children

MDPI

Article

Association Between Performance Metrics of Game-Based Training Content and Reading and Mathematics Learning Difficulties in Children with ADHD

Seon-Chil Kim [1,*], Na-Yeong Kong [2], Sun-Young Lee [1] and Hyun-Suk Lee [3]

[1] Department of Biomedical Engineering, School of Medicine, Keimyung University, Daegu 42601, Republic of Korea; syl982008@gmail.com
[2] Department of Psychiatry, School of Medicine, Keimyung University, Daegu 42601, Republic of Korea; kp8937@kmu.ac.kr
[3] Department of Speech-Language Pathology, Daegu University, Gyeongsan 38453, Republic of Korea; lhskss@hanmail.net
* Correspondence: chil@kmu.ac.kr; Tel.: +82-10-4803-7773

Highlights

What are the main findings?

- Gameplay performance metrics were associated with attentional and executive-function characteristics, particularly visual selective attention, interference control, sustained attention, and inattentive symptoms in children with ADHD.
- The lower-performance group demonstrated higher CLDQ reading scores and significantly lower BASA:CT reading and mathematics T-scores than the higher-performance group.

What are the implications of the main findings?

- Game-based performance metrics may serve as digital behavioral indicators reflecting attention- and learning-related characteristics in children with ADHD.
- These metrics may have potential as supplementary screening and referral-support indicators, although further validation in larger samples is required before clinical application.

Abstract

Background/Objectives: This study investigated whether performance metrics derived from digital games played by children with attention-deficit/hyperactivity disorder (ADHD) are related to various attention indicators, learning difficulties, and academic performance and evaluated their potential as supplementary indicators of learning difficulties. Methods: Forty children aged 5–12 years diagnosed with ADHD were included. Decision time, excess touch count, last touch time, and answer rate were analyzed across four games with poor performance on those metrics during a four-week intervention. Associations between the metrics and the Comprehensive Attention Test (CAT), Korean ADHD Rating Scale (K-ARS), Colorado Learning Difficulties Questionnaire (CLDQ) reading and mathematics scores, and Basic Academic Skills Assessment: Comprehensive Test (BASA:CT) reading and mathematics T-scores were evaluated. Results: Among the 10 games, C3, C5, C7, and C9 consistently showed relatively poorer performance across the four predefined metrics. Gameplay performance metrics from the selected games were associated with visual selective attention, interference control, and sustained attention on the CAT, while excess touch count and last touch time showed nominally significant associations with K-ARS inattention scores. Following Benjamini–Hochberg false discovery rate (FDR) adjustment across 64 correlation tests, 12 of the 28 nominally significant associations remained statistically significant ($q < 0.05$). The lower-performance group had higher CLDQ reading scores and significantly lower BASA:CT reading and mathematics T-scores than the higher-performance group.

Check for updates

Academic Editors: Gyongyi Osser and Carmen Delia Nistor-Cseppento

Received: 4 November 2026
Revised: 23 November 2026
Accepted: 23 November 2026
Published: 25 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

After adjusting for age and sex, the composite gameplay difficulty score was positively associated with CLDQ reading ($B = 0.27$, $p = 0.010$) and mathematics ($B = 0.20$, $p = 0.032$) scores and negatively associated with BASA:CT reading ($B = -1.18$, $p = 0.016$) and mathematics T-scores ($B = -0.74$, $p = 0.036$). Conclusions: Rather than serving as a standalone diagnostic test for learning difficulties, this game-based tool may be developed as a supplementary screening or referral-support indicator to help identify children with ADHD, who can benefit from comprehensive academic assessment, standardized testing, and individualized educational support. However, further validation in independent samples is required for establishing the clinical utility and diagnostic accuracy of this approach.

Keywords: ADHD; game-based digital content; learning difficulties; digital behavioral indicators; reading; mathematics; early screening

1. Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a common childhood neurodevelopmental disorder characterized by core symptoms of inattention, hyperactivity, and impulsivity [1]. These can interfere with sustained task engagement, response inhibition, information retention, and self-regulation, potentially leading to poor academic achievement, incomplete tasks, inconsistent performance, and reduced motivation to learn [2]. In addition, children with ADHD frequently exhibit comorbid conditions, including anxiety, mood disorders, and learning difficulties, indicating that ADHD symptoms alone may not fully account for their academic difficulties [3].

Learning difficulties are a group of neurodevelopmental conditions characterized by persistent difficulties in the acquisition and use of reading, writing, or mathematical skills, despite intellectual abilities within the normal range and adequate educational opportunities [4]. ADHD and learning difficulties have both been associated with difficulties in cognitive and executive functions, including working memory, inhibitory control, and cognitive flexibility. These overlapping cognitive characteristics suggest that attention-related and learning-related difficulties may involve shared cognitive vulnerabilities [4,5]. In particular, inattention and impulsivity in children with ADHD may manifest as reading errors, spelling mistakes, omission of steps in mathematical problem solving, and task avoidance. Consequently, underlying deficits in academic skills may be misinterpreted as or masked by the behavioral symptoms of ADHD [6,7].

Therefore, when evaluating academic difficulties in children with ADHD, distinguishing between a performance deficit (i.e., due to ADHD-related symptoms, the child is unable to consistently demonstrate knowledge or skills that have already been acquired) and an academic skill deficit (foundational skills in reading, writing, or mathematics have not been adequately acquired) is important [8]. Performance deficits may vary with fluctuations in attentional control or impulsivity, whereas academic skill deficits require direct and systematic educational interventions that target the affected academic domain. However, conventional approaches, including clinical interviews, parent- and teacher-report questionnaires, and standardized assessments, require trained professionals and considerable assessment time, making repeated evaluations of all children with ADHD challenging in both clinical and school settings [9]. Accordingly, supplementary assessment tools that can efficiently identify the characteristics associated with learning difficulties without replacing comprehensive standardized assessments for learning difficulties are needed [10].

Game-based digital content, used as a training tool for children with ADHD, has the advantage of maintaining children's engagement through intrinsic motivation while continuously collecting fine-grained behavioral data generated during task performance.

In addition to the overall scores, detailed performance information, such as response time, repetitive or unnecessary responses, time required to complete a task, and accuracy, can be quantified. These metrics may provide useful information for examining characteristics related to attentional control, response regulation, and information processing [11].

Accordingly, the present study focused on performance metrics, including decision time, excess touch count, last touch time, and answer rate, automatically collected while children with ADHD engaged with game-based content as part of a digital training intervention. These metrics may quantitatively capture final game scores as well as response efficiency and behavioral characteristics during task performance. We hypothesized that these performance metrics would be associated with attention and cognitive function as well as learning difficulties and basic academic performance in reading and mathematics. If such associations are demonstrated, gameplay logs may have potential utility beyond simple training scores as digital behavioral indicators reflecting individual differences in attentional control and impulsivity [12].

Previous studies on game-based content for children with ADHD and learning difficulties have primarily focused on changes in attention, memory, or academic achievement following cognitive training and on the usability of the content [13,14]. Similar studies involving children with co-occurring ADHD and learning difficulties have compared cognitive performance before and after extended periods of computerized cognitive training, and serious-game studies have primarily evaluated improvements in attention following interventions [15,16]. Previous research on serious games for children with learning difficulties has included small-scale studies focusing on usability, technology acceptance, and feedback from children and educators, rather than directly evaluating improvements in cognitive performance [17]. Although recent systematic reviews of artificial intelligence and adaptive learning technologies have generally reported positive findings, no studies have been judged to have a low risk of bias, and moderate or greater methodological limitations have been identified [18]. Previous research has predominantly examined whether the use of engaging game-based tools during training improves functional outcomes [13], whereas studies directly investigating the association between gameplay performance data and learning difficulties or academic performance remain limited.

The game-based digital content used in this study was developed and implemented considering its potential applications in both clinical and school settings. The existing content was configured for use in school-based field evaluations without requiring substantial modifications. The program was designed to administer 10 games over approximately 12 min and automatically recorded information, including content type, task sequence, progress, interval-level scores, performance time, session number, and total score. In the school setting, gameplay logs were evaluated along with classroom attention, self-regulated learning ability, and learning motivation to examine the feasibility of implementation in general classroom environments. In this study, we investigated how gameplay performance data from children with ADHD who participated in the clinical study were associated with attention and cognitive function, learning difficulties in reading and mathematics, and basic academic performance.

The present study differs from previous research in three key respects. First, rather than using newly developed and unvalidated games, we investigated game-based content developed for clinical training of children with ADHD and explored its potential extension as a digital behavioral assessment approach that may support the early identification of children at risk of learning difficulties. Second, rather than relying on simple total scores or mean changes before and after the intervention, this study aimed to use detailed gameplay log metrics, including decision time, excess touch count, last touch time, and answer rate, to distinguish between ADHD-related training performance and its association with learning

difficulty-related academic difficulties. Accordingly, this study aimed to investigate the associations of digital gameplay performance metrics with attention and cognitive function, ADHD symptoms, parent-reported reading and mathematics learning difficulties, and basic academic performance of children with ADHD. Third, we evaluated the potential utility of gameplay performance data as digital behavioral indicators reflecting learning difficulties and basic academic performance characteristics in children with ADHD.

2. Materials and Methods

2.1. Study Participants

The participants were children aged 5–12 years who had been clinically diagnosed with ADHD according to the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition [19]. The inclusion criteria were as follows: voluntary written informed consent from a legal guardian and assent from the child; ability to comply with the assessments and instructions of the principal investigator; and intellectual functioning within the normal range, as assessed using the Korean Wechsler Intelligence Scale for Children–Fifth Edition [20]. The exclusion criteria included the presence of medical or psychiatric conditions other than ADHD that could affect study participation; physical limitations that interfered with the use of digital devices; markedly impaired motor function, including seizure disorders; color vision deficiency; having a family member enrolled in or participating in the same study; or any other condition that, in the investigator's judgment, made the child unsuitable for participation.

Although the eligibility criteria outlined included children aged 5 to under 12 years, no participants younger than 6 years were enrolled in the study. Thus, the K-WISC-V was administered only to children within its standardized age range.

To assess learning difficulty-related characteristics in all participants, caregivers completed the Colorado Learning Difficulties Questionnaire (CLDQ) [21]. The CLDQ was used to assess reading and mathematics learning difficulties among the participants and was not used as a criterion for participant inclusion or exclusion.

During the study design phase, a target sample size of 40 participants was established based on previous studies and practical feasibility of participant recruitment. The target sample size was not determined based on an a priori power analysis. Instead, a sensitivity analysis was performed using G*Power version 3.1.2 to determine the minimum detectable effect size for the exploratory correlation analysis given the predetermined sample size [22]. Assuming a two-sided significance level of $\alpha = 0.05$, a statistical power $(1 - \beta)$ of 0.80, and a total sample size of 40 participants, the minimum detectable effect size for Pearson correlation analysis was estimated to be an absolute correlation coefficient of $|r| = 0.42$.

The overall study flow is presented in Figure 1.

This prospective study was conducted over a 3-month period, from 26 May to 30 August 2026, at the Department of Psychiatry of K University Hospital. Before data collection and assessment, the study protocol was approved by the Keimyung University Dongsan Hospital Institutional Review Board (IRB No. 2025-03-002-003) and registered with the Clinical Research Information Service (CRIS No. KCT0010850).

The informed consent form and participant information sheet included information regarding the study objectives, participation procedures, potential risks and benefits, confidentiality, and restrictions of the collected data to research purposes only. Baseline data collection and study procedures were initiated after obtaining consent from participants and their legal guardians. Participants were selected based on a comprehensive evaluation that included clinical interviews conducted by a psychiatrist and psychological assessments including attention and intelligence tests.

Although no direct adverse effects were anticipated in the participating children, the use of digital devices may cause frustration, headaches, or dizziness. Therefore, device use was limited to less than 30 min using a program that monitored the usage time.

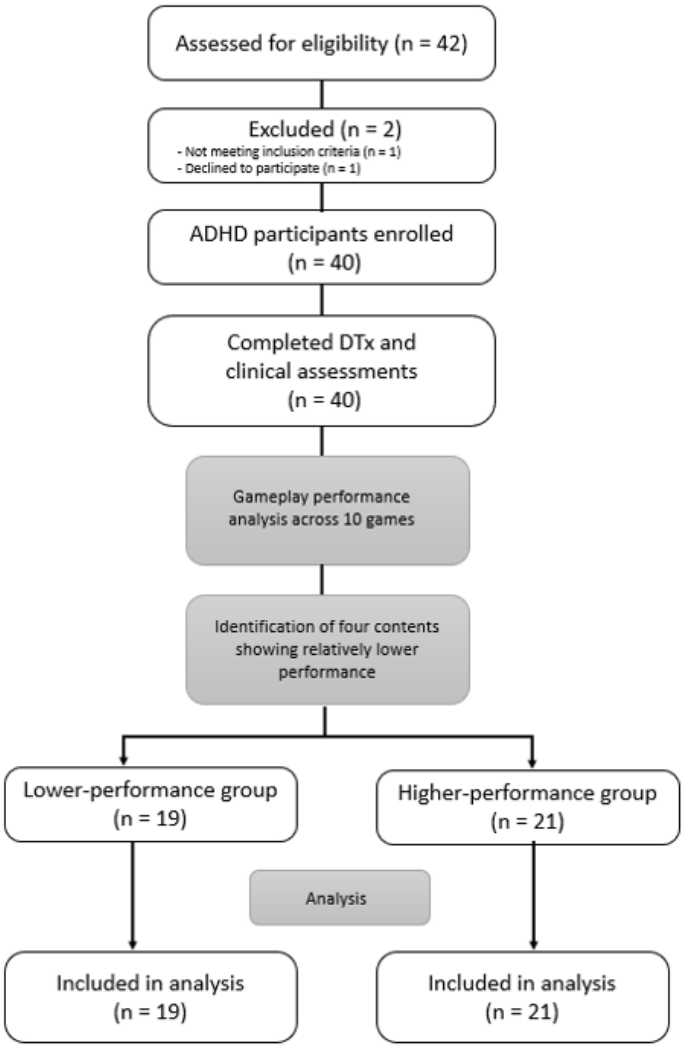

Figure 1. Flow diagram of study participants. ADHD: attention-deficit/hyperactivity disorder; DTx: digital therapeutics.

2.2. Digital Intervention Tool and Procedures

The game-based digital content used in this study was developed to improve attention and cognitive function in children with ADHD and implemented as an adjunctive therapeutic intervention. Performance characteristics were examined in relation to attention and cognitive function, ADHD symptoms, and learning difficulty-related characteristics [23].

The original clinical trial included an intervention group ($n = 20$) and a waitlist control group ($n = 20$). Participants in the waitlist control group subsequently received the same game-based digital therapeutic intervention. Accordingly, all 40 participants received and completed the four-week intervention. For the present study, gameplay performance metrics were collected during Week 4 of each participant's intervention, regardless of the original group allocation, and analyzed collectively to examine their associations with attention- and learning-related outcomes.

The ADHD digital intervention consisted of game-based content requiring various cognitive functions (InderTech, Version 2.0, Korea). Objective performance data for each participant were automatically recorded during gameplay. In this study, four gameplay

performance metrics—decision time, excess touch count, last touch time, and answer rate—were selected for analysis from the performance metrics collected across 10 games. Decision time and last touch time were used to reflect the temporal characteristics of task performance, excess touch count was used to indicate the frequency of unnecessary additional responses, and the answer rate was used to reflect task accuracy. Decision time was calculated as the time interval between the first and last screen touches. Excess touch count was calculated by subtracting the number of touches required for correct responses from the total number of screen touches. Last touch time was defined as the elapsed time until the final touch required to complete each task. The total touch count was excluded from the analysis because its interpretation could vary depending on the task structure and gameplay mechanics of each game. Each of the four metrics was analyzed separately to determine whether the performance characteristics observed during the specific games were associated with attention and cognitive function, ADHD symptoms, and learning difficulty-related characteristics.

The digital intervention was adaptively configured according to each participant's performance pattern, including accuracy, response speed, and error characteristics. Based on these performance characteristics, the combination of games, difficulty level, feedback, and training schedule was automatically individualized for each participant [24].

Game difficulty was adaptively adjusted based on each participant's gameplay performance. Difficulty levels increased when participants demonstrated better performance, whereas the current difficulty level was maintained when performance was insufficient for progression.

During Week 1, all participants completed the same 10 games, each comprising 10 levels, to assess their initial performance characteristics. The training content and difficulty levels were adjusted based on quantitative indicators, including the performance evaluation scores obtained during the digital training sessions. Subsequently, in weeks 2–4 (three weeks), training was conducted according to the protocol established for each training program. For the analyses in the present study, gameplay performance metrics collected during the final week of the four-week intervention (Week 4) were obtained from four selected games (C3, C5, C7, and C9). Based on the initial performance data obtained during Week 1, individual patterns of accuracy, response speed, and errors (e.g., omissions and incorrect responses) were analyzed, and the game combination, difficulty level, and feedback were subsequently adjusted accordingly. During each game session, performance data including accuracy, response speed, and error characteristics were automatically recorded and analyzed. These data were used to derive gameplay performance metrics including decision time, total touch count, excess touch count, last touch time, and answer rate. The training process was also quantitatively controlled using performance outcomes derived from the digital games, which were consistent with the metrics used during the assessment. Considering the individualized adjustments implemented during weeks 2–4, these metrics were considered indicators of gameplay performance within the adaptive training environment rather than standardized measurements obtained under identical training conditions across participants.

Basic clinical and demographic characteristics, including age, sex, intelligence quotient (IQ), and information on pharmacological treatment, were collected from all participants. The participating children received game-based digital content as an adjunctive intervention for ADHD while continuing their existing pharmacological treatment. The intervention effects were evaluated using pre- and post-intervention assessments. Participants were permitted to continue their ADHD medications that had been prescribed before study enrollment; however, initiation of new medications during the study period was not allowed. Concomitant medications permitted during the study included methylphenidate- and

atomoxetine-based medications, which were maintained within the treatment regimens prescribed prior to study participation.

2.3. Participant Assessment Tools

To assess attention and cognitive function and ADHD symptom characteristics of participants, the Comprehensive Attention Test (CAT) and Korean ADHD Rating Scale (K-ARS) were used, respectively. The CLDQ and Basic Academic Skills Assessment: Comprehensive Test (BASA:CT) were used to assess learning difficulty-related characteristics.

First, the CAT was used to assess attention and cognitive function in children with ADHD. The CAT consists of six subtests that assess attention and cognitive function: visual selective attention, auditory selective attention, sustained attention to response, flanker, divided attention, and working memory [25]. In the present study, four CAT subtests—visual selective attention, auditory selective attention, flanker, and sustained attention to response—were included in the analyses to examine their associations with gameplay performance and learning-difficulty-related characteristics. For each subtest, the sensitivity factor, which reflects the ability to discriminate between the target and nontarget stimuli, was used in the analysis. A sensitivity factor of ≥2 was interpreted as indicating effective discrimination between target and nontarget stimuli [26].

Second, the K-ARS was used to assess behavioral symptoms of ADHD. The K-ARS is an 18-item parent- and teacher-report measure comprising two domains: inattention and hyperactivity/impulsivity. In the present study, the K-ARS was completed by the parents [27]. Each item is rated on a 4-point Likert scale ranging from 0 (never or rarely) to 3 (very often), yielding a total score ranging from 0 to 54, with higher scores indicating greater ADHD symptom severity [28]. In the present analysis, we examined the associations of the K-ARS inattention subscale and total scores with gameplay performance and learning difficulty-related characteristics.

In addition, ADHD symptom severity was clinically assessed by a psychiatrist using the Clinical Global Impression–Severity (CGI-S) scale during participant screening.

Third, the CLDQ was used to assess participants' learning difficulty-related characteristics [21]. In this study, the original CLDQ items were translated into Korean and reviewed by the research team to ensure that the meaning of each item was preserved. The review focused on semantic consistency with the original questionnaire, the clarity of the Korean wording, and the comprehensibility of the items for parent respondents. However, the Korean version used in this study did not undergo a separate formal psychometric validation process. The CLDQ is a parent-report measure that comprises five domains: reading, mathematics, social cognition, social anxiety, and spatial [29]. The reading and mathematics subscales, which were relevant to the objectives of the present study, were used. The reading domain consisted of six items, whereas the mathematics domain comprised five items, including three items from the original scale and two additional items [30]. Accordingly, 11 items, including 6 reading items and 5 mathematics items, were included in the analysis (Figure 2). Each item was rated on a 5-point Likert scale ranging from 1 to 5. The subscale scores were calculated as the mean of the corresponding items, with higher scores indicating greater learning difficulty in the respective domain [31]. In this study, CLDQ scores were not used as criteria for the clinical diagnosis of learning difficulties, but rather as continuous measures of the severity of learning difficulty-related characteristics. Internal consistency of the CLDQ subscales was assessed using Cronbach's α in the study sample; Cronbach's α values were 0.84 for reading and 0.81 for mathematics.

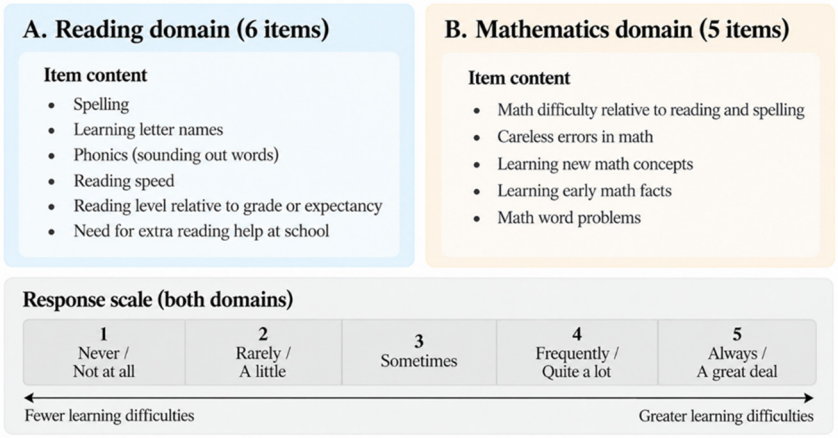

Figure 2. Reading (6 items) and mathematics (5 items) domains of the Colorado Learning Difficulties Questionnaire used in this study.

Finally, BASA:CT was used as a standardized assessment of children's basic academic performance to complement the parent-reported CLDQ. To maintain consistency with the learning domains assessed using the CLDQ, the reading and mathematics domains of the BASA:CT were included in the analysis. Performance in each domain was analyzed using T-scores calculated from grade-specific normative references corresponding to each participant's school grade, with higher T-scores indicating better academic performance in the corresponding domain [32]. BASA:CT results were not used as criteria for the clinical diagnosis of learning difficulties; instead, they were treated as continuous measures reflecting individual differences in participants' reading and mathematics academic performance. Thus, parent-reported learning difficulty-related characteristics, assessed using the CLDQ, and basic academic performance, assessed using the standardized BASA:CT, were analyzed as complementary measures.

2.4. Digital Content Performance Characteristics

To evaluate the performance characteristics during training, the performance on each game was quantitatively analyzed, and the following metrics were collected: decision time (s), touch count (n), excess touch count (n), last touch time (s), and answer rate (%). All gameplay performance metrics used in the analysis were derived from data collected during Week 4.

Higher values for decision time, excess touch count, and last touch time indicated a relatively poor performance, and a lower answer rate indicated a relatively poor performance. Considering that the total touch count could vary depending on the task structure and number of responses required for each game, its direction could not be consistently interpreted as indicating poorer performance. Therefore, total touch count was included only in the descriptive statistics and was excluded from game selection and subsequent statistical analyses.

To identify the games showing characteristics of relatively poor performance, gameplay data from all 10 games were exploratorily compared. Four metrics for which the direction of poorer performance could be consistently interpreted—decision time, excess touch count, last touch time, and answer rate—were applied equally to game selection. The 10 games were ranked according to their mean values for each performance metric. The decision time, excess touch count, and last touch time were ranked from highest to lowest, whereas the answer rate was ranked from lowest to highest. The games that were consistently ranked among the four poorest-performing games across all four metrics were defined as relatively low-performance games and selected for subsequent analyses. The total touch

count was excluded from the game-selection criteria because it could vary according to the task structure and the required number of responses for each game, rendering consistent interpretation of its direction as poorer performance difficult. Game selection was based exclusively on gameplay performance data; the CAT, K-ARS, and CLDQ results were not used in the selection process.

Because the games differed in their mechanics, cognitive demands, and inherent task difficulty, these differences were not separately adjusted for during the exploratory game-selection process. Accordingly, the observed differences in gameplay performance were interpreted as relative performance characteristics within the specific game tasks rather than as standardized comparisons of task difficulty or cognitive ability.

For each selected game, the mean and standard deviation of each performance metric were calculated across participants for descriptive comparison and game selection only. These group-level summary statistics were not used as observations in the subsequent correlation analyses. Instead, participant-level gameplay performance values obtained during Week 4 were used to examine associations with the corresponding clinical assessment measures.

For subsequent analyses, a participant-level composite gameplay difficulty score was calculated based on the selected contents. For each content, the four performance metrics were standardized as z-scores across the participants [33]. To ensure a consistent direction across all performance metrics, the sign of the z-score for the answer rate was reversed, such that higher values for all standardized metrics indicated a relatively poorer gameplay performance. The mean of the four standardized metrics was then calculated for each game to derive a game-specific gameplay difficulty score. These game-specific scores were subsequently averaged across the four selected games to obtain a composite gameplay difficulty score for each participant. Thus, a higher composite gameplay difficulty score indicated poorer gameplay performance. To compare learning difficulty-related characteristics according to the participants' relative levels of gameplay performance, participants were classified into a lower-performance group and a higher-performance group based on the median composite gameplay difficulty score. Participants whose scores were equal to the median were assigned to the higher-performance group. This classification resulted in 19 participants in the lower-performance group and 21 participants in the higher-performance group. This grouping was not based on a clinical diagnosis or an established clinical cut-off but was used solely as an analytical classification to distinguish the relative levels of gameplay performance within the study sample. Because the median-based classification was intended for exploratory group comparisons, the composite gameplay difficulty score was also analyzed as a continuous variable in multiple linear regression models to examine its associations with learning difficulty-related characteristics without dichotomizing the score.

2.5. Types of Game-Based Digital Content

As shown in Figure 3, although the games administered to the children with ADHD differed in their specific task characteristics, they were designed to involve a range of cognitive demands related to attentional control and executive function. Accordingly, the performance on these 10 games required cognitive processes related to selective attention, sustained attention, response inhibition, interference control, and working memory. Therefore, we evaluated whether gameplay performance metrics reflecting these cognitive characteristics are associated with objective measures of attention and cognitive function. In particular, we analyzed the associations between gameplay performance metrics and the results of the visual selective attention, auditory selective attention, flanker, and sustained attention to response tasks in the CAT. Additionally, to evaluate the associations between gameplay

performance characteristics and ADHD symptoms, gameplay performance metrics were analyzed in relation to the K-ARS inattention subscale and total scores.

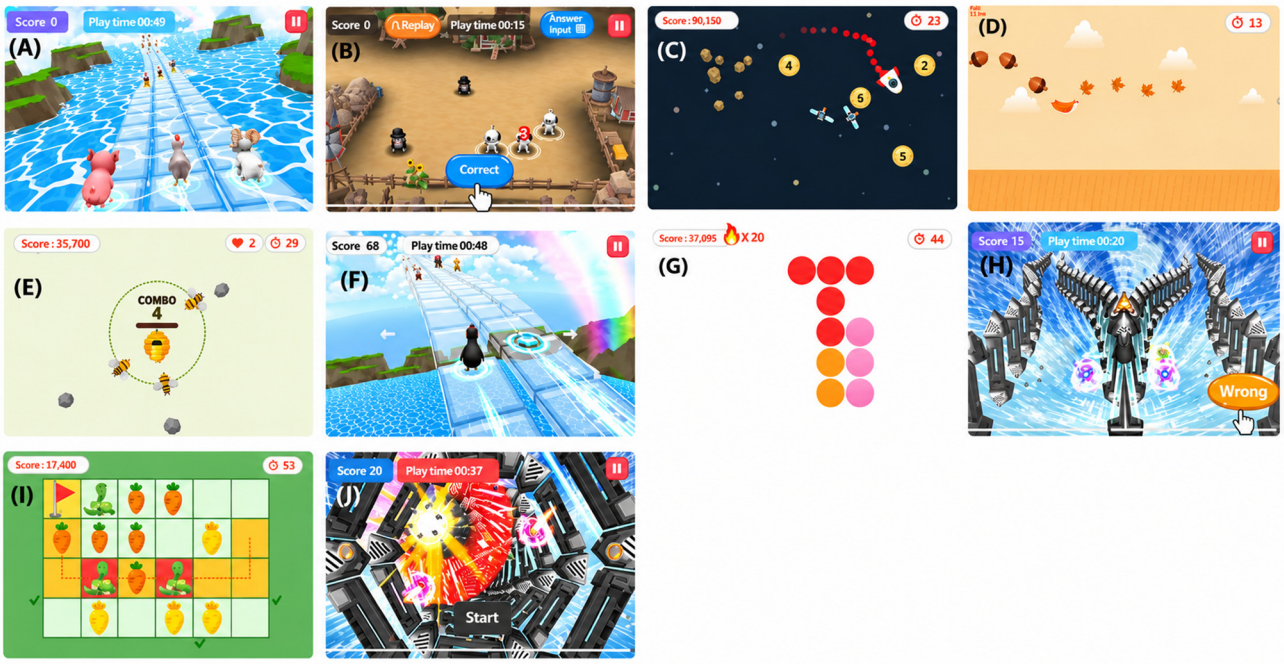

Figure 3. Representative screens of the 10 game-based digital content interventions used by children with ADHD. (**A**) C1: A game in which participants touch the screen to rearrange the order of animals on a glass bridge so that they match the target animal shown moving ahead. (**B**) C2: A game in which participants listen to the type of animal presented through an auditory cue at the beginning of the game, count the number of corresponding animals, and enter the correct number. (**C**) C3: A time-limited game in which participants guide a spaceship according to a presented numerical sequence while avoiding flying obstacles. (**D**) C4: A time-limited game in which participants repeatedly touch the screen to control the position of a flying bird and respond to the presented tasks and targets. (**E**) C5: A game in which participants distinguish flowers from rocks through visual observation to help a bee survive. (**F**) C6: A game in which participants touch the screen to change the position of an animal while attempting to maintain the required position. (**G**) C7: A game in which circles of the same color are grouped into sets, and players tap the sets in ascending order of the number of circles, from the fewest to the most. (**H**) C8: A game in which participants simultaneously control two spaceships independently while avoiding various obstacles. (**I**) C9: A game in which players remember the initially presented locations of snakes and draw a path for a rabbit to collect carrots and reach the flag while avoiding the snakes. (**J**) C10: A game in which two spaceships move in parallel, and players tap the screen to rotate the spaceships clockwise to collect the presented items.

The digital intervention content provided to participants was originally developed as training content targeting cognitive processes, such as exploration and observation, selective memory, selective attention, visuoperceptual memory, sustained attention, and target-related memory [23]. The intervention was designed in a game-based format and incorporated characters and storylines to enhance children's motivation and engagement. The participants progressed through the games by solving tasks with varying difficulty levels within the storyline [34]. The content was designed to engage memory, attention, and broader cognitive functions during problem solving. Progression to subsequent stages required the participants to achieve predefined quantitative performance criteria. Accordingly, their performance during gameplay could be objectively quantified using behavioral data automatically generated during task completion, enabling the quantitative assessment of individual performance characteristics.

2.6. Statistical Analysis

Descriptive statistics were used to summarize the participants' general characteristics and main study variables. Continuous variables are presented as mean ± standard deviation, and categorical variables are presented as frequencies and percentages. To characterize gameplay performance across all 10 digital therapeutics games, the mean and standard deviation for each game were calculated for decision time, total touch count, excess touch count, last touch time, and answer rate. The four games were selected based on their consistently lower performance across the predefined gameplay metrics, independent of their associations with clinical outcomes.

Pearson correlation analyses were performed to evaluate the associations between gameplay performance characteristics of the selected games and attention and cognitive function, and ADHD symptom characteristics in children with ADHD. For each participant, decision time, excess touch count, last touch time, and answer rate were separately averaged across the four selected games (C3, C5, C7, and C9) to obtain participant-level representative values for the correlation analyses. For the CAT, scores from visual selective attention, auditory selective attention, flanker, and sustained attention to response tasks were included in the analyses. For the K-ARS, the inattention subscale score and total score were analyzed. In addition, Pearson correlation analyses were performed to evaluate the associations between learning difficulty-related characteristics and attention and cognitive function, ADHD symptoms, and gameplay performance characteristics. The CLDQ reading and mathematics scores and BASA:CT reading and mathematics T-scores were each correlated with the CAT measures, K-ARS scores, and participant-level representative values of the gameplay performance metrics described above.

An independent-samples *t*-test was performed to compare the CLDQ reading and mathematics scores and BASA:CT reading and mathematics T-scores between the relatively lower- and higher-performance groups, which were classified based on the median composite gameplay difficulty score. The effect sizes for between-group differences were calculated using Cohen's d [35].

In addition, multiple linear regression analyses were performed to evaluate the associations between the composite gameplay difficulty score and learning difficulty-related characteristics, after adjusting for age and sex. The CLDQ reading and mathematics scores and BASA:CT reading and mathematics T-scores were entered separately as dependent variables. The composite gameplay difficulty score was included as the primary independent variable, and age and sex were included as covariates. Age and sex were selected as covariates because of their potential associations with cognitive development, gameplay performance, and academic functioning. Considering the relatively small sample size, a parsimonious adjustment strategy was adopted to limit model complexity and reduce the risk of overfitting. IQ was primarily assessed to determine study eligibility by confirming that participants possessed sufficient intellectual functioning to complete the assessments and digital training tasks. Therefore, IQ was not included as an additional covariate in the primary regression models. Although medication information was also collected, it was not included in the adjusted models. The assumptions of multiple linear regression were assessed using residual-versus-fitted plots for linearity and homoscedasticity, Q-Q plots for residual normality, variance inflation factors for multicollinearity, and Cook's distance for influential observations. Regression results are presented as unstandardized regression coefficients (B), standard errors (SE), standardized regression coefficients (β), 95% confidence intervals (CIs) for B, *p*-values, and adjusted R^2 values [36]. All statistical analyses were performed using SPSS software (version 23.0; IBM Corp., Armonk, NY, USA). All statistical tests were two-sided, and nominal statistical significance was defined as $p < 0.05$. Given the exploratory nature of the correlation analyses and the number of statistical com-

parisons, an additional sensitivity analysis was conducted using the Benjamini–Hochberg false discovery rate (FDR) procedure. The FDR adjustment was applied collectively to the 64 correlation tests reported in Section 3. The original Pearson correlation coefficients and nominal *p*-values, together with the corresponding FDR-adjusted q-values, are presented in Section 3. Associations with $q < 0.05$ were considered statistically significant after FDR adjustment. Results were interpreted with consideration of both nominal and FDR-adjusted statistical significance. Considering the individualized nature of the intervention, potential differences in training difficulty and cumulative training exposure were not separately controlled for in these analyses, and the observed associations were interpreted within the context of the adaptive training environment.

3. Results

3.1. Characteristics of the Study Participants

The general characteristics of the study participants, baseline ADHD symptom severity, and post-intervention clinical and cognitive assessment results are presented in Table 1. A total of 40 children with ADHD were included in the analysis, with a mean age of 8.74 ± 1.62 years. Of these, 29 (72.5%) were boys and 11 (27.5%) were girls. The mean score on the Clinical Global Impression–Severity (CGI-S) scale, assessed by a psychiatrist before the intervention, was 3.6 ± 0.6.

Table 1. Participant characteristics, baseline ADHD symptom severity, and post-intervention assessment results.

Variable	Total ($N = 40$)
Age (in years)	8.74 ± 1.62
Sex, *n* (%)	
Male	29 (72.5)
Female	11 (27.5)
School grade, *n* (%)	
Grade 1	7 (17.5)
Grade 2	9 (22.5)
Grade 3	11 (27.5)
Grade 4	8 (20.0)
Grade 5	5 (12.5)
Grade 6	0 (0.0)
CAT (sensitivity factor)	
Visual selective attention	2.31 ± 0.82
Auditory selective attention	2.18 ± 0.79
Flanker	1.96 ± 0.76
Sustained attention to response	1.84 ± 0.71
K-ARS	
Inattention	9.6 ± 4.8
Total score	18.9 ± 8.2
CLDQ	
Reading	3.6 ± 0.6
Mathematics	2.8 ± 0.5
BASA:CT (T-score)	
Reading	48.4 ± 2.6
Mathematics	44.2 ± 2.5
Composite	46.3 ± 2.6
CGI-S score at baseline	3.6 ± 0.6

Values are presented as mean ± standard deviation (SD) or *n* (%). CAT: Comprehensive Attention Test; K-ARS: Korean ADHD Rating Scale; CLDQ: Colorado Learning Difficulties Questionnaire; BASA:CT: Basic Academic Skills Assessment: Comprehensive Test.

Following the intervention, the mean sensitivity factors for visual selective attention, auditory selective attention, the Flanker task, and the sustained attention to response task on the Comprehensive Attention Test (CAT) were 2.31 ± 0.82, 2.18 ± 0.79, 1.96 ± 0.76, and 1.84 ± 0.71, respectively. The mean inattention subscale and total scores on the Korean ADHD Rating Scale (K-ARS) were 9.6 ± 4.8 and 18.9 ± 8.2, respectively. The mean reading and mathematics scores on the Colorado Learning Difficulties Questionnaire (CLDQ) were 3.6 ± 0.6 and 2.8 ± 0.5, respectively. The mean reading, mathematics, and composite T-scores on the Basic Academic Skills Assessment: Comprehensive Tests (BASA:CT) were 48.4 ± 2.6, 44.2 ± 2.5, and 46.3 ± 2.6, respectively.

3.2. Gameplay Performance Characteristics and Selection of Relatively Low-Performance Games

Gameplay performance characteristics across all 10 digital intervention games are presented in Table 2. A comparison of gameplay performance across the 10 games showed that the same four games (C3, C5, C7, and C9; Figure 3) consistently ranked among the four poorest-performing games across all four metrics: decision time, excess touch count, last touch time, and answer rate (Table 2). Based on this consistent performance pattern, C3, C5, C7, and C9 were selected as relatively low-performance games and were included in subsequent analyses. Among these games, C7 showed the most pronounced pattern of relatively poor performance across all four metrics, with the longest decision time (3.06 ± 0.91 s), highest excess touch count (7.3 ± 3.4), longest last touch time (5.01 ± 1.22 s), and lowest answer rate (65.9 ± 13.1%).

Table 2. Gameplay performance across the 10 digital therapeutic games.

Content	Decision Time (s)	Touch Count (*n*)	Excess Touch Count (*n*)	Last Touch Time (s)	Answer Rate (%)
C1	1.82 ± 0.51	214.3 ± 5.1	2.4 ± 1.7	3.21 ± 0.74	86.4 ± 8.2
C2	1.94 ± 0.56	25.1 ± 5.4	2.7 ± 1.9	3.38 ± 0.81	84.7 ± 9.1
C3	2.91 ± 0.83	31.8 ± 7.2	6.8 ± 3.1	4.82 ± 1.16	68.5 ± 12.4
C4	1.76 ± 0.48	23.7 ± 4.8	2.2 ± 1.6	3.12 ± 0.69	88.1 ± 7.6
C5	2.74 ± 0.79	30.5 ± 6.9	6.1 ± 2.8	4.57 ± 1.08	71.2 ± 11.8
C6	2.03 ± 0.61	25.8 ± 5.7	3.0 ± 2.0	3.45 ± 0.86	83.3 ± 9.5
C7	3.06 ± 0.91	33.1 ± 7.5	7.3 ± 3.4	5.01 ± 1.22	65.9 ± 13.1
C8	1.89 ± 0.54	24.9 ± 5.2	2.6 ± 1.8	3.29 ± 0.77	85.5 ± 8.7
C9	2.83 ± 0.85	31.2 ± 7.1	6.5 ± 3.0	4.69 ± 1.13	69.7 ± 12.2
C10	1.97 ± 0.58	25.4 ± 5.5	2.9 ± 1.9	3.41 ± 0.82	84.0 ± 9.0

Higher decision times, excess touch counts, and last touch times indicate poorer gameplay performance, and lower answer rates indicate poorer performance. Touch count was considered a descriptive behavioral metric because its interpretation may vary according to the task structure.

Relatively low-performance games based on participant-level analysis are shown in Figure 4. These games generally required participants to observe and remember the presented patterns, infer or reproduce the same patterns within a limited time, and perform tasks involving calculation, memory, observation, and attention. A common characteristic across the four selected games was the requirement to process, retain, and predict patterns, suggesting that performance on these games may involve multiple cognitive processes related to attentional control and executive function.

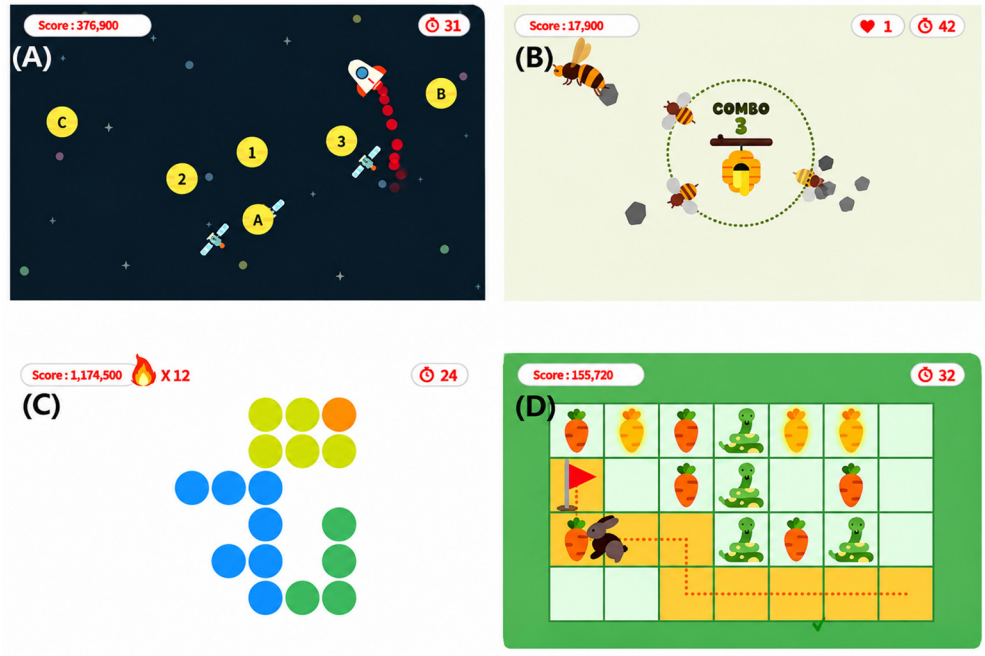

Figure 4. Game-based digital interventions identified as relatively low-performance games among children with ADHD. (**A**) C3: A time-limited game in which participants avoid objects presented on the screen by anticipating their movements and patterns. Successful performance requires processing and predicting movement patterns within a limited time. (**B**) C5: A game in which participants distinguish flowers from rocks, requiring them to remember and discriminate between identical and similar visual forms. (**C**) C7: A game requiring participants to continuously retain and track changes in colors and patterns. (**D**) C9: This task requires spatial memory and route-planning abilities to remember the initially presented locations of snakes and plan a path to reach the target while avoiding the snakes and collecting carrots.

3.3. Correlations of Gameplay Performance Metrics with CAT and K-ARS

The correlations between the gameplay performance metrics of the selected games and the CAT and K-ARS scores are presented in Table 3. The CAT visual selective attention task showed a significant negative correlation with decision time ($r = -0.55$, $p < 0.001$) and a significant positive correlation with answer rate ($r = 0.44$, $p = 0.004$). The flanker task exhibited significant negative correlations with decision time ($r = -0.45$, $p = 0.004$), excess touch count ($r = -0.39$, $p = 0.013$), and last touch time ($r = -0.35$, $p = 0.027$). The sustained attention to response task demonstrated significant negative correlations with excess touch count ($r = -0.38$, $p = 0.016$) and last touch time ($r = -0.50$, $p = 0.001$). By contrast, the auditory selective attention task did not significantly correlate with any of the gameplay performance metrics.

The K-ARS inattention score showed significant positive correlations with excess touch count ($r = 0.49$, $p = 0.001$) and last touch time ($r = 0.37$, $p = 0.019$), and a significant negative correlation with answer rate ($r = -0.34$, $p = 0.032$). By contrast, the K-ARS total score did not significantly correlate with any of the gameplay performance metrics. These findings suggest that gameplay performance metrics may be associated with selected measures of attentional control and inattentive symptoms. To account for multiple comparisons, the Benjamini–Hochberg false discovery rate (FDR) procedure was applied to all 64 correlation tests reported in Tables 3–5. As shown in Table 3, five associations remained statistically significant after FDR adjustment ($q < 0.05$): visual selective attention with decision time ($r = -0.55$, $q = 0.0160$) and answer rate ($r = 0.44$, $q = 0.0284$), Flanker performance with decision time ($r = -0.45$, $q = 0.0284$), sustained attention with last touch time ($r = -0.50$, $q = 0.0160$), and K-ARS inattention scores with excess touch count ($r = 0.49$, $q = 0.0160$).

Table 3. Correlations between gameplay performance metrics and selected CAT and K-ARS measures, with FDR-adjusted q-values.

Assessment	Task/Subscale	Decision Time	Excess Touch Count	Last Touch Time	Answer Rate
CAT	Visual selective attention task	−0.55 ** (<0.001)	−0.30 (0.060)	−0.25 (0.120)	0.44 ** (0.004)
	FDR q-value	**0.0160**	0.1038	0.1422	**0.0284**
	Auditory selective attention task	−0.29 (0.069)	−0.31 (0.052)	−0.30 (0.060)	0.28 (0.080)
	FDR q-value	0.1077	0.1038	0.1038	0.1191
	Flanker task	−0.45 ** (0.004)	−0.39 * (0.013)	−0.35 * (0.027)	0.25 (0.120)
	FDR q-value	**0.0284**	0.0520	0.0720	0.1422
	Sustained attention to response task	−0.25 (0.120)	−0.38 * (0.016)	−0.50 ** (0.001)	0.30 (0.060)
	FDR q-value	0.1422	0.0539	**0.0160**	0.1038
K-ARS	Inattention	0.30 (0.060)	0.49 ** (0.001)	0.37 * (0.019)	−0.34 * (0.032)
	FDR q-value	0.1038	**0.0160**	0.0608	0.0759
	Total score	0.26 (0.105)	0.25 (0.120)	0.30 (0.060)	−0.29 (0.069)
	FDR q-value	0.1400	0.1422	0.1038	0.1077

Values are presented as Pearson correlation coefficients (r), with nominal (unadjusted) *p*-values in parentheses and corresponding false discovery rate (FDR)-adjusted q-values in separate rows. Bold q-values indicate statistical significance after FDR adjustment ($q < 0.05$). * $p < 0.05$ and ** $p < 0.01$ indicate nominal statistical significance before FDR adjustment.

Table 4. Correlations between CLDQ reading and mathematics scores and selected CAT, K-ARS, and gameplay performance metrics, with FDR-adjusted q-values.

Assessment	Task/Subscale	CLDQ Reading	CLDQ Mathematics
CAT	Visual selective attention task	−0.32 * (0.044)	0.26 (0.105)
	FDR q-value	0.1006	0.1400
	Auditory selective attention task	−0.13 (0.424)	−0.10 (0.541)
	FDR q-value	0.4377	0.5496
	Flanker task	−0.40 * (0.011)	−0.39 * (0.013)
	FDR q-value	0.0503	0.0520
	Sustained attention to response task	−0.45 ** (0.004)	−0.36 * (0.023)
	FDR q-value	**0.0284**	0.0669
K-ARS	Inattention	0.43 ** (0.006)	0.35 * (0.027)
	FDR q-value	**0.0349**	0.0720
	Total score	0.26 (0.105)	0.15 (0.355)
	FDR q-value	0.1400	0.3735
Gameplay performance	Decision time	0.26 (0.105)	0.31 (0.052)
	FDR q-value	0.1400	0.1038
	Excess touch count	0.46 ** (0.003)	0.43 ** (0.006)
	FDR q-value	**0.0284**	**0.0349**

Table 4. *Cont.*

Assessment	Task/Subscale	CLDQ Reading	CLDQ Mathematics
Gameplay performance	Last touch time	0.24 (0.136)	0.29 (0.069)
	FDR q-value	0.1583	0.1077
	Answer rate	−0.49 ** (0.001)	−0.30 (0.060)
	FDR q-value	**0.0160**	0.1038

Values are presented as Pearson correlation coefficients (r), with nominal (unadjusted) *p*-values in parentheses and corresponding false discovery rate (FDR)-adjusted q-values in separate rows. Bold q-values indicate statistical significance after FDR adjustment ($q < 0.05$). * $p < 0.05$ and ** $p < 0.01$ indicate nominal statistical significance before FDR adjustment.

Table 5. Correlations between BASA:CT reading and mathematics T-scores and selected CAT, K-ARS, and gameplay performance metrics, with FDR-adjusted q-values.

Assessment	Task/Subscale	Reading T-Score	Mathematics T-Score
CAT	Visual selective attention task	0.30 (0.060)	0.25 (0.119)
	FDR q-value	0.1038	0.1422
	Auditory selective attention task	0.15 (0.356)	0.08 (0.624)
	FDR q-value	0.3735	0.6240
	Flanker task	0.38 * (0.016)	0.42 ** (0.007)
	FDR q-value	0.0539	**0.0373**
	Sustained attention to response task	0.36 * (0.023)	0.38 * (0.016)
	FDR q-value	0.0669	0.0539
K-ARS	Inattention	−0.40 * (0.011)	−0.28 (0.080)
	FDR q-value	0.0503	0.1191
	Total score	−0.26 (0.105)	−0.21 (0.193)
	FDR q-value	0.1400	0.2167
Gameplay performance	Decision time	−0.20 (0.215)	−0.34 * (0.032)
	FDR q-value	0.2372	0.0759
	Excess touch count	−0.45 ** (0.004)	−0.22 (0.173)
	FDR q-value	**0.0284**	0.1977
	Last touch time	−0.18 (0.267)	−0.25 (0.119)
	FDR q-value	0.2896	0.1422
	Answer rate	0.29 (0.069)	0.34 * (0.032)
	FDR q-value	0.1077	0.0759

Values are presented as Pearson correlation coefficients (r), with nominal (unadjusted) *p*-values in parentheses and corresponding false discovery rate (FDR)-adjusted q-values in separate rows. Bold q-values indicate statistical significance after FDR adjustment ($q < 0.05$). * $p < 0.05$ and ** $p < 0.01$ indicate nominal statistical significance before FDR adjustment.

3.4. Correlations of Learning Difficulty-Related Characteristics with CAT, K-ARS, and Gameplay Performance Metrics

The correlations between CLDQ reading and mathematics scores and CAT, K-ARS, and gameplay performance metrics are presented in Table 4. The CLDQ reading score showed significant negative correlations with the CAT visual selective attention task ($r = -0.32$, $p = 0.044$), flanker task ($r = -0.40$, $p = 0.011$), and sustained attention to response task ($r = -0.45$, $p = 0.004$). The CLDQ mathematics score exhibited significant negative correlations with the flanker task ($r = -0.39$, $p = 0.013$) and sustained attention to response task ($r = -0.36$, $p = 0.023$). By contrast, the auditory selective attention task was not significantly correlated with either the CLDQ reading or mathematics scores.

The K-ARS inattention score was significantly and positively correlated with both the CLDQ reading ($r = 0.43$, $p = 0.006$) and mathematics ($r = 0.35$, $p = 0.027$) scores. By contrast, the K-ARS total score was not significantly correlated with either reading or mathematics scores.

Among the gameplay performance metrics, excess touch count showed significant positive correlations with both the CLDQ reading ($r = 0.46$, $p = 0.003$) and mathematics ($r = 0.43$, $p = 0.006$) scores. Answer rate was negatively correlated with both the CLDQ reading ($r = -0.49$, $p = 0.001$) and mathematics ($r = -0.30$, $p = 0.060$) scores, although only the correlation with the reading score reached statistical significance. Decision time and last touch time were not significantly correlated with either the CLDQ reading or mathematics score. After FDR adjustment, five of the 20 correlations presented in Table 4 remained statistically significant ($q < 0.05$). Specifically, CLDQ reading scores remained significantly associated with sustained attention performance ($r = -0.45$, $q = 0.0284$), K-ARS inattention scores ($r = 0.43$, $q = 0.0349$), excess touch count ($r = 0.46$, $q = 0.0284$), and answer rate ($r = -0.49$, $q = 0.0160$). In addition, excess touch count remained significantly associated with CLDQ mathematics scores ($r = 0.43$, $q = 0.0349$).

The correlations of the BASA:CT reading and mathematics T-scores with CAT, K-ARS, and gameplay performance metrics are shown in Table 5. The BASA:CT reading T-score showed significant positive correlations with the CAT flanker task ($r = 0.38$, $p = 0.016$) and sustained attention to response task ($r = 0.36$, $p = 0.023$). Similarly, the BASA:CT mathematics T-score showed significant positive correlations with the flanker task ($r = 0.42$, $p = 0.007$) and sustained attention to response task ($r = 0.38$, $p = 0.016$). By contrast, visual and auditory selective attention tasks were not significantly correlated with reading or mathematics T-scores.

The K-ARS inattention score was negatively correlated with both the BASA:CT reading ($r = -0.40$, $p = 0.011$) and mathematics ($r = -0.28$, $p = 0.080$) T-scores, although only the correlation with the reading T-score reached statistical significance. The K-ARS total score was not significantly correlated with either the reading or mathematics T-score.

Among the gameplay performance metrics, excess touch count was negatively correlated with both the BASA:CT reading ($r = -0.45$, $p = 0.004$) and mathematics ($r = -0.22$, $p = 0.173$) T-scores, although only the correlation with the reading T-score was statistically significant. Decision time was negatively correlated with both the BASA:CT mathematics ($r = -0.34$, $p = 0.032$) and reading ($r = -0.20$, $p = 0.215$) T-scores, with only the correlation with the mathematics T-score reaching statistical significance. Answer rate showed positive correlations of a similar magnitude with the BASA:CT mathematics ($r = 0.34$, $p = 0.032$) and reading ($r = 0.29$, $p = 0.069$) T-scores, although only the correlation with the mathematics T-score reached statistical significance. Last touch time was not significantly correlated with the reading or mathematics T-scores. After FDR adjustment, two of the 20 associations presented in Table 5 remained statistically significant ($q < 0.05$). Specifically, BASA:CT mathematics T-scores remained positively associated with Flanker task performance

(r = 0.42, q = 0.0373), while BASA:CT reading T-scores remained negatively associated with excess touch count (r = −0.45, q = 0.0284).

3.5. Comparison of Learning Difficulty-Related Characteristics According to Gameplay Performance Level

Comparisons of CLDQ reading and mathematics scores between the relatively lower-performance group ($n = 19$) and the higher-performance group ($n = 21$), classified based on the median composite gameplay difficulty score, are presented in Table 6. The CLDQ reading score was significantly higher in the lower-performance group (3.8 ± 0.60) than in the higher-performance group (3.4 ± 0.54; $p = 0.033$), with a between-group effect size of Cohen's d = 0.70. The CLDQ mathematics score was 3.0 ± 0.52 in the lower-performance group and 2.7 ± 0.48 in the higher-performance group; however, the difference between the groups was not statistically significant ($p = 0.065$, Cohen's d = 0.60).

Table 6. Comparison of CLDQ and BASA:CT scores between the lower- and higher-gameplay performance groups.

Measure	Subscale	Lower-Performance Group ($N = 19$)	Higher-Performance Group ($N = 21$)	*p*-Value	Effect Size
CLDQ	Reading	3.8 ± 0.60	3.4 ± 0.54	0.033 *	0.70
	Mathematics	3.0 ± 0.52	2.7 ± 0.48	0.065	0.60
BASA:CT	Reading T-score	47.1 ± 2.10	49.6 ± 2.48	0.002 *	1.08
	Mathematics T-score	43.3 ± 2.47	45.0 ± 2.20	0.027 *	0.73

Values are presented as mean ± standard deviation (SD). Between-group differences were assessed using independent-samples *t*-tests. Effect sizes are reported as Cohen's d * $p < 0.05$.

The BASA:CT reading T-score was significantly lower in the lower-performance group (47.1 ± 2.10) than in the higher-performance group (49.6 ± 2.48; $p = 0.002$), with a between-group effect size of Cohen's d = 1.08. Similarly, the BASA:CT mathematics T-score was significantly lower in the lower-performance group (43.3 ± 2.47) than in the higher-performance group (45.0 ± 2.20; $p = 0.027$, Cohen's d = 0.73), indicating a moderate between-group difference.

3.6. Association Between the Composite Gameplay Difficulty Score and Learning Difficulty-Related Characteristics

In the multiple linear regression analyses adjusted for age and sex, the composite gameplay difficulty score was significantly positively associated with the CLDQ reading score (B = 0.27, 95% CI = 0.07–0.47; β = 0.29, $p = 0.010$), and the model explained approximately 13% of the variance in reading scores (adjusted $R^2 = 0.13$). The composite gameplay difficulty score was also significantly positively associated with the CLDQ mathematics score (B = 0.20, 95% CI = 0.02–0.38; β = 0.26, $p = 0.032$), with the model explaining approximately 9% of the variance in mathematics scores (adjusted $R^2 = 0.09$).

By contrast, for the BASA:CT, the composite gameplay difficulty score was significantly negatively associated with the reading T-score (B = −1.18, 95% CI = −2.13 to −0.23; β = −0.30, $p = 0.016$), and the model explained approximately 12% of the variance (adjusted $R^2 = 0.12$). The composite gameplay difficulty score was also significantly negatively associated with the mathematics T-score (B = −0.74, 95% CI = −1.43 to −0.05; β = −0.19, $p = 0.036$), with the model explaining approximately 8% of the variance (adjusted $R^2 = 0.08$) (Table 7). Diagnostic assessments of the regression models did not reveal substantial violations of linearity, residual normality, or homoscedasticity. No evidence of problematic multicollinearity or highly influential observations was identified.

Table 7. Associations between composite gameplay difficulty score and CLDQ and BASA:CT scores.

Outcome	B	SE	β	95% CI for B	*p*-Value	Adjusted R^2
CLDQ reading	0.27	0.10	0.29	0.07–0.47	0.010 *	0.13
CLDQ mathematics	0.20	0.09	0.26	0.02–0.38	0.032 *	0.09
BASA:CT reading T-score	−1.18	0.47	−0.30	−2.13–−0.23	0.016 *	0.12
BASA:CT mathematics T-score	−0.74	0.34	−0.19	−1.43–−0.05	0.036 *	0.08

B = unstandardized regression coefficient; SE = standard error; β = standardized regression coefficient; CI = confidence interval. Models were adjusted for age and sex. * $p < 0.05$. The predictor was the composite gameplay difficulty score.

4. Discussion

This study selected C3, C5, C7, and C9 from the 10 digital games completed by participants because they consistently showed relatively poorer performance across decision time, excess touch count, last touch time, and answer rate. We then explored whether the performance characteristics of the selected games were associated with attention and cognitive function, learning difficulty-related characteristics in reading and mathematics, and basic academic performance. In particular, C7 showed the most pronounced pattern of relatively low performance across all four metrics, with the longest decision and last touch times, the highest excess touch count, and the lowest answer rate.

The selected games included tasks, such as moving in a specified sequence under time constraints, discriminating visual shapes, sequentially matching colors and shapes, and remembering the initially presented locations of obstacles to plan a route to a target while avoiding them. Therefore, successful performance on these tasks may require selective and sustained attention as well as visual search, working memory, sequential processing, interference control, and response inhibition. Gameplay performance metrics from the selected games exhibited nominally significant associations with performance on the CAT visual selective attention, flanker, and sustained attention to response tasks. In addition, excess touch count and last touch time demonstrated nominally significant associations with the K-ARS inattention scores. By contrast, no significant associations were observed with the K-ARS total score, indicating that performance on the selected games may be more closely related to specific vulnerabilities in attentional control and executive function relevant to learning than to overall ADHD symptom severity.

Associations were also observed between several gameplay performance metrics, learning difficulties, and basic academic performance. Excess touch count positively correlated with both CLDQ reading and mathematics scores, and the answer rate exhibited a nominally significant negative correlation with the CLDQ reading score. Similar patterns were observed for the BASA:CT, a standardized measure of basic academic performance. A higher excess touch count was associated with a lower reading T-score, and a longer decision time was associated with a lower mathematics T-score. In addition, the answer rate demonstrated a nominally significant positive correlation with the mathematics T-scores. In the analyses adjusted for age and sex, the composite gameplay difficulty score was significantly associated with both CLDQ reading and mathematics scores, with the models explaining 13% and 9% of the variance, respectively. By contrast, the composite gameplay difficulty score was significantly negatively associated with the BASA:CT reading and mathematics T-scores, with the models explaining 12% and 8% of the variance, respectively. Furthermore, the relatively lower-performance group had a significantly higher CLDQ reading score and significantly lower BASA:CT reading and mathematics T-scores than the higher-performance group, demonstrating a consistent direction of differences in group comparisons. These findings suggest that repetitive erroneous responses, inefficient task performance, and lower accuracy during gameplay may reflect both differences in game proficiency and partially attentional and executive-function demands related to learning difficulties and basic

academic performance in reading and mathematics among children with ADHD [37,38]. Notably, 12 of the 28 nominally significant correlations remained statistically significant after FDR adjustment, supporting the observed associations between gameplay performance metrics and selected measures of attention and learning difficulties in children with ADHD. Although the absolute differences in BASA:CT T-scores were modest (2.5 points for reading and 1.7 points for mathematics), the observed patterns suggest that gameplay performance metrics may provide complementary information about academic functioning.

Complementary patterns were observed between the parent-reported CLDQ and the BASA:CT. Greater gameplay difficulty was associated with higher learning-difficulty scores on the CLDQ and with lower academic performance T-scores on the BASA:CT. However, owing to the exploratory nature of the present analyses and the limited sample size, these findings should not be interpreted as evidence of the diagnostic validity of gameplay performance metrics.

Unlike previous game-based studies, which focused on improvements in symptoms or academic achievement following intervention, the significance of this study lies in its focus on the potential utility of behavioral data automatically accumulated during gameplay as indicators of learning-difficulty risk. Gameplay logs enable the quantitative assessment of response time, excess responses, and accuracy within a relatively short period without relying solely on subjective observations by examiners [39]. Therefore, if these findings are replicated in independent samples, gameplay performance metrics may serve as supplementary screening indicators to help prioritize children with ADHD, who require comprehensive standardized assessments of reading and mathematics in school or clinical settings. However, considering the limited explanatory power of this study, gameplay scores alone should not be used to determine the presence of a learning difficulty. Instead, they may be more appropriately used as supplementary referral-support indicators in conjunction with parent and teacher reports, clinical interviews, and standardized academic achievement assessments [40].

This study has several limitations. First, the sample size was relatively small, comprising 40 participants from a single center, and no external validation was performed. Second, the selection of C3, C5, C7, and C9 and the subsequent analyses were conducted using the same sample, which may have resulted in overfitting to sample-specific characteristics. Third, although the Benjamini–Hochberg FDR procedure was applied to account for multiple comparisons, the exploratory nature of the correlation analyses, the limited sample size, and the possibility of residual false-positive findings should be considered when interpreting the results. Fourth, the lower- and higher-performance groups were defined using the sample median rather than an established clinical cutoff value. Although this approach facilitated exploratory comparisons of relative gameplay performance, dichotomizing a continuous score may result in information loss and sample-dependent group classification. The composite gameplay difficulty score was also examined as a continuous variable in regression analyses, providing complementary evidence beyond the median-based group comparisons. Nevertheless, the median-based grouping should not be interpreted as a clinically validated classification of learning difficulties. Fifth, the CLDQ is a parent-reported measure of learning difficulties, and the Korean translation used in this study has not undergone a separate psychometric validation. In addition, the participants did not undergo formal clinical diagnosis of learning difficulties. Given the exploratory nature of the study and the relatively small sample size, the regression analyses were adjusted for age and sex to limit model complexity. However, potential residual confounding related to medication status, IQ, and ADHD symptom severity cannot be excluded and warrants further investigation in larger studies. Furthermore, both gameplay data and the CLDQ and BASA assessments were obtained at the end of the four-week intervention. Although all participants initially completed the same set of games, subsequent individualized adjustments in game combinations, difficulty

levels, and feedback may have resulted in differences in cumulative training exposure and task demands across participants. Consequently, Week 4 gameplay performance may reflect not only individual differences in attention and learning-related functioning but also variations in prior training experience and individualized task conditions. Because these factors were not separately controlled for in the present correlation analyses, their potential confounding effects cannot be excluded. Therefore, the observed associations should be interpreted as exploratory relationships within the adaptive training environment rather than as evidence that gameplay metrics independently reflect underlying cognitive abilities or pre-intervention learning characteristics. Future studies should prespecify the selected games and performance metrics, incorporate standardized task conditions and detailed training exposure measures, and evaluate them in independent clinical and school-based cohorts using standardized assessments of reading, writing, and mathematics, together with expert determination of learning difficulties. Such studies should also assess sensitivity, specificity, area under the receiver operating characteristic curve, test–retest reliability, and clinically meaningful cutoff values.

5. Conclusions

In the present study, C3, C5, C7, and C9 among the game-based digital content administered to children with ADHD consistently demonstrated relatively poorer performance in terms of decision time, excess touch count, last touch time, and answer rate. The performance metrics derived from these games were associated with visual selective attention, interference control, sustained attention, and inattentive symptoms. Furthermore, the composite gameplay difficulty score remained significantly associated with both parent-reported reading and mathematics learning difficulties and standardized basic academic performance in reading and mathematics after adjusting for age and sex.

These findings suggest that the selected games may have potential as digital behavioral tasks that reflect both attentional and executive-function demands as well as learning-difficulty-related characteristics in children with ADHD. Accordingly, rather than serving as a standalone diagnostic test for learning difficulties, this game-based tool may have the potential to be developed as a learning difficulty risk-screening and referral-support tool to help identify children with ADHD who may require comprehensive academic assessment and to facilitate referral for standardized testing and individualized educational support. However, to establish this approach as a clinical screening tool, large independent samples and multicenter external validation using formal learning difficulty diagnoses and standardized academic assessments as reference standards are required.

Author Contributions: Conceptualization, S.-C.K.; formal analysis, S.-C.K.; writing—original draft preparation, S.-C.K.; writing—review and editing, S.-C.K.; data curation, S.-C.K.; software, S.-C.K.; project administration, S.-C.K.; investigation, N.-Y.K.; resources, N.-Y.K.; methodology, N.-Y.K.; validation, N.-Y.K.; supervision, N.-Y.K.; data curation, S.-Y.L.; formal analysis, S.-Y.L.; investigation, S.-Y.L.; methodology, S.-Y.L.; resources, S.-Y.L.; visualization, S.-Y.L.; validation, S.-Y.L.; writing—review and editing, H.-S.L.; data curation, H.-S.L.; investigation, H.-S.L.; resources, H.-S.L.; methodology, H.-S.L.; formal analysis. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was prepared as part of the project "2024 AI-Based Brain Disease Digital Medical Device Demonstration Initiative," which was selected by the National IT Industry Promotion Agency (NIPA) under the Ministry of Science and ICT. It was conducted under the auspices of Daegu City, with support from the Daegu Digital Innovation Promotion Agency (DIP) as the lead organization. This research was supported by the Advanced Network for Cultivating Human Resources and Opportunities in Regions (ANCHOR) program through the Daegu ANCHOR Center, funded by the Ministry of Education (MOE) and the Daegu Metropolitan City, Republic of Korea (2026-ANCHOR-03-002).

Institutional Review Board Statement: Approval was obtained from the Institutional Bioethics Review Committee before data collection and evaluation (Keimyung University Dongsan Hospital; Institutional Review Board no. 2025-03-002-003, approval on 26 May 2025).

Informed Consent Statement: Participants and their guardians provided written informed consent to participate in the study.

Data Availability Statement: The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADHD	attention-deficit/hyperactivity disorder
CAT	Comprehensive Attention Test
K-ARS	Korean ADHD Rating Scale
CLDQ	Colorado Learning Difficulties Questionnaire
BASA:CT	Basic Academic Skills Assessment: Comprehensive Test (BASA:CT)

References

1. Shen, F.; Zhou, H. Advances in the etiology and neuroimaging of children with attention deficit hyperactivity disorder. *Front. Pediatr.* **2024**, *12*, 1400468. [CrossRef] [PubMed]
2. Kofler, M.J.; Soto, E.F.; Singh, L.J.; Harmon, S.L.; Jaisle, E.M.; Smith, J.N.; Feeney, K.E.; Musser, E.D. Executive function deficits in attention-deficit/hyperactivity disorder and autism spectrum disorder. *Nat. Rev. Psychol.* **2024**, *3*, 701–719. [CrossRef] [PubMed]
3. Tam, L.Y.C.; Taechameekietichai, Y.; Allen, J.L. Individual child factors affecting the diagnosis of attention deficit hyperactivity disorder (ADHD) in children and adolescents: A systematic review. *Eur. Child Adolesc. Psych.* **2025**, *34*, 1469–1496. [CrossRef] [PubMed]
4. Viesel-Nordmeyer, N.; Reuber, J.; Kuhn, J.T.; Moll, K.; Holling, H.; Dobel, C. Cognitive profiles of children with isolated and comorbid learning difficulties in reading and math: A meta-analysis. *Educ. Psychol. Rev.* **2023**, *35*, 34. [CrossRef]
5. Sadozai, A.K.; Sun, C.; Demetriou, E.A.; Lampit, A.; Munro, M.; Perry, N.; Boulton, K.A.; Guastella, A.J. Executive function in children with neurodevelopmental conditions: A systematic review and meta-analysis. *Nat. Hum. Behav.* **2024**, *8*, 2357–2366. [CrossRef] [PubMed]
6. Trane, F.E.; Willcutt, E.G. Attention-deficit/hyperactivity disorder and academic functioning: Reading, writing, and math abilities in a community sample of youth with and without ADHD. *Res. Child Adolesc. Psychopathol.* **2023**, *51*, 583–596. [CrossRef] [PubMed]
7. Tsantilas, D.; Ilie, A.; Waldon, J.; McGonnell, M.; Corkum, P. The contributions of cognitive abilities to the relationship between ADHD symptoms and academic achievement. *Brain Sci.* **2022**, *12*, 1075. [CrossRef] [PubMed]
8. Horowitz-Kraus, T.; Cirino, P.T.; Cutting, L.E.; Hughes-Berheim, S.S.; Wilkey, E.D.; Barnes, M.A.; Rosch, K.S.; Church, J.A. The interplay between executive function and learning disabilities: Developmental cognitive neuroscience perspectives on reading, math, and ADHD. *Dev. Cogn. Neurosci.* **2026**, *79*, 101722. [CrossRef] [PubMed]
9. Cibrian, F.L.; Monteiro, E.M.; Lakes, K.D. Digital assessments for children and adolescents with ADHD: A scoping review. *Front. Digit. Health* **2024**, *6*, 1440701. [CrossRef] [PubMed]
10. Fletcher, J.M.; Miciak, J. Assessment of specific learning disabilities and intellectual disabilities. *Assessment* **2024**, *31*, 53–74. [CrossRef] [PubMed]
11. Kim, J.S.; Jeong, Y.J.; Kim, S.J.; Jun, S.J.; Park, J.Y.; Hoe, H.S.; Song, J.H. A novel approach using serious game data to predict the WISC-V processing speed index in children with attention-deficit/hyperactivity disorder: Machine learning study. *JMIR Ser. Games* **2025**, *13*, e73408. [CrossRef] [PubMed]
12. Stamatis, C.A.; Heusser, A.C.; Simon, T.J.; Ala'ilima, T.; Kollins, S.H. Real-time cognitive performance metrics derived from a digital therapeutic for inattention predict ADHD-related clinical outcomes: Replication across three independent trials of AKL-T01. *Transl. Psych.* **2024**, *14*, 328. [CrossRef] [PubMed]
13. Lin, J.; Chang, W.R. Effectiveness of serious games as digital therapeutics for enhancing the abilities of children with attention-deficit/hyperactivity disorder (ADHD): Systematic literature review. *JMIR Ser. Games* **2025**, *13*, e60937. [CrossRef] [PubMed]

14. Marinelli, C.V.; Nardacchione, G.; Trotta, E.; Di Fuccio, R.; Palladino, P.; Traetta, L.; Limone, P. The effectiveness of serious games for enhancing literacy skills in children with learning disabilities or difficulties: A systematic review. *Appl. Sci.* **2023**, *13*, 4512. [CrossRef]
15. Farias, A.C.; Cordeiro, M.L.; Felden, E.P. Attention–memory training yields behavioral and academic improvements in children. *Neuropsychiatr. Dis. Treat.* **2017**, *13*, 1761–1769. [CrossRef] [PubMed]
16. García-Redondo, P.; García, T.; Areces, D.; Núñez, J.C.; Rodríguez, C. Serious games and their effect improving attention in students with learning disabilities. *Int. J. Environ. Res. Public Health* **2019**, *16*, 2480. [CrossRef]
17. Yildirim, O.; Surer, E. Developing adaptive serious games for children with specific learning difficulties: A two-phase usability and technology acceptance study. *JMIR Ser. Games* **2021**, *9*, e25997. [CrossRef] [PubMed]
18. Paglialunga, A.; Melogno, S. The effectiveness of artificial intelligence-based interventions for students with learning disabilities: A systematic review. *Brain Sci.* **2025**, *15*, 806. [CrossRef] [PubMed]
19. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; American Psychiatric Association: Arlington, VA, USA, 2013. [CrossRef]
20. Kwak, K.; Jang, S. *Korean Wechsler Intelligence Scale for Children–Fifth Edition (K-WISC-V)*; Inpsyt: Seoul, Republic of Korea, 2019.
21. Willcutt, E.G.; Boada, R.; Riddle, M.W.; Chhabildas, N.; DeFries, J.C.; Pennington, B.F. Colorado Learning Difficulties Questionnaire: Validation of a parent-report screening measure. *Psychol. Assess.* **2011**, *23*, 778. [CrossRef] [PubMed]
22. Faul, F.; Erdfelder, E.; Buchner, A.; Lang, A.G. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* **2009**, *41*, 1149–1160. [CrossRef] [PubMed]
23. Kim, S.C.; Lee, S.Y.; Lee, S.W. Eye-tracking-based digital therapeutic for ADHD in children: Development and preliminary evaluation of an adaptive learning guide. *Healthcare* **2026**, *14*, 1486. [CrossRef] [PubMed]
24. Kim, S.C.; Kong, N.Y.; Lee, S.Y.; Lee, S.W. Data-driven mechanistic analysis of digital therapeutic–assisted training and evaluation of personalized protocol effects in children with attention deficit/hyperactivity disorder. *Front. Psych.* **2026**, *17*, 1794425. [CrossRef] [PubMed]
25. Yoo, H.K.; Lee, J.; Kang, S.H.; Park, E.H.; Jung, J.; Kim, B.N.; Jung-Woo, S.; Park, T.W.; Kim, B.; Young-Sik, L. Standardization of the comprehensive attention test for the Korean children and adolescents. *J. Korean Acad. Child Adolesc. Psych.* **2009**, *20*, 68–75.
26. Luo, J.; Li, F.; Wu, Y.; Liu, X.; Zheng, Q.; Qi, Y.; Huang, H.; Xu, G.; Liu, Z.; He, F.; et al. A mobile device-based game prototype for ADHD: Development and preliminary feasibility testing. *Transl. Psychiatry* **2024**, *14*, 251. [CrossRef] [PubMed]
27. Park, J.I.; Shim, S.H.; Lee, M.; Jung, Y.E.; Park, T.W.; Park, S.H.; Im, Y.J.; Yang, J.C.; Chung, Y.C.; Chung, S.K. The validities and efficiencies of Korean ADHD rating scale and Korean child behavior checklist for screening children with ADHD in the community. *Psych. Investig.* **2014**, *11*, 258. [CrossRef] [PubMed]
28. Kim, K.M.; Ha, M.; Lim, M.H.; Kwon, H.J.; Yoo, S.J.; Kim, E.; Paik, K.C. The symptom trajectory of attention-deficit hyperactivity disorder in Korean school-age children. *Psych. Investig.* **2018**, *15*, 470–475. [CrossRef] [PubMed]
29. Fefe, E.P.; Tigani, X.; Michou, M.; Efthymiou, V.; Christou, A.I.; Darviri, C.; Charalampopoulou, M.; Apostolakis, E.; Papadodima, S.; Kanaka-Gantenbein, C.; et al. Validation of the Greek version of the Colorado learning difficulties questionnaire. *Children* **2025**, *12*, 1623. [CrossRef] [PubMed]
30. Patrick, K.E.; McCurdy, M.D.; Chute, D.L.; Mahone, E.M.; Zabel, T.A.; Jacobson, L.A. Clinical utility of the Colorado learning difficulties questionnaire. *Pediatrics* **2013**, *132*, e1257ve1264. [CrossRef] [PubMed]
31. Pritchard, A.E.; Koriakin, T.; Carey, L.; Bellows, A.; Jacobson, L.; Mahone, E.M. Academic testing accommodations for ADHD: Do they help? *Learn. Disabil.* **2016**, *21*, 67–83. [CrossRef] [PubMed]
32. Kim, D.; Baek, H.; Sohn, E.; Lee, J. An Integrated Study on the Effectiveness and Pedagogical Processes of One-on-One Intervention for Slow Learners through BASA: CT and Keyword Network Analysis. *Korean J. Elem. Educ.* **2026**, *37*, 489–517. [CrossRef]
33. Andrade, C. Z scores, standard scores, and composite test scores explained. *Indian J. Psychol. Med.* **2021**, *43*, 555–557. [CrossRef] [PubMed]
34. Kim, S.C. The effectiveness of story-and quiz-based games in digital interventions for ADHD: A comparative approach. *Appl. Sci.* **2025**, *15*, 4334. [CrossRef]
35. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Front. Psychol.* **2013**, *4*, 863. [CrossRef] [PubMed]
36. Schober, P.; Vetter, T.R. Linear regression in medical research. *Anesth. Analg.* **2020**, *132*, 108. [CrossRef] [PubMed]
37. Gioia, A.R.; Miciak, J.; Peng, P.; Gallagher, M.W.; Williams, M.W.; Farrell, A.; Salentine, C.; Boada, C.; Dragoi, I.; Harmouch, S.; et al. Behavioral attention and academic achievement: A comprehensive meta-analysis. *Child Neuropsychol.* **2026**, *32*, 943–981. [CrossRef] [PubMed]
38. Cortés Pascual, A.; Moyano Muñoz, N.; Quílez Robres, A. The relationship between executive functions and academic performance in primary education: Review and meta-analysis. *Front. Psychol.* **2019**, *10*, 1582. [CrossRef] [PubMed]

39. Panagiotidis-Arrizabalaga, A.; Lavalle, A.; Teruel, M.A.; Sanchis, J.; Ruiz-Robledillo, N.; Costa-López, B. Serious game–based attention assessment for attention-deficit/hyperactivity disorder in children and adolescents: Preliminary cross-sectional convergent validation study. *JMIR Ser. Games*. **2026**, *14*, e80441. [CrossRef] [PubMed]
40. National Joint Committee on Learning Disabilities. Comprehensive assessment and evaluation of students with learning disabilities: A paper prepared by the National Joint Committee on Learning Disabilities. *Learn. Disabil. Q.* **2011**, *34*, 3–16. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.