Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Scaffolding Causal Understanding via Parameter Exploration in Newtonian Mechanics

Jean Paul PIERRET ROBLES[a*], Nonoka AIKAWA[b], Kento KOIKE[c,d] & Takahito TOMOTO[e]
[a]*Graduate School of Information Science, Chiba Institute of Technology, Japan*
[b]*Integrated Education Center, Meikai University, Japan*
[c]*Faculty of Engineering, Tokyo University of Science, Japan*
[d]*Research Liaison Centre, Chiba Institute of Technology, Japan*
[e]*Faculty of Innovative Information Science, Chiba Institute of Technology, Japan*
*jp.pierret09@gmail.com

Abstract: It is crucial to consider parameters such as force and motion when solving physical problems. However, novices often struggle (a) what parameters define a given state and (b) what kind of state transition occurs when a certain operator is applied. Therefore, we created a learning environment for learners to engage in exploration based on the relationship between inputs and outputs. Thus, this study uses error-based simulation to help students learn this exploration through the relationship between parameter inputs and outputs, which is a type of computer-based simulation environment.

Keywords: problem-solving, search space, error-based simulation, error exploration

1. Introduction

Learning is an activity that occurs in a domain that a student has not yet mastered. Concerning learning, problem-solving occurs in an unmastered domain; therefore, arriving at the correct answer in the first attempt is difficult. In such cases, using operators exploratively to transition from the initial state presented in the problem to the required final state is necessary, seeking a solution through trial and error. To conduct such search activities, understanding (a) what parameters define a given state and (b) what kind of state transition occurs when a certain operator is applied is essential.

For example, in a diagram drawing problem in mechanics, the correct answer for an object at rest on the ground is to draw the downward force of gravity from the center and an equal upward normal force pushing back from the ground. The object remains stationary in the vertical direction when the upward and downward forces are equal. To provide a comprehensive explanation of this series of causal relationships, the net force in the vertical direction becomes zero when the upward and downward forces are balanced. If the net force in the vertical direction is zero, and there is no initial velocity, the velocity of the object is zero. Therefore, the object does not move and remains at rest. Learners must construct an appropriate search space for themselves to construct this type of causal understanding, i.e., they must understand which parameters should be considered in a state and how a given operator affects those parameters.

In this study, we extend the framework of the Interactive Learning Environment (ILE). In a conventional ILE, a learner interacts with the learning environment. In other words, the learner observes the feedback that the learning environment outputs in response to their own input parameters and exploratively forms a model of the relationship between input and output. However, as Mizoguchi (2024) also points out, it is not easy for learners to construct an appropriate search space. When a learner is unable to form a search space, they may reach an impasse in forming a model of the relationship between input and output. For example, in

ILEs such as ThinkerTools and EBS (White, 1993; Hirashima, 1998), when force is the input parameter and behavior is the output, learners can only directly observe the relationship between force and behavior. In this situation, if they do not sufficiently understand the concepts of acceleration and velocity, which are necessary for model formation, the model cannot be formed appropriately. Therefore, this study extends the ILE framework to require a task where learners observe the chain of causal relationships that exists between the parameters they input and those the learning environment outputs. In this paper, we develop and evaluate a learning support environment that applies this method to diagram-drawing problems in the field of mechanics.

2. Related Work

2.1 Error-Based Simulation

It is important that students learn proactively and constructively; particularly, errors serve as crucial opportunities to reflect on their states of understanding. By visualizing learners' errors, error-based simulation (EBS) provides them with an opportunity to reflect on their comprehension. This error visualization fosters awareness by presenting learners with the consequences that arise from their specific mistakes. "Awareness of an error" is not about a learner acknowledging a mistake after being shown the correct answer by others; rather, it is about the learner noticing for themselves that the result of their answer is strange. In other words, learners understand why and how their responses are incorrect. EBS has been proposed as a method to achieve this error visualization (Hirashima et al., 1998). EBS is a type of ILE aimed at assisting learners understand the relationship between input and output.

An example of EBS is illustrated in Figure 1. Figure 1 depicts a learner answering the problem, "Draw the forces acting on an object at rest on a table." The normal force from the floor—the phenomenon of the object remaining stationary—and gravity acting from the center of the object are visualized when the learner provides the correct answer (Figure 1, top-right). However, if the learner makes a mistake, such as drawing only the force of gravity, the phenomenon of the object sinking through the desk can be visualized (Figure 1, bottom right). The learner can recognize that their prior knowledge, such as the idea that "an inanimate object like a desk does not push back," is incorrect when shown a phenomenon that is vastly different from the stationary state the learner expected. Furthermore, EBS is expected to assist learners in constructing a mental model for converting force into position. This mental model enables one to mentally simulate how a target (parameter) changes due to an applied operation (operator) during a search activity. In the error-exploration activities of physics learning, a well-constructed mental model is crucial. This is because if a mental model is built within the learner, they can use the equation $F = ma$ to test and discover for themselves the cause-and-effect (input–output) relationship, such as "What happens to acceleration if I increase the force?"

However, conventional EBS has a limited ability to foster an understanding of causal series. For example, typical mechanics EBS systems (Imai et al., 2008; Aikawa et al., 2024a) aid in understanding the relationship between force and motion; however, they do not discuss how a difference in force affects the difference in acceleration, how a difference in acceleration affects the difference in velocity, or how this ultimately leads to a difference in behavior. Although some studies have dealt with the relationship between force and acceleration, Yamada et al. (2016) research examined the relationship between force and acceleration and force and velocity; however, it did not address how a difference in force leads to a difference in acceleration or velocity. Aikawa et al. (2024b) developed an EBS to understand the relationship between force and acceleration and between acceleration and velocity; however, this research does not provide a mechanism for learning the causal series of force, acceleration, and velocity.

2.2 Error Visualization Model

The error visualization model is illustrated in Figure 1. In the error visualization model, three essential parameters were identified for effective error visualization. The first is visibility, which represents the degree of difference between the learner's behavior and the correct behavior. The second is reliability, which represents the validity of an EBS based on a learner's answer. The third is suggestiveness, which represents the difference between the correct behavior and the EBS appropriately implies the difference between the correct answer and the learner's answer.

Within visibility, there are specific conditions for creating awareness of an error, known as the criteria for error visualization (CEV). CEV are two conditions that enable EBS to provide learners with this awareness (Hirashima & Horiguchi 2016; Ueno et al. 2019). Crucial to providing this awareness is the existence of a qualitative difference between the behavior generated by the EBS and the correct behavior. This qualitative difference indicates a distinction between properties. In the context of CEV, a qualitative difference refers to a difference in the qualitative values (+, 0, −) between the behavior based on the learner's answer and the correct behavior. There are two such criteria, CEV-1 and CEV-2.

- CEV-1: A condition in which there is a qualitative difference in the velocity of the behavior generated from the learner's incorrect and correct answers.
- CEV-2: A condition in which there is a qualitative difference in the first derivative of velocity (i.e., acceleration) between the behavior generated from the incorrect answer and that of the correct answer.

A learner can become aware of their error if at least one of these CEV is met. Previously, Hirashima and Horiguchi (2016) proposed a method for enhancing visibility by altering the parameters of physical phenomena. Specifically, they adjusted parameters, such as the size of the object or the angle of inclination, to satisfy the CEV.

However, even when the CEV is satisfied, large parameter fluctuations can make the simulation feel artificial to the learner. Therefore, to account for reliability, Hirashima and Horiguchi (2016) worked to satisfy only one CEV without making significant parameter changes. This successfully prevented the simulator from feeling contrived.

Furthermore, to enhance suggestiveness, Hirashima and Horiguchi (2016) have presented EBS that can correctly point out the cause of an error, which is identified by comparing the learner's answer with the correct one and referencing predefined rules about error causes.

These attempts focused on refining visualization to enable learners to engage in error exploration activities. However, Hirashima and Horiguchi (2016) have also pointed out that there is a trade-off between visibility and reliability.

2.3 Research Gaps and Objectives

By demonstrating the difference between a strange and correct simulation, EBS prompts learners to become aware of their errors. However, EBS does not provide a mechanism for learners to become aware of all the parameters that should be fully considered. Therefore, this research introduces an "observation activity" to facilitate error exploration, where learners explore all relevant parameters. In the error exploration activity, learners observed the difference between correct behavior and behavior based on their answers. Subsequently, we have them conduct observations that map their input to the behavior to understand how their answers lead to the generation of strange behaviors. In this activity, the learner considers why and how it is wrong compared to the correct simulation by observing the differences and mapping the correspondence between their answer and behavior. By having learners ponder the meaning of a strange simulation and its differences, they are expected to explore and correct their errors proactively. Subsequently, to observe the differences and map their answers to the behavior, learners searched within a search space composed of the parameters presented by the simulation (e.g., position and velocity).

Even in conventional EBS, it is thought that learners who can engage in appropriate trial and error implicitly perform the type of exploration activity described above. However, it is assumed that some learners engage in random exploration without considering the meaning of the generated simulation and its differences, and make haphazard corrections to match the correct motion. Therefore, this study designed and developed a learning support environment that explicitly presents an error exploration activity based on the learner’s answer as a task.

In this study, we aimed to facilitate and understand learners’ exploratory activities. In this study, we addressed the following research questions:

- RQ1: Can learners understand the forces acting on an object through error exploration?
- RQ2: Is the understanding of the difference between correct behavior and behavior based on one’s answers promoted by supporting the learner’s error exploration?
- RQ3: Is mapping the answer to the behavior promoted by supporting the learner’s error exploration?

In this study, we developed a system that facilitates the observation of differences and correspondences between answers and behavior when a learner provides an incorrect answer to a problem. An evaluation experiment was conducted to answer these questions.

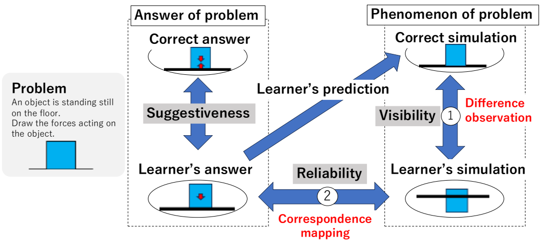

Figure 1. Error visualization model. (Based on the model proposed by Hirashima et al., 1998)

3. Proposed Method

The error exploration activity consists of two main processes: first, observing the "correct simulation" and the "simulation based on the learner’s answer" (observation of differences, Figure 2(1)), and second, linking the "learner’s answer" to the "simulation based on the learner’s answer" (correspondence between the answer and the behavior, Figure 2(2)). The learner engages in the error exploration activity through a cycle of observing differences and then mapping the correspondence between their answers and behavior. Specifically, we frame the observation and correspondence processes as a task by having the learners break down and consider the parameters of their answers stepwise. By having the learner observe differences and map the correspondence between their answers and behavior in this way, it is expected that they will understand the strangeness of their simulation, connect their answers to it, and comprehend their errors. The error-exploration activity focused on enabling learners to discover parameters within the search space heuristically during problem-solving.

To observe differences in error exploration activity, we used the measurement tool from Ueno et al. (2019). A measurement tool is a framework that adds visual information by visualizing measurements of values such as Force, Acceleration, and Velocity using metaphors such as meters or gauges, thereby highlighting qualitative changes (in line with CEV-1). We framed the observation of differences as a task by having the learner use this measurement tool in the simulation.

Figure 2 illustrates the learning activity for learners who cannot perform the error-exploration activity independently. First, the learners were instructed to observe the presented simulations for differences. After this initial observation, they considered where the differences might exist. Finally, they conducted a focused observation to confirm whether a difference existed. By guiding the learner through this phased process of observing and then considering the differences, we facilitated the correspondence between their answers and behavior.

What is crucial in mapping the correspondence between the answer and behavior is enabling learners to sequentially examine the parameters from the simulation back to the solution. This is because the simulation visualized the motion as illustrated in Figure 3.

Specifically, the simulation is the result of a causal chain: the applied force determines the acceleration, which in turn determines the velocity, which, in turn, determines the change in position over time. Because there are many parameters between the learner's answer (force) and the simulation's behavior (change in position), it is vital to understand them step-by-step to map the answer to the behavior properly.

To help learners examine these parameters step-by-step, it is necessary to provide an explorable environment and navigate through them as needed. However, because of the wide variety of possible incorrect answers and problems, the specific parameters to be examined can vary significantly. Therefore, preparing a predefined sequence for each pattern is difficult. To solve this problem, we use a Constraint Structure (Hirashima et al., 1995), as illustrated in Figure 4. A constraint structure represents the background of a problem as a network of quantitative relationships. The order in which the parameters were examined was determined based on the structure. Specifically, it includes equations for Position, Velocity, and Acceleration, as well as equations for calculating the total combined force and partial combined force from the individual force components drawn by the learner.

Figure 4 depicts a concrete example of the step-by-step examination of these parameters. The problem involves drawing the forces acting on an object in uniform motion on a frictionless plane. The qualitative values (+, 0, −) of each parameter are determined by the learner's answer and the correct answer. Once the values are determined, a difference appears between the learner's parameters and the correct parameters. Specifically, there were changes in rightward Force, Acceleration, Velocity, and Position. Thus, a step-by-step order was determined to examine the parameters backward, from the change in position to the resulting velocity, to its cause in acceleration, and to the root error in the applied force.

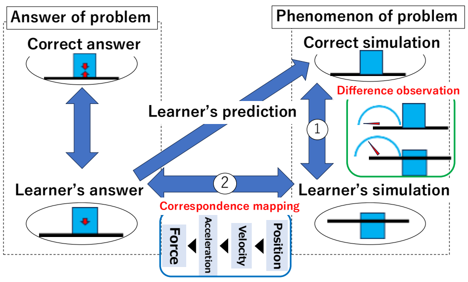

Figure 2. Difference observation and correspondence mapping. (Based on the model proposed by Hirashima et al., 1998)

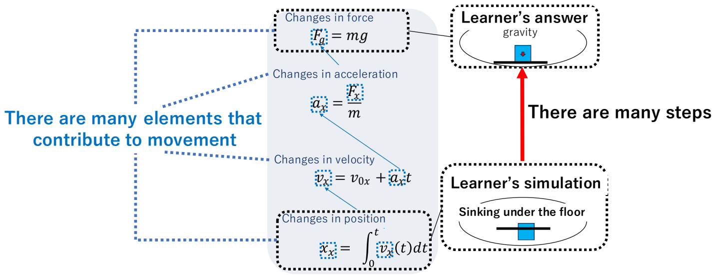

Figure 3. Various parameters in the movement.

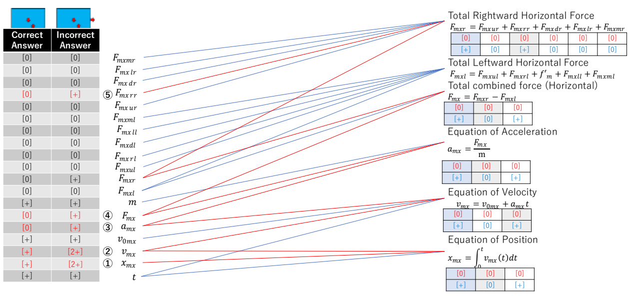

Figure 4. Sequence of parameter exploration guided by the learner's answer.

4. Proposed System

This section describes the workflow of our proposed system. Figure 5(a) depicts a diagram of the problem. When given the problem, "Draw the forces acting on two objects stacked vertically," the learner draws a Force diagram for the upper and lower objects based on the problem statement. In this example, the learner's answer is incorrect, showing only gravity and the normal force for both objects. Figure 5(b) depicts the simulation confirmation screen. On this screen, the learner confirms the correct simulation generated based on their answer. In the example illustrated in Figure 5(b), the objects are stationary in both simulations. The process up to this point in Figures 5(a) and (b) is the same as the tasks a learner performs in a conventional EBS.

Figures 5(c) and (d) are the screens where the learner engages in error exploration. In the incorrect answer in Figure 5(a), the force of the top object pushing on the bottom is missing, and the corresponding normal force from the bottom object is insufficient. To address these errors, the learner is prompted to use a measurement tool capable of measuring parameters such as Velocity, Acceleration, Total combined force, and partial combined force, to identify the source of the error. In Figure 5(c), the learners are asked if they think there is a difference between the correct simulation and their own. The question was, "Do you think there is a difference in the Partial combined force on the blue object in the vertical direction? " Answering this question transitioned the user to the screen, as illustrated in Figure 5(d). As illustrated in Figure 5(d), the learner was asked to select the measurement tool to investigate the differences they hypothesized. The learner, suspecting a difference in the "Partial combined force on each object," chooses to measure it using the "Partial Combined Force Meter." By selecting "Place," the measurement tool is placed on the simulation.

When a learner's answer contains multiple errors to explore, they may not be able to remember all the errors they have found, making it difficult to apply their findings when correcting the answer as part of the problem-solving process. Therefore, we implemented an Error Exploration Notebook (Figure 6) that summarized the areas explored by the learner. The notebook displayed an image of the force diagram, an image of the simulation generated from the learner's answer, and a record of whether a difference was found during error exploration. This allows learners to review the differences in their answers, reflect on their findings, and think about where to explore next. Furthermore, by referencing notebooks when correcting answers, modifications can be made based on exploration activities.

(a) Input *(b)* Observe simulations *(c)* Difference Selection *(d)* Measurement Tool Selection

Figure 5. System screens.

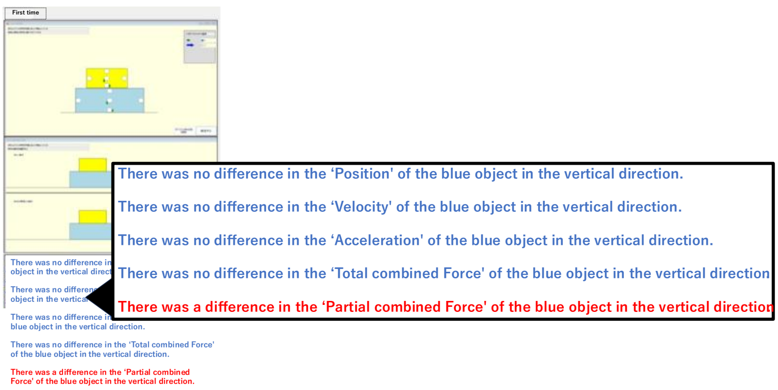

Figure 6. Error exploration notebook.

5. Evaluation

In this study, we established the following three hypotheses corresponding to RQ1–RQ3:

- H1: Learners can accurately draw diagrams of forces acting on an object.
- H2: After observing differences in motion, learners will be able to identify the differences in each element.
- H3: After observing the differences in motion, learners become aware of the causal relationships between position, velocity, acceleration, and force.

To evaluate the learning effectiveness of the proposed system and verify H1–H3, we conducted a preliminary evaluation experiment. The participants were 13 university students from the Faculty of Information Science who had previously studied mechanics. They were randomly assigned to an experimental group (n = 7) and a control group (n = 6). The experimental group used the developed system, whereas the control group used conventional EBS. The experiment consisted of a pre-test, a session using the system, and a post-test. The tests addressed problems related to mechanics.

Participants were given drawing and presumptive thinking tests. The Drawing Test measured the learning effect on mechanics diagram drawing problems and was conducted to verify H1(Aikawa 2024a). The Presumptive Thinking Test was conducted to measure a learner's ability for causal inference. This test is based on the principle that genuine causal inference requires more than simply noticing covariation between two events; it requires an understanding of the underlying causal mechanism that explains the "why" and "how" of the relationship (Klahr et al., 2019). Accordingly, the test evaluates a learner's ability to compare different conditions to identify what variable has changed and to explain the mechanism by which that change affects the outcome. This test was conducted to verify H2 and H3.

The Drawing Test consisted of five problems that required drawing the forces acting on an object, with a maximum score of five points (one point per question). This confirms that learners can draw forces accurately. The Presumptive Thinking Test consists of three problem-solving tasks, where, as illustrated in Figure 7, learners demonstrate a certain phenomenon and are asked to describe the differences and the reasons for those differences based on cause and effect. This confirms whether learners can observe the differences and map the correspondence between an answer and its behavior.

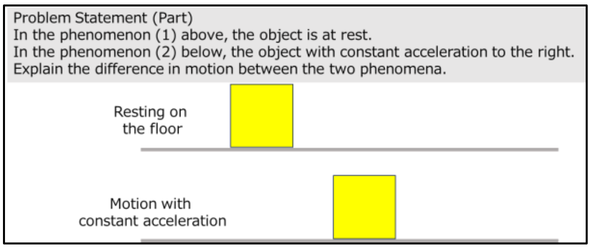

Figure 7. The presumptive thinking test.

5.1 Drawing Test

To verify H1, Table 1 shows the pre- and post-test results of the Drawing Test. When Cohen's d was calculated for the effect size, the experimental group had a medium effect size, whereas the control group had almost no effect. Furthermore, Analysis of Variance (ANOVA) showed no interaction effect, no significant difference between the groups, and no significant difference ($p < .05$) between the pre- and post-tests.

This can be attributed to the ceiling effect in the drawing tests. Looking at Table 1, the average pre-test scores for both groups were 3.8 or higher out of a maximum of 5 points, suggesting that the participants already had a high level of proficiency before the intervention. The learning effectiveness of conventional EBS has been recognized in previous studies. However, because the average pre-test score was 3.8 or higher, the measurement was conducted with very little room for improvement, which is presumed to be why a statistically significant difference was unlikely to appear. Thus, H1 is not supported.

These findings address RQ1, which sought to determine if learners can understand the forces acting on an object through error exploration. As the analysis demonstrated, the non-significant results of the Drawing Test (meaning H1 was not supported) were attributed to a ceiling effect. Therefore, due to this methodological limitation, the current experiment could not provide a conclusive answer to RQ1.

Table 1. *Pre/Post results and effect size for the drawing test*

	Pre-test	Post-test	Effect size
	Medium (SD)	Medium (SD)	
Experimental Group	3.86 (0.99)	4.43 (1.05)	0.56 (medium)
Control Group	3.83 (1.21)	4.00 (0.82)	0.16 (very small)

5.2 Presumptive Thinking Test

To verify H2, Table 2 shows the results for the mention rate of differences between the pre- and post-tests of the Presumptive Thinking Test. One point was awarded for mentioning each of the following: position, velocity, acceleration, total net force, partial net force, and force, thus evaluating whether or not learners could identify the appropriate parameters. The maximum possible score was six points.

In the experimental group, the effect size was large for the position and acceleration, medium for the velocity and total net force, small for the force, and no effect for the partial net force. For the control group, the effect size was medium for force, whereas there was almost no effect on velocity and acceleration, and no effect on position, total net force, or partial net force. An Analysis of Variance (ANOVA) for each parameter showed a significant trend for the interaction effect of the acceleration parameter. Furthermore, a test of simple main effects revealed a significant trend between the pre- and post-tests for the experimental group and between the experimental and control groups at the post-test stage. Additionally, for the force parameter, there was a significant difference ($p < .05$) between the pre- and post-tests.

We will now discuss the reasons for the difference in effect sizes for the position and acceleration parameters and the reason for the significant difference in the force parameter between the experimental and control groups. Unlike the conventional EBS system of the control group, the error exploration system of the experimental group allowed learners to explore differences in motion parameters. Therefore, the effect size of the parameter mentioned rate increased more in the experimental group than it did in the control group.

RQ2 asked whether supporting learners' error exploration promotes their understanding of the difference in behavior between the correct simulation and the one based on their answer. The results of the Presumptive Thinking Test address this question, showing that the experimental group had a significantly higher rate of identifying differences in parameters like position and acceleration compared to the control group. This finding supports H2. Therefore, it is suggested that the proposed system effectively promoted learners' understanding of the differences in motion, thus providing an affirmative answer to RQ2.

To verify H3, Table 3 shows the results for the mention rate of causal relationships in the pre-test and post-test of the Presumptive Thinking Test. One point was awarded if the learner mentioned a cause-and-effect relationship for a difference in the behavior of the presented phenomenon. The maximum possible score was six points.

For the experimental group, the effect size was medium for the relationships of velocity →position, acceleration→position, and force→position. In the control group, the effect size was small for acceleration and force velocities.

We discuss the reason for the difference in the mention rate of causal relationships between the experimental and control groups. Unlike the conventional EBS system of the control group, the error exploration system of the experimental group prompted learners to consider the relationship between force and motion stepwise, starting from their answers. Therefore, it was hypothesized that learners would be able to construct mental models of the phenomenon.

RQ3 asked whether supporting learners' error exploration promotes the mapping between their answer and the resulting behavior. The findings revealed that the experimental group showed a greater tendency than the control group to mention the causal relationships between position, velocity, acceleration, and force, a result that supports H3. This suggests that by being supported in their error exploration, learners were better able to understand how their answer (the force they input) leads to the resulting behavior (the motion) through a causal chain. Therefore, this process promoted the mapping between their answer and the behavior, addressing RQ3 affirmatively.

Table 2. *Results of statements for the difference between pre-/post-movements in the presumptive thinking test*

Movement difference	Pre-test	Post-test	Effect size
	Experimental Group		
Position	0.00 (0.00)	0.44 (0.50)	1.27 (large)
Velocity	0.11 (0.31)	0.33 (0.47)	0.56 (medium)
Acceleration	0.11 (0.31)	0.44 (0.50)	0.80 (large)
Total combined force	0.00 (0.00)	0.11 (0.31)	0.50 (medium)
Partial combined force	0.00 (0.00)	0.00 (0.00)	-
Force	0.33 (0.47)	0.22 (0.42)	0.25 (small)
	Control Group		
Position	0.00 (0.00)	0.00 (0.00)	-
Velocity	0.33 (0.47)	0.40 (0.49)	0.14 (very small)
Acceleration	0.20 (0.40)	0.13 (0.34)	0.18 (very small)
Total combined force	0.00 (0.00)	0.00 (0.00)	-
Partial combined force	0.00 (0.00)	0.00 (0.00)	-
Force	0.27 (0.44)	0.07 (0.56)	0.56 (medium)

Table 3. *Results of statements for the causal relationship between pre-/post-movements in the presumptive thinking test*

Causal relationship	Pre-test	Post-test	Effect size
	Experimental Group		
Velocity → Position	0.00 (0.00)	0.19 (0.39)	0.69 (medium)
Acceleration → Position	0.00 (0.00)	0.14 (0.35)	0.58 (medium)
Acceleration → Velocity	0.00 (0.00)	0.10 (0.29)	0.46 (small)
Force → Position	0.00 (0.00)	0.19 (0.39)	0.69 (medium)
Force → Velocity	0.57 (0.49)	0.57 (0.49)	0.00 (very small)
Force → Acceleration	0.10 (0.29)	0.24 (0.43)	0.39 (small)
	Control Group		
Velocity → Position	0.00 (0.00)	0.00 (0.00)	-
Acceleration → Position	0.00 (0.00)	0.00 (0.00)	-
Acceleration → Velocity	0.17 (0.37)	0.11 (0.31)	0.24 (small)
Force → Position	0.06 (0.23)	0.06 (0.23)	0.00 (very small)
Force → Velocity	0.39 (0.49)	0.56 (0.50)	0.32 (small)
Force → Acceleration	0.11 (0.31)	0.11 (0.31)	0.00 (very small)

6. Conclusion

Because problem-solving in learning occurs in an unmastered domain, arriving at the correct answer on the first attempt is difficult. In such cases, it is important to understand (a) the definition of the parameters defining a given state and (b) the state transitions caused by applying an operator to move from the problem's initial state to the final state. Therefore, this study extends the ILE framework to help learners constructively learn this search space. Specifically, we propose a method that, when learners err, helps them understand which parameters their input affected and the resulting behavior. This method requires learners to engage in error exploration. Through trial and error, they must observe the chain of influence between parameters that explains the causal relationship between their answer and the resulting behavior. We developed and evaluated a learning support system applying this method to mechanics diagram drawing problems.

The experimental results indicate that learners using the developed system focused on parameters related to the differences between motion and force. Furthermore, their statements linking a phenomenon's behavior to forces suggest that they mapped the phenomenon to the behavior.

For future work, we will consider representing the search space as a map, using the learner's error exploration activities as a starting point, to allow learners to experience the scientific hypothesis-testing cycle. Visualizing the search space as a map is expected to help learners plan their exploration strategies.

Acknowledgements

This work was supported by JSPS KAKENHI, Grant Numbers JP25K21361 and JP25K21362.

References

Mizoguchi, R. (2024). Wakaranai joutai no moderu-ka [Modeling of "not understanding" state]. Journal of the Japanese Society for Artificial Intelligence, *39*(2), 103-110.

White, B. Y. (1993). ThinkerTools: Causal Models, Conceptual Change, and Science Education. *Cognition and Instruction*, *10*(1), 1–100.

Hirashima, T., Horiguchi, T., Kashihara, A., & Toyoda, J. (1998). Error-based simulation for error-visualization and its management. *International Journal of Artificial Intelligence in Education, 9,* 17–31.

Imai, I., Tomoto, T., Horiguchi, T., & Hirashima, T. (2008). Tyugaku-rika ni okeru wo motita jugyo-zissen: "Newton ni tyosen" project [A classroom practice of error-based simulation to improve pupils' understanding of mechanics: The "Challenge to Newton!" project]. *Transactions of Japanese Society for Information and Systems in Education, 25*(2), 194–203.

Aikawa, N., Maeda, S., Mogi, T., Koike, K., Tomoto, T., Imai, I., Horiguchi, T., & Hirashima, T. (2024a). Classroom implementation of an auxiliary problem presentation system for mechanics adapted to learners' errors. *Research and Practice in Technology Enhanced Learning, 19,* Article 26.

Yamada, A., Shinohara, T., Horiguchi, T., Hayashi, Y., & Hirashima, T. (2016). Tashiten Error-Based Simulation no sekkei - kaihatsu to jikkenteki-hyouka [Design and Development of Multi Viewpoints Error-Based Simulation and Experimental Evaluation]. *Transactions* of The Institute of Electronics, Information and Communication Engineers D, J99-D (12), 1158–1161.

Aikawa, N., Hiramoto, C., Tomoto, T., Horiguchi, T., & Hirashima, T. (2024b). Learning Support System for Relationships Among Forces, Accelerations, Velocities, and Movements in Error-Based Simulation. *HCI International 2024: Human-Computer Interaction* (Lecture Notes in Computer Science, vol. 14691, pp.3–15). Springer.

Hirashima, T., & Horiguchi, T. (2016). Creation of cognitive conflict by error-visualization: Error-based simulation and its practical use in science learning. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, Design, and Technology* (pp. 1-33). Springer.

Ueno, U., Tomoto, T., Horiguchi, T., & Hirashima, T. (2019). A support system for learning physics in which students identify errors using measurements displayed by a measurement tool. *Workshop Proceedings of the International Conference on Computers in Education (ICCE 2019)*, 426–434.

Hirashima, T., Azuma, S., Kashihara, A., & Toyoda, J. (1995). A formulation of auxiliary problems and its evaluation. *Proceedings of the International Conference on Artificial Intelligence in Education (AIED)*, 186–193.

Klahr, D., Zimmerman, C., & Matlen, B. J. (2019). Improving students' scientific thinking. In J. Dunlosky & K. A. Rawson (Eds.), The Cambridge handbook of cognition and education (pp. 67–99). Cambridge University Press.