Hindawi
Scientific Programming
Volume 2022, Article ID 6972215, 9 pages
https://doi.org/10.1155/2022/6972215

Research Article

Comparing Middle School Students' Scientific Problem-Solving Behavior in Hands-On Manipulation Performance Assessment: Terms by Eye-Tracking Analysis

Shiyi Zang,[1,2] Pingting Lin,[1,2] Xinyu Chen,[1,2] Yi Bai,[1,2] and Huihua Deng[1,2]

[1]*Key Laboratory of Child Development and Learning Science of Ministry of Education, Southeast University, Nanjing, China*
[2]*Institute of Child Development and Education, Southeast University, Nanjing, China*

Correspondence should be addressed to Huihua Deng; 230208222@seu.edu.cn

Received 3 June 2022; Revised 16 June 2022; Accepted 20 June 2022; Published 14 July 2022

Academic Editor: Lianhui Li

Copyright © 2022 Shiyi Zang et al. 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.

This study specifically designed an eye-tracking supported scientific problem-solving assessment: hands-on manipulation task system to explore the differences in visual attention and cognitive processes between high and low science achievement groups. Thirteen students with high science achievement and fourteen students with low science achievement participated. Students needed to complete the hands-on manipulation assessment, consisting of three modules, including selecting experimental equipment, experimental design, and building the experimental model. Behavioral and eye movement data were collected during the process. The results showed that the high science achievement group allocated more visual attention to the hands-on manipulation task, acquired more information through visual fixation, and assigned more attention to the key area. In module three of the hands-on manipulation task, the high science achievement group transformed from a visual channel to a tactile channel, and they generated more hands-on behaviors depending on the experimental area. Furthermore, the results showed a high correlation between students' eye movement behavior and the performance of scientific problem-solving assessments. Eye movement behavior could predict students' performance in scientific problem-solving. The average fixation duration and the average fixation duration of the interest area were two significant determining parameters. The implications of the experimental results for front-line science education, curriculum designers, and science assessment were also discussed.

1. Introduction

China's educational reform in recent years emphasized the importance of cultivating students' ability to solve scientific problems. According to the Primary School Science Curriculum Standards issued by the Ministry of Education in 2017, primary school students should master the necessary scientific knowledge and scientific research methods, develop the ability to handle practical scientific problems, and participate in public affairs. In the face of increasingly complex scientific problems, the ability to quickly retrieve and integrate essential information to form practical scientific problems solving paths and take actions along a practical path is crucial [1]. In PCAP scientific academic assessment content framework, three competencies in students' continuous knowledge learning were mentioned: scientific inquiry, scientific reasoning, and problem-solving. Thinking skills that students need include creativity, critical thinking, problem-solving, and metacognition [2]. Problem-solving is the cognitive processing of transforming a known state into a goal state when the problem solver has no obvious solution [3]. Newell and Simon [4] believed that problem-solving began with information retrieval to construct the mental representation of external problems in well-structured problems. By comparing the difference between the target state and the known state, effective operations are performed to successfully transform the given state of the system into the target state. The OECD [1] proposed the framework for problem-solving that involves four main processes: "exploration and understanding," "representing and formulating," "planning and executing," and "monitoring and reflecting." Traditional assessment measures like standardized tests are unsuitable for assessing multiple core competencies such as problem-solving competence. The following four methods are suggested for assessment: project assessment, performance assessment, group assessment, and portfolio assessment [5]. The reasoning could be strengthened when the relevant behavior is observed multiple times in multiple environments, and its performance takes into account cognitive, motivational, ethical, and emotional aspects [6]. Therefore, it is necessary to integrate evidence from various assessment sources. Redecke [7] stressed that, according to the core literacy to determine different evaluation ways, computer-based testing, online testing, simplified game, and electronic portfolio played a crucial role in the core literacy assessment. Meanwhile, the school curriculum evaluation should strengthen the technical development and application of authentic situation assessment. The Program for International Student Assessment (PISA) emphasized three characteristics of the problem-solving definition: "authentic situation" rather than the abstract problem. The "nonobvious" of scientific problem-solving reflected the nonroutine of problem solutions. The "interdisciplinary assessment" draws on a wide range of problem-solving knowledge [1].

The level description of problem-solving is not based on the specific thought process. It focuses on situational complexity, authenticity, and strategies. Authentic situation tasks require presenting situations that reflect real challenges in an individual's life and work, where students use scientific knowledge and skills, independent judgment, and creative collaboration to solve problems. The design of problem-solving tasks in [8] had carried out the computer management and human-computer interactive evaluation based on. An interactive task was added in which participants were designed to manipulate variables to determine the impact on the results [1]. NAEP has added interactive computer tasks and hands-on task sections since 2009. In hands-on tasks, students have to design scientific inquiry by themselves, select materials, and conduct an inquiry to solve problems and explain scientific phenomena [9]. Adopted the human-computer interaction mode, and the computer dynamically collected students' data to complete authentic situation tasks. However, as PISA tried to avoid repetition with other literacy evaluations when designing the assessment of problem-solving ability, thus such questions rarely involve the subject background knowledge, which restricted the practical application of problem-solving competency assessment. The "authenticity assessment" is always associated with performance assessment. The competency to use knowledge can only be assessed when completing a job in response to a specific task by using our knowledge and skills. This evaluation method can fully reflect students' competency, learning achievements, and extensive knowledge understanding [10]. According to the "authenticity evaluation" model, Chang [10] designed three different question forms to investigate students' scientific cognition level when studying ninth-grade students' scientific literacy. These studies found that the "authenticity evaluation" model with different question forms was an effective method to evaluate students' scientific literacy, among which "hands-on manipulation" significantly influences students' evaluation results. However, cultivating students' hands-on manipulation competency is often neglected due to time problems. Therefore, educators can improve their understanding of complex concepts by designing authentic situation models to reduce the gap between students' knowledge learning and real life.

The development of big data and information technology [11–14] provides a reliable supporting tool for this research. Many studies have found that some eye movement parameters [15–17], such as fixation duration, fixation duration percentage, fixation frequency, regression frequency, and average pupil diameter, provided critical evidence to explore the psychological process of students. Tai [18] compared the eye movement differences in problem-solving behaviors of participants with different professional backgrounds in scientific evaluation and found that the eye movement differences were related to the participant's professional level on a specific topic. Tsai et al. [19] investigated the eye movement differences of students with different professional backgrounds in PISA online assessment, and the study found that students majoring in science allocated more attention to critical areas and conducted deeper processing. Hu et al. [20] studied the information processing strategies of students in solving PISA interactive and analytical questions, which required self-discovery of information or provided a large amount of information in the questions. Students with good performance had a longer fixation duration in both questions. They tended to apply their previously constructed problem patterns to solve current problems, meanwhile adopting plan-driven forward reasoning to obtain more effective actions. Underperforming students switched their gaze points frequently between areas of interest and adopted "goal-result" backward reasoning, which required constant searching to reduce the difference between the current state and the target state. Krstić et al. [21] explored the students' eye movement patterns in PISA reading evaluation and found that students with high scores paid more attention to relevant information in texts and pictures, while students with low scores lingered in texts and pictures instead of extracting essential information. Kaller et al. [22] found evidence to distinguish the processes using eye-tracking technology. The study showed that the initial fixation change and fixation duration did not change with the problem structure, suggested that the early eye movements generate mental representations of the starting state and the target state, and compared the two states to determine structural differences. The problem structure did not affect the initial fixation change. The final gaze shifted before the problem structure affected the execution phase, suggesting that fine processing was related to the final fixation change. In recent studies, Mobile Eye Tracker (MET) has been conducted to compare the visual attention characteristics widely, people's cognitive process in authentic situation tasks allows participants to move their heads freely, and their behaviors are not strictly restricted. For instance, participants can walk in the park [23] and complete the experimental task [24]. The applications of MET meet the needs of different topic formats in the evaluation, especially the hands-on evaluation format. In educational evaluation, using MET to study problem-solving helps reveal participants' strategies for acquiring information, making it possible for procedural evaluation. Bryan et al. [25] applied MET to the International Adult Competency Assessment (PIAAC). The study selected a computer-based technology assessment containing literacy, numeracy, and problem-solving modules, where participants may use a pen or calculator and interact verbally with the researcher. Off-screen eye movements were captured with MET to see how participants read and recognized information during problem-solving. In the experiment of physical concept learning, Chien et al. [24] proposed that hands-on manipulation was essential for laboratory learning. The previous studies showed the following: (1) More professional people showed more refined visual behaviors when performing tasks, paid more attention to areas related to actual tasks, and paid less attention to redundant areas [26]. (2) Studies on grouping participants according to their performance found that participants with higher achievement could not only effectively use previous knowledge [27] and identify task-related areas but also adopt more effective analysis strategies to gain relevant information [20]. (3) Eye movement indicators could predict students' evaluation performance. Studies have shown that successful problem solvers had smaller pupils on easy problems while larger pupils on complex problems [28]. Other studies have found that average fixation duration was the best predictor; students with longer average fixation duration showed deeper cognitive processing [27].

2. Purpose

This study designed a set of scientific problem-solving evaluation tasks in the authentic situation, including the paper evaluation and hands-on manipulation assessment. After completing the scientific knowledge task module presented on the computer, participants were required to complete three scientific problem-solving hands-on manipulation tasks. According to their scientific knowledge concept evaluation and previous comprehensive scores, the participants were divided into high and low groups. The differences in eye movement and cognitive process among different groups were explored. Eye movement tracking technology was used to record and collect the eye movement behavior of each participant, such as their average fixation time, fixation time percentage, average pupil diameter, fixation times in each task module, and the average fixation time and fixation time percentage in different areas of interest in each task module. The following research questions can be used as guidelines in pursuing this goal:

(1) Would the high science achievement group perform better than the low science achievement group in the hands-on manipulation assessment?

(2) When performing the hands-on manipulation assessment, would the high science achievement group allocate more visual attention than the low science achievement group?

(3) What is the relationship between students' scientific problem-solving performance and eye movement behavior?

(4) Could the eye movement behavior be used to predict students' performance in scientific problem-solving?

3. Methods

3.1. Participants. In this study, thirty-nine students from grade seven in Taicang No. 1 Middle School, Jiangsu Province, China, were randomly selected. The number of participants in the preliminary experiment was 7. The number of valid participants in the formal experiment was 27, and 5 were invalid participants, who were rejected because the rate of eye movement data did not reach 90%. There was no significant difference in age between the high and low science achievement groups ($n = 27$; between ages 12 and 13; $M = 12.66$; $SD = 1.09$; 15 males and 12 females). The students who participated in the assessment have studied the core science concepts covered. The scores of students' entrance science examination and core concept evaluation were weighted as the basis for grouping. The average score of 27 students was 56.78, and there was an extremely significant difference in the weighted average score between the two groups of students, $t_{(1,\ 25)} = 5.97$, $p < 0.001$. Twenty-seven students were divided into high and low science achievement groups based on their weighted average scores. There were 13 students in the high science achievement group but 14 students in the low science achievement group.

3.2. Design of Scientific Problem-Solving: Hands-On Manipulation Assessment. This study developed a science problem-solving assessment, "Energy Efficient House," to evaluate students' core science concept of "heat absorption and heat dissipation." Participants were required to make experimental hypotheses, build experimental platforms, obtain measured data, and form preliminary conclusions according to the experimental result. The hands-on manipulation assessment consisted of three task modules, which correspond to the problem-solving framework "exploration and understanding," "representing and formulating," and "planning and execution" proposed by OECD [1]. Students were required to read computer materials, learn assistance manuals, and operate experimental instruments.

Module one was the selection of experimental equipment, in which participants chose experimental equipment suitable according to the various equipment provided by the laboratory. Module two was the experimental design. Participants made the experimental hypothesis of the "relationship between color and heat absorption ability" and designed the experimental scheme based on the problem-solving orientation. Module three was to build the experimental model; after completing the preliminary experimental design and equipment selection, participants needed to use experimental materials, sensors, and computer software to build an experimental model and then verify the hypothesis of "the relationship between object's color and heat absorption ability" by using the measured data.

The assessment was designed based on the ECD (Evidenced-Centered Design) model and explored the students' grasp of "heat energy" as the specific form of energy and conversion principle. Authoritative international science problem-solving assessments such as PISA, TIMSS, and NAP are based on paper or online text formats, while our research attempted to study hands-on science problem-solving processes.

The evaluation information presented on the computer was in well-constructed mixed text and composed of written web pages with paragraphs, tables, charts, and graphs organized in a mutually supportive and coherent manner. The computer was also equipped with experimental software adapted to the heat sensor, which was convenient for reading data from the sensor.

The experimental materials included a heating lamp, three metal heat absorption rods of the same material in different colors, an iron rack, a clamp, three-hole support, a pen, and a heat sensor. The usage of experimental equipment and sensors was described in the auxiliary manual, which was convenient for students to read. A team of four science education designers, consisting of two science teachers with master's degrees in science education, a science education researcher with a doctoral degree, and a science education professor, was responsible for designing the overall science problem-solving assessment.

The online scientific problem-solving evaluation assessment platform runs on the Linux server. The core of the scientific problem-solving evaluation system was programmed in PHP and MySQL to process large data sets and analytical programs efficiently.

3.3. The Encoding Process. The response of all participants in the scientific problem-solving assessment was recorded in the database for analysis. Meanwhile, the experimenter recorded participants' performance in the hands-on assessment for analysis. Three experimenters were all graduate students majoring in "educational technology" and "applied psychology." The experts' team developed a standard coding system and experimenter manual.

The coding system carried out detailed coding guidelines for the evaluation criteria of each question, such as how the principal tester scores the answers of each item and examples of answers that middle school students may have. Coding guidelines were developed according to the content of each module to ensure reliability and coding quality between encoders.

3.4. Eye Movement Analysis. Eye-tracking glasses 2 W were the instrument for participants to operate flexibly in computer-based scientific problem-solving assessment. The mobile tracking device designed and manufactured by the German company SMI recorded each student's eye movements at a 60 Hz sampling rate. It was a noninvasive system, as a standard pair of glasses weight 68 grams, with less impact on the psychological and physical burden of the students. The device was specially designed for the study of dynamic eye movement, and it provided maximum peripheral perception and binocular visual positioning to capture the natural gaze behavior of the students during the hands-on manipulation assessment. A standard three-point calibration and validation procedures were completed before the formal assessment. The following eye movement parameters were further analyzed from the original data to record participants' completion duration, total fixation duration (TFD), average fixation duration (AFD), fixation duration percentage (FDP), and Z score of pupil diameter (PDZ). BeGaze^{TM} software was used to analyze further each student's eye movement parameters in scientific problem-solving assessment. Division of AOI in hands-on manipulation assessment is shown in Figure 1.

4. Results

4.1. Different Hands-On Assessment Performance between High and Low Science Achievement Groups. The independent sample t-test result for the high and low groups of students showed that the high science achievement group students significantly outperformed the low science achievement group in the total score of the hands-on task, $t_{(1,\ 25)} = 5.97$, $p < 0.001$, the first module, $t_{(1,\ 25)} = 3.34$, $p < 0.01$, and the second module, $t_{(1,\ 25)} = 4.36$, $p < 0.001$. However, there was no significant difference in the score of the third module, $t_{(1,\ 25)} = 1.25$, $p > 0.05$.

4.2. Differences in Eye Movement Behavior between High and Low Science Achievement Groups in Hands-On Manipulation Assessment. Two-factor repeated measure ANOVA results among three modules in the assessment and the eye movement indexes furthermore conducted the simple effect analysis of the three modules and different scientific achievements, as shown in Tables 1–3. Tests of within-subjects effects on the module are shown in Table 1, and tests of between-subjects effects on group are shown in Table 2. The results of the simple effect analysis of different group students on FDP, AFD, and PDZ are shown in Table 3.

The main effect of the three modules on the FDP was highly significant, $F_{(2,\ 50)} = 11.11$, $p < 0.001$, $\eta^2 = 0.31$. The main effect of scientific literacy on FDP was significant, $F_{(1,\ 25)} = 7.95$, $p < 0.01$, $\eta^2 = 0.24$. The edge of interaction between scientific literacy level and hands-on assessment was significant, $F_{(2,\ 50)} = 2.53$, $p = 0.090$, $\eta^2 = 0.09$. During three modules of hands-on manipulation assessment, the result revealed that the high science achievement group allocated more visual attention to the task and more access to information through visual gaze than the low scientific achievement group, including $\mathrm{FDP}_{\mathrm{module1}}$, $F_{(1,\ 25)} = 1.87$, $p > 0.05$, $\mathrm{AFD}_{\mathrm{module1}}$, $F_{(1,\ 25)} = 1.32$, $p > 0.05$, $\mathrm{PDZ}_{\mathrm{module1}}$, $F_{(1,\ 25)} = 8.21$, $p < 0.01$. $\mathrm{FDP}_{\mathrm{module2}}$, $F_{(1,\ 25)} = 6.42$, $p < 0.05$, $\mathrm{AFD}_{\mathrm{module2}}$, $F_{(1,\ 25)} = 5.10$, $p < 0.05$, $\mathrm{PDZ}_{\mathrm{module2}}$, $F_{(1,\ 25)} = 6.99$, $p < 0.05$. $\mathrm{FDP}_{\mathrm{module3}}$, $F_{(1,\ 25)} = 6.92$, $p < 0.05$,

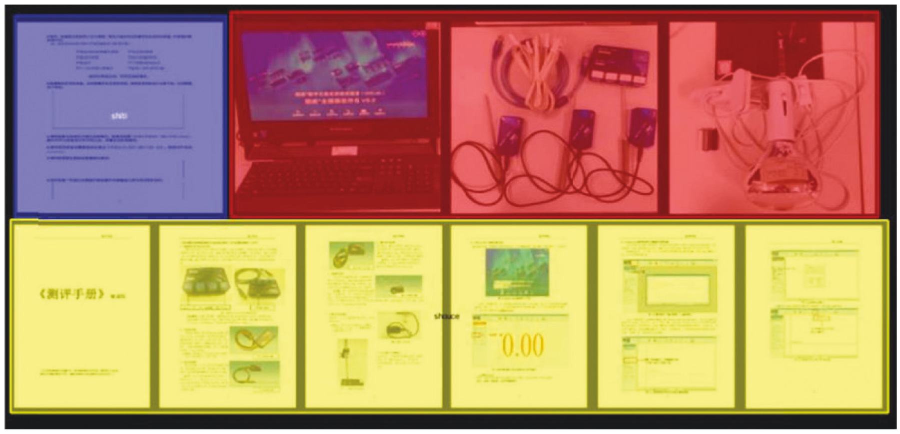

Figure 1: Division of AOI in hands-on manipulation assessment.

Table 1: Tests of within-subjects effects on the module.

	Source	Sum of squares	*df*	Mean square	*F*	η^2
	Module	603.60	2	301.80	11.11^{***}	0.31
FDP	Module * group	137.64	2	77.58	2.53	0.09
	Error	1358.32	50	27.17		
	Module	304559.91	1.36	224371.70	41.55^{***}	0.62
AFD	Module * group	20634.92	1.36	15201.91	2.82	0.10
	Error	183264.20	33.94	5400.49		
	Module	47.49	1.04	45.51	592.11^{***}	0.96
PDZ	Module * group	0.58	1.04	0.56	7.26^{*}	0.24
	Error	1.85	24	0.08		

Table 2: Tests of between-subjects effects on group.

	Source	Sum of squares	*df*	Mean square	*F*	η^2
FDP	Group	683.26	1	683.26	7.95^{**}	0.24
	Error	2149.97	25	86.00		
AFD	Group	61848.75	1	61848.75	6.30^{*}	0.20
	Error	245413.36	25	9816.53		
PDZ	Group	$1.03 * 10^{-12}$	1	1	0.16	0.01
	Error	$1.54 * 10^{-13}$	23	$0.03 * 10^{-12}$		

Table 3: Results of the simple effect analysis of different group students on FDP, AFD, and PDZ.

	AOI	Sum of squares	*df*	Mean square	*F*
	Selection of experimental equipment	49.28	1	49.28	1.87
FDP	Experimental design	214.91	1	214.91	6.42^{*}
	Build the experimental model	556.70	1	556.70	6.92^{*}
	Selection of experimental equipment	1503.70	1	1503.70	1.32
AFD	Experimental design	58363.37	1	58363.37	5.10^{*}
	Build the experimental model	22616.60	1	22616.60	4.96^{*}
	Selection of experimental equipment	0.25	1	0.25	8.21^{**}
PDZ	Experimental design	0.33	1	0.33	6.99^{*}
	Build the experimental model	0.01	1	0.01	1.89

$AFD_{module3}$, $F_{(1,\ 25)} = 4.96$, $p < 0.05$, and $PDZ_{module3}$, $F_{(1,\ 25)} = 1.89$, $p > 0.05$.

4.3. Differences of AOI in Eye Movement Behavior between High and Low Science Achievement Groups in Hands-On Manipulation Assessment. The two-factor repeated measurement ANOVA and multiple comparison test conducted on the FDP of AOI between the high and low groups and the key areas of each module were shown in Table 4. In module one, the FDP of the high group was significantly higher than the low group on the key auxiliary manual area, $F_{(1,\ 25)} = 10.02$, $p < 0.01$. In module two, the FDP of the high group was significantly higher than the low group on the key

Table 4: The key area of each module in hands-on manipulation assessment.

Stage of module	Auxiliary manual area	Worksheet area	Experimental area
Module one: Selection of experimental equipment	+		
Module two: Experimental design		+	
Module three: Build the experimental model			+

worksheet area, $F_{(1,\ 25)} = 7.40$, $p < 0.05$. In module three, the FDP of the high group was significantly higher than the low group on the key experimental area, $F_{(1,\ 25)} = 8.73$, $p < 0.01$.

4.4. The Correlation and Regression Analysis of Students' Eye Movement Behaviors and Scientific Problem-Solving Performance. The results of the Pearson correlation analysis, which was conducted to explore the correlation between students' eye movement behavior and scientific problem-solving performance, indicated that students' scientific problem-solving performance was highly correlated with their eye movement behavior, including cFDP ($r = 0.56$, $p < 0.01$), aAFD ($r = 0.77$, $p < 0.001$), cAFD ($r = 0.86$, $p < 0.001$), and bPDZ ($r = -0.97$, $p < 0.001$). In addition, a stepwise multiple regression analysis was used to assess which eye movement indicators could best predict students' scientific problem-solving performance. The results showed that aAFD and bAFD were the most significant predictors of eye movement when participants completed the scientific problem-solving task, as shown in Table 5, $T = -2.43$, $p < 0.05$, $\beta = -0.66$; $T = 2.31$, $p < 0.05$, $\beta = 0.48$.

The regression equation of Model one was

$$Y = 16.78 - 21.75\text{aPDZ} + 37.62c\text{PDZ} - 0.22\text{aAFD} + 0.05\text{bAFD} + 0.07\text{cAFD}. \quad (1)$$

The study compared the prediction result with students' hands-on manipulation performance; the prediction success rate for high science achievement students was 84.6%, but 78.6% for low science achievement students, and the total prediction success rate was 81.5%. In conclusion, the AFD and PDZ in the three modules of hands-on manipulation assessment could better predict the students' scientific performance. What is more, if the AFD of students was lower in module one but higher in module two, the student's scientific performance may be better.

4.5. The Correlation and Regression Analysis of AOI Index and Scientific Problem-Solving Performance of Students. The results of Pearson correlation analysis showed that students' AOI was highly correlated with their scientific problem-solving performance, including a2AFD($r = 0.81$, $p < 0.001$), b1AFD($r = 0.62$, $p < 0.001$), b3AFD($r = 0.63$, $p < 0.001$), and c3AFD($r = 0.7$, $p < 0.001$), as shown in Table 6. In addition, a stepwise multiple regression analysis was used to assess which AOI indicators could best predict students' science problem-solving performance. The results showed that b1AFD and c2AFD were the most significant predictors of AOI when participants completed the scientific problem-solving task, $T = 2.92$, $p < 0.01$, $\beta = 0.75$; $T = -2.60$, $p < 0.05$, $\beta = -0.60$.

Table 5: Standard multiple regression for eye movement indexes on scientific achievements.

Parameter	B	SE	β	T	Tolerance	VIF
Constant	16.78	37.60		0.45		
aPDZ	−21.75	13.20	−0.43	−1.65	0.40	2.52
cPDZ	37.62	26.73	0.37	1.41	0.40	2.47
aAFD	−0.22	0.09	−0.66	−2.43∗	0.36	2.75
bAFD	0.05	0.02	0.48	2.31∗	0.64	1.56
cAFD	0.07	0.04	0.43	1.70	0.43	2.32

∗a stands for module 1, b stands for module 2, and c stands for module 3.

Table 6: Standard multiple regression for AOI on scientific achievements.

Parameter	B	SE	β	T	Tolerance	VIF
Constant	82.36	18.60		4.43 ∗ ∗ ∗		
a1AFD	−0.08	0.05	−0.49	−1.68	0.27	3.67
a2AFD	−0.09	0.13	−0.23	−0.66	0.18	5.49
a3AFD	0	0.02	0	0.01	0.72	1.39
b1AFD	0.04	0.02	0.75	2.92∗ ∗	0.35	2.88
b2AFD	0.01	0.02	0.09	0.42	0.49	2.06
b3AFD	0.03	0.03	0.19	0.88	0.48	2.09
c1AFD	−0.01	0.02	−0.07	−0.37	0.59	1.70
c2AFD	−0.10	0.04	−0.60	−2.60∗	0.43	2.35
c3AFD	0.05	0.03	0.38	1.54	0.38	2.66

∗a1 stands for the auxiliary manual area in module 1, a2 stands for the worksheet area in module 1, a3 stands for the experimental area in module 1, and so on.

The regression equation of Model two was

$$Y = 82.36 - 0.08\text{a1AFD} - 0.09\text{a2AFD} + 0.04\text{b1AFD} + 0.01\text{b2AFD} + 0.03\text{b3AFD} - 0.01\text{c1AFD} - 0.10\text{c2AFD} + 0.05\text{c3AFD}. \quad (2)$$

The study compared the prediction results with students' hands-on manipulation performance; the prediction success rate of Model two for high science achievement students was 69.2%, but 92.9% for low science achievement students, and the total prediction success rate was 81.5%. Model two could explain 61% of the variability in student performance during the same overall success rate as Model one. Therefore, Model two could better predict the students' scientific performance. According to the significance index, if the AFD on the worksheet was higher in module two experimental design, while the AFD on the auxiliary manual was lower in module two experimental model building, the students' scientific performance may be better.

5. Discussion

The comparison of eye movement behavior between the high and low science achievement groups.

The study showed that the high science achievement group performed better than the low science achievement group in the hands-on manipulation performance, while eye movement behavior also showed that FDP and AFD were higher in the high science achievement group. One possible explanation was that high science achievement students allocated more visual attention in the hands-on manipulation assessment, more concentrated, and acquired information through visual gaze to a greater extent, while low science achievement students were more distracted by external factors when completing the task. Goldhammer et al. [29] pointed out that, in problem-solving tasks, the more time students spend on the task, the better they perform. In addition, they found the effect of time depended on the difficulty and intensity of the task, especially in the more difficult tasks; investing time in cognitive activities had a positive predictive effect on the successful completion of the task. The result of this study was consistent with previous studies that high science achievement students had higher fixation duration and were more concentrated.

The total fixation counts of high scientific achievement students were lower in module two experimental design and module three experimental model building. High science achievement students allocated more fixation duration and average fixation duration and fewer fixation counts in the cognitive processing, which were consistent with previous studies that students with longer average fixation duration tend to show deeper cognitive processing and better learning results. Students with high science achievement were more likely to be aware of the tasks' difficulties, especially representation, refinement, and execution. Thus, they took longer to think profoundly and achieve higher accuracy. These findings provided great support to several previous studies suggesting that high-level problem-solving students had deeper cognitive processing of task-related areas. They tended to apply their previously constructed mental models to solve problems, while low-level problem-solving students have not established a correct mental model due to their shallow cognitive processing of the critical areas in a problem-solving task. Thus, they reevaluated the consistency of the established model with the problem task by using the visual information of the auxiliary areas [30].

In the hands-on manipulation assessment, the high science achievement group had a higher PDZ than the low science achievement group, consistent with previous studies that individuals showed more significant pupil dilation when handling complex tasks [31]. Pupil dilation reflected individuals' continuous information processing, and challenging tasks would elicit more cognitive effort. In this study, the average pupil diameter of students was maximized in module three, which was the most challenging module. However, our results indicated no statistically significant difference between the high and low science achievement groups in module three, and one possible explanation was that students in both groups made more cognitive efforts in module three.

5.1. Comparison of AOI between High and Low Science Achievement Groups. The results showed that students with high science achievement assigned more visual attention to the auxiliary manual in module one, which was the key area of this module. Meanwhile, they allocated more attention to the key experimental area in module three, consistent with previous research that high-level students paid more attention to the task-related areas. In the hands-on manipulation assessment, the worksheet area and auxiliary manual area belong to the visual mode information of text or pictures, while the experimental area belongs to the tactile channel mode. The comparison between the two groups revealed that the high science achievement students had a statistically significant change in the problem-solving process from relying on visual channel mode to tactile channel mode. Particularly in module three, high science achievement students relied more on experimental areas and produced more hands-on behaviors.

The fixation sequence of the AOI revealed that high science achievement students showed more continuous cognitive processing in the key areas. In module three, the low scientific achievement students' gaze of interaction frequently switched between the key experimental area and auxiliary areas. The eye movement index reflected that the low science achievement students allocated more fixation duration in auxiliary manual area, which meant they could have recognition difficulty in experimental model building. The findings of students' cognitive processing were consistent with previous studies, and high-level problem solvers could deeper conduct cognitive processing of task-related areas [32].

5.2. Eye Movement Behavior Predicted Students' Scientific Problem-Solving Performance. The regression prediction model showed that the AFD and the AFD of AOI were the best indicators to predict students' performance in scientific problem-solving performance assessment. The finding added to previous evidence that eye movement was highly correlated with mental effort, reading processes, computer-based assessment of performance, and the construction of mental models. Students may perform better if they engage in deeper cognitive processing on relatively difficult tasks. In solving systematic experimental problems, if students could think profoundly and construct psychological models gradually, their evaluation performance may be better. These findings have implications for how eye-tracking-supported hands-on manipulation science problem-solving assessment could be used to improve science learning in junior high school students. In order to cultivate and improve students' scientific literacy, educators should pay more attention to students' cognitive process and deep thinking in science teaching and design high-level cognitive activities to achieve this goal. For instance, to establish a complete practice, encourage students to explore learning. The practice should be set up as realistic as possible, regarded as solving problems in real life. When designing the practiced mask, we could start from the eight links of "asking questions, making assumptions, making plans, collecting evidence, processing information, drawing conclusions, expressing and communicating, and reflecting and evaluating."

In the practice of student inquiry, educators could use prominent hints to guide students to think deeply about key information and difficult problems rather than repeatedly thinking superficially. The guidance and intervention in practice would shape students' thinking and behavior patterns, relieve their anxiety when meeting practical problems, and improve students' problem-solving ability. However, solving practical problems depends not only on the ideas and skills of the problem solver but also on the difficulty of the problem itself. Therefore, the significance of setting up practical exploration lies in improving students' problem-solving ability on the one hand and guiding students to see the boundaries and limitations of problem-solving and recognize the future development direction of a specific field on the other hand, which has more excellent value in improving students' cognitive level and scientific literacy.

6. Conclusions

This study specifically designed an eye-tracking supported problem-solving assessment hands-on manipulation task system to further explore the differences in visual attention and cognitive progress between high and low science achievement groups in the scientific problem-solving process. The research combined synchronous computer and pupil 60 Hz tracking system technology with a hands-on manipulation assessment system to explore the students' cognitive processes in hands-on operation. The system included static concept evaluation problems and dynamic hands-on problems instead of being limited to reading and solving cognitive model exploration in static problems. These results are also very encouraging as they demonstrate some important findings. (1) Students in the high science achievement group performed better than those in the low science achievement group in solving scientific problems assessment, consistent with their eye movement behavior. (2) The results of ANOVA in eye movement indicator showed that, in the three modules of hands-on manipulation assessment, students with high science achievement had more fixation and concentration on the task and acquired more information through visual fixation. In contrast, students with low science achievement were more likely to be distracted by external factors when finishing the assessment. (3) ANOVA results in the AOI index showed that students with high science achievement paid more attention in key areas during different modules. In addition, students with high science achievement had a more noticeable change from relying on visual channel mode to relying on tactile channel mode in the problem-solving process. Particularly in module three, the execution phase of problem-solving, students with high science achievement paid more attention to the experimental area, which produced more operation behavior. These results were consistent with their eye movement behavior. (4) Eye movement behavior was highly correlated with scientific problem-solving performance. (5) Eye movement behavior could predict students' performance in scientific problem-solving. The AFD and the AFD of AOI were significant determining parameters.

The highlights of this work are summarized as follows:

(1) This study specifically designed an eye-tracking supported scientific problem-solving assessment: hands-on manipulation task system to explore the differences in visual attention and cognitive processes between high and low science achievement groups.
(2) The result showed that the high science achievement group allocated more visual attention to the hands-on manipulation task, acquired more information through visual fixation, and assigned more attention to the key area than the low science achievement groups.
(3) The study also found that the average fixation duration and the average fixation duration of the AOI were two significant determining parameters for predicting students' performance in scientific problem-solving. [33–39]

Data Availability

The dataset used to support the findings of the study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] N. J. Oecd Publishing, *PISA 2012 Assessment and Analytical Framework: Mathematics, reading, Science, Problem Solving and Financial Literacy*, p. 264, OECD Publishing (NJ3), Paris, 2013.
[2] S. E. Fahyan, "Assessment and teaching of 21st century skills," Edited by P. Griffin, B. McGaw, and E. Care, Eds., Springer, Cham, 2014.
[3] R. E. Mayer, "The promise of multimedia learning: using the same instructional design methods across different media," *Learning and Instruction*, vol. 13, no. 2, pp. 125–139, 2003.
[4] H. A. Simon and A. Newell, "Human problem solving: the state of the theory in 1970," *American Psychologist*, vol. 26, no. 2, pp. 145–159, 1971.
[5] M. Kalantzis, B. Cope, and A. Harvey, "Assessing multiliteracies and the new basics," *Assessment in Education: Principles, Policy & Practice*, vol. 10, no. 1, pp. 15–26, 2003.
[6] D. S. E. Rychen and L. H. E. Salganik, *Key Competencies for a Successful Life and Well-Functioning Society*, Hogrefe & huber, Cambridge, 2003.
[7] C. Redecker, "The use of ICT for the assessment of key competences," Publications Office of the European Union, Brussels, 2013.
[8] PISA results, *PISA 2012 Results: Creative Problem Solving*, Oecd, Paris, 2014.

[9] N. Board, *Science Framework for the 2011 National Assessment of Educational Progress*, National Assessment Governing Board, New York, 2014.

[10] S. N. Chang and M. H. Chiu, "The development of authentic assessments to investigate ninth graders? Scientific literacy: in the case of scientific cognition concerning the concepts of chemistry and physics," *International Journal of Science and Mathematics Education*, vol. 3, no. 1, pp. 117–140, 2005.

[11] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *Journal of Industrial Information Integration*, vol. 26, no. 9, Article ID 100289, 2022.

[12] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, pp. 174988–175008, 2020.

[13] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 8, no. 99, pp. 154663–154670, 2020.

[14] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II," *Complexity*, vol. 2020, no. 6, 24 pages, Article ID 3853925, 2020.

[15] Y. Tang and J. Su, "Eye movement prediction based on adaptive BP neural network," *Scientific Programming*, vol. 2021, Article ID 4977620, 9 pages, 2021.

[16] X. Li, Y. Zhou, and Y. He, "The fusion of eye movement and piezoelectric sensing technology assists ceramic art process optimization and mechanical characterization," *Journal of Sensors*, vol. 2021, Article ID 9748335, 11 pages, 2021.

[17] X. Jinjing, "Research on flipped classroom of university curriculum using eye movement analysis and LSTM neural network," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 2559864, 10 pages, 2022.

[18] R. H. Tai, J. F. Loehr, and F. J. Brigham, "An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments," *International Journal of Research and Method in Education*, vol. 29, no. 2, pp. 185–208, 2006.

[19] P. Y. Tsai, T. T. Yang, and H. C. She, "Explore college students' cognitive processing during scientific literacy online assessments with the use of eye tracking technology," in *Proceedings of the 2015 IEEE International Conference on Advanced Learning Technologies*, pp. 303-304, Hualien, Taiwan, July 2015.

[20] Y. Hu, B. Wu, and X. Gu, "Learning analysis of k-12 students' online problem solving: a three-stage assessment approach," *Interactive Learning Environments*, vol. 25, no. 2, pp. 262–279, 2017.

[21] K. Krstić, A. Šoškić, V. Ković, and K. Holmqvist, "All good readers are the same, but every low-skilled reader is different: an eye-tracking study using PISA data," *European Journal of Psychology of Education*, vol. 33, no. 3, pp. 521–541, 2018.

[22] C. P. Kaller, B. Rahm, K. Bolkenius, and J. M. Unterrainer, "Eye movements and visuospatial problem solving: identifying separable phases of complex cognition," *Psychophysiology*, vol. 46, no. 4, pp. 818–830, 2010.

[23] F. Y. Yang, M. J. Tsai, and G. L. Chiou, "Instructional suggestions supporting science learning in digital environments based on a review of eye tracking studies," *Educational Technology & Society*, vol. 21, no. 2, pp. 28–45, 2018.

[24] K. P. Chien, C. Y. Tsai, H. L. Chen, W. H. Chang, and S. Chen, "Learning differences and eye fixation patterns in virtual and physical science laboratories," *Computers & Education*, vol. 82, pp. 191–201, 2015.

[25] M. Bryan, A. P. Bayliss, F. Piers, P. E. Engelhardt, E. S. Gareth, and B. Francesca, "Observing response processes with eye tracking in international large-scale assessments: evidence from the OECD PIAAC assessment," *European Journal of Psychology of Education*, vol. 33, no. 3, pp. 1–16, 2018.

[26] A. Gegenfurtner, E. Lehtinen, and R. Säljö, "Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains," *Educational Psychology Review*, vol. 23, no. 4, pp. 523–552, 2011.

[27] S. C. Chen, H. C. She, M. H. Chuang, J. Y. Wu, J. L. Tsai, and T. P. Jung, "Eye movements predict students' computer-based assessment performance of physics concepts in different presentation modalities," *Computers & Education*, vol. 74, no. 3, pp. 61–72, 2014.

[28] C. J. Wu, C. Y. Liu, C. H. Yang, and Y. C. Jian, "Eye-movements reveal children's deliberative thinking and predict performance on arithmetic word problems," *European Journal of Psychology of Education*, vol. 36, no. 1, pp. 91–108, 2020.

[29] F. Goldhammer, J. Naumann, A. Stelter, K. Tóth, H. RölkeRölke, and E. Klieme, "The time on task effect in reading and problem solving is moderated by task difficulty and skill: insights from a computer-based large-scale assessment," *Journal of Educational Psychology*, vol. 106, no. 3, pp. 608–626, 2014.

[30] K. A. Ericsson and J. H. Moxley, "Experts' superior memory: from accumulation of chunks to building memory skills that mediate improved performance and learning," in *The SAGE handbook of applied memory*, pp. 404–420, Sage, Newcastle upon tyne, 2014.

[31] P. S. Huang and H. C. Chen, "Gender differences in eye movements in solving text-and-diagram science problems," *International Journal of Science and Mathematics Education*, vol. 14, no. S2, pp. 327–346, 2015.

[32] E. M. Reingold, N. Charness, M. Pomplun, and D. M. Stampe, "Visual span in expert chess players: evidence from eye movements," *Psychological Science*, vol. 12, no. 1, pp. 48–55, 2001.

[33] S. D. Newman and G. Pittman, "The tower of London: a study of the effect of problem structure on planning," *Journal of Clinical and Experimental Neuropsychology*, vol. 29, no. 3, pp. 333–342, 2007.

[34] T. Purpose and P. O. Pisa, *Beyond Pisa 2015: A Longer-Term Strategy Of Pisa*, Pisa, 2020.

[35] J. Rahm, H. C. Miller, L. Hartley, and J. C. Moore, "The value of an emergent notion of authenticity: examples from two student/teacher-scientist partnership programs," *Journal of Research in Science Teaching*, vol. 40, no. 8, pp. 737–756, 2003.

[36] N. Ruh, B. Rahm, J. M. Unterrainer, C. Weiller, and C. P. Kaller, "Dissociable stages of problem solving (ii): first evidence for process-contingent temporal order of activation in dorsolateral prefrontal cortex," *Brain and Cognition*, vol. 80, no. 1, pp. 170–176, 2012.

[37] M. A. Ruiz-Primo, "Informal formative assessment: the role of instructional dialogues in assessing students' learning," *Studies In Educational Evaluation*, vol. 37, no. 1, pp. 15–24, 2011.

[38] O. J. Solheim and P. H. Uppstad, "Eye-tracking as a tool in process-oriented reading test validation," *International Electronic Journal of Environmental Education*, vol. 4, no. 1, pp. 153–168, 2011.

[39] G. Wiggins, "Moving to modern assessments," *Phi Delta Kappan*, vol. 92, no. 7, p. 63, 2011.