, plan-organize, organization of materials, and monitoring. The combination of these two indexes generates the Global Executive Composite (GEC) score. Each factor transformed an original score into a T-score with a mean (M) of 50 and standard deviation (SD) of 10. Higher scores, as assessed by parents, indicate more severe impairment. For this study, a Korean version of the questionnaire, translated and approved by the author of the questionnaire, was utilized, and it was obtained from a publisher authorized by the author.

4.3.2 ADHD Symptoms. The severity of ADHD symptoms was measured using the ADHD Rating Scale (ADHD RS). The ADHD RS is an ADHD symptom severity scale designed by DuPaul according to the DSM-IV Diagnostic Criteria of Mental Disorders [17]. The ADHD RS consists of two factors–Inattention and Hyperactivity-impulsivity–and each factor is evaluated through nine items. Inattention refers to an inability to maintain attention. Hyperactivity-impulsivity is defined as excessive action or irritability and impulsiveness. The ADHD RS test is measured by parents or teachers rating the child on a 4-point scale ranging from 0 to 3. Higher scores

Table 1: Demographic data of study participants.

Characteristics	Group		$\chi 2$ (df)	p
	Intervention group (n = 67)	Control group (n = 65)		
Gender				
Boy	34	38	0.792 (1)	0.373
Girl	33	27		
Age				
6	28	23	0.571 (1)	0.450
7	39	42		

on the ADHD RS indicate higher levels of ADHD symptom severity. In this study, a standardized scale was used for Korean children [57].

4.4 Data Analysis

Initially, an independent t-test or Mann-Whitney U-test was conducted to assess homogeneity based on the general characteristics and initial outcome results of both groups. To facilitate this, the normality of the data was assessed using Kolmogorov-Smirnov or Shapiro-Wilk tests. Second, Cronbach's alpha for each factor was verified. The reliability of an assessment is considered acceptable when Cronbach's alpha exceeds 0.7 [28]. Depending on the normality of the data results, either Repeated Measures Analysis of Variance or Generalized Estimating Equation (GEE) was used to analyze the differences between the two groups at two time points with a 6-week interval. That is, when the data did not follow a normal distribution, it was analyzed using GEE, and the family of models was selected based on the characteristics of the data. Lastly, we examined whether there were changes in executive function and ADHD symptoms in children classified as ADHD-suspected in the intervention group. Depending on the normality test results, the data were analyzed using either a paired t-test or the Wilcoxon signed ranks test. Prior research indicated that children were considered potential ADHD cases if their total ADHD RS score, as assessed by their parents, was 19 or higher, and if their teachers' assessment yielded a score of 17 or higher [22]. Categorization of the ADHD-suspected group was determined based on the ADHD RS scores evaluated by the homeroom teacher. Statistical analyses were conducted using SPSS 26.0, and a p-value < 0.05 was considered statistically significant. The feedback received from parents and teachers was analyzed through qualitative summaries.

5 RESULTS

5.1 Manipulation Check

The task completion rates were checked to confirm whether the children in the intervention group actually used the prototype for the 6 weeks. Children are able to perform tasks they agreed to conduct with ForME, which are 1 to 4 tasks per day. Each child's task completion rate was analyzed on a daily basis. Consequently, the 67 children assigned to the intervention group demonstrated an average task completion rate of 56.2% per day. Additionally, more than one task per day was performed for an average of 31.8 days during the 42-day ForME trial period.

5.2 Demographic Data Results

All participants whose data were analyzed were first-year elementary school students. Of the 67 participants in the intervention group, 34 were male and 33 were female (Table 1). There were 28 6-year-olds and 39 7-year-olds. Of the 65 control group students, 38 were male and 27 were female, of which 23 were 6-year-olds and 42 were 7-year-olds. Participants in the intervention and control groups showed no statistically significant differences in terms of their demographic variables, as shown in Table 1.

5.3 Homogeneity between Two Groups

When the normality test on the initial outcome was conducted using the Kolmogorov-Smirnov and Shapiro-Wilk tests, the normality assumption was not satisfied. Therefore, the Mann-Whitney U-test was performed on the intervention and control groups to assess homogeneity. Table 2 shows the homogeneity of outcomes between the two groups (p > 0.05).

5.4 Outcome Analysis and Results

First, the Cronbach's alpha values for each of the sub-factors in the outcomes exceeded 0.7, indicating their reliability (BRIEF: BRI = 0.896, MI = 0.933, GEC = 0.943; ADHD RS (parents): Inattention = 0.836, Hyperactivity/Impulsive = 0.770, Total ADHD RS = 0.883, ADHD RS (teacher): Inattention = 0.962, Hyperactivity/Impulsive = 0.950; Total ADHD RS = 0.974). Then, Kolmogorov-Smirnov and Shapiro-Wilk tests were conducted; however, the dependent variables in this study did not meet the normality assumption. Therefore, GEE was used to explore the effects of executive function and ADHD symptoms. In particular, it was analyzed as gamma GEE with log link, considering the characteristics of the data corresponding to continuous variables and being skewed to the left. The detailed statistical analysis results are presented in Table 3. The results are described with a focus on the interaction effect between group and time on outcomes. The interaction effect provides important information, suggesting that the differences between the two groups changed over time.

Initially, the interpretation of the sub-factors of executive function was as follows. GEE analyses indicated a significant group x time interaction for BRI (β = 0.052; Wald $\chi^2{}_1$ = 6.463; p = 0.011) and GEC (β = 0.044; Wald $\chi^2{}_1$ = 4.427; p = 0.035) but not for MI (β = 0.035; Wald $\chi^2{}_1$ = 2.736; p = 0.098). In other words, the use of ForME helped improve the BRI and GEC in children's executive function. Group differences in the patterns of change over time

Table 2: Homogeneity of outcomes between intervention and control groups.

Variables	Group		Z	p
	Intervention group (n = 67) Mean ± SD	Control group (n = 65) Mean ± SD		
BRIEF [a]				
BRI	44.21 ± 8.22	41.95 ± 6.40	-1.564	0.118
MI	44.58 ± 8.88	43.97 ± 8.12	-0.415	0.678
GEC	44.24 ± 8.71	43.02 ± 7.49	-0.754	0.451
ADHD RS [b] (by parents)				
Inattention	2.04 ± 2.93	2.12 ± 2.45	-1.014	0.311
Hyperactivity/Impulsive	2.13 ± 2.70	1.94 ± 2.22	-0.072	0.942
Total ADHD RS	4.18 ± 5.31	4.06 ± 4.18	-0.691	0.490
ADHD RS (by teacher)				
Inattention	4.21 ± 5.99	3.29 ± 6.06	-1.764	0.078
Hyperactivity/Impulsive	3.66 ± 5.53	3.06 ± 5.30	-1.001	0.317
Total ADHD RS	7.87 ± 11.21	6.35 ± 10.98	-1.349	0.177

[a] BRIEF: Behavior Rating Inventory of Executive Function, BRI: Behavioral Regulation Index, MI: Metacognition Index, GEC: Global Executive Composite, [b] ADHD RS: Attention-Deficit Hyperactivity Disorder Rating Scale

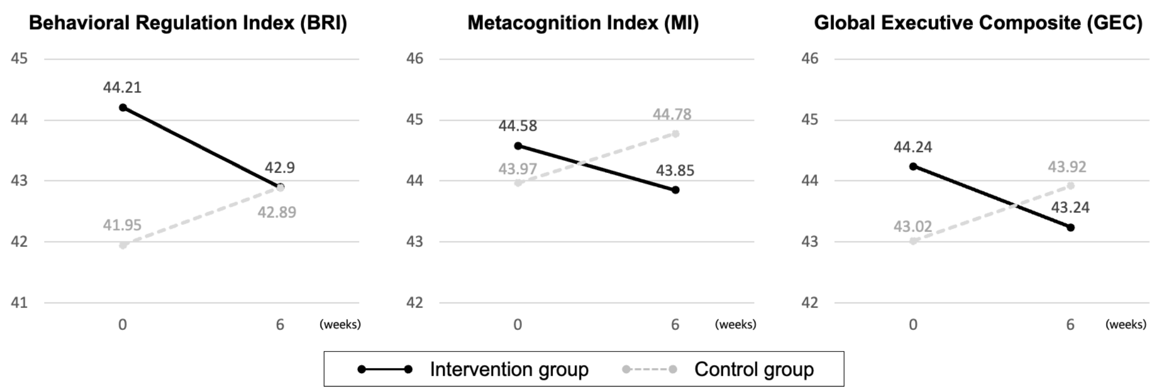

Figure 4: Changes in the Intervention and Control Groups from Pre-test to Post-test in Terms of Executive Function.

are shown in Figure 4. In the intervention group, scores for the sub-factors of executive function generally decreased, while in the control group, an overall increase in scores can be observed.

Next, interpretation of the ADHD RS as assessed by the parents was as follows. GEE analyses indicated a significant group x time interaction for Inattention (β = 0.346; Wald $\chi^2{}_1$ = 5.414; p = 0.020) but not for Hyperactivity/Impulsive (β = 0.300; Wald $\chi^2{}_1$ = 2.782; p = 0.095) and Total ADHD RS (β = 0.250; Wald $\chi^2{}_1$ = 3.181; p = 0.075). Among the ADHD symptoms evaluated by parents, Inattention indicates that the use of ForME significantly improved the child's symptoms. Differences in group patterns of change over time are shown in Figure 5. In the intervention group, there were no significant variations in the three sub-factors. In fact, the mean value of the Total ADHD RS did not change over the course of 6 weeks. In contrast, in the control group, the sub-factors increased over time, indicating worsening of symptoms.

The interpretation of the ADHD RS as assessed by teachers was as follows. GEE analyses indicated a significant group x time interaction for Inattention (β = 0.394; Wald $\chi^2{}_1$ = 11.368; p < 0.001), Hyperactivity/Impulsive (β = 0.286; Wald $\chi^2{}_1$ = 5.668; p = 0.017), and Total ADHD RS (β = 0.378; Wald $\chi^2{}_1$ = 10.651; p = 0.001). Among the ADHD symptoms evaluated by teachers, ForME usage significantly improved the respective symptoms in both Inattention and Hyperactivity/Impulsivity in children. Figure 6 illustrates the differences between the intervention and control groups. In the intervention group, the values of the sub-factors decreased over time, indicating an improvement in symptoms. Conversely, the control group showed an increase in factor values over time.

5.5 Additional Analysis on the Intervention Group

5.5.1 Differences between parents and teachers' evaluation. In this study, the same ADHD RS scale was used to compare the parents' and teachers' evaluations of one student. Based on previous studies, we screened children suspected of having ADHD using the total ADHD RS scores as evaluated by parents and teachers [22].

Table 3: Effects of using ForME on BRIEF and ADHD RS.

Variables	Group	Time Pre-test Mean ± *SD*	 Post-test Mean ± *SD*	Sources	Regression coefficient	*SE* [e]	Wald $\chi 2$ (*df*)	*p*
BRIEF [a]								
BRI	IG (n = 67) [C]	44.21 ± 8.22	42.90 ± 8.42	Group	-0.052	0.029	3.185 (1)	0.074
	CG (n = 65)	41.95 ± 6.40	42.89 ± 6.58	Time	0.015	-0.059	4.065 (1)	**0.044**
				G x T [d]	0.052	0.021	6.463 (1)	**0.011**
MI	IG (n = 67)	44.58 ± 8.88	43.85 ± 9.16	Group	-0.014	0.033	0.174 (1)	0.676
	CG (n = 65)	43.97 ± 8.12	44.78 ± 8.95	Time	-0.017	0.013	1.557 (1)	0.212
				G x T	0.035	0.021	2.736 (1)	0.098
GEC	IG (n = 67)	44.24 ± 8.71	43.24 ± 8.99	Group	-0.028	0.032	0.764 (1)	0.382
	CG (n = 65)	43.02 ± 7.49	43.92 ± 8.04	Time	-0.023	0.014	2.737 (1)	0.098
				G x T	0.044	0.021	4.427 (1)	**0.035**
ADHD RS [b] **(by parents)**								
Inattention	IG (n = 67)	2.04 ± 2.93	2.16 ± 2.28	Group	-0.143	1.863	0.590 (1)	0.442
	CG (n = 65)	2.12 ± 2.45	2.77 ± 3.29	Time	-0.047	0.113	0.177 (1)	0.674
				G x T	0.346	0.149	5.414 (1)	**0.020**
Hyperactivity /Impulsive	IG (n = 67)	2.13 ± 2.70	2.01 ± 2.43	Group	-0.168	0.172	0.957 (1)	0.328
	CG (n = 65)	1.94 ± 2.22	2.40 ± 2.54	Time	-0.100	0.139	0.523 (1)	0.469
				G x T	0.300	0.180	2.782 (1)	0.095
Total ADHD RS	IG (n = 67)	4.18 ± 5.31	4.18 ± 4.27	Group	-0.087	0.182	0.228 (1)	0.633
	CG (n = 65)	4.06 ± 4.18	5.17 ± 5.37	Time	-0.008	0.110	0.005 (1)	0.941
				G x T	0.250	0.140	3.181 (1)	0.075
ADHD RS (by teacher)								
Inattention	IG (n = 67)	4.21 ± 5.99	2.78 ± 5.04	Group	0.053	0.225	0.056 (1)	0.812
	CG (n = 65)	3.29 ± 6.06	3.68 ± 6.34	Time	-0.306	0.959	10.162 (1)	**0.001**
				G x T	0.394	0.117	11.368 (1)	**<0.001**
Hyperactivity /Impulsive	IG (n = 67)	3.66 ± 5.53	2.48 ± 4.76	Group	0.018	0.205	0.007 (1)	0.932
	CG (n = 65)	3.06 ± 5.30	3.20 ± 5.55	Time	-0.286	0.107	7.086 (1)	**0.008**
				G x T	0.286	0.120	5.668 (1)	**0.017**
Total ADHD RS	IG (n = 67)	7.87 ± 11.21	5.24 ± 9.56	Group	-0.024	0.231	0.011 (1)	0.916
	CG (n = 65)	6.35 ± 10.98	6.88 ± 11.49	Time	-0.307	0.104	8.803 (1)	**0.003**
				G x T	0.378	0.116	10.651 (1)	**0.001**

[a] BRIEF: Behavior Rating Inventory of Executive Function, BRI: Behavioral Regulation Index, MI: Metacognition Index, GEC: Global Executive Composite, [b] ADHD RS: Attention-Deficit Hyperactivity Disorder Rating Scale, [c] IG: Intervention group, CG = Control group, [d] G x T: Group x Time, [e] SE: Standard error [*] Statistically significant values are indicated in bold.

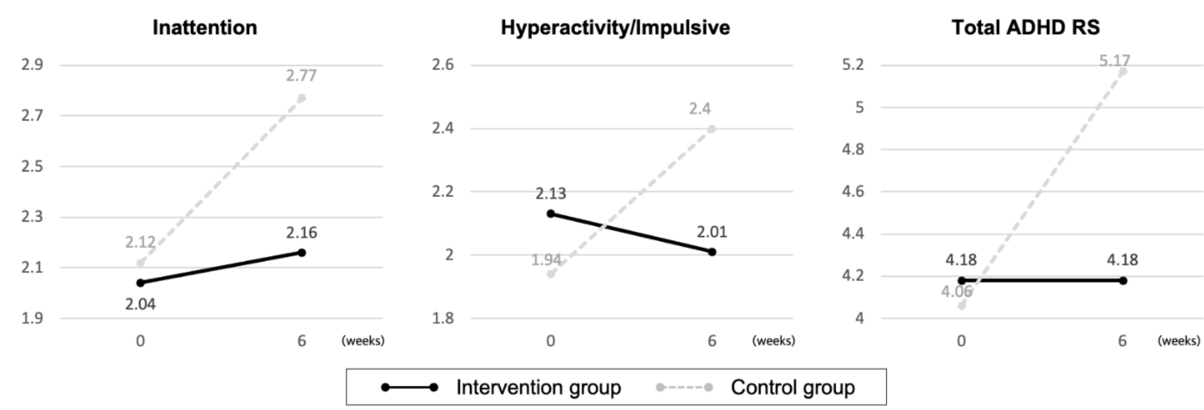

Figure 5: Changes in the Intervention and Control Groups from Pre-test to Post-test in Terms of ADHD Symptoms by Parents.

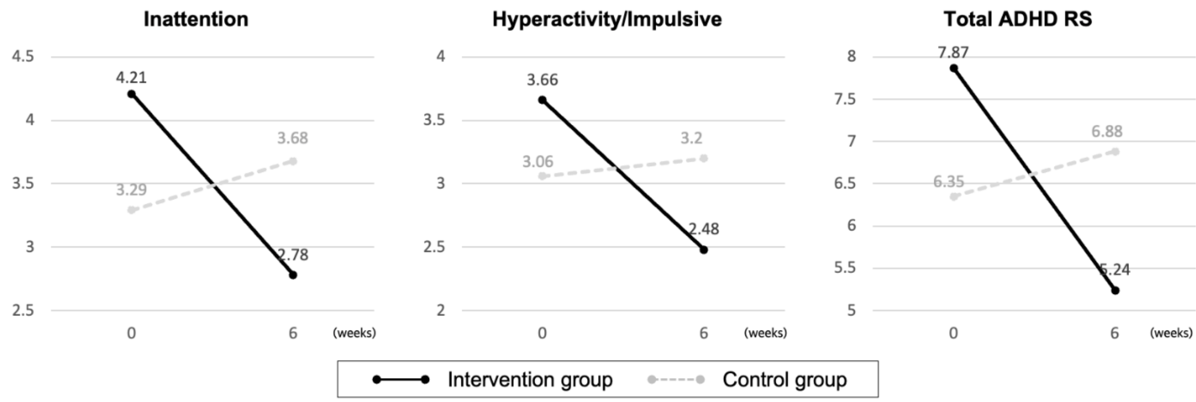

Figure 6: Changes in the Intervention and Control Groups from Pre-test to Post-test in Terms of ADHD Symptoms by teacher.

Table 4: Comparison of ADHD RS scores between parents and teachers for the same child within the intervention group (n = 67).

Variables	Parents	Teacher	Negative ranks		Positive ranks		*Z*	*p*
	Mean ± *SD*	Mean ± *SD*	Mean rank	Sum of ranks	Mean rank	Sum of ranks		
ADHD RS [a]								
Inattention	2.04 ± 2.93	4.21 ± 5.99	16.77	335.50	30.67	889.50	-2.763	**0.006**
Hyperactivity/Impulsive	2.13 ± 2.70	3.66 ± 5.53	16.33	343.00	30.19	785.00	-2.345	**0.019**
Total ADHD RS	4.18 ± 5.31	7.87 ± 11.21	19.84	436.50	32.77	1048.50	-2.638	**0.008**

Children were classified as possible ADHD cases if their parents' assessment of the total ADHD RS score was 19 or higher and their teachers' evaluations yielded a score of 17 or higher [22]. Looking at the scores evaluated by the parents, 2 out of 67 students in the intervention group were classified as ADHD-suspected and 10 by teachers' evaluation. We were able to determine whether there were differences in evaluations between parents and teachers for the same child through the ADHD RS assessment.

As an additional analysis, we also examined whether there were statistical differences between the two evaluators' assessments. In this process, we compared the 67 students in the intervention group and their baseline status (at week 0), as evaluated by parents and teachers. The sub-factor normality of ADHD RS was not met; therefore, a Wilcoxon signed ranks test was performed. The results are shown in Table 4.

[a] ADHD RS: Attention-Deficit Hyperactivity Disorder Rating Scale
[*] Statistically significant values are indicated in bold.

According to these data, we can infer that parents' and teachers' opinions on all three sub-factors of ADHD RS were different from each other. The Wilcoxon signed ranks test indicated that Inattention was statistically significantly rated higher by teachers than by parents Z = -2.763, p = 0.006. Similarly, Hyperactivity/impulsivity reported by teachers was also significantly greater than what parents reported, as evidenced by Z = -2.345 and p = 0.019. Furthermore, the Total ADHD RS score reported by teachers was significantly elevated compared to the scores reported by parents, as indicated by Z = -2.638 and p = 0.008.

Therefore, it is possible that each child could have very different evaluation results depending on who gave them, even though the evaluation was under the same scale.

5.5.2 Changes in Suspected of ADHD due to the Use of ForME.. Teachers can provide more objective assessments when evaluating a single child because, unlike parents, they have other children to compare to in their surroundings [35]. Among the 67 children in the intervention group, 10 were classified by the teachers as having suspected ADHD. We aimed to determine whether ForME had a positive impact on these 10 children through a pre-post comparison.

When normality testing was conducted using the Shapiro-Wilk test, the Hyperactivity/Impulsive and Total ADHD RS evaluations by parents did not meet the normality assumption. Therefore, the results for these two sub-factors were analyzed using the Wilcoxon signed ranks test (Table 5).

[a] ADHD RS: Attention-Deficit Hyperactivity Disorder Rating Scale

Other sub-factors were analyzed using a paired t-test (Table 6).

[a] BRIEF: Behavior Rating Inventory of Executive Function, BRI: Behavioral Regulation Index, MI: Metacognition Index, GEC: Global Executive Composite, [b] ADHD RS: Attention-Deficit Hyperactivity Disorder Rating Scale [*] Statistically significant values are indicated in bold.

As shown in Table 6, statistically significant differences were found in the sub-factors of ADHD RS assessed by teachers in relation to suspected ADHD. Inattention (p = 0.022), Hyperactivity/Impulsive (p = 0.045), and Total ADHD RS (p = 0.031) decreased

Table 5: Changes in outcomes using ForME in suspected ADHD (n = 10) analyzed using the Wilcoxon signed-rank test.

Variables	Time		Negative ranks		Positive ranks		Z	*p*
	Pre-test Mean ± *SD*	Post-test Mean ± *SD*	Mean rank	Sum of ranks	Mean rank	Sum of ranks		
ADHD RS [a] (by parents)								
Hyperactivity/Impulsive	4.5 ± 3.60	3.7 ± 2.79	4.75	28.50	5.50	16.50	-0.715	0.475
Total ADHD RS	8.7 ± 6.91	7.7 ± 4.08	5.25	21.00	3.75	15.00	-0.425	0.671

Table 6: Changes in outcomes using ForME in suspected ADHD (n = 10) analyzed using paired t-test.

Variables	Time		*t*	*p*
	Pre-test Mean ± *SD*	Post-test Mean ± *SD*		
BRIEF [a]				
BRI	46.70 ± 8.62	44.20 ± 8.62	1.884	0.092
MI	48.10 ± 8.90	45.70 ± 6.93	1.577	0.149
GEC	47.50 ± 8.92	45.00 ± 7.78	2.197	0.056
ADHD RS [b] (by parents)				
Inattention	4.20 ± 3.68	4.00 ± 2.16	0.294	0.775
ADHD RS (by teacher)				
Inattention	15.50 ± 5.82	10.70 ± 8.51	2.773	**0.022**
Hyperactivity/Impulsive	14.40 ± 5.4	9.70 ± 8.42	2.330	**0.045**
Total ADHD RS	29.90 ± 10.84	20.40 ± 16.75	2.545	**0.031**

significantly. However, as shown in both Tables 5 and 6, no significant changes were observed in the other factors.

The use of ForME over a 6-week period for suspected ADHD showed a statistically significant improvement in teacher evaluations, whereas it was not evaluated as such in parent assessments.

5.6 Feedback on ForME from Parents and Teachers

After summarizing the online feedback from both parents and teachers, we were able to gauge the effectiveness of the ForME program conducted over six weeks.

First, parents were deeply impressed by their children's efforts to control their behavior. Comments such as "Even if our son didn't do well, it was heartwarming to see him make an effort" and "I felt that my daughter has grown a bit while performing daily tasks with ForME characters" were received. Children's attempts to engage with ForME and perform daily tasks were seen as fresh and challenging.

However, there were instances in which some children lacked time or forgot about task completion, resulting in less meticulous task execution. Parents commented, "*Our son doesn't have enough time for brushing his teeth for 3 minutes. I wish a little more time was allocated*" and "*My child attempted to pack the bag but found it difficult, so I, as a mother, helped.*" Nonetheless, the parents demonstrated patience and provided support to help their children perform tasks independently.

However, in some cases, the program's utility was not evident. Some parents expressed, "Our daughter already had good daily habits, so she didn't need this program" and "I'm not sure how our child benefits from this program." We interpreted that the program might not have been effective for children who deviated significantly from the target audience of this study. However, considering the larger number of positive impacts on children, the overall outcome favored benefits over drawbacks.

Teachers played a significant role in this program. Initially, they had to link each class's students' accounts. Additionally, teachers who wished to do so could monitor and manage their students' task performance. In practice, some teachers assisted students in task selection and even provided gift cards based on task completion rates.

Homeroom teachers gained a deeper understanding of the children through the ForME program. One teacher commented, "We couldn't know how our students spend their time at home from school. With the ForME program, we could learn how our class children spend their day." We believe that gaining insight into a child's school and home life makes teaching more effective.

The program was notably beneficial for children facing difficulties in managing their daily lives. One comment highlighted, "Some parents care well, and their children grow up healthily, but there are also children on the opposite end. This program could be helpful for children facing difficulties."

However, there were several opinions expressed such as, "The program's intention is good, but it is challenging for schoolteachers to manage children's lives at home," and "With so many students to manage, I found it quite difficult to track individual student behavior using the prototype as the workload became difficult to manage." This situation calls for a thorough examination of how this program can operate effectively in educational settings.

6 DISCUSSION

6.1 The need for a mental health support system involving schools

6.1.1 Differences between parents' and teachers' evaluations of children. In this study, parents and teachers used the ADHD RS questionnaires as a screening tool. In the actual analysis, the cut-off criteria were set differently depending on the evaluator based on previous research [22]. However, as shown in Table 4, the degree to which parents and teachers evaluated the same child showed significant differences. Of the 67 intervention group participants, 2 (3%) of the parents' responses met the ADHD screening criteria, whereas the teachers assessed that 10 (15%) of them met the criteria.

Previous research has emphasized the possibility of differences in evaluation results between parents and teachers [3]. This is because parents primarily observe children's characteristics in everyday life, whereas teachers mainly observe them in task-oriented situations, which can lead to significant variations in outcomes. In addition, a High ADHD RS score does not necessarily mean that children meeting these criteria will all be diagnosed with ADHD.

However, without schools' opinions, it seems unlikely that children who are potentially experiencing difficulties with their mental health will be able to receive any necessary support. It is difficult for parents to have the chance to look at their child as others do, as they would spend most of their time with only their child. However, teachers are accustomed to taking care of a group of children, which enables them to see the differences in each child's reactions to the same situation [44]. Therefore, school engagement in the process of screening children for potential mental health problems can be significantly beneficial for early diagnosis of these patients.

Therefore, it can be argued that there is demand for a program involving schools that considers the benefits of early diagnosis or treatment of ADHD.

6.1.2 Establishment of a system providing equal opportunities without bias concerns. The conversational agent ForME that we developed is an assistive technology that supports young ADHD patients, for which this research is being carried out as one of the supporting works. Therefore, conversation structures and interaction methods were designed to be suitable for children with ADHD.

However, this study did not impose any significant restrictions on participant eligibility, which means that children who did not have any mental health-related difficulties were allowed to participate. It was possible to have a restriction that only children experiencing difficulties or ADHD symptoms could participate, but there were concerns about creating unfair negative perceptions of those who applied for the program. Children who had already been diagnosed with other mental disorders were not eligible to participate. However, most children from the participating schools were provided with equal opportunities to participate in the program.

Consequently, a diverse group of children with various backgrounds participated in the study. Some children were suspected of having ADHD; however, many children with good executive function also participated in this study. Nonetheless, the target subjects of this study were children with suspected ADHD. As shown in Table 6, it was estimated that there was some degree of effectiveness in children suspected of having ADHD. In addition, as shown in Table 3, the intervention provided in this study positively influenced the participating children.

One of the strengths of this study is that it addressed areas relevant not only to children with ADHD but also to typically developing children. In other words, the program aimed to create a positive impression not only for children with mental health difficulties, but also for those without, making it more widely applicable within school settings. Within educational settings, it is imperative that mental health initiatives in schools be inclusive and target all children regardless of their mental health status [36, 49]. Therefore, in a school-involved system, it is preferable to provide inclusive opportunities rather than screen child participants for specific target diseases.

6.1.3 Design direction of systems considering stakeholders' context. Prototypes were developed using a user-centered design approach involving iterative planning, design modifications, and deployment [65]. In this study, we utilized an application with a conversational agent for children, a mobile monitoring application for parents, and a management web page for teachers.

In this study, we sought to integrate different usage environments, such as family units of children and parents, with the educational context of teachers and schools. While we did identify challenges and dissatisfaction among teachers during the use of the web page, the significant benefits when teachers have a deep understanding of the children are clear.

In our future research, we aim to alleviate the burden on teachers through thoughtful planning. Teachers identified in this study that tracking their students' ForME engagement and life/activity patterns was an arduous task. Moreover, the web page provided to the teachers in this study was equipped only with a monitoring function, which meant that it lacked practical functions that would help teachers with relevant management or intervention. Future studies will reflect the feedback from our teacher stakeholders and improve the web page's functionalities so that it can better support the role of teachers in the entire system. For example, allowing them to access a list of students with their ratings based on their ForME usage, instead of tracking each student individually on the current iteration of the web page, may help alleviate their workload and concerns. We plan to introduce additional features that allow parents to monitor and encourage their children, enabling management to take place within each household.

6.2 Improving children's functions with ForME

6.2.1 Conversational agents for neurodevelopmental disorders. Recently, researchers have shown an increased interest in providing therapeutic interventions using conversational agents for patients with neurodevelopmental disorders [13, 59, 61]. The broad category of neurodevelopmental disorders encompasses conditions such as intellectual developmental disorders, global developmental delay, autism spectrum disorder, and ADHD. In particular, a significant number of studies have explored training in the communication and social domains of ASD children [61, 63].

This trend is believed to have developed because of technical advancements that now allow input and output interaction modalities to be applied as/through speech [61]. Nevertheless, some difficulties remain in the interaction process between young children and agents owing to limitations in natural language processing, etc. [62]. In addition, other issues such as delays in language development in patients make the child-agent interaction even more difficult.

In this study, we allowed children to experience multiple exchanges of turns in an interaction with higher accuracy by reconstructing a chatbot builder service that has already been commercialized. The ability to engage in multiple turns of conversation with a conversational agent is anticipated to be a significant advantage in the scenario set up incorporating therapeutics. In particular, we applied self-instruction training to the conversation scenario between a conversational agent and children, with the expectation that this approach can facilitate the application of cognitive-behavioral therapy to children.

Numerous non-pharmacological treatment methods have already demonstrated effectiveness in the treatment of neurodevelopmental disorders. However, due to the considerable time and effort required for patients to perceive the effects of non-pharmacological treatments, pharmacological therapy is often chosen. The integration of therapeutic approaches tailored to children with mental health issues into scenarios involving conversations with conversational agents is expected to transform them into safe and effective therapeutic tools.

6.2.2 Changes in executive function. ForME allows children to choose their own daily tasks and time to proceed. Through this, children experience the process of selecting and planning out their goals and how they can achieve them. ForME helps children kick-start and complete their daily tasks when needed. This enables children to experience a sense of achievement and reflection on their goals. The repetition of this process makes it possible for them to become accustomed to routines in their day-to-day lives.

BRIEF is an assessment tool that can measure children's behavior based on their parents' observations and evaluations. Therefore, it can be assumed that performing daily tasks with ForME made a positive change in the children's behavior, which the parents observed.

Specifically, according to the findings from the analysis of the intervention and control groups shown in Table 3, this study confirmed that the overall executive function was improved in children who used the conversational agent. There was a statistically significant improvement in BRI, which consists of inhibition, shifting, and emotional control, as well as in GEC, which can examine behavioral regulation and overall executive function. Although it was not statistically significant in MI, which oversees meta-cognition consisting of initiation, working memory, plan-organize, organization of materials, and monitoring, the results showed that the MI value improved with the use of ForME (see Figure 4).

In the pre-post comparison of the ADHD-suspected group (Table 6), no statistically significant differences were detected in any BRI, MI, and GEC items. However, the results suggested a marginal statistical significance in GEC ($p = 0.056$). They also showed improvements in values in all aspects. This observation may support the hypothesis that ForME has a positive effect on executive function in the ADHD-suspected group. However, further research with an adequate number of participants is necessary to investigate whether it is possible to enhance executive function in children with ADHD or in those suspected of having ADHD. Nevertheless, it can be inferred that interventions through ForME can also have a positive influence on children with well-developed executive function.

6.2.3 Changes in ADHD symptoms. As mentioned in 6.1.1, the evaluation of a child by parents and teachers showed a significant difference. Accordingly, the assessment of the ADHD RS by the two evaluators can be considered different.

First, among the sub-factors evaluated by parents, Inattention was the only factor that showed statistically significant results (Table 3).

As shown in Figure 5, the findings revealed a remarkable increase in Inattention in the control group, whereas the increase in the intervention group's inattention was minimal. Although Hyperactivity/impulsive was not statistically significant, there was a noticeable increase in the control group and a decrease in the intervention group. Additionally, the Total ADHD RS value of the control group increased significantly, while that of the intervention group did not change; its statistical value showed marginal significance ($p = 0.075$). This indicates that the control group that did not use ForME showed deterioration in executive function over a period of six weeks, but the Intervention group assisted by ForME maintained a similar degree of function over the same period.

In the teacher evaluation, the intervention group showed improved functionality in all ADHD RS sub-factors for the six weeks of using ForME, while the control group showed deterioration in all for the same period, from which statistically significant differences were observed (Table 3 and Figure 6).

In conclusion, based on the opinions of the two evaluators combined, it was confirmed that the use of ForME helped improve participants' inattention, among other functions, and the results also suggested that it can have an effect on impulsiveness.

However, the two evaluator groups–parents and teachers–provided conflicting ideas. It is assumed that being at the beginning of the semester greatly affects teachers' judgments. Korean schools start their new grade/semester in March, which means that the study was conducted three months after a child entered school. In other words, this study was conducted during a time when children who had just entered school were trying to adapt to the new environment. It is possible that the adaptation phase in school and the ForME-assisted period coincided for the children, resulting in a greater effect. Therefore, one possible implication is that children can greatly benefit from the process of structuring their daily routines by themselves with the help of ForME, especially during the first few months after entering elementary school.

In the pre-post comparison of the ADHD-suspected group (Table 5 and Table 6), all ADHD RS sub-factors evaluated by the teacher showed statistical significance. However, when evaluated by parents, the same categories were not found to be statistically significant. However, we identified a general tendency of a decrease in the scores of the ADHD RS sub-factors in the parents' evaluation.

Given the relatively small sample size of children identified as potentially having ADHD in this study, further research with a larger cohort is warranted. However, the presence of statistically significant changes in the ADHD sub-factors assessed by teachers suggests that there may be an effect of ForME.

7 LIMITATIONS AND FUTURE RESEARCH

This study had the following limitations.

First, the exact mental health status of the children could not be determined. We only conducted a screening survey and could not have the diagnosis confirmed by a doctor. Additional analysis was performed only on the ADHD-suspected groups, and it is unknown how many of them actually might have been diagnosed with ADHD. Future studies should be conducted in the form of clinical research on children diagnosed with ADHD.

Second, many participants dropped out of the study. This study did not place any significant restrictions on participant eligibility. In addition, the entire process was conducted remotely. Therefore, the number of participants had decreased mid-way through. A research design that can provide participants with a more specific guide to minimize the number of dropouts is needed.

Lastly, in this study, we only collected feedback from teachers and parents; thus, we were unable to specifically ascertain the experiences of child users. It is essential to incorporate a research design that encompasses the perspectives of the key users in this study–children, parents, and teachers–to better understand their experiences.

In future studies, we will conduct clinical research on children with ADHD by introducing a process of connecting children who are screened in the questionnaire evaluation with doctors. Undertaking this project will allow us to gain a better understanding of the clinical effects of the conversational agent we developed.

8 CONCLUSION

In this study, a six-week experiment was conducted to determine whether a conversational agent that helps perform daily tasks can improve executive function and relieve ADHD symptoms in children. The conversational agent we developed assists children in selecting and performing their own daily tasks and to get them accustomed to the course of actions through repetition. The findings from this study suggest that the conversational agent, which was developed based on evaluations by parents and teachers, can benefit children who use it in a way that improves executive function and alleviates ADHD symptoms. Schools were involved in this study in terms of recruitment, management, and evaluation of participants, and as a result, a mental health support system for children can be achieved through collaboration with schools.

The findings of this study have a number of practical implications.

First, it was possible to see how teachers and parents responded differently to children's ADHD evaluations. Based on these differences, we argue that there is a definite need for a new mental health support system that involves schools.

Second, this study confirmed that the conversational agent was effective not only for children suspected of having ADHD but also for those who are not. Children's improved executive function and healthy routines built in their daily lives are expected to have a positive impact on their lives.

Third, the empirical findings provide a positive impression of the field of education and mental health studies using new technologies. In collaboration with the Office of Education, we were able to introduce a system that targets a large population of children. Overall, this study strengthens the notion that technology can enrich children's lives.

Ultimately, we aim to argue for the substantial involvement of schools and families in creating school mental health systems for children. We intend to bolster this claim by seeking robust evidence in future research.

ACKNOWLEDGMENTS

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI22C0775).

REFERENCES

[1] Achraf Ammar, Khaled Trabelsi, Michael Brach, Hamdi Chtourou, Omar Boukhris, Liwa Masmoudi, Bassem Bouaziz, Ellen Bentlage, Daniella How, Mona Ahmed, Patrick Mueller, Notger Mueller, Omar Hammouda, Laisa Liane Paineiras-Domingos, Annemarie Braakman-jansen, Christian Wrede, Sophia Bastoni, Carlos Soares Pernambuco, Leonardo Mataruna, Morteza Taheri, Khadijeh Irandoust, Aïmen Khacharem, Nicola L Bragazzi, Jana Strahler, Jad Adrian, Albina Andreeva, Jordan M Glenn, Nicholas T Bott, Faiez Gargouri, Lotfi Chaari, Hadj Batatia, Samira C khoshnami, Evangelia Samara, Vasiliki Zisi, Parasanth Sankar, Waseem N Ahmed, Gamal Mohamed Ali, Osama Abdelkarim, Mohamed Jarraya, Kais El Abed, Wassim Moalla, Mohamed Romdhani, Asma Aloui, Nizar Souissi, Lisette Van GemertPijnen, Bryan L Riemann, Laurel Riemann, Jan Delhey, Jonathan Gómez-Raja, Monique Epstein, Robbert Sanderman, Sebastian Schulz, Achim Jerg, Ramzi Al-Horani, Taysir Mansi, Mohamed Jmail, Fernando Barbosa, Fernando Ferreira-Santos, Boštjan Šimunič, Rado Pišot, Saša Pišot, Andrea Gaggioli, Piotr Zmijewski, Stephen J Bailey, Jürgen Steinacker, Karim Chamari, Tarak Driss and Anita Hoekelmann. 2021. Effects of home confinement on mental health and lifestyle behaviours during the COVID-19 outbreak: Insight from the ECLB-COVID19 multicenter study. Biology of sport. 38, 1. 9-21. https://doi.org/10.5114/biolsport.2020.96857

[2] Adele Diamond and Daphne S. Ling. 2019. Review of the evidence on, and fundamental questions about, efforts to improve executive functions, including working memory. in Jared M. Novick, and others (eds), Cognitive and Working Memory Training: Perspectives from Psychology, Neuroscience, and Human Development. Oxford Academic.

[3] Akın Tahıllıoğlu, Öznur Bilaç, Taciser Uysal and Eyüp Sabri Ercan. 2021. Who predict ADHD with better diagnostic accuracy?: Parents or teachers?. Nordic Journal of Psychiatry. 75, 3. 214-223. https://doi.org/10.1080/08039488.2020.1867634

[4] Albert W. Wienen, Maruschka N. Sluiter, Ernst Thoutenhoofd, Peter de Jonge and Laura Batstra. 2019. The advantages of an ADHD classification from the perspective of teachers. European Journal of Special Needs Education. 34, 5. 649-662. https://doi.org/10.1080/08856257.2019.1580838

[5] Andrea B. Martinez, Melissa Co, Jennifer Lau and June S. L. Brown. 2020. Filipino help-seeking for mental health problems and associated barriers and facilitators: a systematic review. Social Psychiatry and Psychiatric Epidemiology. 55. 1397-1413. https://doi.org/10.1007/s00127-020-01937-2

[6] Anouck I. Staff, Barbara J. van den Hoofdakker, Saskia van der Oord, Rianne Hornstra, Pieter J. Hoekstra, Jos W. R. Twisk, Jaap Oosterlaan and Marjolein Luman. 2021. Effectiveness of specific techniques in behavioral teacher training for childhood ADHD: A randomized controlled microtrial. Journal of Clinical Child & Adolescent Psychology. 50, 6. 763-779. https://doi.org/10.1080/15374416.2020.1846542

[7] Bruce F. Pennington and Sally Ozonoff. 1996. Executive functions and developmental psychopathology. Journal of child psychology and psychiatry. 37, 1. 51-87. https://doi.org/10.1111/j.1469-7610.1996.tb01380.x

[8] Cassandra B. Romine and Cecil R. Reynolds. 2005. A model of the development of frontal lobe functioning: findings from a meta-analysis. Applied neuropsychology. 12, 4. 190-201. https://doi.org/10.1207/s15324826an1204_2

[9] Chamberlain C. Diala, C. Muntaner, C. Walrath, K. Nickerson, T LaVeist and P Leaf. 2001. Racial/ethnic differences in attitudes toward seeking professional mental health services. American Journal of Public Health. 91, 5. 805. https://doi.org/10.2105/ajph.91.5.805

[10] Clancy Blair and Rachel Peters Razza. 2007. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Child development. 78, 2. 647-663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

[11] Cynthia A. Riccio and Hilary Gomes. 2013. Interventions for executive function deficits in children and adolescents. Applied Neuropsychology: Child. 2, 2. 133-140. https://doi.org/10.1080/21622965.2013.748383

[12] Desiree W. Murray, Brooke S. G. Molina, Kelly Glew, Patricia Houck, Andrew Greiner, Dalea Fong, James Swanson, L. Eugene Arnold, Marc Lerner, Lily Hechtman, Howard B. Abikoff and Peter S. Jensen. 2014. Prevalence and characteristics of school services for high school students with attention-deficit/hyperactivity disorder. School Mental Health. 6. 264–278. https://doi.org/10.1007/s12310-014-9128-6

[13] Do Eun Park, Junghan Lee, Jeehyun Han, Jinwoo Kim and Yee Jin Shin. 2023, A Preliminary Study of Voicebot to Assist ADHD Children in Performing Daily Tasks. International Journal of Human–Computer Interaction. 1-14. https://doi.org/10.1080/10447318.2023.2169530

[14] Do Eun Park, Yee-Jin Shin, EunAh Park, In Ae Choi, Woo Yeon Song and Jinwoo Kim. 2020. Designing a Voice-Bot to Promote Better Mental Health: UX Design for Digital Therapeutics on ADHD Patients. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI'20). Association for Computing Machinery, Honolulu, HI, USA, 1-8. https://doi.org/10.1145/3334480.3382948

[15] Doeun Park, Jinwoo Kim, and Yee Jin Shin. 2023. Conversational Agent for Supporting Self-regulation in Children's Daily Activities in the Pandemic Era. Archives of Design Research. 36, 2. 7-23. https://doi.org/10.15187/adr.2023.05.36.2.7

[16] Donald H. Meichenbaum and Joseph Goodman. 1971. Training impulsive children to talk to themselves: a means of developing self-control. Journal of abnormal psychology. 77, 2. 115-126 https://doi.org/10.1037/h0030773

[17] George J. DuPaul, Thomas J. Power, Arthur D. Anastopoulos and Robert Reid. 1998. ADHD Rating Scale—IV: Checklists, norms, and clinical interpretation. Guilford Press. New York, NY.

[18] Georgia Kokkalia, Athanasios S. Drigas, Alexandra Economou and Petros Roussos. 2019. School Readiness From Kindergarten to Primary School. Int. J. Emerg. Technol. Learn. 14, 11. 4-18. https://doi.org/10.3991/ijet.v14i11.10090

[19] Gerard A. Gioia, Peter K. Isquith, Steven C. Guy and Lauren Kensworthy. 2015. Brief 2: Behavior rating inventory of executive function.

[20] Guilherme Polanczyk, Maurício Silva de Lima, Bernardo Lessa Horta, Joseph Biederman and Luis Augusto Rohde. 2007. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. American journal of psychiatry. 164, 6. 942-948. https://doi.org/10.1176/ajp.2007.164.6.942

[21] Helene J. Polatajko, Angela D. Mandich, Linda T. Miller and Jennifer J. Macnab. 2001. Cognitive orientation to daily occupational performance (CO-OP) part II the evidence. Physical & Occupational Therapy in Pediatrics. 20, 2-3. 83-106. https://doi.org/10.1080/J006v20n02_06

[22] Jae-Won Kim, Ki-Hong Park and Min-Jeong Choi. 2004. Screening for attention deficit/hyperactivity disorder in community mental health services for children. Journal of Korean Neuropsychiatric Association. 200-208.

[23] Jane E. Barker, Andrei D. Semenov, Laura Michaelson, Lindsay S. Provan, Hannah R. Snyder and Yuko Munakata. 2014. Less-structured time in children's daily lives predicts self-directed executive functioning. Frontiers in psychology. 5. 593. https://doi.org/10.3389/fpsyg.2014.00593

[24] Jenny S. Radesky, Heidi M. Weeks, Rosa Ball, Alexandria Schaller, Samantha Yeo, Joke Durnez, Matthew Tamayo-Rios, Mollie Epstein, Heather Kirkorian, Sarah Coyne and Rachel Barr. 2020. Young children's use of smartphones and tablets. Pediatrics. 146, 1. 1-8. https://doi.org/10.1542/peds.2019-3518

[25] Josef Perner and Birgit Lang. 2000. Theory of mind and executive function: Is there a developmental relationship?, Understanding other minds: Perspectives from developmental cognitive neuroscience. Oxford University Press. Oxford, NY.

[26] Judith R. Harrison, Kimberly Vannest, John Davis, and Cecil Reynolds. 2012. Common problem behaviors of children and adolescents in general education classrooms in the United States. Journal of Emotional and Behavioral Disorders. 20, 1. 55-64. https://doi.org/10.1177/1063426611421157

[27] Judy Illes, and Barbara J. Sahakian (eds). 2011. Development of the adolescent brain. Oxford handbook of neuroethics. Oxford University Press. Oxford, NY.

[28] Jum C. Nunnally. 1978. An Overview of Psychological Measurement. In: Wolman, B.B. (eds) Clinical Diagnosis of Mental Disorders. Springer, Boston, MA. https://doi.org/10.1007/978-1-4684-2490-4_4

[29] Karen J. Miller. 2005. Executive functions. Pediatric annals. 34, 4. 310-317. https://doi.org/10.3928/0090-4481-20050401-12

[30] Katya Rubia, Anna B. Smith, Michael J. Brammer and Eric Taylor. 2003. Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. Neuroimage. 20, 1. 351-358. https://doi.org/10.1016/S1053-8119(03)00275-1

[31] Laura A. Nabors, Matthew W. Reynolds and Mark D. Weist. 2000. Qualitative evaluation of a high school mental health program. Journal of Youth and Adolescence. 29, 1. 1-13. https://doi.org/10.1023/A:1005129403974

[32] Lea Auhuber, Mandy Vogel, Nico Grafe, Wieland Kiess andTanja Poulain. 2019. Leisure activities of healthy children and adolescents. International journal of environmental research and public health. 16, 12. 2078; https://doi.org/10.3390/ijerph16122078

[33] Maggie E. Toplak, Richard F. West and Keith E. Stanovich. 2013. Practitioner review: Do performance-based measures and ratings of executive function assess the same construct?. Journal of child psychology and psychiatry. 54, 2. 131-143. https://doi.org/10.1111/jcpp.12001

[34] Maria Moldavsky and Kapil Sayal. 2013. Knowledge and attitudes about attention-deficit/hyperactivity disorder (ADHD) and its treatment: the views of children, adolescents, parents, teachers and healthcare professionals. Current psychiatry reports. 15, 8. 377. https://doi.org/10.1007/s11920-013-0377-0

[35] Maria Moldavsky, Carla Groenewald, Victoria Owen and Kapil Sayal. 2013. Teachers' recognition of children with ADHD: Role of subtype and gender. Child and adolescent mental health. 18, 1. 18-23. https://doi.org/10.1111/j.1475-3588.2012.00653.x

[36] Mark D. Weist, Mark A. Sander, Christine Walrath, Benjaman Link, Laura Nabors, Steve Adelsheim, Elizabeth Moore, Jenni Jennings and Kristine Carrillo. 2005. Developing principles for best practice in expanded school mental health. Journal of Youth and Adolescence. 34. 7-13. https://doi.org/10.1007/s10964-005-1331-1

[37] Max Doan, Franceli L. Cibrian, Agnes Jang, Nihar Khare, Sean Chang, Aiyuan Li, Sabrina Schuck, Kimberley D. Lakes and Gillian R. Hayes. 2020. CoolCraig: A smart watch/phone application supporting co-regulation of children with ADHD. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI'20). Association for Computing Machinery, Honolulu, HI, USA, 1-7.

[38] Melissa L. Danielson, Rebecca H. Bitsko, Reem M. Ghandour, Joseph R. Holbrook, Michael D. Kogan and Stephen J. Blumberg. 2018. Prevalence of parent-reported ADHD diagnosis and associated treatment among US children and adolescents, 2016. Journal of Clinical Child & Adolescent Psychology. 47, 2. 199-212. https://doi.org/10.1080/15374416.2017.1417860

[39] Mina Fazel, Vikram Patel, Saji Thomas, Wietse Tol. 2014. Mental health interventions in schools in low-income and middle-income countries. The Lancet Psychiatry. 1, 5. 388-398. https://doi.org/10.1016/S2215-0366(14)70357-8

[40] Minha Hong, Young Sook Kwack, Yoo-Sook Joung, Soyoung Irene Lee, Bongseog Kim, Seok Han Sohn, Un-Sun Chung, Jaewon Yang, Soo-Young Bhang, Jun-Won Hwang, Hyung-yun Choi, In Hwan Oh, Yeon Jung Lee and Geon Ho Bahn. 2014. Nationwide rate of attention-deficit hyperactivity disorder diagnosis and pharmacotherapy in K orea in 2008–2011. Asia-Pacific Psychiatry. 6, 4. 379-385. https://doi.org/10.1111/appy.12154

[41] Nicholas P. Allan, Laura E. Hume, Darcey M. Allan, Amber L. Farrington, Christopher J. Lonigan. 2014. Relations between inhibitory control and the development of academic skills in preschool and kindergarten: a meta-analysis. Developmental psychology. 50, 10. 2368. https://doi.org/10.1037/a0037493

[42] Philip David Zelazo, Clancy B. Blair and Michael T. 2016. Executive Function: Implications for Education. NCER 2017-2000. National Center for Education Research.

[43] Rae Thomas, Sharon Sanders, Jenny Doust, Elaine Beller and Paul Glasziou. 2015. Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. Pediatrics. 135, 4. e994-e1001. https://doi.org/10.1542/peds.2014-3482

[44] Ross W. Greene, Sara K. Beszterczey, Tai Katzenstein, Kenneth Park and Jennifer Goring. 2002. Are students with ADHD more stressful to teach? Patterns of teacher stress in an elementary school sample. Journal of emotional and behavioral disorders. 10, 2. 79-89. https://doi.org/10.1177/10634266020100020201

[45] Russell A. Barkley. 1997. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychological bulletin. 121, 1. 65-94.

[46] Russell A. Barkley. 2011. The important role of executive functioning and self-regulation in ADHD. Journal of Child Neuropsychology. 113, 21. 41–56.

[47] Russell A. Barkley. 2015. Executive functioning and self-regulation viewed as an extended phenotype: Implications of the theory for ADHD and its treatment. In R. A. Barkley (Ed.), Attention-deficit hyper- activity disorder: A handbook for diagnosis and treatment (pp. 405– 434). The Guilford Press.

[48] Russell A. Barkley. 2020. Executive functions: What they are, how they work, and why they evolved. Guilford Press. New York, NY.

[49] S. Andrew Garbacz, Yen Lee, Garret J. Hall, Elizabeth A. Stormshak and Laura Lee McIntyre. 2021. Initiating Family–School Collaboration in School Mental Health through a Proactive and Positive Strengths and Needs Assessment. School Mental Health. 13, 4. 667-679. https://doi.org/10.1007/s12310-021-09455-5

[50] Sarah-Jayne Blakemore and Suparna Choudhury. 2006. Development of the adolescent brain: implications for executive function and social cognition. Journal of child psychology and psychiatry. 47, 3-4. 296-312. https://doi.org/10.1111/j.1469-7610.2006.01611.x

[51] Sharon Wigal, Phillip Chappell, Donna Palumbo, Shannon Lubaczewski, Sara Ramaker, and Richat Abbas. 2020. Diagnosis and treatment options for preschoolers with attention-deficit/hyperactivity disorder. Journal of child and adolescent psychopharmacology. 30, 2. 104-118. https://doi.org/10.1089/cap.2019.0116

[52] Sophie Leitch, Emma Sciberras, Brittany Post, Bibi Gerner, Nicole Rinehart, Jan M. Nicholson and Subhadra Evans. 2019. Experience of stress in parents of children with ADHD: A qualitative study. International journal of qualitative studies

on health and well-being. 14, 1. 1690091. https://doi.org/10.1080/17482631.2019.1690091

[53] Susan dosReis, Crystal L. Barksdale, Andrea Sherman, Kate Maloney and Alice Charach. 2010. Stigmatizing experiences of parents of children with a new diagnosis of ADHD. Psychiatric Services. 61, 8. 811-816. https://doi.org/10.1176/ps.2010.61.8.811

[54] Susan Young and J. Myanthi Amarasinghe. 2010. Practitioner Review: Non-pharmacological treatments for ADHD: A lifespan approach. Journal of Child Psychology and Psychiatry. 51, 2. 116-133. https://doi.org/10.1111/j.1469-7610.2009.02191.x

[55] Tena Velki , Zvonimir Užarević and Snježana Dubovicki. 2019. Self-Evaluated ADHD Symptoms as Risk Adaptation Factors in Elementary School Children. Društvena istraživanja. 28, 3. 503-522. https://doi.org/10.5559/di.28.3.07

[56] Thomas E. Brown. 2002. DSM-IV: ADHD and executive function impairments. Advanced studies in Medicine. 2, 25. 910-914.

[57] Yu-Kyoung So, Joo-Sun Noh, Young-Shin Kim, Seon-Gyu Ko and Yun-Joo Koh. 2002. The reliability and validity of Korean parent and teacher ADHD rating scale. Journal of the Korean Neuropsychiatric Association. 41, 2. 283-289.

[58] Zsofia K. Takacs and Reka Kassai. 2019. The efficacy of different interventions to foster children's executive function skills: A series of meta-analyses. Psychological bulletin. 145, 7. 653. https://doi.org/10.1037/bul0000195

[59] Doeun park, Myounglee Choo, Bohyun Jin, Un Sun Chung, Jinwoo Kim, Junghan Lee, and Yee Jin Shin. 2023. Utilizing a Conversational Agent to Promote Self-efficacy in Children: A Pilot Study on Low Cognitive Ability Children with Attention Deficit Hyperactivity Disorder. In Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI'23). Association for Computing Machinery, Hamburg, Germany, 1-7. https://doi.org/10.1145/3544549.3585887

[60] Eunbin Kang, and Youn Ah Kang. 2023. Counseling chatbot design: The effect of anthropomorphic chatbot characteristics on user self-disclosure and companionship. International Journal of Human–Computer Interaction. 1-15. https://doi.org/10.1080/10447318.2022.2163775

[61] Fabio Catania, Micol Spitale, and Franca Garzotto. 2023. Conversational agents in therapeutic interventions for neurodevelopmental disorders: a survey. ACM Computing Surveys. 55, 10. 1-34. https://doi.org/10.1145/3564269

[62] Fabio Catania, Tanya Talkar, Franca Garzotto, Benjamin R. Cowan, Thomas F. Quatieri, and Satrajit Ghosh. 2023. Multimodal Conversational Agents for People with Neurodevelopmental Disorders. In Proceedings of the 25th International Conference on Multimodal Interaction (ICMI'23). Association for Computing Machinery, Paris, France, 824-825. https://doi.org/10.1145/3577190.3617133

[63] Inha Cha, Sung-In Kim, Hwajung Hong, Heejeong Yoo, and Youn-kyung Lim. 2021. Exploring the use of a voice-based conversational agent to empower adolescents with autism spectrum disorder. In Proceedings of the 2021 CHI conference on human factors in computing systems (CHI'21). Association for Computing Machinery, Yokohama, Japan, 1-15. https://doi.org/10.1145/3411764.3445116

[64] Robert J. Moore, and Raphael Arar. 2019. Conversational UX design: A practitioner's guide to the natural conversation framework. Morgan & Claypool. Association for Computing Machinery. NewYork, NY, USA. https://doi.org/10.1145/3304087

[65] Yvonne J G Korpershoek, Sander Hermsen, Lisette Schoonhoven, Marieke J Schuurmans, and Jaap C A Trappenburg. 2020. User-centered design of a mobile health intervention to enhance exacerbation-related self-management in patients with chronic obstructive pulmonary disease (Copilot): mixed methods study. Journal of medical Internet research, 22, 6. e15449. https://doi.org/10.2196/15449