ELSEVIER
Learning and Instruction 16 (2006) 199−212
Learning and Instruction
www.elsevier.com/locate/learninstruc

Supporting students' learning with multiple representations in a dynamic simulation-based learning environment

Jan van der Meij*, Ton de Jong

University of Twente, Faculty of Behavioural Sciences, Department of Instructional Technology, PO BOX 217, 7500 AE Enschede, The Netherlands

Abstract

In this study, the effects of different types of support for learning from multiple representations in a simulation-based learning environment were examined. The study extends known research by examining the use of dynamic representations instead of static representations and it examines the role of the complexity of the domain and the learning environment. In three experimental conditions, the same learning environment, that of the physics topic of moments, was presented with separate, non-linked representations (S-NL condition), with separate, dynamically linked representations (S-DL condition), and with integrated, dynamically linked representations (I-DL condition). The learning environment was divided into low complexity and high complexity parts. Subjects were 72 students from middle vocational training (aged 16−18). Overall, the I-DL condition showed the best learning performance. Subjects in the I-DL condition, compared to the S-NL condition, showed better learning results on posttest items measuring domain knowledge. A trend in favor of the I-DL condition compared to the S-NL condition was found on posttest items measuring subjects' ability to translate between different representations. A subjective measure of experienced difficulty showed that subjects in the I-DL condition experienced the learning environment as easiest to work with. The complexity of the learning environment and domain interacted with the effects of the experimental conditions. Differences between conditions were only found on the test items that corresponded to the high complexity part of the learning environment.

Keywords: Multiple representations; Instructional technology; Scientific discovery learning; Simulation-based learning environments

1. Introduction

Many learning materials offer multiple representations. Textbooks, for example, often use photographic images or diagrams to illustrate and explain parts of the text. In early computer-based learning environments, texts and images were applied in the same way as in textbooks, namely as static images. Therefore, research on representations in textbooks was also valid for these early computer-based learning environments. In modern, computer-based learning environments many dynamic representations are available, including audio, video, animations, dynamically changing graphs and tables (Lowe, 2003), and interactive dynamic visuals. This development poses new challenges and

* Corresponding author. Tel.: +31 53 489 3598/4359; fax: +31 53 489 2849.
E-mail addresses: j.vandermeij@utwente.nl (J. van der Meij), a.j.m.dejong@utwente.nl (T. de Jong).

0959-4752/$ - see front matter
doi:10.1016/j.learninstruc.2006.03.007

opportunities and calls for a new line of research to study the implications for learning using these multiple dynamic representations (Ploetzner & Lowe, 2004).

1.1. Benefits of multiple representations

Different types of (dynamic) representations exist, and combining different representations in one interface may have several advantages (e.g., Ainsworth & van Labeke, 2004). First, each representation can show specific aspects of the domain to be learned. Different types of representations may be useful for different purposes, as they differ in their representational and computational efficiency (Larkin & Simon, 1987). Text and pictures, for example, are good representations to present the context of a problem. Diagrams are well suited for presenting qualitative information, and graphs, formulas, and numeric representations can be used to show quantitative information. Graphs, in particular, are important tools in enabling learners to predict relationships between variables and to show the nature of these relationships (McKenzie & Padilla, 1984). It is expected that learners benefit from the properties of each representation and that this will lead to a deeper understanding of the subject being taught (Ainsworth, Bibby, & Wood, 1997; de Jong et al., 1998; Seufert, 2003; van Labeke & Ainsworth, 2001).

A second benefit of a multi-representational learning environment is that one representation can constrain the interpretation of another representation. An animation, for example, can constrain the interpretation of a graph. There is a strong tendency among learners to view graphs as pictures rather than as symbolic representations (Kaput, 1989; Mokros & Tinker, 1987). When the animation shows a car riding up a hill with constant power, it constrains the interpretation of the speed shown in a line graph. The animation can show learners that the line graph is not representing a valley but the speed of the car; learners can see that the car slows down going up the hill and that it accelerates going down the hill. The purpose of the constraining representation is not to provide new information but to support the learners' reasoning about the less familiar representation (Ainsworth, 1999).

A third advantage of the use of multiple representations is that by translating between representations, learners build abstractions that may lead to a deeper understanding of the domain (Ainsworth & van Labeke, 2004).

1.2. Problems with multiple representations

When learning with multiple representations, learners are faced with four tasks. First, they have to understand the syntax of each representation. They must learn the format and operators of the representations. For example, the format of a graph would include attributes such as labels, number of axes, and line shapes. Examples of graph operators are finding the gradients of lines, minima and maxima, and intercepts (Ainsworth et al., 1997). Second, learners have to understand which parts of the domain are represented. In a simulation about a car in motion, for example, the learner has to relate the slope of the line in a speed–time graph to the right property of the moving car. A relevant question would be: Does the line represent the acceleration of the car or does it represent the speed of the car? In addition, the operators of one representation are often used inappropriately to interpret a different representation. This results in common mistakes such as viewing a graph as a picture (see Mokros & Tinker, 1987). Third, learners have to relate the representations to each other if the representations (partially) present the same information. We define relating as linking the surface features of different representations. When, for example, a numerical representation and a graph have to be related, learners must find the corresponding variables in both representations. Fourth, learners have to translate between the representations. We define translating as having to interpret the similarities and differences of corresponding features of two or more representations.

A first problem that learners may encounter when learning with multiple representations is that they have difficulties relating different representations. This problem is related to the split-attention problem as studied by Chandler and Sweller (1991) and Mayer and Moreno (1998). When learning with separate representations, learners are required to relate disparate sources of information, which may generate a heavy cognitive load that may leave less resources for actual learning (Sweller, 1988, 1989).

Second, a number of studies have reported problems that novices have in translating between representations. Tabachneck, Leonardo, and Simon (1994) reported that novices learning with multiple representations in economics did not attempt to translate information between line graphs and written information. Experts, in contrast, tied graphical and verbal representations closely together. Similar results were reported by Kozma (2003), who reviewed experimental and naturalistic studies that examined the role of multiple representations in understanding science. He looked at the differences between expert chemists and chemistry students in their representational skills and in their use of representations in science laboratories. Experts coordinated features within and across multiple representations to reason about their research. Students, on the other hand, had difficulty moving across or connecting multiple representations, so their understanding and discourse were constrained by the surface features of individual representations.

1.3. Types of support

The problems that were mentioned in the preceding section in regard to relating and translating representations traditionally (with static representations) are approached by integrating representations. Dynamic presentations offer the possibility to connect representations not only by integrating them, but also by linking them so that a change in one representation is concurrent with a change in another representation.

1.3.1. Integrating

One way to make relations between representations explicit for the learner is to physically integrate the representations (e.g., Chandler & Sweller, 1991). Multiple representations, when integrated, appear to be one representation showing different aspects of the domain. Through integration, relations between the representations are directly shown to the learner. Integrating representations also supports learners in the translation process. Having all related elements in the same place makes it easier to interpret the similarities and differences of corresponding features. Several studies conclude that learning with integrated representations leads to better knowledge than learning with representations that are not integrated (Ainsworth & Peevers, 2003; Chandler & Sweller; Mayer & Moreno, 1998; Tabbers, Martens, & Van Merriënboer, 2000). Different results were found by Bodemer, Ploetzner, Feuerlein, and Spada (2004). They compared learning from a situation in which learners had to integrate representations themselves to learning from a non-integrated format and a pre-integrated format. Bodemer et al. found that learners in the integrated condition did not learn more than learners in the non-integrated condition, unless learners had to actively integrate the representations themselves. In their study, the active integration was done with static representations, whereas the learning environment contained dynamic representations.

1.3.2. Dynamic linking

A second way to make the relation between different representations explicit for the learner, in the case of dynamic representations, is to provide the learner with dynamic linking (Ainsworth, 1999). With dynamically linked representations, actions performed on one representation are automatically shown in all other representations. If a learner, for example, changes the value of a force in a numerical representation, the corresponding representation of the force in an animation is updated automatically. It is expected that dynamic linking helps the learner to establish the relationships between the representations (e.g., Kaput, 1989; Scaife & Rogers, 1996). An environment using multiple linked representations can facilitate novices' learning even if their understanding of symbolic expressions draws heavily on an incomplete or inaccurate knowledge of the domain (Kozma, Russell, Jones, Marx, & Davis, 1996). Some literature, however, also mentions disadvantages of dynamic linking. Ainsworth, for example, asserts that a constructivist approach to education might argue that dynamic linking leaves a learner too passive in the process. Dynamic linking may discourage reflection on the nature of the translations, leading to a failure by the learner to construct the required understanding (p. 133). Another problem with dynamic linking might be that with multiple dynamically changing representations, learners need to attend to and relate changes that occur simultaneously in different regions of various representations, which may lead to cognitive overload (see Lowe, 1999).

2. Research questions

The goal of this study was to determine if integrating and/or linking dynamic multiple representations has an effect on learning outcomes. This was examined in a (simulation based) learning environment with dynamic representations. Most of the studies on integrating representations (Chandler & Sweller, 1991; Mayer & Moreno, 1998; Tabbers et al., 2000) investigated the integration of one diagram and text. Only the environment used by Ainsworth and Peevers (2003) consisted of more than two representations. These representations were also more complex than the representations used in the other studies. The study by Bodemer et al. (2004) suggested that the complexity of the domain and learning environment might influence the effects of integrating representations on learning. They found effects of integrating representation in a low complexity environment, but not in an environment of a higher complexity. In this study, we took complexity into account by dividing the learning environment into low and high complexity parts.

The context of the study was a guided discovery simulation-based learning environment called 'Moment.' Learners studied the topic of 'moments' by means of multiple representations of an open-end spanner tightening a bolt and of a crane hoisting a load. Three versions of the same simulation-based learning environment were compared: a learning environment with separate, non-linked representations (S-NL condition), a learning environment with separate, dynamically linked representations (S-DL condition), and a learning environment with integrated, dynamically linked representations (I-DL condition). We did not include a fourth (integrated, non-linked) condition, because integration and non-linking cannot be combined (see Section 3.2). All learning environments contained the same content. We expected that the S-DL and I-DL learning environments would lead to better learning results than the S-NL learning environment. We expected that the I-DL learning environment would lead to the best learning results.

With regard to complexity, we expected a positive effect of integration as long as the interface did not become too complex. We expected larger differences between conditions in the high complexity part of the learning environment compared to the low complexity part, because in the high complexity part, more representations are presented simultaneously and/or more variables are introduced, which leads to more complex representations and relations. This would mean that support in the form of integration or linking would be more necessary and have a larger impact (Lowe, 1999). It could, however, also be the case that the integration of representations hinders learning when the representation becomes too complex (see Bodemer et al., 2004). In that case, subjects in the I-DL condition would not perform as well as subjects in the S-DL condition. To assess this effect, we measured the subjectively experienced complexity of the different parts of the learning environment.

3. Method

3.1. Subjects

Subjects were Dutch students from four middle vocational training schools. The subjects were between 16 and 18 years old. They were all taking a course in mechanical engineering. One hundred twenty-eight subjects started the experiment; 36 subjects missed the session of working with the learning environment and two subjects did not take part in the pretest, which resulted in 90 subjects participating in all three phases of the experiment. Subjects who worked with the learning environment were randomly assigned to one of the three experimental conditions.

This paper reports analyses done with 72 subjects. From the set of 90 subjects participating in all phases of the experiment (pretest, working with the learning environment, and posttest), 18 subjects were removed from the sample because they did not work through all five progression levels in the learning environment and, therefore, did not explore all parts of the simulated domain. Table 1 shows how the subjects were distributed over conditions and schools.

3.2. Materials

3.2.1. Computers

The experiments were conducted in computer classrooms with IBM compatible Pentium III 450 MHz processors and 256 MB RAM computers. During the experiments, all subject actions with the computer program were logged automatically.

Table 1
Distribution of subjects per condition

School	Condition			Row total
	S-NL	S-DL	I-DL	
1 (m/f)	4 (4/0)	7 (7/0)	6 (6/0)	17 (17/0)
2 (m/f)	5 (5/0)	4 (4/0)	5 (5/0)	14 (14/0)
3 (m/f)	7 (6/1)	6 (4/2)	5 (3/2)	18 (13/5)
4 (m/f)	8 (8/0)	7 (7/0)	8 (8/0)	23 (23/0)
Column total	24 (23/1)	24 (22/2)	24 (22/2)	72 (67/5)

m = male, f = female.

3.2.2. SimQuest learning environment Moment

Subjects worked with the learning environment, Moment, that was built in the authoring environment SimQuest (de Jong, van Joolingen, Veermans, & van der Meij, 2005; van Joolingen & de Jong, 2003). Learners studied the physics topic of moments. The learning environment is based on guided discovery learning (de Jong & van Joolingen, 1998). Because it contains a simulation model that is not directly visible to the learner, the learner has to engage in discovery activities in order to learn about the properties of this model, and the learner is guided in the discovery process by 'cognitive tools' such as model progression, assignments, and explanations. Learners explore the simulation model by manipulating values of (input) variables and observing the behavior of other (output) variables. It is expected that learners acquire a deeper understanding of the domain by understanding the relations between the variables, and are able to transfer their knowledge to similar 'problems' in other (real) situations.

The learning environment has five progression levels. Table 2 gives an overview of these levels.

Learners start exploring a specific aspect of the domain by choosing an assignment from the menu. When opening an assignment, a corresponding simulation interface opens. Each assignment starts with a short description of an aspect of the domain, asks the learner to explore this aspect, and asks the learner to answer a question about it.

Fig. 1 shows an example of an assignment with corresponding simulation interface.

The simulation interface contains a maximum of five representations: (1) diagrammatic representation, (2) concrete representation, (3) numerical representation,[1] and (4, 5) two graphs (moment–force and moment–arm or moment–force and moment–height). Fig. 2 shows an example of an S-NL and the corresponding I-DL simulation interface.

Learners can manipulate the input variables in all types of representations either by using the provided sliders (concrete and diagrammatic representations) or by using the arrow keys (numerical representation). If a learner manipulates a slider or an arrow key, the corresponding changes are shown in the representations in real time. So, if a learner moves the force-slider, the element representing force is updated continuously and immediately, as is the change in moment. Learners can compare situations by moving the slider back and forth between different states of the simulation.

Representations 1, 2, and 3 are representations that are usually found in textbooks. These are the basic types of representations presented to learners studying the domain of moments. They all support learners in getting insight into the domain from different perspectives. The concrete representation is a learner-controlled animation that provides the learner with a context for the simulated task. This representation links the learning material to a real life experience. In the first four levels (see Table 2) the concrete representation is an animation of an open-end spanner, because most of the learners in the target group have experience using this tool. In the fifth level of the learning environment, a hoisting crane is introduced, because this gave us the opportunity to introduce a new variable (height) and because we wanted to provide a new concrete context to give learners the opportunity to apply their knowledge in a new situation. In addition, this representation is less learner-controlled; it changes over time after pressing a start button. The diagrammatic representation helps learners go beyond the concrete situation to a more abstract understanding of the relation between the variables involved. By providing this type of representation, it is expected that learners can use their acquired understanding in new situations. Both the concrete and diagrammatic representations present the domain in a qualitative way. The numerical representation gives a quantitative view of the variables involved. The contribution of this representation is in showing the values of the variables to support the numerical relations between the variables. The graphs are provided to help learners predict relationships between the variables. In the graphs, any 'pictorial' similarity to the represented domain has disappeared; therefore, graphs represent the domain in a more abstract way than do the concrete and diagrammatic representations (Bernsen, 1994). The graphs, however, give the learner more direct information about the relations between the variables than do the other representations.

The representations in the S-NL learning environment are not linked. Within a representation, changing an input variable (e.g., length) leads to a real time update of an output variable (moment). However, changing values of variables in one representation does not lead to changes in the other representations. In this learning environment, learners need to relate representations themselves. Changing values of variables in one representation in the S-DL learning environment, by contrast, leads to changes in all representations. When, for example, a learner changes the value

[1] In progression level 1, the numerical representation contains sliders with the following indications: minimum, zero, and maximum. The values of input variables can be changed by these sliders.

Table 2
Overview of progression levels of learning environment

	Level				
	1	2	3	4	5
Complexity	Low	Low	High	High	High
Representations[a]	1–3	1–3	1–5	1–3	1–5
Context	Spanner	Spanner	Spanner	Spanner	Crane
Number of variables	3	3	3	6	4
Qualitative/quantitative	Qualitative	Quantitative	Quantitative	Quantitative	Quantitative
Number of assignments	7	7	6	3	7

[a] See Fig. 2.

of the force in the numerical representation, not only the value of the moment in the numerical representation but also the force and moment in all other representations change accordingly. In the I-DL learning environment, the diagrammatic, concrete, and numerical representations (representations 1, 2, and 3 shown in Fig. 2) are integrated. These representations are placed 'on top of each other,' resulting in one representation showing these three representations in an integrated format. Of the five representations used, only the diagrammatic, concrete, and numerical representations could be integrated. The diagrammatic and realistic representations could easily be integrated because they share the same spatial properties. The numerical representations were also integrated because they could be placed near the objects of the other two representations. But, because the formats of the graphs differ from the diagrammatic and concrete representations, they could not be integrated and, therefore, are represented separately in this learning environment. We chose to dynamically link the graphs and the integrated representation, because the other representations are also dynamically linked by integrating them into one representation.

In all learning environments, color coding is used to indicate similar variables. Force is colored red, length is colored green, and moment is colored blue in all but the concrete representation.

The learning environment has low and high complexity parts (see Table 2). In the low complexity part, learners explore moment caused by force and length by investigating the behavior of moment on a bolt caused by a force on an open-end spanner. They do this in a qualitative (level 1) and a quantitative (level 2) way. Levels 3–5 form the high complexity part of the learning environment. In level 3, a moment–force and a moment–arm graph are introduced. In level 4, a second force is introduced, resulting in more complex representations. In this level, the graphs are not used in order to avoid cognitive overload. Therefore, the number of representations in level 4 is three. In level

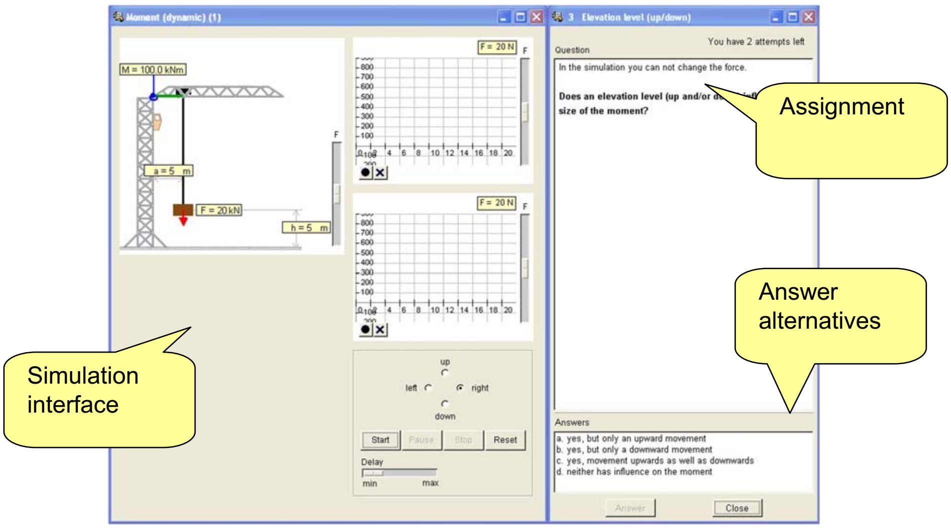


Fig. 1. Example of Moment assignment (representations from the I-DL condition).

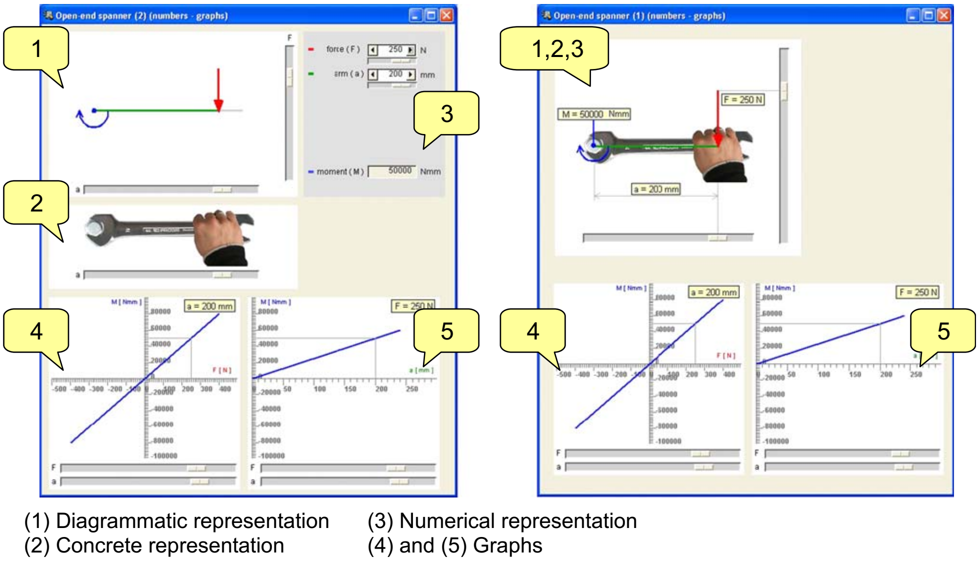


Fig. 2. Example of simulation interface. Left: representations from the S-NL and S-DL conditions. Right: representations from the I-DL condition.

5, learners explore moment caused by force, length and height by investigating the moment on a hoisting crane caused by a load. In this level there are four variables and five representations.

3.2.3. Tests and questionnaires

A paper-and-pencil pretest assessed the subjects' prior domain knowledge. It was administered a week before the subjects worked with the learning environment. A paper-and-pencil posttest was administered directly after they worked with the learning environment. Both tests consisted of 38 multiple-choice items with four answer possibilities, divided into three item types: 7 items on subject matter content, 17 items with transfer problems, and 14 items on the translation between representations. The domain items tested the subjects' domain knowledge. The content of the items was analogous to the content of the learning environment. Transfer items tested the ability of the subjects to apply their acquired knowledge in new situations: new contexts and relations between variables that were not asked for in the learning environment but could be derived from the domain knowledge. These items were included because the goal of scientific discovery learning is not only to help subjects acquire domain knowledge but also to enable them to apply their knowledge in new situations. The representation items tested the subjects' ability to relate and translate between different representational formats.

For each pretest and posttest item, a subject received a score of 1 if the item was answered correctly or a score of 0 if the answer was incorrect. The maximum score was 38. Fig. 3 shows examples of three test items.

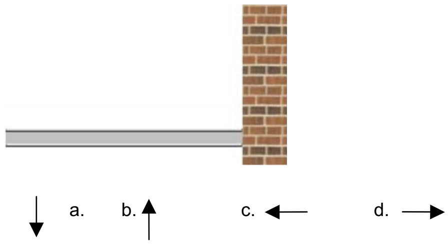
The posttest differed slightly from the pretest in that it contained minor changes in the item order and the order of the answer alternatives. Because subjects did not know which items were changed, they could not rely on a memory strategy.

An electronic questionnaire based on Swaak's S.O.S. scale (Swaak, 1998) was used to assess the subjects' opinions on the learning environment and the domain. The questionnaire contained seven questions (Q1–7) that asked subjects to rate topic, simulation, and assignment complexity. In particular it asked subjects to score the topic as easy, average, or difficult (Q1), whether they found working with the simulation easy, average, or difficult (Q2), and whether they found the assignments clear (Q3) and useful (Q4) (yes or no). Additionally, subjects were asked if they could always find the arm (Q5), force (Q6), and moment (Q7) in the simulation (yes or no). The questionnaire was given five times to subjects while they worked with the learning environment, after the last assignment of each progression level. Subjects had to complete the questionnaire before they could continue. For the first two questions in the electronic questionnaire, a three point scale was used. Possible answers were: easy, average, or difficult, which were coded as 1, 2, or 3, respectively. For the other five questions, a two point scale was used. Possible answers were: yes or no, which were coded as 1 or 2, respectively.

1. If you tighten a bolt with an open-end spanner, then where is the moment the largest?

a. At the bolt
b. Between the hand and the bolt
c. At the hand
d. At the end of the open-end spanner

2. What is the direction of the force if the moment is positive?



3. In the picture you see a hand exercising a negative force on an open-end spanner.

Which of the following figures is the right reproduction of length, force, and moment?

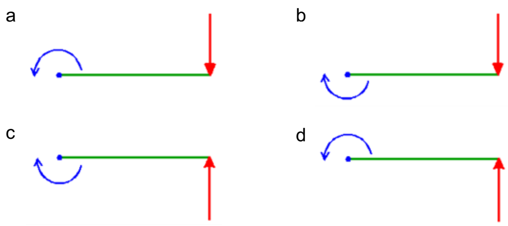


Fig. 3. Examples of (1) domain, (2) transfer, and (3) representation items.

3.3. Procedure

The experiments were held at the four participating schools and consisted of three experimental sessions: pretest, working with the learning environment, and posttest.

The pretest session lasted a maximum of 45 min. Subjects were informed about the experiment and were told that the test measured their prior knowledge of force, arm, and moment. Subjects were asked whether they were already familiar with the term ‘moment,’ and got a brief description if they were not. Subjects were asked to answer all test items, even if they were unsure about the right answer.

The learning environment session took place a week after the pretest session and lasted a maximum of 1 h. Subjects were randomly assigned to one of the three conditions using their seating placement. They did not know beforehand in which condition they were going to be placed. At the start of the session, the subjects were told that their task was to learn with the learning environment. They worked on their own and could question the teacher or experiment leader about operating the learning environment. The experiment leader gave a short introduction on how to control the learning environment. The electronic questionnaire had to be filled in five times while working with the learning environment. The subjects were asked to work through all of the progression levels and were asked to do all of the assignments. When they were ready, they could ask to do the posttest.

The posttest took place directly after the learning environment session. The subjects could work a maximum of 45 min on this test. The subjects were not allowed to use the learning environment during the test and were asked to fill in all test items, even if they were unsure about the right answer.

4. Results

4.1. Pretest and posttest

The overall mean score on the pretest was 22.15 out of 38 multiple-choice items (SD = 3.97). These data indicate that the subjects had some prior knowledge in the domain. The overall mean score on the posttest was 25.00 out of 38 multiple-choice items (SD = 4.52). Table 3 shows the means and standard deviations of the scores on the three different item types in the pretest and posttest.

A repeated measures ANOVA showed that the overall posttest score of the 72 subjects was significantly better than the overall pretest score ($F(1,71) = 30.90$, $p < .01$). Repeated measures ANOVAs showed that the posttest scores on domain and representation items were significantly better than the pretest scores on these item types ($F(1,71) = 17.96$, $p < .01$ and $F(1,71) = 36.34$, $p < .01$). Scores on transfer posttest items were not better than pretest scores ($F(1,71) = 2.08$, $p = .15$). Table 4 shows the means and standard deviations of the pretest and posttest scores for the three different item types per condition.

One way ANOVAs showed no significant differences between the experimental conditions on pretest domain scores, transfer scores, and representation scores ($F(2,69) = .76$, $p = .47$; $F(2,69) = .922$, $p = .40$; $F(2,69) = .24$, $p = .79$). This means that subjects in the experimental conditions did not differ in prior knowledge. One way ANOVAs showed no significant relations between overall pretest scores and schools and overall pretest scores and gender ($F(3,68) = .98$, $p = .41$ and $F(1,70) = 1.91$, $p = .17$).[2] Therefore, there was no need to correct for these variables.

One way ANOVAs on the posttest scores showed the following results. A significant difference was found between conditions on domain item scores ($F(2,69) = 3.23$, $p < .05$). Tukey HSD post hoc analyses showed that subjects in the I-DL condition scored significantly better than those in the S-NL condition. Differences between the other conditions were not significant. No difference was found between conditions on transfer item scores ($F(2,69) = .26$, $p = .78$). A trend was found for the difference between conditions on representation item scores ($F(2,69) = 2.63$, $p = .08$). Tukey HSD post hoc analyses showed that the trend on representation items was in favor of the I-DL condition compared to the S-NL condition.

4.2. Posttest scores based on complexity

The learning environment was divided into low and high complexity parts. Based on this distinction, we divided the corresponding posttest items into low and high complexity categories. Twenty-three posttest items corresponded to the low complexity part and 15 items corresponded to the high complexity part. A one way ANOVA showed a significant difference between the experimental conditions on the overall scores on the posttest items corresponding to the high complexity part of the learning environment ($F(2,69) = 3.37$, $p < .05$). However, Tukey HSD post hoc analyses did not show where the differences were found. No difference was found on the low complexity part. Comparing the posttest scores between conditions for the different item types (domain, transfer, and representation) based on complexity, a one way ANOVA showed a significant difference between the experimental conditions on the high complexity domain item posttest scores ($F(2,69) = 1.54$, $p < .05$). Tukey HSD post hoc analyses showed that subjects in the I-DL condition scored significantly better than subjects in the S-NL condition. No significant differences were found between conditions on the other item types.

[2] For all 126 subjects: $F(3,122) = .46$, $p = .71$ and $F(1,124) = .35$, $p = .56$.

Table 3
Means and standard deviations of pretest and posttest scores

	Pretest			Posttest		
	Mean	SD	%	Mean	SD	%
Domain items (max. 7)	4.94	1.28	71	5.63	1.17	80
Transfer items (max. 17)	9.68	2.10	57	10.18	2.27	60
Representation items (max. 14)	7.53	2.10	54	9.18	2.39	66
Total (max. 38)	22.15	3.97	58	25.00	4.52	66

$n = 72$.

4.3. Experienced domain complexity

The experienced domain complexity was measured by the questionnaire question: "I find the topic at this moment: easy, average, or difficult." The question appeared at the end of each of the five progression levels. We calculated a mean score for the five answers given (see Table 5, question 1). A one way ANOVA showed a significant difference between the experimental conditions on experienced domain complexity ($F(2,69) = 3.45$, $p < .05$). Tukey HSD post hoc analyses showed that subjects in the I-DL condition experienced the domain as easier than subjects in the S-DL condition. Differences between the other conditions were not significant.

Subjects experienced the domain complexity differently throughout the learning environment. If subjects scored higher on 'easy,' the complexity was rated as low. If subjects scored higher on 'average' and 'difficult,' the complexity was rated as high. Progression levels 1 and 2 were experienced as low complexity. Progression levels 3–5 were experienced as high complexity. Based on this experienced complexity, we calculated the means of scores from the first and second appearances of the questionnaire (domain experienced as low complexity) and the means of scores from the third, fourth and fifth appearances of the questionnaire (domain experienced as high complexity). A one way ANOVA on these mean scores showed a significant difference between the experimental conditions on the high complexity part of the learning environment ($F(2,69) = 4.11$, $p < .05$). Tukey HSD post hoc analyses showed that subjects in the I-DL condition experienced the domain as easier than subjects in the other conditions in the high complexity part of the learning environment. No significant differences were found between the conditions in the low complexity part ($F(2,69) = .83$, $p = .44$).

4.4. Experienced learning environment complexity

The experienced learning environment complexity was measured by the questionnaire question: "I find working with the simulation at this moment: easy, average, or difficult." The question appeared at the end of each of the five

Table 4
Means (standard deviations) of pretest and posttest scores per condition

	Condition					
	S-NL		S-DL		I-DL	
Pretest						
Domain items (max. 7)	5.20	(1.14)	4.83	(1.27)	4.79	(1.44)
Transfer items (max. 17)	9.21	(2.60)	9.88	(1.99)	9.96	(1.57)
Representation items (max. 14)	7.29	(2.05)	7.59	(2.26)	7.71	(2.05)
Total (max. 38)	21.71	(4.36)	22.29	(4.20)	22.46	(3.41)
Posttest						
Domain items (max. 7)	5.25	(1.39)	5.58	(1.21)	6.08	(.72)
Transfer items (max. 17)	10.25	(2.42)	9.92	(2.22)	10.37	(2.24)
Representation items (max. 14)	8.54	(2.22)	8.95	(2.69)	10.04	(2.05)
Total (max. 38)	24.04	(4.94)	24.46	(4.51)	26.50	(3.83)

$n = 72$.

Table 5
Means (standard deviations) of questionnaire answers

	Condition					
	S-NL		S-DL		I-DL	
Question 1[a]	1.65	(.41)	1.78	(.47)	1.46	(.38)
Question 2[a]	1.55	(.36)	1.62	(.47)	1.23	(.32)
Question 3[b]	1.19	(.23)	1.40	(.31)	1.09	(.19)
Question 4[b]	1.18	(.35)	1.38	(.39)	1.26	(.33)
Question 5[b]	1.13	(.22)	1.15	(.23)	1.03	(.09)
Question 6[b]	1.17	(.27)	1.17	(.26)	1.03	(.13)
Question 7[b]	1.14	(.23)	1.20	(.28)	1.03	(.13)

[a] $n = 72$.
[b] $n = 69$ (three subjects were removed because they answered the question less than three times).

progression levels. We calculated a mean score for the five answers given (see Table 5, question 2). A one way ANOVA showed a significant difference between the experimental conditions in experienced complexity when working with the learning environment ($F(2,69) = 6.87$, $p < .01$). Tukey HSD post hoc analyses showed that subjects in the I-DL condition experienced working with the learning environment as easier than subjects in the linked and separate conditions. Differences between the other conditions were not significant.

A one way ANOVA on these mean scores showed a significant difference between the experimental conditions in the high complexity part of the learning environment ($F(2,69) = 6.80$, $p < .01$). Tukey HSD post hoc analyses showed that subjects in the I-DL condition experienced working with the learning environment as easier than subjects in the other conditions in the high complexity part of the learning environment. A trend was found between the conditions in the experienced low complexity part ($F(2,69) = 2.72$, $p = .07$). Tukey HSD post hoc analyses showed that subjects in the I-DL condition experienced working with the learning environment as easier than subjects in the S-DL condition. This confirms our presumption that progression levels 1 and 2 had a low complexity and that levels 3–5 had a high complexity.

4.5. Usefulness of assignments and finding variables

A mean score for each of questions 3–7 was calculated for all appearances of the questionnaire (see Table 5). One way ANOVAs showed a significant difference between conditions on questions 3 (I find the assignments clear: yes, no) and 7 (I can find the moment in the simulation everywhere: yes, no) ($F(2,66) = 9.42$, $p < .01$; $F(2,66) = 3.86$, $p < .05$).[3] Tukey HSD post hoc analyses showed that subjects in the S-DL condition experienced the assignments as less clear than subjects in the other conditions and that subjects in the I-DL condition could find the moment more frequently than those using the linked version, but not more frequently than those using the separate version. No differences were found between conditions on questions 4 (I find the assignments useful: yes, no), 5 (I can find the arm in the simulation everywhere: yes, no), and 6 (I can find the force in the simulation everywhere: yes, no).

4.6. Subjects' interaction with the learning environment

Table 6 shows the means and standard deviations of the time subjects spent on working with the learning environment, the number of assignments they did, and the calculated time spent per assignment. A one way ANOVA showed a significant difference between the experimental conditions on the average time spent working with the simulation ($F(2,69) = 4.23$, $p < .05$). Tukey HSD post hoc analyses showed that subjects in the S-DL condition worked for a significantly shorter duration in the environment than subjects in the other conditions. One way ANOVAs on the number of assignments done and time per assignment showed no differences between conditions.

Because time on task may have affected posttest scores, ANCOVAs were performed with time on task as covariate, showing similar results as the reported ANOVAs. A significant difference was found between conditions on domain item scores ($F(2,68) = 4.35$, $p < .05$). No difference was found between conditions on transfer item scores

[3] Three subjects were removed from the sample because they answered the questionnaire less than three times.

Table 6
Means (standard deviations) of total time, number of assignments, and time per assignment

	Condition						Total	
	S-NL		S-DL		I-DL			
Total time (min)	39.71	(6.71)	33.79	(7.72)	40.17	(10.53)	37.89	(8.85)
Assignments done (max. 31)	29.71	(2.14)	28.54	(3.64)	30.04	(1.65)	29.43	(2.66)
Time per assignment (s)	80.34	(12.70)	71.62	(16.38)	80.08	(19.64)	77.34	(16.76)

($F(2,68) = .04$, $p = .96$). A trend was found for the difference between conditions on representation item scores ($F(2,68) = 2.58$, $p = .08$).

5. Discussion

The aim of this study was to examine ways to support learners in the translation between different representations in a simulation-based learning environment. Three versions of the same simulation-based learning environment were compared: a learning environment with separate, non-linked representations (S-NL condition), a learning environment with separate, dynamically linked representations (S-DL condition), and a learning environment with integrated, dynamically linked representations (I-DL condition). We expected that dynamic linking would free the subjects from mentally relating the representations and, therefore, we expected to find a larger learning effect for the S-DL learning environment compared to the S-NL version. We also predicted that the I-DL learning environment would lead to the best learning results as long as the integrated representations were not too complex for the subjects.

Overall, we found that subjects learned from working with the learning environment. Posttest scores were significantly better than pretest scores, but only on domain and representation items. We found that dynamic linking alone (S-DL condition) did not lead to better learning outcomes than non-linking. We found that subjects in the I-DL condition had the best scores on posttest domain items. They scored significantly better than subjects in the S-NL condition, but not better than those in the S-DL condition. A trend was found for representation items. The trend was again in favor of the I-DL condition, but again only in comparison with the S-NL condition.

As expected, complexity of the learning environment interacted with the effects of the experimental conditions. The differences seen on the domain items were only found on items that corresponded to the high complexity part of the learning environment, but not on the items that corresponded to the low complexity part. The contingency that the integrated representation could become too complex when more variables were introduced was not supported by our data. To the contrary, subjects in the I-DL condition experienced level 4 of the learning environment, where a second force was introduced, as easier than subjects in the other conditions. It looks like finding the relations between representations in the separate and linked conditions was more complex than the complexity of the integrated representation. In the S-NL and S-DL conditions, the subjects had to relate nine variables that were presented separately. Linking these variables helped the subjects to find their relations, but not enough to experience this as less difficult than the I-DL condition.

The fact that we did not find better results for dynamic linking in comparison with non-linking seems to be in contrast with other studies reporting positive effects of linking representations (e.g., Kozma, 2003; Kozma et al., 1996; Tsui & Treagust, 2003; Wu, Krajcik, & Soloway, 2001). There are, however, two issues that make these studies different from ours. First, in our study the S-DL condition differed from the S-NL condition only by the presence of *dynamic* linking. Apart from dynamic linking, we used color coding to relate representations, but this color coding was present in both the S-DL and S-NL conditions. Other studies (e.g., Kozma et al.) combined different ways to relate representations, including dynamic linking, and did not examine the effect of dynamic linking alone. Taken together, these results may suggest that in our case, color coding may have been sufficient and, because of that, the dynamic linking had no additional effect.

A second aspect of our learning environments that could have helped subjects to relate and translate representations could have been the instructional support in our environments. Instructional support, such as the assignments that we provided the subjects with, may highlight the correspondence between related representations (see e.g., Ardac & Akaygun, 2004). The assignments in our learning environments were the same for all conditions.

Where dynamic linking did not lead to better learning effects, integration plus linking did. We found significant differences on domain items, which indicates that the learners learn the domain better if the representations are integrated (and linked). This is in accordance with Chandler and Sweller (1991), who found that integrating instruction led to better learning results than separate instruction, as long as the materials chosen were unintelligible without mental integration. The trend found on representation items may indicate that integrating representations supports learners in relating different representations.

The results on the transfer items were not as expected. An important motivation to use multiple representations is that they should encourage learners to construct a deeper knowledge of a domain (e.g., Ainsworth, 1999; Petre, Blackwell, & Green, 1998). Petre et al. asserted that having to make the mental transference between representations (and possibly between paradigms) forces reflection beyond the boundaries and details of the first representation and an anticipation of correspondences in the second. The deeper level of cognitive processing can reveal glitches that might otherwise have been missed. We expected that by using multiple representations, subjects could transfer their knowledge of the domain presented in the learning environment to other, comparable, situations. However, we did not find significant differences on transfer items between the pretest and posttest. This could possibly be explained by the fact that subjects worked with the learning environment for a short period of time (the average learning time was 38 min) and therefore did not explore the domain deeply. Subjects, therefore, did not obtain enough insight into the relations between the domain variables to be able to transfer their knowledge to new situations. A second aspect could be that in the learning environments, the domain was presented in specific contexts. Representation 2 (see Fig. 2) showed the physics system under study; an open-end spanner with a hand tightening a bolt or a hoisting crane. It showed the domain in a way it would appear in a real-world situation. This representation was meant to constrain the interpretation of the other representations. A drawback could have been that subjects related the domain to the presented contexts too much and were therefore not able to transfer their new knowledge to other contexts.

Although their pretests and posttests contained transfer items, Wu et al. (2001), Ainsworth et al. (1997), and van Labeke and Ainsworth (2002) did not look explicitly for learning effects of (dynamically) linked multiple representations on these items. Tsui and Treagust (2003), doing research on genetics reasoning, found that some but not all subjects scored better on transfer items. When analyzing their interviews, they found that the subjects who did not improve showed no mindful interaction with the multiple representations in their learning environment. According to Tsui and Treagust, learners' mindfulness in interacting with multiple representations is a theme which appeared to be crucial in the development of genetics reasoning and in the transfer of that reasoning to new problem situations.

Like, for example, Petre et al. (1998), we believe that having to make mental translations between representations is a good way to acquire deeper knowledge in a domain. It is worthwhile to further investigate the effects of different types of support when offering learners multiple representations. Integrating representations looks promising, but, as Lowe (2004) also asserts, additional support is probably needed to let learners have mindful interaction with the representations.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education*, *33*, 131–152.

Ainsworth, S., & van Labeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction*, *14*, 241–255.

Ainsworth, S. E., Bibby, P. A., & Wood, D. J. (1997). Information technology and multiple representations: new opportunities—new problems. *Journal of Information for Teacher Education*, *6*, 93–104.

Ainsworth, S. E., & Peevers, G. J. (2003). *The interaction between informational and computational properties of external representations on problem-solving and learning*. Paper presented at the 25th Annual Conference of the Cognitive Science Society, Boston, Massachusetts, USA.

Ardac, D., & Akaygun, S. (2004). Effectiveness of multimedia-based instruction that emphasizes molecular representations on students' understanding of chemical change. *Journal of Research in Science Teaching*, *41*, 317–338.

Bernsen, N.O. (1994). *Modality theory in support of multimodal interface design*. Paper presented at the AAAI Spring Symposium on Intelligent Multi-Media Multi-Modal Systems, Stanford.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, *14*, 325–341.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*, 293–332.

de Jong, T., Ainsworth, S., Dobson, M., van der Hulst, A., Levonen, J., & Reimann, P., et al. (1998). Acquiring knowledge in science and mathematics: the use of multiple representations in technology based learning environments. In M. van Someren, P. Reimann, H. Boshuizen, & T. de Jong (Eds.), *Learning with multiple representations* (pp. 9–41). Oxford: Elsevier Science.

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*, 179–201.

de Jong, T., van Joolingen, W. R., Veermans, K., & van der Meij, J. (2005). Authoring discovery learning environments: in search for re-usable components. In J. M. Spector, & D. A. Wiley (Eds.), *Innovations in instructional technology: Essays in honor of M. David Merrill* (pp. 11–28). Mahwah, NJ: Erlbaum.

van Joolingen, W. R., & de Jong, T. (2003). Simquest: authoring educational simulations. In T. Murray, S. Blessing, & S. Ainsworth (Eds.), *Authoring tools for advanced technology educational software: Toward cost-effective production of adaptive, interactive, and intelligent educational software* (pp. 1–31). Dordrecht: Kluwer Academic Publishers.

Kaput, J. J. (1989). Linking representations in the symbol systems of algebra. In S. Wagner, & C. Kieran (Eds.), *Research issues in the learning and teaching of algebra* (pp. 167–194). Hillsdale, NJ: Erlbaum.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, *13*, 205–226.

Kozma, R. B., Russell, J., Jones, T., Marx, N., & Davis, J. (1996). The use of multiple, linked representations to facilitate science understanding. In S. Vosniadou, R. Glaser, E. DeCorte, & H. Mandel (Eds.), *International perspectives on the psychological foundations of technology-based learning environments* (pp. 41–60). Hillsdale, NJ: Erlbaum.

van Labeke, N., & Ainsworth, S. (2001). *Applying the deft framework to the design of multi-representational instructional simulations*. Paper presented at the AIED'2001—10th International Conference on Artificial Intelligence in Education, San Antonio, Texas.

van Labeke, N., & Ainsworth, S. E. (2002). Representational decisions when learning population dynamics with an instructional simulation. In S. A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems. Proceedings of the 6th international conference ITS 2002* (pp. 831–840). Berlin: Springer-Verlag.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–99.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, *14*, 225–244.

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphs. *Learning and Instruction*, *13*, 157–176.

Lowe, R. K. (2004). Interrogation of a dynamic visualisation during learning. *Learning and Instruction*, *14*, 257–274.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*, 312–320.

McKenzie, D., & Padilla, M. (1984). *Effect of laboratory activities and written simulations on the acquisition of graphing skills by eighth grade students*. Unpublished manuscript, New Orleans, LA.

Mokros, J. R., & Tinker, R. F. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, *24*, 369–383.

Petre, M., Blackwell, A. F., & Green, T. R. G. (1998). Cognitive questions in software visualization. In J. Stasko, J. Domingue, M. Brown, & B. Price (Eds.), *Software visualization: Programming as a multi-media experience* (pp. 453–480). MIT Press.

Ploetzner, R., & Lowe, R. (2004). Dynamic visualizations and learning. *Learning and Instruction*, *14*, 235–241.

Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work? *International Journal of Human-Computer Studies*, *45*, 185–213.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, *13*, 227–237.

Swaak, J. (1998). *What-if: Discovery simulations and assessment of intuitive knowledge*. Enschede, The Netherlands: Twente University.

Sweller, J. (1988). Cognitive load during problem-solving: effects on learning. *Cognitive Science*, *12*, 257–285.

Sweller, J. (1989). Cognitive technology: some procedures for facilitating learning and problem-solving in mathematics and science. *Journal of Educational Psychology*, *81*, 457–466.

Tabachneck, H. J. M., Leonardo, A. M., & Simon, H. A. (1994). How does an expert use a graph? A model of visual and verbal inferencing in economics. In A. Ram, & K. Eiselt (Eds.), *16th annual conference of the cognitive science society* (pp. 842–847). Hillsdale, NJ: Erlbaum.

Tabbers, H.K., Martens, R.L., & Van Merriënboer, J.J.G. (2000). *Multimedia instructions and cognitive load theory: Split-attention and modality effects*. Paper presented at the National Convention of the Association for Educational Communications and Technology, Long Beach, CA.

Tsui, C. Y., & Treagust, D. F. (2003). Genetics reasoning with multiple external representations. *Research in Science Education*, *33*, 111–135.

Wu, H. K., Krajcik, J. S., & Soloway, E. (2001). Promoting understanding of chemical representations: students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching*, *38*, 821–842.