

Article

Not peer-reviewed version

Game-based Assessment of Students' Digital Literacy Using Evidence-Centered Game Design

[Jiayuan Li](#) , Jie Bai , [Sha Zhu](#) ^{*} , [Harrison Hao Yang](#) ^{*}

Posted Date: 4 December 2023

doi: 10.20944/preprints202312.0193.v1

Keywords: digital literacy; digital game-based assessment; ECGD; AHP; assessment model



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Game-Based Assessment of Students' Digital Literacy Using Evidence-Centered Game Design

Jiayuan Li ¹, Jie Bai ¹, Sha Zhu ^{1,*} and Harrison Hao Yang ^{2,*}

¹ Central China Normal University, 430079 Wuhan, China; 971477999@qq.com (J.L.); 961186618@qq.com (J.B.)

² State University of New York at Oswego, 13126 Oswego, USA

* Correspondence: dr.zhusha@qq.com (S.Z.); harrison.yang@oswego.edu (H.H.Y.)

Abstract: This study measured secondary students' digital literacy using a digital game-based assessment system that was designed and developed based on the Evidence-Centered Game Design (ECGD) approach. A total of 188 secondary students constituted the valid samples in this study. Fine-grained behavioral data generated from students' gameplay processes were collected and recorded with the assessment system. The Delphi method was used to extract feature variables related to digital literacy from the process data, and the Analytic Hierarchy Process (AHP) method was used to construct the measurement model. The assessment results of the ECGD-based assessment had a high correlation with standardized test scores, which have been shown to be reliable and valid in prior large-scale assessment studies.

Keywords: digital literacy; digital game-based assessment; ECGD; AHP; assessment model

1. Introduction

The rapid development of new generations of information technology, such as artificial intelligence and big data, has had a profound impact on education. As a result, education is currently undergoing a crucial period of digital transformation [1]. A key strategy for bridging the digital divide and advancing the digital transformation of education is the enhancement of students' digital literacy [2]. Digital literacy comprises the comprehensive ability to effectively use information technology to access and exchange information, create digital content, and adhere to ethical norms in a digital society [3]. It is particularly crucial to improve the digital skills and critical thinking of the digital natives born in the 21st century, who are inherently familiar with digital media and resources. The conversion of the digital divide into digital opportunities is of great significance in promoting the advancement of digital education.

Digital literacy assessment is crucial for accurately gauging students' level of digital literacy and is thus a fundamental prerequisite for enhancing their digital literacy skills [4]. However, the current methods used to assess students' digital literacy have several limitations in terms of evaluation content and tools, making it a challenge to comprehensively, accurately, and objectively evaluate students' proficiency. First, evaluation content to date has primarily focused on traditional "test question-answer" paradigms, emphasizing students' basic cognitive abilities such as digital knowledge and application skills. Consequently, such an approach fails to evaluate higher-order thinking skills, such as the ability to use digital technology to analyze and solve problems, engage in innovation and creativity, and address ethical and moral issues in digital society. Second, the evaluation tools predominantly rely on standardized tests and self-reported scales [5]. While these methods provide a summarized evaluation of students' digital literacy, they are not suited to measuring implicit cognitive abilities and thinking processes [6]. Although several studies have introduced situational task assessment and portfolio assessment as alternative tools for process evaluation in recent years, the data generated from these assessments still fall short of providing comprehensive evidence for students' digital literacy performance [7]. Given the significance of

digital literacy for fostering innovative talent and preparing students to tackle future opportunities and challenges, it is imperative to address the pressing issue of how to accurately evaluate students' digital literacy.

The application of the ECGD theory in assessment design has the potential to overcome the limitations of the current methods for the evaluation of students' digital literacy. In the era of artificial intelligence, it is possible to comprehensively and non-invasively record students' online activities. Researchers have thus begun to explore assessments based on process data, guided by the ECGD theory [8]. The ECGD concept emphasizes the use of complex tasks to elicit students' ability and performance. During student evaluation, process data are collected through gamified tasks that are engaging and interactive. This data extraction process captures evidence that reflects students' ability and performance. Building upon this foundation, this study proposes a new method for evaluating students' digital literacy based on the ECGD paradigm. The research team has developed a game-based assessment tool to measure students' digital literacy, which collects fine-grained procedural data generated by students during the task completion process. The Delphi method is employed to determine the evidence that reflects primary and secondary school students' digital literacy, and the analytic hierarchy process is used to assign weights to the extracted characteristic variables. These weights serve as the scoring criteria for measuring students' digital literacy. Furthermore, the research team has conducted a practical analysis to validate the proposed method [9]. The results of this study will contribute to the iterative optimization of each aspect of the evaluation method, ultimately providing a more reliable and effective approach to accurately assess students' digital literacy levels.

2. Literature Review

2.1. Digital Literacy

The ideological source of digital literacy can be traced back to the 1995 book *Being Digital*, by American scholar Nicholas Negroponte. He pointed out that people should become masters of digital technology, able to use digital technology to adapt, participate in, and create digital content for learning, work, and life [10]. With the continuous development of society, the concept of digital literacy has been continuously extended and expanded, in an evolutionary process that can be roughly divided into three stages:

2.1.1. Stage 1: The Period Spanning from the 1990s to the Early 21st Century

The concept of digital literacy was first proposed by Paul Gilster, an American scholar of space science and technology, in his 1997 monograph, "Digital Literacy". He defined digital literacy as "the ability to understand information, and more importantly, the ability to evaluate and integrate various formats of information that computers can provide" [11]. In 2004, the Israeli scholar Yoram Eshet-Alkalai pointed out that digital literacy is regarded as a necessary survival skill in the digital era. It includes the ability to use software and operate digital devices; various complex cognitive, motor, sociological, and emotional skills used in the digital environment; and the ability to perform tasks and solve complex problems in the digital environment [12]. Thus, during Stage 1, people focused on the digitalization of the living environment and the basic skills required to meet the challenges of digitalization. The concept of digital literacy in this period mainly emphasized the ability to understand, use, evaluate, and integrate the digital resources and information provided by computers.

2.1.2. Stage 2: The Period Spanning from the Beginning of the 21st Century to the 2010s

With the real arrival of the digital age, information technology as represented by the Internet accelerated the development of the digital society, and countries around the world began to pay attention to the core ability of digital literacy. In 2003, the digital horizon report issued by the New Zealand Ministry of Education pointed out that digital literacy is a kind of "life skill" that supports the innovative development of ICT in industrial, commercial, and creative processes. Learners need

to acquire confidence, skills, and discrimination in order to use information technology in an appropriate way [13]. In 2009, Calvani proposed that digital literacy consists of the ability to flexibly explore and respond to new technological situations; the ability to analyze, select, and critically evaluate data and information; the ability to explore technological potential; and the ability to effectively clarify and solve problems [14]. In Stage 2, the concept of digital literacy emphasized the correct and appropriate use of digital tools, not limited to computers, and attached importance to innovation and creation, emphasizing that users are not only the users of digital technology, but also the producers of digital content.

2.1.3. Stage 3: The Period Spanning from the 2010s to the Present

With the rapid development of artificial intelligence, big data, the Internet of things, and other information technologies, digital literacy has drawn increasing attention at the national level. At the same time, a range of international organizations have also begun to develop concepts and frameworks of digital literacy, such that its conceptual development can be said to have entered a period of “contention of a hundred schools of thought”. For example, in 2018, UNESCO defined digital literacy as “the ability to access, manage, understand, integrate, communicate, evaluate and create information safely and appropriately through digital technologies for employment, decent jobs and entrepreneurship” [15]. In 2021, China proposed that “digital literacy and skills are the collection of a series of qualities and abilities such as digital acquisition, production, use, evaluation, interaction, sharing, innovation, security and ethics that citizens in a digital society should have in their study, work and life” [16]. During this period, the concept of digital literacy focused more on the unity of humans’ own development and social advancement, the humanistic attributes of digital literacy, and ethics, morality, laws, and regulations.

Based on the above points of view, our research team posits that digital literacy is a multifaceted construct comprising an individual’s awareness, ability, thinking and cultivation to properly use information technology to acquire, integrate, manage, and evaluate information; understand, construct, and create new knowledge; and to discover, analyze, and solve problems [17]. Digital literacy comprises four dimensions: information awareness and attitude (IAA), information knowledge and skills (IKS), information thinking and behavior (ITB), and information social responsibility (ISR) [18].

2.2. *The Digital Literacy Assessment*

Initially, the evaluation of students’ digital literacy was mainly based on classical test theory (CTT) [19,20]. Researchers compiled or adapted digital literacy evaluation tools (mainly self-reported scales) through literature review and expert consultation, and used them to measure students’ digital literacy level. This method mainly determines students’ digital literacy level based on their reports. Thus, the results are greatly affected by subjective factors in the participants, and thus there is a need for improvement in the evaluation accuracy of this method [21–24]. In view of the limitations of CCT and the strong subjectivity of the evaluation results of self-reported scales, the academic community began to explore the use of item response theory (IRT) in the evaluation of students’ digital literacy [25]. At this stage, evaluation tools were mainly based on standardized test questions. The evaluation of students’ digital literacy based on item response theory provides a unified standard for measuring both the subjects’ digital literacy level and the statistical parameters of the items. This approach successfully addresses issues such as the difficulty of the estimation of various parameters depending too much on the sample, and the subjectivity of the evaluation results, thus effectively enhancing the accuracy of the evaluation results. For example, Zhu developed a standardized test measuring students’ digital literacy, comprising 37 multiple-choice questions using a Rasch model based on IRT, yielding a more accurate and objective tool for the evaluation of students’ digital literacy [7]. Later, Nguyen and Habók also adopted a similar approach and compiled a digital literacy test comprising self-reported scales and standardized test questions, in which multiple-choice questions were used to measure the students’ digital knowledge [6]. However, relying solely on such self-reported scales and standardized tests is still problematic for meeting the need to evaluate high-order thinking

abilities such as computational thinking, digital learning, and innovation in digital literacy [26,27]. Moreover, such tools yield summative evaluations, and the evaluation results cannot well reflect the actual level of students' digital literacy [28,29].

2.3. Game-Based Assessment of Digital Literacy Based on ECGD

Mislevy proposed the ECGD approach [30]. Based on ECD, this approach incorporates game-based tasks into assessment design. ECD is a systematic method that guides the design of educational evaluations [31]. It emphasizes the construction of complex task situations, obtaining multiple types of procedural data, and achieving evidence-based reasoning [32]. At present, ECD has been widely used in international large-scale assessment programs, such as PISA, NAEP, and ATC21S, which were designed and developed based on the ECD evaluation framework. In addition, ECD is also widely used in the evaluation of core literacy, 21st-century skills, computational thinking, data literacy, logical reasoning ability, scientific literacy, and other forms of high-level thinking. Some researchers have also used ECD to assess students' digital literacy. For example, Avdeeva evaluated the digital skills dimension of students' digital literacy using a method based on ECD [33]. They verified the advantages of ECD theory in terms of accuracy and other aspects, as compared to project response theory. Zhu conducted an in-depth analysis of ECD theory, were the first to put forward the idea of students' digital literacy evaluation based on ECD, and initially constructed an evaluation method for students' digital literacy driven by theory and data [18].

The advantage of game-based assessment tasks is that one can create an environment of richness, playability, and simulation. Doing so can effectively reduce anxiety in the assessment process and improve participation, and is suitable for assessing students' higher-order thinking abilities [34]. In addition, a game-based assessment task can also establish a task set with multiple difficulty levels to assess students' performance in different situations, so as to effectively distinguish the assessment results [35]. A game-based assessment task can be a continuous or cyclic process, which can obtain students' procedural behavior data [36]. The ECGD conceptual framework is the basis of evaluation design, including three main models: the student model, task model, and evidence model. The student model defines the knowledge, skills, and abilities (KSAs) to be measured. The evidence model describes how to update the information for student variables in the task model based on the test-takers' performance in the task. The task model describes how to structure different kinds of situations to evoke student performance to obtain data. Evaluation methods based on ECGD emphasize the creation of complex and realistic tasks to arouse students' performance on KSA, which helps to reflect students' digital literacy in real scenes. Students' game processes can also generate rich and complex process data, which can provide rich evidence to reflect students' KSA.

To date, many researchers have used ECGD to evaluate higher-order thinking abilities. For example, Chu and Chiang developed a game-based assessment tool based on ECGD to measure scientific knowledge and skills [37]. The results show that task-related behavior characteristics can well predict students' mastery of overall skills. Bley developed a game-based assessment system that uses the ECGD method to measure the internal capabilities of enterprises [38]. The results showed that learners' entrepreneurial ability and cognitive performance in tasks can be accurately measured in this way. However, there have been few empirical studies on digital literacy assessment using ECGD. Although our research team has previously developed a game-based assessment of students' digital literacy based on ECGD [39], the research only proposed a conceptual framework and designed a game-based assessment tool for students' digital literacy, but did not verify its effectiveness using empirical measurement data. Therefore, this study aimed to evaluate secondary school students' digital literacy based on the conceptual framework of ECGD, so as to verify the effectiveness of our ECGD-based assessment approach.

3. Methods

3.1. Participants

This study was conducted with five classes of a middle school located in Wuhan Economic Development Zone, China. The selected school and classes were chosen for the following reasons: 1) the school highly prioritizes the cultivation of students’ digital literacy, and has actively participated in prior digital literacy assessments for students; 2) the students were familiar with the regulations and processes of computer-based assessment and showed enthusiasm for participating in the game-based assessment. A total of 210 seventh-grade students, comprising 114 boys and 96 girls, participated in this study. The students’ ages ranged from 12 to 15 years old. Prior to participation, all participants were provided with information regarding the study’s purpose and were required to sign a formal consent form in order to participate.

3.2. Instruments

3.2.1. A Digital Game-Based System for Assessing Students’ Digital Literacy

In order to induce performance relative to the students’ digital literacy ability model, this study uses the narrative game “Guogan’s Journey to A Mysterious Planet”, developed by our research team to evaluate students’ digital literacy [39]. The game comprises 13 tasks involving the four dimensions of digital literacy. Students are given two chances to complete each task. If a student finishes a task incorrectly twice, the system presents a “Pass card” and begins the next task. During gameplay, students different gain gold coins according to their task performance. Students can choose whether to click a “Help” button during the gameplay process; doing so costs them coins. They can also click a “Return” button to return to the previous page to confirm information.

Table 1 shows the details of the game-based assessment tasks, including the task type, the corresponding dimension of digital literacy, and the observed variables. The observed variables include: completion time (the time period elapsed from the player starting the task to completing it), thinking time (the total time the mouse cursor stayed in different areas of the interface during the students’ answering process), correctness (whether the task was completed correctly or not), answer times (whether students completed the task successfully on the first try), help times (number of times the “Help” button was clicked), return times (number of times the “Return” button was clicked), similarity (the similarity of the action sequence with the reference sequence), and efficiency (the efficiency of the action sequence).

Table 1. Description of the game-based assessment tasks.

Task	Task Type	Dimension	Observed Variables
Tasks 1,3,6,10,11	Multiple-choice question	IAA	completion time, thinking time, correctness, answer times, help times, return times
Tasks 2,13		ISR	
Task 7		IKS	
Task 8		ITB	
Task 4	Maze	ITB	completion time, thinking time, correctness, answer times, help times, return times, similarity, efficiency
Task 5	Dragging question	ITB	
Task 9	Matching question	ITB	
Task 12	Sorting question	IKS	

3.2.2. Standardized Test

To verify the results of the game-based assessment test, this study used the digital literacy standardized test designed in our previous study, comprising 24 multiple-choice questions and multiple-answer questions (i.e., multiple choice questions with more than one correct answer). This test has been validated many times in China’s large-scale student digital literacy assessment project,

and has been demonstrated to have good reliability and validity, difficulty, discrimination, and other indicators, with high standard values [40].

3.3. Data Collection

With the help of the research team and school administrators, the students from the five classes participated in the digital literacy assessment in the designated computer laboratory. The students were required to complete the game assessment and standardized test within 40 minutes. Specifically, the assessment procedure comprised three steps. First, before the assessment, the information technology teacher informed the students of the purpose of the evaluation and emphasized the operating rules, browser settings, and other precautions, and distributed the assessment hyperlink to the students through the teacher's computer. Second, upon accessing the assessment hyperlink, students were required to fill in their personal information and complete the digital literacy standardized test. Finally, upon submitting their responses to the standardized test, students were automatically redirected to the digital literacy game-based assessment system, where they completed the game-based assessment tasks according to the situational sequence.

3.4. Data Storage

The game-based assessment system uses xAPI to record the process data generated during the assessment. xAPI is a standard for describing and exchanging learning experiences. It records the behavior (verb), the object of the behavioral operation (object), the tool used (tool), and the timestamp of the occurrence of the student's (actor) behavior within a context (context), with the task serving as the core. The xAPI data collection framework is employed to characterize the click behavior of students when they complete tasks using specific format statements. These statements are subsequently placed in the learning recording system (LRS) to facilitate the real-time tracking, collection, and storage of students' click data. To generate statements, xAPI specifies the format object representation (JSON) of JavaScript. Figure 1 shows an example of xAPI-based data stored in JSON format.

```

{ "actor_id":1652264602039,
  "context":
    [ {"question_id":"6",
      "task":
        [ {"verb":"enter",
          "object":"",
          "tool":"",
          "timestamp":"1652844561574"},
          {"verb":"start_time",
            "object":"3075",
            "tool":"",
            "timestamp":"1652844556485"},
            {"verb":"input",
              "object":"在网络发言时要遵守网络法律",
              "tool":"",
              "timestamp":"16528449625489"},
              {"verb":"end_time",
                "object":"1",
                "tool":"",
                "timestamp":"1652846512555"},
                {"verb":"current_coins_record",
                  "object":"80",
                  "tool":"",
                  "timestamp":"1652847612485"}],
        },
      ],
    "actor_data":
      { "qid":232,
        "school":"",
        "grade":"6",
        "class":"5",
        "name":"盖 XX",
        "gender":"1"}
    }

```

Figure 1. An example of xAPI-based data stored in JSON format.

3.5. Data Pre-Processing and Analysis

The game-based assessment data of 210 students were collected and matched with their standardized test scores based on the xAPI-based process data. Because the data collected in this study were redundant, inconsistent, and even noisy, it was difficult to meet the experimental requirements. Therefore, the data were cleaned and preprocessed as follows. 1) Missing data processing. When the proportion of missing student data samples was relatively high, it was necessary for the data to be eliminated. In this study, the samples with a high rate of missing fields were eliminated, and the samples with low missing rates were filled by the mean method. A total of 10 samples were eliminated for this reason. 2) Abnormal data processing. Abnormal data were those generated by students who failed to adhere to the assessment rules for their operations, as this has an impact on the accuracy of the model prediction. A total of 12 samples were eliminated due to human errors, repeated answers by the same student, or too short of a response time. Ultimately, a total of 22 samples were eliminated, and the remaining 188 samples were used as experimental data.

In this study, the Delphi method was used to determine the characteristic variables and construct the evaluation model of students' digital literacy. An "expert advisory group" was established, composed of 14 scholars, teachers, and researchers in the field of digital literacy assessment. They were invited to complete a survey, determine the feature variables of each task in relation to students' digital literacy, and rank these feature variables according to their importance. Specifically, this study conducted three rounds of expert consultations. In each round, a questionnaire was distributed to the experts and their feedback was collected, analyzed, and summarized. The feature variables were revised and iterated based on the experts' opinions. After the third round of consultation, the experts' opinions were generally consistent and met the expected requirements. Then, the weights of the characteristic variables were determined using the analytic hierarchy process. The analysis steps were as follows: 1) a judgment matrix was built by comparing the importance of the characteristic variables; 2) the weight of the characteristic variable of each task was calculated; and 3) the consistency of the judgment matrix was verified [41].

In order to explore the relationship between item difficulty and students' ability, this study generated an item-person map using a Rasch model. This is a graphical representation of person-abilities and item-difficulties, drawn based on the equal measures (logits) of the raw item difficulties and raw person scores [42]. The item-person map is divided into four quadrants, in which person estimates and item estimates are distributed on the left and right sides, respectively, based on person-ability and item-difficulty estimates [43]. Generally, the persons in the upper left quadrant show better abilities, implying that the easier items were not difficult enough for them. Meanwhile, the items on the upper right show higher difficulty, suggesting that they are beyond the students' ability level [7]. Rasch model analysis includes the following three steps: 1) computing the unidimensionality; 2) calculating person and item reliability coefficients; and 3) generating an item-person map [44].

4. Results

4.1. Construction of the Evaluation Model for Digital Literacy

4.1.1. Determination of the Feature Variables

After three rounds of expert opinions, the feature variables of each task were finally formed, as shown in Table 2. Return times was regarded as an inapplicable characteristic variable for all tasks, while completion time, thinking time, and correctness were identified as characteristic variables of tasks 1, 2, 3, 11, and 13. With regard to tasks 5, 9, and 12, all the feature variables other than return times were considered to be relevant variables for assessing students' digital literacy.

Table 2. The feature variables of each task as identified by experts.

Task \ Feature Variables	CT	TT	CO	HT	AT	EF	SI	RT
Task 1	√	√	√	×	×	×	×	×
Task 2	√	√	√	×	×	×	×	×
Task 3	√	√	√	×	×	×	×	×
Task 4	√	√	×	√	×	√	√	×
Task 5	√	√	√	√	√	√	√	×
Task 6	√	√	√	√	√	×	×	×
Task 7	√	√	√	√	√	×	×	×
Task 8	√	√	√	×	√	×	×	×
Task 9	√	√	√	√	√	√	√	×
Task 10	√	√	√	√	√	×	×	×
Task 11	√	√	√	×	×	×	×	×
Task 12	√	√	√	√	√	√	√	×

Task 13	√	√	√	×	×	×	×	×
---------	---	---	---	---	---	---	---	---

NOTE: CT: completion time; TT: thinking time; CO: correctness; HT: help times; AT: answer times; EF: efficiency; SI: similarity; RT: return times.

4.1.2. Calculation of the Weight of Feature Variables

After determining the feature variables, this study employed the analytic hierarchy process to calculate the weights of the feature variables of each task. First, a hierarchical structure model of each task's feature variables was constructed. The "expert advisory group" then assessed the relative importance of each task's feature variables by ranking them on a scale from 1 to 9. A judgment matrix was then constructed to calculate the weight of each task's feature variables. The average random consistency index RI was in the range of 0 to 1.59, and the calculated consistency ratio Cr was less than 0.1. All judgment matrices thus met the consistency requirements [45]. Table 3 shows the calculation results for the weights of the feature variables of each task. The weight calculation results indicate that CO had the largest weight proportion across all tasks, while the feature variables with the smallest weights varied across different tasks. For example, the feature variable with the smallest weight in tasks 1, 2, 3, 4, 6, 7, 8, 10, 11, and 13 was TT, while the feature variables with the smallest weights in tasks 5, 9, and 12 were AT, HT, and CT.

Table 3. Calculation results of the weights of the feature variables for each task.

Task	Feature Variables	CT	TT	CO	HT	AT	EF	SI
Task 1		0.254	0.114	0.632	/	/	/	/
Task 2		0.263	0.119	0.618	/	/	/	/
Task 3		0.276	0.118	0.606	/	/	/	/
Task 4		0.162	0.130	/	0.266	/	0.291	0.151
Task 5		0.138	0.112	0.220	0.089	0.106	0.168	0.167
Task 6		0.104	0.069	0.422	0.173	0.232	/	/
Task 7		0.110	0.071	0.432	0.181	0.206	/	/
Task 8		0.215	0.131	0.441	/	0.213	/	/
Task 9		0.131	0.106	0.242	0.103	0.113	0.149	0.156
Task 10		0.139	0.074	0.427	0.190	0.170	/	/
Task 11		0.241	0.118	0.641	/	/	/	/
Task 12		0.101	0.109	0.262	0.094	0.114	0.168	0.152
Task 13		0.237	0.126	0.637	/	/	/	/

4.1.3. Construction of the Assessment Model

Based upon the weights of the digital literacy evaluation indicators for primary and secondary students as obtained in our previous study [40], and also upon the weighting results obtained from the above analysis, the linear mathematical expression of the digital game-based digital literacy assessment model was calculated as follows:

$$Y = 0.295 * Y1 + 0.170 * Y2 + 0.289 * Y3 + 0.246 * Y4 \quad (1)$$

Where,

$$Y1 = 0.20 * (B1 * 0.254 + B2 * 0.114 + B3 * 0.632) + 0.20 * (B1 * 0.276 + B2 * 0.118 + B3 * 0.606) + 0.20 * (B1 * 0.104 + B2 * 0.069 + B3 * 0.422 + B4 * 0.173 + B5 * 0.232) + 0.20 * (B1 * 0.139 + B2 * 0.074 + B3 * 0.427 + B4 * 0.190 + B5 * 0.170) + 0.20 * (B1 * 0.241 + B2 * 0.118 + B3 * 0.641) \quad (2)$$

$$Y2 = 0.5 * (B1 * 0.110 + B2 * 0.071 + B3 * 0.432 + B4 * 0.181 + B5 * 0.206) + 0.5 * (B1 * 0.101 + B2 * 0.109 + B3 * 0.262 + B4 * 0.094 + B5 * 0.114 + B6 * 0.168 + B7 * 0.152) \quad (3)$$

$$Y3 = 0.25 * (B1 * 0.162 + B2 * 0.130 + B4 * 0.266 + B6 * 0.291 + B7 * 0.151) + 0.25 * (B1 * 0.138 + B2 * 0.112 + B3 * 0.220 + B4 * 0.089 + B5 * 0.106 + B6 * 0.168 + B7 * 0.167) + 0.25 * (B1 * 0.215 + B2 * 0.131 + B3 * 0.441 + B5 * 0.213) + 0.25 * (B1 * 0.131 + B2 * 0.106 + B3 * 0.242 + B4 * 0.103 + B5 * 0.113 + B6 * 0.149 + B7 * 0.156) \quad (4)$$

$$Y4 = 0.5 * (B1 * 0.263 + B2 * 0.119 + B3 * 0.618) + 0.5 * (B1 * 0.237 + B2 * 0.126 + B3 * 0.637) \quad (5)$$

In the above expressions, Y1, Y2, Y3, and Y4 represent the four dimensions of digital literacy, namely IAA, IKS, ITB, and ISR; B1-B7 represents the values of the seven feature variables, namely completion time, thinking time, correctness, help times, answer times, efficiency, and similarity.

Among these, the values for thinking time and completion time needed to be processed and then weighted. According to the experts' suggestions, the data for thinking time and completion time were processed as follows. 1) Reasonable time ranges for students' thinking times and answering times were determined based on the corresponding time distributions. 2) The duration of thinking times and completion times were divided into four parts (10%, 40%, 70%, and 100% of the duration of the longest time). 3) Values were assigned to different durations, according to the principle that the longer the duration, the lower the score. For example, if a student's thinking time fell within the top 10% of the reasonable time range, then a value of 1 was assigned; if a student's thinking time fell within 100% of the reasonable time range, then a value of 1/4 was assigned. Additionally, a value of 1 or 0 was assigned to the variable of correctness depending on whether the student answered correctly or incorrectly; a value of 1 was assigned to the variable of help only if the students did not click the "Help" button, otherwise 0 was assigned; a value of 1 was assigned to the variable of answer times only if the student answered correctly the first time, otherwise 0 was assigned. The values of similarity and efficiency were calculated using the Levenshtein distance.

4.2. Analysis of the Game-Based Assessment Tasks

This study analyzed the data for the students' digital literacy game-based assessment tasks using a Rasch model. As shown in Figure 3, on the right side, the difficulty of the items decreases from top to bottom; similarly, students' ability decreases from top to bottom on the left side. Item difficulty covered about five logit, while the students' ability covered seven logit. Among them, the most difficult task was task 9, while the easiest task was task 4. The average value of the students' ability in digital literacy was slightly higher than that of the items' difficulty, indicating that the difficulty of the game-based assessment tasks were appropriate to the students' actual level of digital literacy.

Figure 2. The person-item map.

4.3. Description of the Digital Game-Based Assessment Results

4.3.1. Analysis of overall results for the game-based assessment

Based on the above evaluation model, the average scores for students' overall digital literacy and its four individual dimensions were calculated. The results (see Figure 2) show that the students' overall digital literacy performance was good, with an average score of 70.14. However, the four dimensions of digital literacy were unbalanced. Specifically, the average scores for IAA, ITB, and ISR were more than 70%, while the average score for IKS was less than 50%. Of these, the average score for ISR was the highest (87.5%), indicating that students had commendable information ethics and exhibited a strong adherence to digital laws and regulations. The average score for IKS was the lowest (48.9%), indicating that the students had limited knowledge of information science and lacked proficiency in using digital technology to solve practical problems. The average score for ITB was 78.6%, suggesting that the students had the ability to decompose, abstract, and summarize complex problems. However, they encountered challenges in identifying the rules and characteristics of implicit information when problem-solving. The students' performance in the dimension of IAA was slightly lower than that in the ITB dimension, with an average score of 71.6%. This indicates that the students had a certain level of sensitivity and judgment to information, but they lacked the awareness of how to use digital technology to address real-world problems.

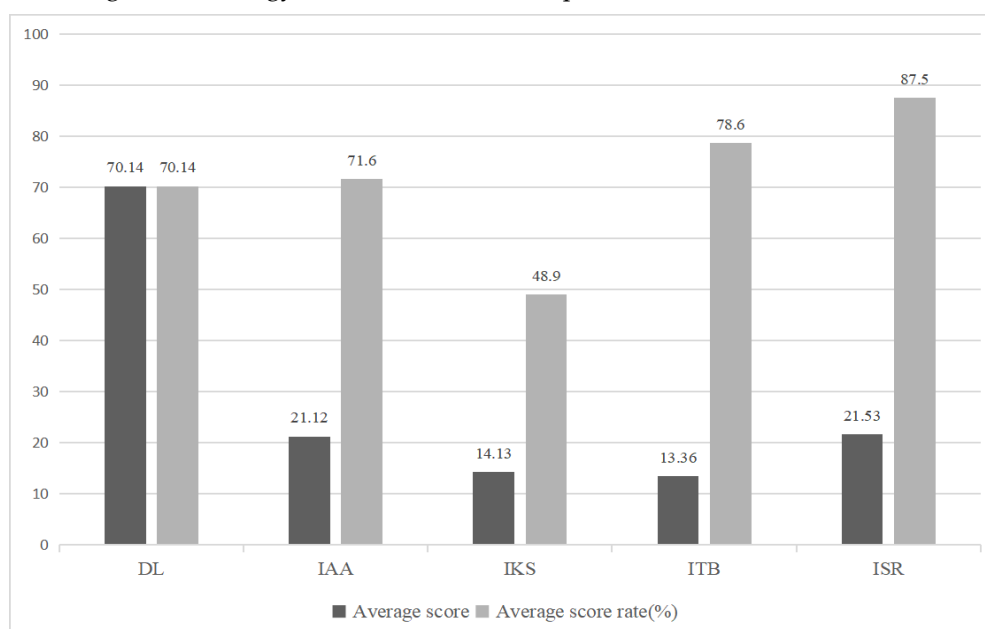


Figure 3. Overall results of the game-based assessment.

4.3.2. Analysis of the Feature Variables of the Game-Based Assessment

The average values of the seven feature variables of each dimension of students' digital literacy were calculated within the context of the game-based assessment tasks. It should be noted that tasks 2 and 13, relevant to ISR, were not mapped to the two feature variables of help times and answer times. Furthermore, the tasks associated with the IAA and ISR dimensions are multiple-choice questions, which do not generate such feature variables as efficiency and similarity.

As shown in Table 4, the average values of CT and TT (from the ITB dimension) were the highest, indicating that students spent the most time (both in terms of completion time and thinking time) on the tasks relevant to the ITB dimension. In terms of the feature variables related to game configuration, the average values of HT and AT (from the IKS dimension) were the highest, indicating that the students tended to frequently click the "Help" button and get the "Pass card", respectively. In terms of the feature variables related to behavior sequence, it is noteworthy that only the tasks associated with the ITB and IKS dimensions encompass drag and drop questions, maze questions, connection

questions, and sorting questions, all of which can collect behavior sequence data. The average values of similarity and efficiency in the ITB dimension were slightly higher than those in the IKS dimension. This implies that students exhibit a higher level of deviation between their action sequence and the reference sequence when responding to a task in terms of IKS. In terms of the feature variables related to answer results, the mean value of CO in the ISR dimension was the highest, while the CO value in the IKS dimension was the lowest.

Table 4. Results for the feature variables of the game-based assessment.

index	IAA	IKS	ITB	ISR
CT	7.468	18.716	23.758	6.87
TT	4.084	4.286	4.374	2.876
CO	71.452	48.168	77.241	81.422
HT	0.847	1.268	0.514	/
AT	0.769	1.068	0.375	/
EF	/	0.807	0.841	/
SI	/	0.587	0.795	/

4.4. Verification of the Digital Game-Based Assessment Results

In this study, Pearson correlation analysis was used to examine the relationship between the scores obtained from the digital literacy standardized test and the results of the digital literacy game-based assessment. The results show that the correlation coefficient was 0.914, suggesting a strong association between the results of the game-based assessment and the standardized test scores. Furthermore, this study revealed a significant association between the outcomes of the game-based assessment and the standardised test scores in connection to the four aspects of digital literacy. Specifically, the correlation coefficients for IAA, IKS, ITB, and ISR were determined to be 0.925, 0.918, 0.878, and 0.897, respectively. These results indicate that the game-based assessment tool is a highly reliable and valid instrument for assessing students’ digital literacy.

5. Discussion and Conclusions

Working from the concept of ECGD, a digital literacy assessment for secondary school students was carried out using the digital literacy game-based assessment system developed by the research team. Specifically, this study assessed students’ digital literacy through the following four steps: (1) establishing an ECGD based evaluation framework (comprising a student model, evidence model, and task model); (2) inducing students’ performance related to digital literacy using the digital literacy game-based assessment system; (3) using the Delphi method to determine feature variables based on the complex and fine-grained procedural behavior data generated in the process of students’ game-playing, and calculating the weights of the feature variables using the analytic hierarchy process (AHP); and (4) constructing an evaluation model to measure students’ digital literacy. The results show that the digital game-based assessment results were consistent with the standardized test scores, indicating that the ECGD-based digital literacy assessment approach is reliable and valid. This assessment can collect abundant evidence for digital literacy, much more so than traditional assessment methods as standardized testing. For example, by analyzing the similarity and efficiency of students’ action sequences, one can analyze the hidden problem-solving processes behind consistent answer results, thus obtaining more objective and richer insight into students’ performance. Therefore, we believe that the ECGD-based assessment method is particularly suitable for evaluating such complex and implicit competencies as digital literacy.

We believe that this study makes some important theoretical and practical contributions. In terms of theory, we put forward the idea of a game-based assessment of students’ digital literacy based on ECGD theory, establishing a new paradigm for the evaluation of students’ digital literacy and other higher-order thinking abilities. Specifically, the idea of a digital literacy game-based assessment based on ECGD can stimulate students’ interest and induce students’ digital literacy-

related performance in situational tasks. Doing so effectively solves the problems affecting current research on the evaluation of students' digital literacy, such as the impact of test anxiety on the objectivity of evaluation and the difficulty of measuring students' implicit thinking ability. In practice, we have built a student digital literacy evaluation model, which provides a new way of mapping process data onto higher-order thinking ability. Specifically, this study focuses on the problem that research to date on student digital literacy evaluation focuses on result data to the detriment of process data. Thus, in this study, the game-based evaluation tool was used to collect fine-grained process data, which was then mined to determine the correlation mechanism linking the digital literacy evaluation index to the key process data. The mapping relationship and weights linking the game evaluation tasks to the characteristic variables were determined through the Delphi method and analytic hierarchy process, and finally the student digital literacy evaluation model was established based on the integration of the result data with the process data, so as to reveal the underlying logic of the relationship between the two, thus accomplishing the accurate evaluation of students' digital literacy.

Based on the overall evaluation results and performance as evaluated by the feature variables of the game, this study found that the secondary school students had a moderate level of digital literacy, and the development of their four dimensions of digital literacy was not balanced. With regard to the score on each dimension, students performed better in the dimensions of IAA, ITB, and ISR, while they performed the worst in the IKS dimension. This suggests that most of the students' information consciousness and attitude are in an immature state of formation, although they have relatively keen attention and judgment to information, can use the methods in the field of computer science in problem-solving processes, and can abide by the code of ethical behavior, laws, and regulations of cyberspace. However, students' mastery of the relevant concepts, principles, and skills of information science and technology is still insufficient. Analysis of the feature variables showed that students' completion times and thinking times were longer in the ITB dimension than the other dimensions, although the score for the ITB dimension ranked second. That is to say, although the students had the ability to abstract and decompose questions, they tended to spend a longer time thinking when completing situational tasks with complex operations. Thus, it is still necessary to cultivate students' ability to discover, analyze, explore, and solve problems, so they can gradually internalize the ability to solve complex and real-world problems. In the IKS dimension, students' responses had low correctness and similarity, indicating that their mastery of scientific knowledge and application skills was not solid. Furthermore, the students were more inclined to click the "Help" button while completing the tasks associated with the IKS dimension. These findings address the necessity and significance of teaching the basic knowledge and skills of information science and technology, which should be the foundation of students' digital literacy education. In the ISR dimension, students not only achieved their highest scores, but also had the shortest completion times and thinking times. This suggests that students possess the ability to swiftly make appropriate decisions regarding digital ethics and morality. Consequently, it can be inferred that students possess a strong awareness in areas such as information security, risk management, and control, as well as intellectual property protection. Overall, the aforementioned research findings align with the outcomes of the large-scale standardised assessment results conducted by the research team in the early stage [40].

This study has some limitations that should be noted. First, its sample size was relatively small, and only one school was selected for assessment. We plan a follow-up study that will expand the sample size to collect samples of primary and secondary school students from all regions of China, so as to verify the effectiveness of the ECGD-based digital game assessment approach with more and richer data, and improve the reliability and representativeness of the research results. Second, this study only collected the clickstream data of students in the game assessment tasks, but did not collect eye movement, video, or other multimodal data. Future studies may make use of multimodal data, such as facial expression, facial posture, and eye movement, all of which could reflect students' performance in terms of digital literacy. Finally, although this study compared the game-based assessment results with standardized test scores via Pearson correlation analysis, the ECGD-based

assessment approach should be further verified in future research with a psychological measurement model, such as the cognitive diagnosis model.

Author Contributions: Writing, methodology, software, J.L.; assessment design, writing, J.B.; conceptualization, writing review and editing, S.Z.; supervision, H.H.Y. All authors have agreed with the publication of the manuscript.

Funding: The work was funded by a grant from the National Natural Science Foundation of China (No. 62107019), the Key project of the special funding for educational science planning in Hubei Province in 2023 (2023ZA032), and the Key Subjects of Philosophy and Social Science Research in Hubei Province of 2022 (22D043).

Ethics Statement: The procedures for human participants involved in this study were consistent with the ethical standards of Central China Normal University and the 1975 Helsinki Declaration. Prior to their involvement in the study, all participants provided their permission by signing a consent form in order to engage in the study.

Data Availability Statement: The datasets presented in this article are not readily available given the confidential nature of the data. Requests to access the datasets should be directed to zhusha@mail.ccnu.edu.cn.

Acknowledgments: This article is an extended version of “Assessing secondary students’ digital literature using an evidence centered game design approach”, which was published in the 16th International Conference on Blended Learning (pp. 214-223) by Springer Nature Switzerland. We appreciate the recommendation of the conference and the permission of the authors to publish the extended version.

Conflicts of Interest: The authors declare that there is no potential conflict of interest.

References

1. Oliveira, K.K.D.S., & de SOUZA, R.A. (2022). Digital transformation towards education 4.0. *Informatics in Education*, 21(2), 283-309.
2. Office of Educational Technology. Launching a digital literacy accelerator [EB/OL].(2022-09) [2022-12-10]. <https://tech.ed.gov/launching-a-digital-literacyaccelerator/>.
3. Zhu, S., Yang, H.H., MacLeod, J., Yu, L., & Wu, D. (2019). Investigating teenage students’ information literacy in China: A social cognitive theory perspective. *The Asia-Pacific Education Researcher*, 28, 251-263.
4. Tinmaz, H., Lee, Y.T., Fanea-Ivanovici, M., & Baber, H. (2022). A systematic review on digital literacy. *Smart Learning Environments*, 9(1), 1-18.
5. Chang, Y. K., Zhang, X., Mokhtar, I.A., Foo, S., Majid, S., Luyt, B., & Theng, Y.L. (2012). Assessing students’ information literacy skills in two secondary schools in Singapore. *Journal of information literacy*, 6(2), 19-34.
6. Nguyen, L.A.T., & Habók, A. (2022). Digital literacy of EFL students: An empirical study in Vietnamese universities. *Libri*, 72(1), 53-66.
7. Zhu, S., Wu, D., Yang, H.H., Wang, Y., & Shi, Y. (2019). Development and validation of information literacy assessment tool for primary students. In 2019 International Symposium on Educational Technology (ISET) (pp. 7-11). IEEE.
8. Rowe, E., Asbell-Clarke, J., & Baker, R.S. (2015). Serious games analytics to measure implicit science learning. *Serious games analytics: Methodologies for performance measurement, assessment, and improvement*, 343-360.1.
9. Zhu, S., Li, J., Bai, J., Yang, H.H., & Zhang, D. (2023). Assessing Secondary Students’ Digital Literacy Using an Evidence-Centered Game Design Approach. In *International Conference on Blended Learning* (pp. 214-223). Cham: Springer Nature Switzerland.
10. Negroponte, N. (1995). *Being Digital*. New York: Vintage Books.
11. Gilster, P., & Glistner, P. (1997). *Digital literacy*. New York: Wiley Computer Pub.
12. Eshet, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of educational multimedia and hypermedia*, 13(1), 93-106.
13. New Zealand Ministry of Education(2003). *Digital Horizons: Learning through ICT*. Wellington: Learning Media.
14. Calvani A, Fini A, Ranieri M. Assessing digital competence in secondary education. *Issues, models and instruments[M]//Leaning M. Issues in Information and Media Literacy: Education, Practice and Pedagogy*. Santa Rosa, California: Informing Science Press, 2009: 153-172.
15. UNESCO Institute for Statistics. A global framework of reference on digital literacy skills for indicator 4.4.2 [EB/OL]. [2023-06-01]. <http://uis.unesco.org/sites/default/files/documents/ip51-global-framework-reference-digital-literacy-skills-2018-en.pdf>.

16. Office of the Central Cyberspace Affairs Commission of China. Action Plan for Enhancing Digital Literacy and Skills for All [EB/OL]. http://www.cac.gov.cn/2021-11/05/c_1637708867754305.htm.
17. Wu, D., Zhu, S., Yu, L.Q., & Yang, S. (2020). Information literacy assessment for middle and primary school students. Beijing, China: Science Press.
18. Zhu, S., Sun, Z., Wu, D., Yu, L., & Yang, H. (2020). Conceptual Assessment Framework of Students' Information Literacy: An Evidence-Centered Design Approach. In 2020 International Symposium on Educational Technology (ISET) (pp. 238-242). IEEE.
19. DeVellis R F. Classical test theory[J]. Medical care, 2006: S50-S59.
20. Ng W. Can we teach digital natives digital literacy?(2012)[J]. Computers & education, 59(3): 1065-1078.
21. Nikou S, Molinari A, Widen G(2020). The interplay between literacy and digital technology: A fuzzy-set qualitative comparative analysis approach[J].
22. Cabero Almenara, J., Barroso Osuna, J.M., Gutiérrez Castillo, J.J., & Palacios-Rodríguez, A.D.P. (2020). Validación del cuestionario de competencia digital para futuros maestros mediante ecuaciones estructurales.
23. Peled, Y., Kurtz, G., & Avidov-Ungar, O. (2021). Pathways to a knowledge society: A proposal for a hierarchical model for measuring digital literacy among israeli pre-service teachers. Electronic Journal of e-Learning, 19(3), pp118-132.
24. Lukitasari M, Murtafiah W, Ramdiah S, et al.(2022). Constructing Digital Literacy Instrument and its Effect on College Students' Learning Outcomes[J]. International Journal of Instruction, 15(2).
25. Colwell, J., Hunt-Barron, S., & Reinking, D. (2013). Obstacles to developing digital literacy on the Internet in middle school science instruction. Journal of Literacy Research, 45(3), 295-324.
26. Porat, E., Blau, I., & Barak, A. (2018). Measuring digital literacies: Junior high-school students' perceived competencies versus actual performance. Computers & Education, 126, 23-36.
27. Öncül, G. (2021). Defining the need: Digital literacy skills for first-year university students. Journal of Applied Research in Higher Education, 13(4), 925-943.
28. Reichert F, Zhang D J, Law N W Y, et al.(2020). Exploring the structure of digital literacy competence assessed using authentic software applications[J]. Educational Technology Research and Development, 68(6): 2991-3013.
29. Bartolomé, J., & Garaizar, P. (2022). Design and Validation of a Novel Tool to Assess Citizens' Netiquette and Information and Data Literacy Using Interactive Simulations. Sustainability, 14(6), 3392. Mislevy R J, Corrigan S, Oranje A; et al. Psychometrics and game-based assessment[J]. Technology and testing: Improving educational and psychological measurement, 2016: 23-48.
30. Mislevy R J, Corrigan S, Oranje A; et al. Psychometrics and game-based assessment[J]. Technology and testing: Improving educational and psychological measurement, 2016: 23-48.
31. Mislevy, R.J., Almond, R.G., & Lukas, J.F. (2003). A brief introduction to Evidence-centered Design. ETS Research Report Series, 2003(1), i-29.
32. von Davier, A.A. (2017). Computational psychometrics in support of collaborative educational assessments. Journal of Educational Measurement, 54(1), 3-11.
33. Avdeeva, S., Rudnev, M., Vasin, G., Tarasova, K., & Panova, D. (2017). Assessing Information and Communication Technology Competence of Students: Approaches, Tools, Validity and Reliability of Results. Voprosy obrazovaniya/Educational Studies Moscow, (4), 104-132.
34. Turan, Z., & Meral, E. (2018). Game-Based versus to Non-Game-Based: The Impact of Student Response Systems on Students' Achievements, Engagements and Test Anxieties. Informatics in Education, 17(1), 105-116.
35. XU, J., & LI, Z. (2021). Game-based psychological assessment. Advances in Psychological Science, 29(3), 394.
36. Shute, V.J., Wang, L., Greiff, S., Zhao, W., & Moore, G. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. Computers in Human Behavior, 63, 106-117.
37. Chu, M.W., & Chiang, A. (2018). Raging Skies: Development of a Digital Game-Based Science Assessment Using Evidence-Centered Game Design. Alberta science education journal, 45(2), 37-47.
38. Bley, S. (2017). Developing and validating a technology-based diagnostic assessment using the evidence-centered game design approach: An example of intrapreneurship competence. Empirical Research in Vocational Education and Training, 9(i), 1-32.
39. Zhu, S., Bai, J., Ming, Z., Li, H., & Yang, H.H. (2022). Developing a Digital Game for Assessing Primary and Secondary Students' Information Literacy Based on Evidence-Centered Game Design. In 2022 International Symposium on Educational Technology (ISET) (pp. 173-177). IEEE.
40. Zhu, S., Yang, H.H., Wu, D., & Chen, F. (2021). Investigating the relationship between information literacy and social media competence among university students. Journal of educational computing research, 59(7), 1425-1449.

41. Wang, X., Chen, T., Zhang, Y., & Yang, H.H. (2021). Implications of the Delphi method in the evaluation of sustainability open education resource repositories. *Education and Information Technologies*, 26, 3825-3844.
42. Hani Syazillah, N., Kiran, K., & Chowdhury, G. (2018). Adaptation, translation, and validation of information literacy assessment instrument. *Journal of the association for information science and technology*, 69(8), 996-1006.
43. Oon, P.T., Law, N., Soojin, K., Kim, S., & Tse, S.K. (2013). Psychometric assessment of ICILS test items on Hong Kong and Korean students: A Rasch analysis. 5th IEA-IRC 2013.
44. Rasch, G. (1993). Probabilistic models for some intelligence and attainment tests. MESA Press, 5835 S. Kimbark Ave., Chicago, IL 60637; e-mail: MESA@uchicago.edu; web address: www.rasch.org; tele.
45. Rosaria de FSM Russo, R., & Camanho, R. (2015). Criteria in AHP: A systematic review of literature. *Procedia Computer Science*, 55, 1123-1132.

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.