Error-based simulation as a thought experimental method for changing “motion implies a force” misconception: An evaluation

Horiguchi, Tomoya
Hirashima, Tsukasa
Hayashi, Yusuke

(Citation)
Journal of Computer Assisted Learning,39(4):1290-1302

(Issue Date)
2023-08

(Resource Type)
journal article

(Version)
Accepted Manuscript

Error-Based Simulation as a Thought Experimental Method for Changing “Motion Implies a Force” Misconception: An Evaluation

Abstract

Background

In learning mechanics, students often believe that “force is exerted on moving objects.” As this misconception called “motion implies a force” (MIF) is difficult to correct, various teaching methods have been proposed, such as showing refutational/explanatory text (Palmer & Flanagan, 1997; Takagaki, 2004), promoting students’ conceptual development (Dekkers & Thijs, 1998) and epistemology-focused instruction (Elby, 2001).

Objectives

The purpose of this research is to empirically examine the effectiveness of an alternative method we have proposed for helping students overcome MIF misconceptions. The method is called the “Error-based simulation (EBS)” that creates a counterexample to students' misconceptions by simulating erroneous phenomena based on students’ incorrect ideas.

Methods

We conducted a two-year experimental implementation of EBS in mechanics classes at a technical college, and investigated the results. Students worked on the pre-test, the exercise with EBS, the post-test, and the delayed post-test.

Results and Conclusions

In both years, MIF-answers (wrong answers due to MIF) in the learning task (problems learned with EBS) significantly decreased from the pre-test to the post-test and the delayed post-test. The decrease in other erroneous answers was not significant. EBS was effective mainly in correcting MIF-answers, and the effect was maintained over the long term. In the transfer task (new problems), the ratio of MIF-answers to all erroneous answers in both the post-test and delayed post-test was much lower than that on the learning task in the pre-test. These results suggest that learning with EBS not only corrected MIF-answers but also promoted the correction of MIF misconceptions. A comparison between our results and those of other methods for correcting MIF misconceptions revealed that the effect of EBS was equal to or greater than that of other methods mentioned above.

Keywords. Motion Implies a Force (MIF) misconception, error-based simulation (EBS), conceptual change, cognitive conflict, counterexample, practical use evaluation

INTRODUCTION

In elementary science education, students must acquire the ability to explain and predict natural phenomena using scientific concepts. However, students often understand natural phenomena with scientifically incorrect concepts called misconceptions. In particular, students who learn physics often have misconceptions even after they are taught the correct scientific concepts (Driver et al., 1985; McCloskey, 1983; Osborne & Freyberg, 1985). Such misconceptions are usually difficult to overcome because they are deeply rooted in the experiences of students' daily lives (Bransford et al., 2000; Mestre, 1994).

The Motion Implies a Force (MIF) misconception is one of such misconceptions in which students believe that "force is exerted on moving objects." Scientifically, a force changes the motion of objects. That is, an object moving at a constant velocity needs no force to keep its motion, while an object at rest needs force to start moving (or a moving object needs force to speed up, slow down or change its direction of motion). Therefore, the "force" in the MIF misconception is not scientific "force," but a preconception undifferentiated from motion and energy. For example, while an object is moving in a straight line on smooth floor at a constant velocity, many students believe that continuing force is being exerted on the object in the direction of motion to keep it moving. In another example, when an object is thrown upward by hand, students often believe that the object is "imprinted" with upward force, and it ascends against gravity until the force is consumed. Students tend to rigidly maintain such MIF misconceptions even after they are taught the scientific concept of force (Caramazza et al., 1981; Clement, 1982; McCloskey, 1983).

Several teaching methods have been proposed to change students' MIF misconceptions. Most of them follow the classical model of conceptual change proposed by Posner et al. (1982), and Strike and Posner (1985) as their theoretical basis. Some methods (Palmer & Flanagan, 1997; Takagaki, 2004) focus on evoking and resolving "cognitive conflict" according to the model by Hashweh (1986), in which Posner's (1985) model was made more "procedural." Some methods (Dekkers & Thijs, 1998; Elby, 2001) are "epistemological" methods that aim to integrate students' intuition with scientific concepts, focusing on their epistemological development. These methods usually first present a phenomenon that cannot be explained with students' misconceptions, and then introduce a scientific concept that can explain the phenomenon. Here, it is crucial that the scientific concept works as a "counterexample" to the students' misconceptions. Therefore, after students observe and measure the phenomenon sufficiently, a teacher explains that the scientific concept is consistent (and students' misconceptions are inconsistent) with the phenomenon. Experiments and computer simulation (with measuring instruments) are often used to assist students' observations (Towne et al., 1993; Wenger, 1987).

However, a common problem with these methods is that their counterexample is usually "the (verbal) explanation of the correct concept" (Palmer & Flanagan, 1997; Takagaki, 2004; Dekkers & Thijs, 1998; Elby, 2001, for example). For example, consider the problem in which an object on a smooth floor is pushed by a hand for an instant and keeps moving at a constant velocity, and suppose that students think force is exerted on the object in the direction of motion. Suppose a teacher says to students that "no force is exerted on the object after the hand parts from it" and that "force does not cause the objects' velocity but their acceleration (i.e., Newton's second law)." This explanation would hardly work as a counterexample unless students could embody the explained concept and connect it to the phenomenon. In this case, presenting the phenomenon using an experiment or simulation is insufficient because students know how the object moves (i.e., uniform motion). Even if the measurement reveals that there is no force in the direction of motion, it is difficult to connect the fact with the abstract concept. In order for the teacher's explanation to work, it is necessary to make a "thought experiment" from which one can infer that "according to Newton's second law, if any force was exerted on the object, it would not have moved at a constant velocity because it would be accelerated by the force."

Error-based simulation (EBS) is an alternative method for presenting a counterexample to students' misconceptions by making a thought experiment. In EBS, the model of the target phenomenon is made based on students' misconceptions; then, the simulation is made based on the model to display an erroneous phenomenon. When observing such an erroneous phenomenon, students who know the correct phenomenon but understand it using the wrong concepts tend to realize that their understanding is incorrect (Hirashima et al., 1998). In the above example, if students believe that a force is being exerted on the object in the direction of motion, EBS shows the object's accelerated motion instead of uniform motion. (e.g., Students are shown a figure of the situation on screen and required to draw the force(s) that they think is(are) being exerted on the object as an arrow with touch operation. When they draw a force in the direction of the object's motion, the acceleration is calculated by substituting the force into the formula of Newton's second law and the object's (accelerated) motion is simulated which is shown to the student.) Such EBS functions as a counterexample to the students' erroneous prediction. In EBS, although the correct concept (i.e., Newton's second law) is used to calculate the behavior of the object (by substituting the incorrect force in the formula), the concept is not directly explained to students. They are expected to discover the correct idea (i.e., no force is exerted on the object in the direction of motion) by themselves through trial and error. In our previous practice in junior high schools, we taught the "normal reaction" in static situations using EBS (Hirashima et al., 2009; Horiguchi et al., 2014). The understanding of students who learned the concept through EBS was considerably superior to that of the students who learned it in the usual way. Additionally, it was suggested that students who saw counterexamples by EBS reformed the idea rooted in their daily experiences and integrated it with scientific concepts.

In this paper, we propose an alternative method for helping students overcome MIF misconceptions by using EBS, and report the results of experiments planned to verify the effect of our method[1].

ERROR-BASED SIMULATION: PROVIDING COUNTEREXAMPLES TO STUDENTS' MISCONCEPTIONS

In this section, we first give a brief review of the literature on MIF misconceptions. Second, we review related work on teaching methods used to correct students' misconceptions. Third, we discuss the problem of these methods. Finally, we introduce the framework of EBS as an alternative method, and compare its features with those of other teaching methods.

"Motion Implies a Force" Misconceptions

Students' intuitive preconceptions and their origin have been investigated in many studies (e.g., Driver et al., 1985; McCloskey, 1983; Osborne & Freyberg, 1985; Caramazza et al., 1981; Clement, 1982; McCloskey, 1983). As for MIF misconceptions, a common explanation is such as: "In the real world, where friction is present, one must push on an object to keep it moving. Since friction is often not recognized as a force by the beginner, the student may believe that continuing motion implies the presence of a continuing force in the same direction, as a necessary cause of the motion (Clement, 1982)." It is, however, not likely that such misconceptions have been inductively derived from students' daily experience because the details of students' response differ depending on the context. For example, some students thought that a "force" is acquired by a pushed object but not by a carried object (McCloskey, 1983). Additionally, students had a variety of idea about how an acquired "force" interacts with gravity (McCloskey, 1983). Therefore, it is more likely that the misconceptions concern subjective models of motion that reflect students' conceptual system used to organize their experience and to deduce prediction with moving objects (Caramazza et al., 1981).

It is often pointed out that MIF misconceptions are similar to pre-Newtonian "impetus theory," which asserts that an object set in motion acquires an "internal force" or "impetus" that maintains the motion, and that a moving object's impetus gradually dissipates (either spontaneously or as a result of external influences), until it slows down and comes to a stop (McCloskey, 1983).

Although the "force" in MIF misconceptions is inappropriate from Newtonian mechanics view, rather than conflicting with scientific force or inertia, it appears to be a concept in which naive intuition about "force" and "(inertial) mass" (which are undifferentiated) is mixed with scientific knowledge about force, mass and momentum students learned in physics class. According to their naive theory, a state of rest (absence of "force") and a state of motion (presence of "force") are qualitatively distinguished (which is inconsistent with Newtonian mechanics). This distinction might come from the naive sense of "passive" and "active."

Related Work on Conceptual Change

Most of the teaching methods for correcting students' misconceptions proposed so far have been based on the classical model of conceptual change proposed by Posner et al. (1982) and Strike and Posner (1985). This model presents that misconceptions change with a "paradigm shift" (Kuhn, 1962) instead of presenting an empirical view that misconceptions change when they contradict observational evidence. That is, misconceptions and scientific concepts are regarded as competitors in "conceptual ecology" (Toulmin, 1972) which includes exemplars, anomalies, analogies and metaphors, epistemological beliefs, and knowledge of other fields. Posner's model (Posner et al., 1982; Strike and Posner, 1985) provides the following pedagogical principles for making conceptual change: (1) making students understand the scientific concept, (2) clarifying that the scientific concept conflicts with misconceptions, and (3) showing that the scientific concept is more plausible and fruitful than misconceptions. This principle was made more "procedural" by Hashweh (1986) in his model, focusing on evoking and resolving "cognitive conflict." For example, some methods were proposed and practiced based on Hashweh's model (Palmer & Flanagan, 1997; Takagaki, 2004). Thereafter, through criticism by other researchers (e.g., Duit, 2003) and modifications made by the authors (Strike & Posner, 1992), Posner's model has evolved more by considering social contexts, aiming to integrate students' intuitions with scientific concepts focusing on their development, and avoiding excessive emphasis on cognitive conflict (Vosniadou & Ioannides, 1998). In some methods, for example, learning context in class was considered, and descriptions of concepts were carefully elaborated on (Dekkers & Thijs, 1998; Elby, 2001). In this section, we outline some of the previous methods that attempted to change students' MIF misconceptions.

In Palmer and Flanagan's (1997) study, students of primary and secondary schools who were judged to have MIF misconceptions by using a pre-test (a problem that was the same as Problem (3) used in our tests), read a "refutational text" in which the correct concepts of force were presented as other students' opinions such as "there is no upward force because the hand is not pushing the ball anymore." After a class discussion based on the text, they worked on a post-test and delayed post-test (each of which included the same problem as the pre-test, and other two problems such as a ball thrown vertically downward and a ball in trajectory motion). They found that, 43% to 57% of them corrected their MIF misconceptions in the post-test, and 35% to 44% of them maintained the correction after one–two weeks in the delayed post-test.

In Takagaki's (2004) study, university students who were judged to have MIF misconceptions by using a pre-test (two problems that were the same as Problem (3) and Problem (6) used in our tests, respectively), read a text designed to trigger and resolve cognitive conflict. In the text, correct ideas of force such as "the direction of motion is not always the same as that of exerted force" and "gravity decelerates the ball moving upward" were explained. They also watched a video clip of the correct motion and the correct force exerted on the object. In a post-test (which was the same as the pre-test), 70.3% of them correctly answered the problem of the ball thrown upward and 43.6% correctly answered the problem of the ball thrown diagonally. Thus, they had corrected their MIF misconceptions.

Dekkers and Thijs' (1998) study was based on the results of an experiment that revealed that university students' concept of force was not a scientific concept but a preconception undifferentiated from motion and energy. They designed a teaching method for promoting students' conceptual development. First, students learned about the interaction between forces using different types of examples. Then, by introducing conventions for drawing force vectors, they were taught scientifically correct and sharable concepts of force. Finally, they discussed the problems of a trolley moving on a floor and a uniformly moving ball in class. As a result, while the average correct score was 15% in the pre-test (including 11 test items on identifying forces exerted on moving objects), it was 57% in the post-test (including 7 test items similar to those in the pre-test). Two months later, in the delayed post-test (similar to the pre-test), the average correct score increased to 79%.

Elby (2001) designed "epistemology-focused instruction" in which students' conceptual development was integrated with their epistemological development. For example, regarding the forces exerted on a car running at a constant velocity, most students tend to believe that the force exerted on the car in the direction of motion is greater than the force exerted by air resistance in the reverse direction. A teacher then assists students in thinking about the acceleration of the car and to connect their ideas to Newton's second law. If the student's understanding was correct, the car would have a positive net force; therefore, it would not move at a constant velocity because its acceleration must be positive according to the formula "F = ma." In this case, the students' intuition contradicts scientific concepts. The teacher then introduces another situation of a dolly on the floor that is pushed and starts to move. In this case, most students believe the dolly accelerates because the pushing force is greater than the friction from the floor, and their intuition does not contradict Newton's second law. Through such instruction, students are able to consider when their intuition is correct/wrong and consistent/inconsistent with scientific concepts and they can thus integrate their intuition with scientific concepts.

Problem of Current Teaching Methods

The commonality among the teaching methods described above is that they involve the integration of students' intuitions based on their daily lives with scientific concepts, and not merely denying their misconceptions. In particular, providing counterexamples of these misconceptions plays an important role. Counterexamples can trigger "cognitive conflict," which can cause a conceptual change in a student (Gagne, 1985; Glynn et al., 1991; Osborne & Freyberg, 1985). Additionally, fewer students who learn with counterexamples have recurrent misconceptions than those who learn through the explanation of correct concepts (Bransford et al., 2000).

However, counterexamples should be carefully designed because students often ignore them or need help to comprehend them to reach the correct understanding (Chinn & Brewer, 1993). To increase their effectiveness, counterexamples should be considered important by students, and some help should be given to lead students to the correct understanding (Chinn & Brewer, 1993).

The teaching methods described above have been carefully designed to integrate students' misconceptions with scientific concepts. By using a set of situations in which a phenomenon can or cannot be explained with students' (erroneous) ideas, their misconceptions are confronted with scientific concepts. However, there is a common problem in using counterexamples; that is, the counterexamples provided by these teaching methods are merely "the (verbal) explanation of the correct concept." As mentioned above, students often think that a force is exerted on an object moving on a smooth floor at a constant velocity. The explanation that "no force is exerted on the object after the hand (that pushed the object for an instant to start motion) parts from it" and "force doesn't cause objects' velocity but cause their acceleration" (in Takagaki's (2004) study, a diagram is shown in which the force exerted on the object at each instant is depicted) would not work as a counterexample unless the students embody the explained concept and connect it to the phenomenon. For this explanation to work, a "thought experiment" needs to be done by students from which they can infer that "according to Newton's second law, if any force was exerted on the object, it could not move at a constant velocity because it would be accelerated by the force." However, a "thought experiment" is difficult for students to do because it would cost much cognitive load. In such a case, technology-assisted learning could help.

Framework of EBS

To solve the problem discussed above, we introduce the framework of Error-based simulation (EBS) that provides technology-based assistance. Figure 1 shows the framework of EBS, in which a simulation is generated by mapping errors in symbolic expressions to erroneous behavior. Such a simulation is called EBS. The difference between EBS and the predicted correct behavior makes students aware of errors and motivates them to correct their errors. If students have misconceptions that are expressed through wrong answers, the erroneous behavior works as a counterexample to their misconception.

[Place Figure 1 here]

We present an example of EBS using the mechanics problem in Figure 2. Students are shown an illustration of a mechanical situation and are required to draw all the forces applied to the objects. Owing to their misconceptions, some students may mention erroneous forces, which is regarded as an externalization of their erroneous ideas. Based on students' drawings, the acceleration of each object is calculated, and the motion is simulated. In the problem shown in Figure 2 (a skater traveling straight on frictionless ice at a constant velocity), for example, students tend to draw a force exerted in the direction of travel. In the EBS, the skater is then shown to accelerate in that direction while traveling at a constant velocity in the correct motion. Such an unnatural phenomenon is expected to be useful as a counterexample to students' erroneous ideas. On observing this, the students are able to correct their errors with high intrinsic motivation. Even when students predict the correct phenomenon but explain it with incorrect concepts, EBS can provide a useful counterexample. For example, when students believe that there is no force except gravity being exerted on a book placed on a table, the book sinks into the table. Similarly, when they believe that a skater's net force is positive, she accelerates instead of maintaining a constant velocity.

Note that the students' answers should express their incorrect idea, and include sufficient information to create a model of a given situation to generate the simulation. Additionally, because the simulator for EBS (called a "robust simulator") explicitly handles the constraints of the model, it understands what constraints are violated (Horiguchi & Hirashima, 2006; Horiguchi et al., 2012). The constraints "two solid objects never overlap" and "the skater maintains a constant velocity" are violated in the above examples. Therefore, it becomes possible to define the criteria for estimating the effect of EBS (e.g., "impact") as a counterexample. We have developed simulators that generate EBS in elementary mechanics and other domains and constructed learning environments in which EBSs are managed from several educational viewpoints (Hirashima et al., 1998; Horiguchi & Hirashima, 2006; Horiguchi et al., 2012; Kunichika et al., 2006; Matsuda et al., 2003).

[Place Figure 2 here]

The EBS framework proposes an alternative method of providing counterexamples based on "simulation and interactivity." In the above example in Figure 2, the EBS simulates and shows the accelerated motion of the object (i.e., the skater) instead of uniform motion. Although the simulation is generated using the correct concept (i.e., the erroneous forward force is substituted into Newton's second law to determine the object's acceleration), the concept itself is not explained. Students are expected to reach the correct idea (i.e., no forward force is exerted on the object) by themselves through trial and error, which makes the simulation show the predicted motion (i.e., uniform motion). That is, while the methods so far provide an explanation of the correct concept, EBS helps students discover the correct concept by themselves. As discovery learning might not be easy for students, the EBS framework usually provides two types of assistance. First, a set of situations are prepared in which a simpler one (in which a misconception is easily isolated and identified) progressively becomes more complex (in which multiple concepts interact with each other). Second, immediate feedback is provided for every wrong answer due to a misconception in the form of the object's erroneous behavior (i.e., the result of the thought experiment is visualized by EBS). Such technology-assisted learning framework would present a promising method for changing students' misconception.

Purpose of Research

In this study, we investigate the usefulness of EBS in correcting students' MIF misconceptions. As we discussed in the previous section, EBS aims to make students discover the correct concept by themselves based on "simulation and interactivity." However, in general, discovery learning might not be easy for students. There are cases in which direct instruction is more effective (Klahr, 2009). Additionally, since EBS often shows phenomena that are impossible in the real world (e.g., a book sinks into a table), students could be confused by it. While our previous research revealed that EBS did not confuse students and was effective in correcting their misconceptions about Newton's third law in static situations (Hirashima et al., 2009; Horiguchi et al., 2014), the current research deals with MIF misconceptions in dynamic situations. Therefore, we focus on whether the effect of EBS in correcting students' MIF misconceptions is at least equal to or greater than those of other methods, without creating any confusion.

EXPERIMENT

We conducted an experiment in class to measure the effect of EBS in correcting students' MIF misconceptions. The result is first evaluated in itself, then compared with those of other methods.

Although the basic framework of this experiment was the same as Hirashima et al. (2017), we conducted the experiment twice (in 2016 and 2017) to confirm the stability of the effect of EBS, and added a delayed post-test to examine whether the effect of EBS was maintained over the long term.

Target Misconception

According to Clement (1982), we classified the MIF misconception as follows.

MIF-(1): Force in the direction of motion is necessary to cause and maintain an object's motion, even when moving at a constant velocity.

MIF-(2): When there is explicit resistance against motion, a force greater than the resistance is necessary to maintain the motion.

MIF-(3): The force in the direction of motion increases/decreases according to the velocity of the motion.

If a student's wrong answer could be explained based on these misconceptions, we termed it "MIF-answer." If the number of students' MIF-answers decrease after learning, their MIF misconception is considered to have been resolved through learning.

Instruments

Learning Environment

In this research, an EBS-based learning environment (hereafter called "the system") was implemented as an Android tablet PC application, in which students worked on a set of problems in mechanics. Figure 3 and 4 show the snapshots of the system. For each problem, students were shown an illustration which consisted of three figures of a moving object at different instants (Figure 3(a) and 3(b)). Students were then required to draw all the forces applied to the object in the situation for each instant. Forces were drawn as arrows. After choosing the magnitude and direction of force from a menu, students tapped the "Create force" button to make the arrow appear on the screen (Figure 4(a)). They then dragged it onto an object to apply the force. For each object, the net force was calculated based on the drawing, and the resulting motion (i.e., EBS) was simulated using Newton's second law. When an EBS was shown, the corresponding correct motion was simultaneously shown for comparison (Figure 4(b)). When the drawing included errors, the object's motion of EBS often contradicted the students' prediction and differed from the correct motion. Such a simulation was expected to work as a counterexample to make students aware of and correct their errors. Students could modify their drawings and simulate the motion repeatedly to arrive at the final (correct) solution.

[Place Figure 3(a) and 3(b) here]

[Place Figure 4(a) and 4(b) here]

Tests

In the experiment, MIF misconceptions MIF-(1), (2), and (3), described above, were targeted. If students' ideas reflected these misconceptions before using EBS and did not reflect them after using EBS, their MIF misconceptions were assumed to have been resolved. We implemented the following six problems in the system:

Problem (1): A skater traveling straight on (frictionless) ice at a constant velocity (in which MIF-(1) is predicted. Students are likely to draw a force in the direction of travel. In the EBS, the skater was shown to accelerate in that direction).

Problem (2): A man who descends in air with a parachute at a constant velocity (in which MIF-(2) is predicted. Students are likely to draw downward and upward forces, where the former is greater than the latter. In the EBS, the man was shown to accelerate downward).

Problem (3): A thrown ball rising vertically upward (in which MIF-(3) is predicted. Students are likely to draw an upward force that gradually decreases as the ball ascends. In the EBS, the ball was shown to accelerate upward).

Problem (4): A spacecraft flies straight through space at a constant velocity (in which MIF-(1) is predicted. Students are likely to make an error isomorphic to that of problem (1)).

Problem (5): An object on a floor with friction shoved horizontally at a constant velocity (in which MIF-(2) is predicted. Students are likely to make an error isomorphic to that of problem (2)).

Problem (6): A thrown ball rising diagonally upward (in which MIF-(3) is predicted. Students are likely to make an error that is isomorphic to that of problem (3)).

These six problems are called the "learning task." Problems (1), (2), and (3) were termed as "basic problems." Further, problems (4), (5), and (6) were the "advanced problems" of problems (1), (2), and (3), respectively. Additionally, we predicted that MIF-(1) would appear in all the problems.

For example, in Problem (2), students could draw only the downward force, which was counted as MIF-(1).

In the pre-test, students solved the six problems of the learning task (which were learned on the system) as a written test. In the post-test, in addition to the six problems, they solved the following four problems of the "transfer task" as a written test. In the delayed post-test, they solved the same problems as the post-test as a written test.

Problem (7): A dolly first descends a slope and then enters a horizontal floor. There is no friction throughout the motion.

Problem (8): A sled accelerates on (frictionless) ice with a continuous horizontal force.

Problem (9): A box is decelerating on a horizontal floor with friction.

Problem (10): An elevator is being lifted up at a constant velocity.

In tests, each problem was presented with an illustration of mechanical situation and a text of question. Each illustration consists of three figures of an object at different instants. Students were required to draw all the forces applied to the object in the situation for each instant (Forces were drawn as arrows).

Force Concept Inventory (FCI)

At the beginning of the experiment, the students were required to complete the force concept inventory (FCI) (Hestenes et al., 1992) to measure their basic understanding of mechanics. The FCI is a test designed to measure students' understanding of Newtonian mechanics. Students who can answer 60% (18 points) of the test correctly are considered to have effective problem-solving ability related to the consistent use of the concepts of Newtonian mechanics (Hestenes et al., 1992).

The FCI is used frequently in Japan (where our experiment was conducted) to measure students' understanding of mechanics. For example, Yasuda et al. (2011) reported that the average score of 16 undergraduate students who attended an introductory lecture on mechanics was 70% (21 points), and the score of 5 of these students was below 60%. Kuroda et al. (2017) reported the average scores of undergraduate students who attended lectures on introductory physics. In the lecture, the average scores before and after the lecture were 41.5% (12.5) and 48.8% (14.6), respectively. With respect to another lecture, the scores were 38.7％ (11.6 points) and 42.9% (12.9 points), respectively (the numbers of students were not clearly described, but charts suggested that there were several dozens of them) (Kuroda et al., 2017). Additionally, Ito (2015) reported that the average score of 68 high-school students in a non-natural science course was 45.3%.

Participants and Procedures

We used the system described above for teaching mechanics at a Japanese technical college in 2016 and 2017. In 2016, thirty-seven third-grade (which corresponds to twelfth grade in U.S.) students attended this class. In 2017, thirty-five third-grade students attended this class. They were 17-18 years old, and had completed the basic course of mechanics.

In both years, the students first completed a standardized test for twenty minutes to measure their basic understanding of mechanics (i.e., the FCI described above). The students then completed the pre-test in seven minutes, and worked on the learning task on the system for seventeen minutes. Each student was given her/his own tablet PC. While they were instructed to work on all the six problems (i.e., learning task), it was up to each student how much time she/he spent on each problem. Additionally, while they were recommended to work on the problems in the original order, it was also up to each student in what order she/he proceeded with the exercise. In working on a problem, they could draw forces and execute simulation as many times they liked. Throughout the exercise, teachers never intervened in students' solving problems. After the exercise with the system, the students completed the post-test in ten minutes. Finally, approximately one month later, the students completed the delayed post-test in ten minutes.

RESULTS AND DISCUSSION

We scored the students' answers on the tests as follows. If an erroneous arrow was drawn due to the MIF misconception, the answer was classified as an "MIF-answer." If the erroneous arrow was not due to the MIF misconception, the answer was classified as "other erroneous answer."

Students' Basic Understanding of Mechanics

Table 1 shows the FCI results of the students who participated in our experiment. The average scores in 2016 and 2017 were 43.4% (13.03 points) and 48.1% (14.44 points), respectively, with no significant differences between the years ($p=.349$, Wilcoxon rank sum test). Therefore, in both years, students with the same level of understanding of mechanics participated in our experiment, and their understanding was insufficient despite having already learned Newtonian mechanics to a certain extent.

[Insert Table 1 here]

Average Scores on the Learning Task

Table 2 shows the average scores on the learning task for the pre-test, post-test, and delayed post-test. The statistical analysis (i.e., Wilcoxon signed rank test with Bonferroni correction) revealed the following.

In 2016, significant differences were observed between the pre-test and post-test ($p=2.151e-06$), and between the pre-test and delayed post-test ($p=3.042e-05$); both their effect sizes were large ($r=1.42$ and 1.03, respectively). Even though a significant difference was observed between the post-test and delayed post-test ($p=0.01214$), the effect size was small ($r=0.390$).

In 2017, significant differences were observed between the pre-test and post-test ($p= 2.122e-06$), and between the pre-test and delayed post-test ($p= 8.763e-06$); both their effect sizes were large ($r=1.83$ and 1.09, respectively). Additionally, even though a significant difference was observed between the post-test and delayed post-test ($p= 3.801e-05$), the effect size was medium ($r=0.720$).

These results suggest that learning with EBS was effective in improving students' understanding of mechanics, and the effect remained after a certain period of time.

[Insert Table 2 here]

Types of Erroneous Answers on the Learning Task

Tables 3 and 4 show the average number of erroneously answered problems in the learning task for the pre-test, post-test, and delayed post-test. The statistical analysis (i.e., Wilcoxon signed rank test with Bonferroni correction) revealed the following.

In 2016, the average number of MIF-answered problems significantly decreased from the pre-test to the post-test (p=1.172e-05), and from the pre-test to the delayed post-test (p=3.048e-04); the effect sizes were large (r=1.42 and 1.11, respectively). Further, even though the average number of MIF-answered problems increased from the post-test to the delayed post-test, the difference was not significant (p=0.8028). On the other hand, the average number of other erroneously answered problems did not significantly change between any pairs of tests (p=0.2050 for pre-test and post-test, p=0.6132 for post-test and delayed post-test, p=0.8172 for pre-test and delayed post-test). Additionally, the ratios of MIF-answered problems to the total number of erroneously answered problems were 73.2%, 61.3%, and 60.5% in the pre-test, post-test, and delayed post-test, respectively.

In 2017, the average number of MIF-answered problems significantly decreased from the pre-test to the post-test (p=4.941e-06), and from the pre-test to the delayed post-test (p=8.748e-05); the effect sizes were large (r=1.65 and 1.00, respectively). Additionally, even though the average number of MIF-answered problems significantly increased from the post-test to the delayed post-test (p=1.72e-03), the effect size was medium (r=0.610). On the other hand, the average number of other erroneously answered problems did not significantly change between any pair of tests (p=0.1417 for pre-test and post-test, p=0.1212 for post-test and delayed post-test, p=1.440 for pre-test and delayed post-test). Additionally, the ratios of MIF-answered problems to the total number of erroneously answered problems were 80.5%, 65.2%, and 71.4% in the pre-test, post-test, and delayed post-test, respectively.

These results suggest that learning with EBS was effective mainly in the resolution of MIF misconceptions and that the effect remained after a certain period of time, whereas our previous research showed only that EBS decreased the number of students' erroneous answers (Hirashima et al., 1998; Hirashima et al., 2009; Horiguchi et al., 2014).

Although the problems of the learning task in tests were the same as those that the students learned with the system, this result indicates the effectiveness of the EBS. If students had answered the post-test questions based on memorized correct answers they observed while learning with the system, both types of erroneous answers (i.e., MIF-answers and other erroneous answers) would have decreased to the same degree, but a difference was observed in both years. Additionally, this improvement mostly remained in the delayed post-test. These facts suggest that learning with EBS caused the students to learn the correct concepts which lead to the correction of errors.

[Insert Table 3 here]

[Insert Table 4 here]

Types of Erroneous Answers on the Transfer Task

Table 5 shows the average number of erroneously answered problems in the transfer task for the post-test and delayed post-test. The statistical analysis (i.e., Wilcoxon signed rank test with Bonferroni correction) revealed the following.

In 2016, between the post-test and delayed post-test, no significant differences were observed in the average number of MIF-answered and other erroneously answered problems ($p=0.140$ and $p=0.450$, respectively). Additionally, the ratios of MIF-answered problems to the total number of erroneously answered problems were 36.4% and 25.1% in the post-test and delayed post-test, respectively.

In 2017, between the post-test and delayed post-test, no significant differences were observed in the average number of MIF-answered and other erroneously answered problems ($p=0.180$ and $p=0.560$, respectively). Additionally, the ratios of MIF-answered problems to the total number of erroneously answered problems were 28.9% and 36.4% in the post-test and delayed post-test, respectively.

These ratios are much lower than those on the pre-tests (73.2% (2016) and 80.5% (2017); note that these data are for the learning task). Students who learned with EBS gave fewer MIF-answers, even in problems they saw for the first time. Additionally, this improvement mostly remained in the delayed post-test. These results suggest that learning with EBS lead to conceptual clarity and the correction of errors.

[Insert Table 5 here]

Comparison with Other Methods

The results presented above show that using EBS to correct students' MIF misconceptions is significantly effective without confusing students. In our previous work, we verified EBS was significantly effective in teaching the balance of forces exerted on objects at rest (Hirashima et al., 2009; Horiguchi et al., 2014). In such static situations, the difference between EBS and the correct simulation was very clear because an object that should keep still erroneously moved. In contrast, in this work, we verified EBS was effective even when objects' correct behavior is motion, which opens up new possibilities of EBS. In this section, by comparing our results with those of related work reviewed in the previous section, we discuss whether the effect of EBS in correcting MIF misconceptions is equal to or greater than that of other methods. First, we make a comparison mainly from a statistical viewpoint, then discuss the features of EBS as a technology-based method compared to other methods.

In Palmer and Flanagan's (1997) study, 61 primary school students and 63 secondary school students were judged to have MIF misconceptions using a pre-test. They then read a "refutational text" (see the section "Related Work on Conceptual Change") and discussed its contents in class. As a result, 43% of primary school students and 57% of secondary school students corrected their MIF misconceptions in the post-test, and 35% of primary school students and 44% of secondary school students maintained the correction after one to two weeks in the delayed post-test. In each of the post-test and delayed post-test, students were judged to have corrected their MIF misconception or to maintain the correction when they had correctly answered at least two of the three problems (the rate of erroneous answers was less than 33%).

On comparing these results with ours, in our experiment, the number of students who erroneously answered two or more of the six problems in the pre-test (the rate of erroneous answers was more than 33%) was 27 in 2016 and 25 in 2017; these students were judged to have MIF misconceptions. Among them, the number of students who gave less than three MIF-answers on the ten problems in the post-test (the rate of erroneous answers is less than 30%) was 19 (70%) in 2016 and 21 (84%) in 2017. In the delayed post-test, the number of students who gave less than three MIF-answers was 20 (74%) in 2016 and 12 (48%) in 2017. They were judged to have corrected their MIF misconceptions or to maintain the correction. Considering that our criteria for MIF misconception resolution are slightly stricter, our method was probably as successful as Palmers and Flanagans' (1997).

In Takagaki's (2004) study, in the pre-test comprising two problems (problem-A and problem-B that were identical to our Problem (3) and Problem (6), respectively), out of 82 university students, eight (9.8%) correctly answered problem-A and four (4.9%) correctly answered problem-B. They read a text designed to trigger and resolve cognitive conflict (see the section "Related Work on Conceptual Change"). In a post-test (which is the same as the pre-test), out of 82 students, 60 (73.2%) correctly answered problem-A and 38 (46.3%) correctly answered problem-B.

On comparing this result with ours, in Takagaki's (2004) study, if students who correctly answered at least one of two problems were judged as not having an MIF misconception, the rate of such students was at most 7.3% (12 out of 164) in the pre-test and 59.8% (98 out of 164) (the percentages may be overestimated because the numbers in the two problems are totaled). On the other hand, in our experiment, the rates of students who correctly answered at least one out of Problem (3) and Problem (6) (i.e., problem-A and problem-B) in the pre-test were 16.7% (nine out of 54) in 2016 and 23.2% (16 out of 69) in 2017. In the post-test, the rates were 62.5% (40 out of 54) in 2016 and 76.1% (51 out of 67) in 2017. Considering that the rate of correct answers in our pre-test was higher than that of Takagaki's, our method was probably as successful as Takagaki's (2004).

In Dekkers and Thijs' (1998) study, university students who had a preconception of force in the preceding experiment participated in the class (four classes each of which had about thirty students) designed for promoting their conceptual development (see the section "Related Work on Conceptual Change"). Sixty-six students were randomly selected for the tests. As a result, while the average correct score was 15% in the pre-test, it was 57% in the post-test. Two months later, in the delayed post-test, the average score of correct responses increased to 79%.

To compare this result with ours, we transformed the test scores shown in Table 2 to average correct scores using the maximum score. In our experiment, the average correct scores were 20.7% in 2016 and 26.7% in 2017 in the pre-test; 63.5% in 2016 and 78.2% in 2017 in the post-test; and 50.5% in 2016 and 58.5% in 2017 in the delayed post-test. Therefore, our method seems to be as successful as that of Dekkers and Thijs' (1998), at least in the post-test. As for the delayed post-test, on the other hand, the average correct score increased significantly in Dekkers and Thijs' (1998) study, while it decreased in our study. This suggests that the students' conceptual development in our study was maintained at a shallower level as compared to Dekkers and Thijs' (1998) method.

Note that an accurate comparison between these methods is impossible because the target students, classroom-related characteristics, contents of tests, and methods of evaluation are different. However, the statistical comparison above suggests that the effectiveness of EBS which provides technology-based assistance was at least equal to or greater than that of other methods in correcting students' MIF misconceptions.

We think two features of EBS brought the success of our method compared to other methods. First, for triggering cognitive conflict, showing an erroneous behavior of objects that reflects students' misconception appears to be better than the (verbal) explanation of the correct concept. An animation would be easier for students to understand intuitively than a text-based explanation. (In fact, the rate of students who corrected MIF misconception in Takagaki's (2004) study was generally higher than that in Palmer and Flanagan's (1997) study. The former used a video clip while the latter not.) EBS is a powerful medium that expresses not only objects' behavior but also students' erroneous idea. Second, while a set of problems and instructions were carefully designed to connect concrete situations and abstract concepts in Dekkers and Thijs' (1998) and Elby's (2001) studies, our method is rather simple in which only the current problem and students' answer are used for generating EBS. Instead, the advantage of our method is its interactivity (i.e., students can try their idea and receive immediate feedback). The fact that such discovery learning succeeded in correcting MIF misconception to some extent indicates a potential of technology-assisted learning in this area, though the level of understanding with EBS should be further investigated.

CONCLUSION

In this paper, we proposed EBS as a method of correcting students' MIF misconceptions and reported the results of a two-year experimental implementation of EBS at a technical college. In both years, EBS not only had an effect on the resolution of MIF misconceptions but also promoted the correction of errors at a conceptual level. These results suggest that EBS does not confuse students, even though it shows unnatural phenomena as counterexamples. The results also suggest that EBS promotes students' discovery of the correct concept and its effectiveness is at least equal to or greater than other methods that involve explaining the concept directly. Therefore, we think the technology-based assistance by EBS is promising method for changing students' misconception.

The number of problems used in the experiment was not large, but the problems covered most typical situations in which MIF misconceptions tend to occur, the resolution of which is central to learning elementary mechanics. Therefore, we believe our results have a certain degree of generality and usefulness.

We point out three limitations of this research. One limitation is that students' conceptual change or conceptual development was measured only with the results of tests. Though we think these results strongly suggest students' conceptual change or development, it would be helpful to adopt a linguistic method such as discourse analysis based on an interview with a teacher or discussion between students.

The other two limitations concern the EBS methodology itself. First, in EBS framework, students' misconceptions should be embodied as objects' behavior. If a phenomenon in EBS doesn't clearly reflect students' error, it would not help them (e.g., suppose a student draws erroneous forces exerted on an object but the net force accidentally becomes correct, her/his error would not be reflected in its behavior). In our experiment, though problems were carefully selected so as to minimize the possibility of such occurrence, we cannot guarantee that such a thing would not happen. In such a case, EBS should be supplemented with other teaching material. Second, in our experiment, students had to express their (erroneous) idea only by drawing arrows which were equally exerted on any objects/agents. Even if a student had an unscientific preconception such as "impetus" or "vital energy," she/he had to translate it into such a uniform "force." Thus, some information about students' misconception might be lost.

Regarding future research, we plan to first confirm these results with a larger sample. Additionally, we plan to investigate the effect of EBS before and after learning mechanics using mathematical formalism. We will adopt the method of interview or class discussion to identify what types of misconception students exactly have (as discussed in “Motion Implies a Force Misconception” subsection), and to clarify the process of how students discover the “(right) theory of force and motion” through the interaction with the system. Furthermore, the level of conceptual change (Hewson & Hewson, 1992) occurring due to EBS should be identified and compared with those of other methods. In particular, from the viewpoint of a “categorical shift” (Chi, 2008), although EBS would effectively make students aware of category mistakes (by embodying misconceptions as concrete phenomena), it should be supplemented with an explanation of the correct category. Therefore, combining EBS with other teaching methods and embedding EBS in pedagogically effective lessons must also be explored.

ENDNOTE

[1] Though we previously reported the partial results of this experiment conducted in 2016 (Hirashima et al., 2017), it was only a one-year result and lacked the data of delayed post-test (of which analysis was ongoing when the conference paper was written). In this study, we confirmed its reproducibility by adding the results of another year, and clarified that the effect of EBS was maintained over a long term by adding the data of the delayed post-test. In particular, only after analyzing the data of the delayed post-test, we could discuss whether the effect of EBS on students' understanding was at the conceptual level or due to memorization.

REFERENCES

Authors (date). Journal Title.

Bransford, J. D., Brown, A. L., & Cocking, R. (2000). *How people learn: Brain, mind, experience, and school*. National Academy Press.

Caramazza, A., McCloskey, M., & Green, B. (1981). Naïve beliefs in "sophisticated" subjects: misconceptions about trajectories of objects, *Cognition*, *9*, 117–123. https://doi.org/10.1016/0010-0277(81)90007-x

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categilical shift. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 61–82). Routledge.

Chinn, C. A., & Brewer, W. F. (1993). Factors that influence how people respond to anomalous data. In *Proceedings of the 15th annual conference of the cognitive science society* (pp. 318–323). Lawrence Erlbaum Associates.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*, 66–71. https://doi.org/10.1119/1.12989

Dekkers, P. J. J. M. & Thijs, G. D. (1998). Making Productive Use of Students' Initial Conceptions in Developing the Concept of Force, *Science Education*, *82*, 31–51. https://doi.org/10.1002/(sici)1098-237x(199801)82:1<31::aid-sce3>3.0.co;2-1

Driver, R., Guesne, E., & Tiberghien, A. (1985). *Children's ideas in science*. Open University Press.

Duit, R. (2003). Conceptual change: A powerful framework for improving science teaching and learning, *International Journal of Science Education*, *25(6)*, 671–688. https://doi.org/10.1080/09500690305016

Elby, A. (2001). Helping physics students learn how to learn. *American Journal of Physics, 69*, 54–64. https://doi.org/10.1119/1.1377283

Gagne, E. D. (1985). *The cognitive psychology of school learning*. Little Brown.

Glynn, S. M., Yeany, R. H., & Britton, B. K. (1991). *The psychology of learning science*. Lawrence Erlbaum Associates.

Hashweh, M. Z. (1986). Toward an explanation of conceptual change, *European Journal of Science Education*, *8*, 229–249. https://doi.org/10.1080/0140528860080301

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30(2)*, 141–158. https://doi.org/10.1119/1.2343497

Hewson, P. W. and Hewson, M. G. (1992). The status of students' conceptions, In R. Duit, F. Goldberg, & H. Neidderer (Eds.), *Research in Physics Learning: Theoretical Issues and Empirical Studies* (pp. 59–73). Institute of Science Education.

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(2), 17–31.

Hirashima, T., Imai, I., Horiguchi, T., & Toumoto, T. (2009). Error-based Simulation to promote awareness of errors in elementary mechanics and its evaluation. In Proceedings of the 14th international conference on artificial intelligence in education (pp. 409–416). Pasadena, CA.

Hirashima, T., Shinohara, T., Yamada, A., Hayashi, Y. & Horiguchi, T. (2017) Effects of Error-Based Simulation as a Counterexample for Correcting MIF Misconception. In Proceedings of the 18th international conference on artificial intelligence in education (pp. 90-101). Springer, Cham.

Horiguchi, T., & Hirashima, T. (2006). Robust simulator: A method of simulating learners' erroneous equations for making Error-based Simulation. Lecture Notes in Computer Science, 4053, 655–665. https://doi.org/10.1007/11774303_65

Horiguchi, T., Hirashima, T., & Forbus, K. D. (2012). A model-building learning environment with Error-based Simulation. In Proceedings of the 26th international workshop on qualitative reasoning. http://www.qrg.northwestern.edu/Resources/qrw_proceedings_index.html

Horiguchi, T., Imai, I., Toumoto, T., & Hirashima, T. (2014). Error-based Simulation for error-awareness in learning mechanics: An evaluation. Journal of Educational Technology & Society, 17(3), 1–13.

Ito, S. (2015). Interactive demonstration lecture in physics with PC-tablet promoting construction of scientific concept. In *Proceedings of 2015 PC Conference* (pp. 305–308).

Klahr, D. (2009). "To Every Thing There is as Season, and a Time to Every Purpose Under the Heavens" What about Direct Instruction, In S. Tobias, & T. H. Duffy (Eds.), *Constructivist Theory Applied to Instruction: Success or Failure?* (pp. 291–310). Taylor and Francis. https://doi.org/10.4324/9780203878842-24

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. The University of Chicago Press.

Kunichika, H., Hirashima, T., & Takeuchi, A. (2006). Visualizing errors for self- correcting discrepancy between thinking and writing. In Proceedings of the 14th international conference on computers in education (pp. 483–490).

Kuroda, K., Mizuno, T., & Miyata, S. (2017). Study on the concept of the students attending several physics classes - for the improvement of the physics education and putting active-learning into practice. *Memoirs of The College of Engineering, Tamagawa University, 52*, 9–20.

Matsuda, N., Takagi, S., Soga, M., Hirashima, T., Horiguchi, T., Taki, H., . . . Yoshimoto, F. (2003). Tutoring system for pencil drawing discipline. In Proceedings of ICCE2003 (pp. 1163–1170). Hong Kong.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner, D., & A. L. Stevens (Eds.), *Mental models* (pp. 299–324). Lawrence Erlbaum Associates.

Mestre, J. P. (1994). Cognitive aspects of learning and teaching science. In S. J. Fitzsimmons & L. C. Kerplelman (Eds.), *Teacher enhancement for elementary and secondary science and mathematics: Status, issues, and problems* (pp. 31–53). National Science Foundation.

Osborne, R., & Freyberg, P. (1985). *Learning in science - the implications of children's science*. Heinemann Publishers Ltd.

Palmer, D. H., & Flanagan, R. B. (1997). Readiness to Change the Conception That "Motion-Implies-Force": A Comparison of 12-Year-Old and 16-Year-Old Students, *Science Education, 81*, 317–331. https://doi.org/10.1002/(sici)1098-237x(199706)81:3<317::aid-sce4>3.0.co;2-g

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66,* 211–227. https://doi.org/10.1002/sce.3730660207

Strike, K. A., & Posner, G. J. (1985). A conceptual change view of learning and understanding. In L. H. T. West & A. L. Pines (Eds.), Cognitive Structure and Conceptual Change (pp. 211–231). Academic Press.

Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. Duschl & R. Hamilton (Eds.), Philosophy of Science, Cognitive Psychology, and Educational Theory and Practice (pp. 147–176). State University of New York Press.

Takagaki, M. (2004). How Undergraduate Students Change Their Preconceptions About Forces, *The JapaneseJournal of Developmental Psychology, 15(2)*, 217–229.

Toulmin, S. (1972). *Human Understanding*. Princeton University Press.

Towne, D. M., de Jong, T., & Spada, H. (1993). *Simulation-based experiential learning*. Springer-Verlag. https://doi.org/10.1007/978-3-642-78539-9

Vosniadou, S., & Ioannides, C. (1998). From conceptual change to science education: A psychological point of view. *Science Education, 20*, 1213–1230. https://doi.org/10.1080/0950069980201004

Wenger, E. (1987). *Artificial intelligence and tutoring systems: Computational and cognitive approaches to the communication of knowledge*. Morgan Kaufmann.

Yasuda, J., Uematsu, H., & Nitta, H. (2011). Validity evaluation of the Japanese translation version of Force Concept Inventory (FCI). *Journal of Physics Education Society of Japan, 59(2)*, 90–95.

Figure 1. Framework of Error-based Simulation.

Note. EBS is generated by mapping errors in symbolic expressions to erroneous behaviors. The difference in behavior expression makes students aware of errors and motivates them to correct the errors. If students have misconceptions expressed in the form of wrong answers, the erroneous behavior works as a counterexample to their misconception.

Figure 2. An example of EBS.

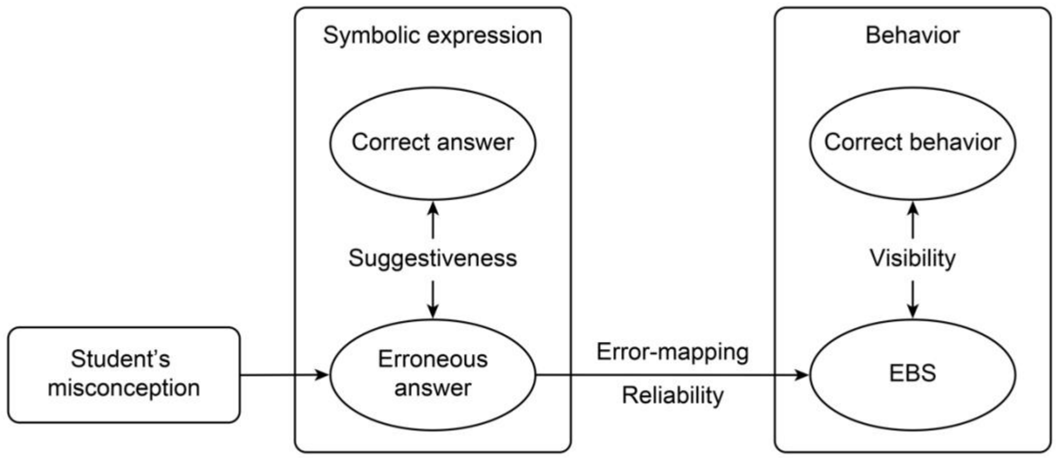
Note. Students are required to draw all the forces applied to a skater traveling straight on (frictionless) ice at a constant velocity. Due to misconceptions, students often draw a force in the direction of travel. In the EBS, the skater is then shown to accelerate in that direction. Such an unnatural phenomenon is expected to be useful as a counterexample to students' erroneous ideas and contributes to the correction of errors with high intrinsic motivation.

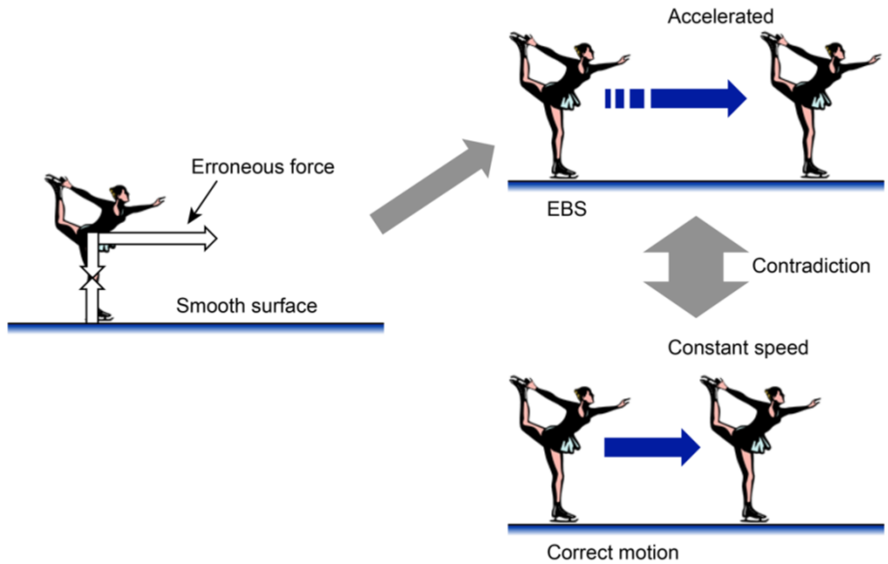
Figure 3(a) and 3(b). A snapshot of the system (1).

Note. An EBS-based learning environment was implemented as an Android tablet PC application. Students were shown an illustration which consisted of three figures of a moving object at different instants. Students were then required to draw all the forces applied to the object for each instant as arrows.

Figure 4(a) and 4(b). A snapshot of the system (2).

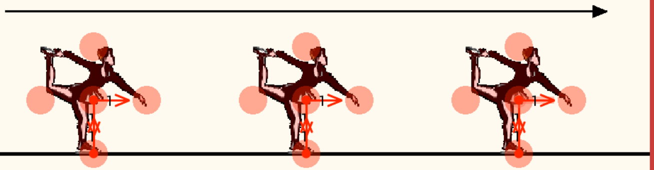
Note. After choosing the magnitude and direction of force from a menu, students tapped the "Create force" button to make the arrow appear on the screen (Figure 4(a)). They then dragged the arrow onto an object to apply the force. For each object, the net force was calculated based on the drawing, and the resulting motion was simulated using Newton's second law. When the drawing included errors, the object's motion often contradicted the students' prediction and differed from the correct motion (Figure 4(b)).


Symbolic expression
Correct answer
Suggestiveness
Student's misconception
Erroneous answer
Error-mapping
Reliability
Behavior
Correct behavior
Visibility
EBS


Erroneous force
Smooth surface
Accelerated
EBS
Contradiction
Constant speed
Correct motion

Force

A skater is traveling straight on ice at a constant velocity. Air resistance is negligible. Draw all forces applying to the skater.



Problem 1

See the correct motion

Delete

Delete all

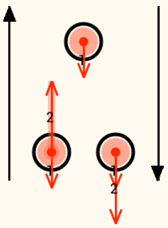
Back

Diagnose

Create force

Force

A ball is thrown upward. It first ascends vertically. Then, it stops at some height. Finally, it descends vertically. Air resistance is negligible. Draw all forces applying to the ball at each moment.



Problem 3

See the correct motion

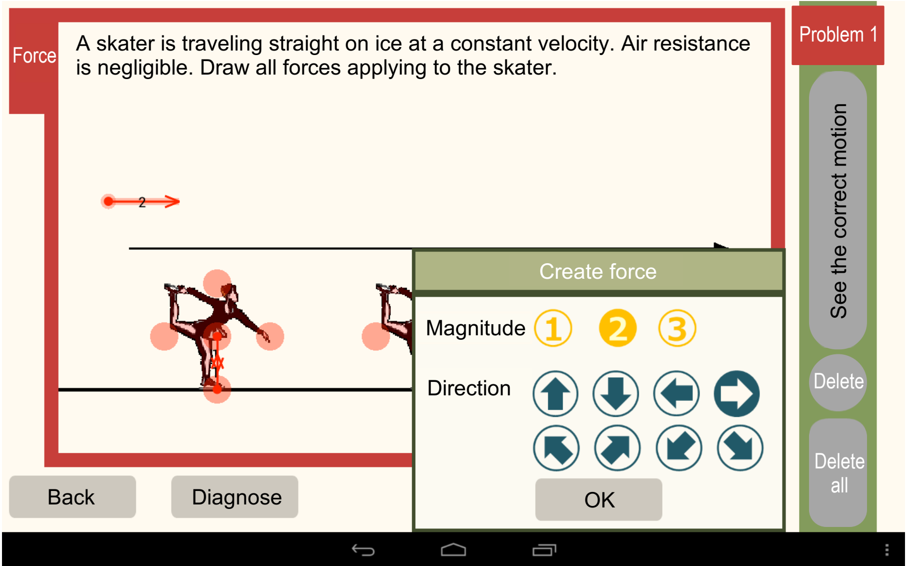
Delete

Delete all

Back

Diagnose

Create force


Force
A skater is traveling straight on ice at a constant velocity. Air resistance is negligible. Draw all forces applying to the skater.
2
Create force
Magnitude
1
2
3
Direction
OK
Back
Diagnose
Problem 1
See the correct motion
Delete
Delete all

Correct motion

Force

See again

Back to drawing

Table 1. Results of participant responses on the force concept inventory (maximum score = 30)

2016		2017		Between 2016 & 2017
M	SD	M	SD	p
13.03	5.90	14.44	6.00	0.349

Table 2. Scores on the learning task (maximum score = 6)

Year	Pretest		Posttest		Delayed posttest		Between pretest & posttest		Between posttest & delayed posttest		Between pretest & delayed posttest	
	M	SD	M	SD	M	SD	p	Cohen's d	p	Cohen's d	p	Cohen's d
2016	1.24	1.48	3.81	2.10	3.03	1.95	$2.15\times10^{-6**}$	1.42	$1.21\times10^{-2*}$	0.39	$3.04\times10^{-5**}$	1.03
2017	1.60	1.80	4.69	1.58	3.51	1.69	$2.12\times10^{-6**}$	1.83	$3.80\times10^{-5**}$	0.72	$8.76\times10^{-6**}$	1.09

Note. p: Bonferroni-corrected p value

Table 3. Numbers of erroneously answered problems on the learning task (1)

Year	Pretest				Posttest				Delayed posttest			
	MIF-answers		Other answers		MIF-answers		Other answers		MIF-answers		Other answers	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
2016	3.50	1.76	1.28	1.07	1.14	1.56	0.72	0.96	1.53	1.80	1.00	1.00
2017	3.54	1.90	0.86	0.96	0.86	1.29	0.46	0.77	1.77	1.65	0.71	0.91

Table 4. Numbers of erroneously answered problems on the learning task (2)

Year	MIF-answers						Other answers					
	Between pretest & posttest		Between posttest & delayed posttest		Between pretest & delayed posttest		Between pretest & posttest		Between posttest & delayed posttest		Between pretest & delayed posttest	
	p	Cohen's d	p	Cohen's d	p	Cohen's d	p	Cohen's d	p	Cohen's d	p	Cohen's d
2016	$1.17 \times 10^{-5**}$	1.42	0.803	0.23	$3.05 \times 10^{-4**}$	1.11	0.205	0.55	0.613	0.29	0.817	0.27
2017	$4.94 \times 10^{-6**}$	1.65	$1.72 \times 10^{-3**}$	0.61	$8.75 \times 10^{-5**}$	1.00	0.142	0.46	0.121	0.30	> 1.0	0.16

Note. p: Bonferroni-corrected p value

Table 5. Numbers of erroneously answered problems on the transfer task

Year	Posttest				Delayed posttest				Between posttest & delayed posttest	
	MIF-answers		Other answers		MIF-answers		Other answers		MIF-answers	Other answers
	M	SD	M	SD	M	SD	M	SD	p	p
2016	0.78	0.86	1.36	1.16	0.53	0.64	1.58	1.46	0.14	0.45
2017	0.61	0.64	1.50	1.15	0.80	0.95	1.40	1.18	0.18	0.56