The Effects of Representational Format in Simulation-Based Inquiry Learning

Bas Kollöffel, Ton de Jong, & Tessa H.S. Eysink

University of Twente, The Netherlands

Abstract

The format of external representations affects learning and problem solving processes. In general, pictorial representations are supposed to lead to superior performance and to reduce cognitive load. However, the effects are domain-specific. The current paper describes a study on the effects of representational format in the domain of combinatorics and probability theory. Three different formats (pictorial, arithmetical, and textual) are contrasted in a pre-test post-test design and evaluated in terms of problem solving facilitation, transfer, cognitive load, and interactivity. It was found that arithmetical representations led to superior performance, whereas pictorial representations led to poorer performance. Moreover, pictorial representations were associated with higher levels of cognitive load.

Introduction

The format of external representations (symbols, diagrams, etcetera) is known to play a critical role in learning and understanding (Ainsworth & Loizou, 2003; Cheng, 1999; Zhang, 1997). In cognitive sciences, external representations are usually classified into two categories: nonverbal (pictorial) and verbal representations. Verbal representations include both natural and formal (e.g., arithmetical) languages (Klein, 2003; Paivio, 1990). In general, pictorially represented material is associated with superior performance compared to textual material (Goolkasian, 2000), for example in deductive reasoning (Bauer & Johnson-Laird, 1993), spatial reasoning (Boudreau & Pigeau, 2001), and sentence verification tasks (Goolkasian, 1996). Many researchers observe that, despite a vast amount of research on the facilitative effects of pictures, not much is known about the cognitive processes underlying these effects (Cheng, 1999; Glenberg & Langston, 1992; Goolkasian, 2000; Scaife & Rogers, 1996; Vekiri, 2002; Winn, 1993). Moreover, some researchers argue that the effects of representations cannot readily be generalized to other domains (Cheng,
Lowe, & Scaife, 2001; Scaife & Rogers, 1996; Zhang, 1997). Leung, Low, and Sweller (1997) argue that the advantage of a particular format in learning is not universal, but depends on a complex interaction between the student’s ability, the nature of the material, and the practice time available. Modern computer technology in instruction increasingly provides learners with interactive representations. In simulation-based instruction for example, learners are enabled to manipulate representations in order to explore the concepts and principles underlying the domain (de Jong & van Joolingen, 1998). Research on the effects of external representations should therefore not only focus on the effects of representational format (pictures, texts, formulas, etcetera), but also on the cognitive processes underlying these effects, including the role of interactivity of representations (Dobson, 1999; Rogers, 1999; Scaife & Rogers, 1996, 2005). The current study concentrates on the relation between the format of interactive external representations in computer-based simulations on the one hand, and learning processes and outcomes on the other.

Representations

Representations and cognition

Leading views in cognitive science, for example dual coding theory (Paivio, 1990), dual channel assumption (Mayer, 2003), and Baddeley’s (1997) visuospatial sketchpad and phonological loop, postulate that external representations are processed, encoded, and stored by two different cognitive systems, one for nonverbal information and one for verbal information. However, not only the cognitive systems but also the properties of the representation are assumed to influence which information is attended to and how people tend to organize, interpret, and remember the information presented. Larkin and Simon (1987) state that the value of representations depends on two factors: informational and computational efficiency. Informational efficiency refers to how representations organize information into data structures. Computational efficiency refers to the ease and rapidity with which inferences can be drawn from a representation. The authors argue that even when a picture and a text are informationally equivalent, meaning that all of the information in one representation can be inferred from the other and vice versa, the picture is often more effective because inferences can be drawn easier and quicker from pictures. Furthermore, they argue that in pictures the explicitness of information and the preservation of topological and structural relations facilitate the extraction of information required for interpretation.

To illustrate different ways of organizing information into data structures, we will use the example of a credit card that is protected by a four-digit PIN-code, 5526. If one does not know the PIN-code, the code can be guessed. Guessing the digits that are in the code is necessary but not sufficient. One also needs to determine the particular order in which these digits occur in the code. Therefore, guessing the code can be conceived as finding a pathway through a four-stage process, starting with guessing the first digit of the code (stage 1), then proceeding to the next stage, that of guessing the second digit, and so on. Succeeding in finding the right pathway stands or falls with selecting the right digit in each stage. Selecting the wrong digit in one or more stages will wreck the attempt to find the correct pathway. Each subsequent stage depends on the outcome of the
previous one and therefore finding the right pathway depends on the outcome of each of the four subsequent stages. There are several ways of representing the stages, the possible and desired outcomes, and pathway through the stages. Below, three possible representations will be shown, a pictorial representation (nonverbal), a textual (verbal-naturalistic), and an arithmetical (verbal-formal).

![Figure 1. Tree diagram representing finding the PIN-code](image)

In Figure 1 finding the PIN-code is represented pictorially as a tree diagram. Tree diagrams can specify every possible outcome and are considered especially effective in assessing the arithmetical probability of various outcomes (Halpern, 1989). In pictures or diagrams, information is indexed by a two-dimensional location in a plane, preserving explicitly information about topological and structural relations (cf. Larkin & Simon, 1987). Figure 1 schematically displays all possible pathways almost literally as a road map (to be read from left to right), with the arrows indicating all possible pathways that could be followed (with the bold arrows indicating the desired pathway). Therefore, the arrows explicitly express the topological relations. Structural relations can be inferred from information like 1) there is only one desired pathway, 2) the ratio of the desired pathway and the total of possible pathways, and 3) in each stage the set from which a digit can be selected is equally big and consisting of nominally equal elements as in the other stages. A drawback of the tree diagram is that the emphasis on topological and structural relations is at the expense of arithmetical power. Rather than explicating the exact probability of finding the PIN-code (that is, the ratio of the number of desired pathways and the total number of possible pathways), a tree diagram affords more an (almost intuitive) impression of this ratio, rather than the exact number.

A second possible way of representing the PIN-code problem is displayed in Figure 2. It is possible to construct a tree diagram like Figure 1 on basis of the text in Figure 2. Reversely, it is possible to create a text like that in Figure 2 on basis of the tree diagram in Figure 1. Therefore, both figures are informationally equivalent. However, the textual representation emphasizes other relational features. In verbal representations like the one in Figure 2, information is organized sequentially, preserving temporal and logical relations (cf. Larkin & Simon, 1987), rather than topological and structural relations.
Figure 2. Text representing finding the PIN-code

The use of natural language facilitates processes of relating information in the text to everyday experiences and situations. On the other hand, problems with the comprehension of text may hamper the problem solving performance (Koedinger & Nathan, 2004; Lewis & Mayer, 1987; Nathan, Kintsch, & Young, 1992). Although informationally equivalent, the pictorial and textual representation may differ with regard to computational equivalence. What distinguishes a text from merely a set of sentences, is that a text is cohesive. Sentences build upon each other and refer to each other. In order to understand the meaning of a text while reading, the reader has to keep certain elements of the text highly activated in working memory (Glenberg, Meyer, & Lindem, 1987). By comparing newly encountered elements with the elements held in working memory, the reader can determine if the elements refer to the same entity. The process of keeping elements activated burdens working memory considerably. Sweller, Van Merriënoer, and Paas state that “any interactions between elements held in working memory themselves require working memory capacity, reducing the number of elements that can be dealt with simultaneously” (1998, p.252). Leung and his colleagues (1997) argue that extensive reading induces heavy cognitive load and harms problem solving. Even though Figure 2 is informationally equivalent to Figure 1, keeping many elements in working memory might be at the expense of computational efficiency.

A third way of representing the PIN-code problem is displayed in Figure 3. From this arithmetical representation it is much harder to (re)construct representations like the tree diagram and the text and vice versa. Therefore, the informational equivalence strongly depends on the knowledge of the viewer about the meaning of arithmetical representations. One needs to know for example the conceptual meaning of the multiplication sign, which indicates that for the pathway the outcomes of the individual stages are not independent (if they were, the multiplication signs needed to be replaced with plus signs). This illustrates exactly the trouble with arithmetical representations. In arithmetical representations the underlying principle or concept is not as explicit as in pictorial and textual representations. Most learners tend to view mathematical symbols (e.g. multiplication signs) purely as indicators of which operations need to be performed on adjacent numbers. They tend to see arithmetical representations as procedures to find a solution, rather than reflections of principles and concepts underlying these procedures (Atkinson, Catrambone, & Merrill, 2003; Cheng, 1999; Greenes, 1995; Nathan et al., 1992; Niemi, 1996; Ohlsson &
Rees, 1991). They, therefore, easily lose sight of the meaning of their actions. If this is the case, processing formal notations becomes an end in itself (Cheng, 1999). Moreover, learning arithmetic procedures without conceptual understanding tends to be error prone, easily forgotten, and not readily transferable (Ohlsson & Rees, 1991).

\[
\rho (\text{PIN} = 5526) = \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10}
\]

Figure 3. Equation representing finding the PIN-code

Using arithmetical representations does not always mean a disadvantage for learners. Leung, Low, and Sweller (1997) showed that textual representations are not always superior to arithmetical representations. In a series of four experiments, they found an advantage of an equation format over a verbal format for more able students. Furthermore, they demonstrated an interaction between acquisition period and format. For short periods of acquisition the verbal format was found to be superior to the equation format, but when the acquisition period was extended, this trend was reversed in favor of the equation format.

Which representation is better?: Evaluating the effects of representational format

Based on the properties of the different representational formats and on the empirical findings, drawing a conclusion in favor of one of the formats is challenging. Below, some guidelines will be formulated for evaluating the effects of representational format and the cognitive processes underlying these effects. These guidelines, which are partly inspired by the external cognition approach (Scaife & Rogers, 1996, 2005), will serve as a framework for the current study of the effects of representational format that will described later in this article. Following Larkin and Simon (1987), baseline for drawing valid conclusions from comparisons of representations is by comparing representations that are informationally equivalent, at least for as far as possible. The criteria that will be used are: facilitation of problem solving, transfer, cognitive load, and interactivity.

Facilitation of problem solving

Facilitation of problem solving refers to the extent to which problem solving is made easier or more difficult when using different representations. Problem solving schemata may contain conceptual, procedural, and situational knowledge (de Jong & Ferguson-Hessler, 1986, 1996). Conceptual knowledge is “implicit or explicit understanding of the principles that govern a domain and of the interrelations between units of knowledge in a domain” (Rittle-Johnson, Siegler, & Alibali, 2001, p.346). Conceptual knowledge develops by establishing relationships between pieces of information or between existing knowledge and new information. Procedural knowledge is “the ability to execute action sequences to solve problems” (Rittle-Johnson
et al., 2001, p.346). Knowledge that enables identification of the problem situational knowledge (de Jong & Ferguson-Hessler, 1996). They describe it as “knowledge about situations as they typically appear in a particular domain” (p.106). Situational knowledge enables learners to sift relevant features out of the problem statement and to supplement information in the statement. Therefore, in order to determine the extent to which problem solving is made easier or more difficult when using different representations it is necessary not only to take into account general knowledge measures, but to differentiate between situational, conceptual, and procedural knowledge.

Transfer

Transfer is the ability to apply problem solving schemata to other problems that may or may not be similar to the original problems where the schemata were acquired. The representational format used in the initial problems may influence the extent to which transfer is possible. For example, earlier we have argued that, when confronted with arithmetical representations (e.g. equations) learners tend to focus on procedures rather than on the principles and concepts underlying these procedures. This may facilitate solving similar problems but may hamper solving dissimilar problems. The properties of different representational formats may therefore affect transfer.

Cognitive load induced by representational format

This criterion refers to the amount of cognitive load induced by differential representations when solving problems and therefore echoes Larkin and Simon’s (1987) concept of computational efficiency. Cognitive load theory (CLT) assumes that the capacity of working memory is limited. The load of working memory is assumed to be affected by the intrinsic nature of the material (intrinsic cognitive load), by the way in which the material is presented (extraneous cognitive load), and by the effort that contributes to the construction of schemas (germane cognitive load) (Paas, Renkl, & Sweller, 2004; Sweller et al., 1998). Extraneous load is considered unnecessary and ineffective because it is imposed by information and activities that do not contribute to the processes of schema construction and automation. Germane load is thought to be effective because it actually fosters these processes. Sweller and his colleagues (1998) argue that extraneous and germane load can be changed by instructional interventions. In line with this, Van Bruggen, Kirschner, and Jochems (2002) suggest that basic questions when assessing the implications of representations from a cognitive load perspective are whether a representation will lead to changes in extraneous and germane cognitive load.

Interactivity

Some electronic learning environments enable learners to interact with representations. In simulation-based instruction for example, learners can manipulate representations in order to explore the concepts and principles underlying the domain. (de Jong & van Joolingen, 1998). Most of these systems offer the possibility to keep track of user actions in the learning environment, including the user’s interactions with representations. The log files, that is the files in which the tracking information is stored, provide insight in the interaction between the learner and the learning environment. Examples of data that can be
retrieved from log files are the time that has been spent in the
learning environment as a whole, but also in specific parts of the
environment, the paths the user followed through the environment,
selections made by the user, settings of variables in simulations,
changes made by the user in these settings, etcetera. These data may be
helpful to analyze differences in user responses to different
representational formats. For example, does representational format
influence exploratory behavior in simulations.

Research question

Representational format may play a critical role in the learning
processes and learning outcomes as well. The focus of the current study
is to determine the effects of representational codes in simulation-
based inquiry learning. Our research question is:

What is the effect of representational format on knowledge construction
processes in learning with interactive representations?

The domain of the instruction is combinatorics and probability theory.
The PIN-code problem given in previous sections is an example of
problems encountered in this domain. The research question will be
answered on the hand of the criteria that were formulated previously:
facilitation of problem solving, transfer, cognitive load, and
interactivity. Results from empirical studies on the effects and the
argument that the effects that have been found cannot readily be
generalized to other domains, hamper the formulation of well-grounded
hypotheses with regard to effects of representations in the current
study, but on basis of the properties of the different representational
formats some expectations will be formulated.

With regard to facilitation of problem solving a distinction is
made between conceptual, procedural, and situational knowledge. First,
the topological and structural relations emphasized in pictorial
representations are both in the domain of conceptual knowledge.
Therefore, pictorial representations are expected to lead to enhanced
levels of conceptual knowledge. Second, in textual representations
information is organized in sequential data structures, emphasizing
temporal and logical relations. The use of natural language facilitates
processes of relating information in the text to everyday experiences
and situations. Therefore, textual representations are expected to
stimulate in particular situational knowledge. Finally, the emphasis on
procedures and operations in arithmetical representations makes it
plausible that this representational format will enhance procedural
knowledge.

With regard to transfer, it is expected that the emphasis of
arithmetical representations on procedures and operations on the one
hand will enhance near transfer, that is solving similar problems, but
on the other hand will hamper far transfer, that is solving dissimilar
problems, because the emphasis on procedures and the abstract nature of
this representational format is at the cost of conceptual and
situational knowledge, which are required for identifying a problem and
for understanding the principles that underlie the dissimilar problem.
It is expected that pictorial representations, because of their
emphasis on topological and structural relations, will enhance far
transfer.
The criterion of cognitive load induced by representational format echoes Larkin and Simon's (1987) concept of computational efficiency. These authors argue that even when a picture and a text are informationally equivalent, pictures are computationally more efficient. Therefore, pictorial representations are expected to reduce cognitive load more than do textual representations. On basis of the findings of Leung and his colleagues (1997) it can be assumed that in most cases textual representations will reduce cognitive load more than arithmetical representations, provided that the texts are not very extensive and that the population of learners does not consist exclusively of high-achievers. Following Van Bruggen and his colleagues (2002) the focus will be on extraneous (ineffective) load and germane (effective) load. Also the amount of invested mental effort (overall load) will be measured.

With regard to interactivity, not much is known about these behavioral aspects in respect to representational format. Therefore, the study of the interactivity will be entirely exploratory. Of interest are the time spent on manipulating representations and the number of manipulations. Different assignments involving manipulating representations will be compared in order to establish whether specific types of assignments interact with representational format with regard to the time spent on manipulating a representation and the number of manipulations.

The current study will use a simulation-based instruction, which allows manipulation of representations. Before the actual experiment is described, the instructional approach will shortly be clarified in the next section.

Simulation-based inquiry learning

In inquiry learning, an exponent of constructivism-based instruction, the focus of instruction is primary on the induction of concepts and principles of a domain (de Jong, in press; Swaak & de Jong, 1996). Learners are invited to discover the properties of the given domain (de Jong & van Joolingen, 1998; van Joolingen, 1993; van Joolingen & de Jong, 1997). A technology that is especially suited for inquiry learning is computer-based simulation. Computer-based simulations contain a model of a system or a process. By manipulating the input variables and observing the resulting changes in output values the learner is enabled to induce the concepts and principles underlying the model (de Jong & Van Joolingen, 1998). Although active learning and meaningful learning are viewed as main characteristics of inquiry learning (Svinicki, 1998), the relation between activity and meaningfulness in learning should be considered with care. Mayer argues that meaningful learning may not simply be the result of behavioral activity per se. He suggests that only specific cognitive activities (e.g. selecting, organizing, and integrating knowledge) may promote meaningful learning (Mayer, 2002, 2004). In order to have learners deploying the required and appropriate cognitive activities and to prevent them from floundering, learners need some level of support. Leaving learners to their own devices in a computer-based simulation is not a very effective and efficient way of learning (de Jong, in press), or, as Mayer puts it: “a formula for educational disaster” (2004, p. 17). Structuring the learning environment may therefore lead to more effective learning (de Jong, in press). De Jong and Van Joolingen (1998) reviewed research publications on scientific inquiry learning
with computer simulation. On basis of their review the authors bring up three instructional measures that hold the promise of positively influencing learning outcomes: providing direct access to domain information, providing learners with assignments, and increasing the model complexity in a gradual way (model progression).

**Experiment**

**Method**

**Participants**

In total, 58 third-grade pre-university education pupils, 27 boys and 31 girls, participated. The average age of the participants was 14.48 years (SD=0.63). The participants had no prior knowledge of combinatorics and probability theory. They attended the experiment during regular school time. Therefore, participation was obligatory. Participants were not compensated financially for their participation.

**Domain**

The chosen domain of the instruction is combinatorics and probability theory. The essence of combinatorics is determining how many combinations can be made of a certain set or subset of elements. In order to determine the number of possible combinations, one also needs to know 1) whether elements may occur repeatedly in a combination (replacement) and 2) whether the order of elements in a combination is of interest. On basis of these two criteria, for so-called problem categories can be distinguished: no replacement-order important; replacement-order important; no replacement-order not important; and replacement-order not important. When the number of possible combinations is known, the probability that one or more combinations will occur in a random experiment can be determined.

**Materials**

The materials used were a simulation-based learning environment and two electronic test environments (one pre-test and one post-test environment).

The pre-test was delivered electronically via the Internet and aimed at measuring prior knowledge. Questions were presented screen-by-screen without the possibility to skip questions or to turn back to previous questions. Learner responses, both on multiple-choice and open-ended items, were collected and recorded electronically. The actual pre-test contained 4 open-ended and 8 multiple-choice questions (see Table 1). The multiple-choice items had twin items in the post-test. These items had the same structure as their counterparts but differed at a superficial level (different names, different order of alternatives).

The post-test aimed at measuring different (sub)types of knowledge. An overview of the 44 post-test items and underlying knowledge types is presented in Table 2.
Like the pre-test environment, the post-test was delivered electronically. Again, questions were presented screen-by-screen without the possibility to skip questions or to turn back to previous questions. Learner responses, both on multiple-choice and open-ended items, were collected and recorded electronically. In cases where learners had to calculate the outcome of an item, a Java-programmed calculator was provided on-screen.

Table 1

Overview of pre-test items

<table>
<thead>
<tr>
<th>Knowledge type</th>
<th>Knowledge subtype</th>
<th>Description</th>
<th>Number and type of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual</td>
<td>C1</td>
<td>Items about relations between problem categories</td>
<td>2 multiple-choice items</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>Items about relations within problem categories</td>
<td>2 multiple-choice items</td>
</tr>
<tr>
<td>Procedural</td>
<td>Easy</td>
<td>Solving simple combinatorial problems</td>
<td>4 open-ended questions</td>
</tr>
<tr>
<td></td>
<td>Near transfer</td>
<td>Solving near transfer problems</td>
<td>4 multiple-choice items</td>
</tr>
</tbody>
</table>

Table 2

Overview of post-test items

<table>
<thead>
<tr>
<th>Knowledge type</th>
<th>Knowledge subtype</th>
<th>Description</th>
<th>Number and type of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual</td>
<td>C1</td>
<td>Items about relations between problem categories</td>
<td>4 multiple-choice items</td>
</tr>
<tr>
<td></td>
<td>C1i</td>
<td>Intuitive items about relations between problem categories</td>
<td>13 multiple-choice items</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>Items about relations within problem categories</td>
<td>4 multiple-choice items</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>Calculations (understanding why)</td>
<td>4 open-ended items</td>
</tr>
<tr>
<td>Procedural</td>
<td>P1</td>
<td>Calculations (understanding how)</td>
<td>2 open-ended items</td>
</tr>
<tr>
<td></td>
<td>P2-near</td>
<td>Solving near transfer problems</td>
<td>8 open-ended items</td>
</tr>
<tr>
<td></td>
<td>P2-far</td>
<td>Solving far transfer problems</td>
<td>4 open-ended items</td>
</tr>
<tr>
<td>Situational</td>
<td>Sit1</td>
<td>Identifying possible and acceptable choices from a set of numbers</td>
<td>1 open-ended item</td>
</tr>
<tr>
<td></td>
<td>Sit1mc</td>
<td>Identifying possible and acceptable choices from a set of numbers</td>
<td>4 multiple-choice items</td>
</tr>
</tbody>
</table>

10
Three simulation-based learning environments, ComBin-Pictorial, ComBin-Arithmetical, and ComBin-Textual, were created with SIMQUEST authoring software. ("ComBin" refers to "combinatorics", the syllable "Bin" refers also to the digital (binary) nature of the learning environment). These environments were basically the same with regard to structure, contents, and assignments. All learners received the same introduction to the domain, assignments, and hypotheses. The set-up of the learning environments is summarized in Table 3.

Table 3

Set-up of the learning environments

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>01. Introduction</td>
<td>Short introduction to probability theory</td>
</tr>
<tr>
<td>02. Individual events</td>
<td>Explanation of individual events</td>
</tr>
<tr>
<td>03. Individual events</td>
<td>Explanation of individual events (continued)</td>
</tr>
<tr>
<td>04. Complex events 1</td>
<td>Introduction to complex events</td>
</tr>
<tr>
<td>05. Variations</td>
<td>Cover story introducing next simulation</td>
</tr>
<tr>
<td>06. Simulation</td>
<td>Exploration of concept of replacement</td>
</tr>
<tr>
<td>07. Combinations</td>
<td>Cover story introducing next simulation</td>
</tr>
<tr>
<td>08. Simulation</td>
<td>Exploration of concept of order</td>
</tr>
<tr>
<td>09. Overview</td>
<td>Overview of 4 problem categories</td>
</tr>
<tr>
<td>10. Complex events 2</td>
<td>Exploration of complex events</td>
</tr>
<tr>
<td>11. Simulation</td>
<td>Exploration of &quot;No replacement-order important&quot;</td>
</tr>
<tr>
<td>12. Simulation</td>
<td>Exploration of &quot;Replacement-order important&quot;</td>
</tr>
<tr>
<td>13. Simulation</td>
<td>Exploration of &quot;No replacement-order not important&quot;</td>
</tr>
<tr>
<td>14. Simulation</td>
<td>Exploration of &quot;Replacement-order not important&quot;</td>
</tr>
<tr>
<td>15. Simulation</td>
<td>Comparing problem categories</td>
</tr>
<tr>
<td>16. End</td>
<td>End</td>
</tr>
</tbody>
</table>

The learning environments only differed with regard to the representational formats used in the simulations. These representational formats were either pictorial, arithmetical, or textual. In the pictorial version (ComBin-Pictorial), input to experiments was done by means of a direct manipulation interface, and output was presented mainly pictorially (see Figure 4).

Before the other ComBin versions are discussed, first it will be described how the simulations are operated. Each simulation was accompanied by a cover story. This was a story about a real-life situation in which combinatorics and probability play a role. The cover story accompanying the simulation in Figure 4 was about predicting the outcome of a foot race.
The probability that a prediction will be correct depends on the total number of runners and the number of runners the prediction is about (is the prediction only about who finishes first, or also about who finishes second, third, etcetera?). In the box at the left hand side of the simulation window (see Figure 4) the learners could set the input values (total number of runners and number of runners the prediction is about). Each time participants set input values they had to press an “Apply”-button before the computer entered these values in the model and calculated and displayed the corresponding output. At the top right of the window, an example of a prediction matching the input values was presented. At the bottom right the output was presented. In the pictorial version of ComBin the output consisted of a tree diagram and the probability that the prediction was correct expressed as a fraction. The learners could change the input values at will and each time the input values were changed, the example and output changed accordingly (after the “apply”-button was pressed). Each version of ComBin contained the same cover stories, but differed with regard to the representational code of the output. In the arithmetical version (ComBin-Arithmetical), learners received output from the experiment as numbers embedded in formulas (see Figure 5).

In the textual version (ComBin-Textual), learners received output from the experiments as numbers embedded in text together with more qualitative textual statements (see Figure 6).
In order to prevent advantages of one version over the others, the examples matched the representational code of the simulation as far as possible. For example, in the case of the foot race, the example indicated runners by the color of their T-shirt in the pictorial version, by the number on the back of the T-shirt in the arithmetical version, and by the runners name in the textual condition.

Each learning environment contained seven simulations. The complexity of the models underlying these simulations increased gradually (model progression). The first simulation (section 06 in the learning environment; see Table 3) gave learners the opportunity to discover the principle of replacement. Three hypotheses were presented and the learner decided by means of the simulation which hypothesis was true. In the second simulation (section 08) the learner had the opportunity to discover the principle of order. Again, from three hypotheses the one that resembled the model underlying the simulation best, needed to be selected. Four simulations (sections 11-14) were dedicated to discovering relations within problem categories (see “Domain”); one simulation for each problem category. The final simulation (section 15) enabled learners to discover relations between the four problem categories. These five simulations (11-15) did not contain hypotheses, but were each time accompanied by an assignment that presented learners with a short cover story that was typical for the principle expressed in the accompanying simulation and then the learners were asked to inquire the relations between the variables in the simulation and to explain these relations (in fact, they were asked to inquire and describe the principle underlying the simulation).

After closing each of these five simulations, a set of six questions with regard to perceived mental effort appeared on the screen. Here, the learners had to indicate on 9-points Likert scales the amount of mental effort with regard to the perceived difficulty of the domain in general (intrinsic load), navigation, task design, and accessibility of information (extraneous load), understanding the simulation (germane load), and the amount of invested mental effort (overall load) (see Table 4). Each time a set of cognitive load appeared, the items of the set were presented in a different order in order to prevent learners from answering in an automatic fashion.
Table 4

Cognitive load items presented to learners after each simulation

<table>
<thead>
<tr>
<th>Type of cognitive load</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL  Intrinsic load</td>
<td>How easy or difficult do you consider probability theory at this moment?</td>
</tr>
<tr>
<td>EL1 Extraneous load</td>
<td>How easy or difficult is it for you to work with the learning environment?</td>
</tr>
<tr>
<td>EL2 Extraneous load</td>
<td>How easy or difficult is it for you to distinguish important and unimportant information in the learning environment?</td>
</tr>
<tr>
<td>EL3 Extraneous load</td>
<td>How easy or difficult is it for you to collect all the information that you need in the learning environment?</td>
</tr>
<tr>
<td>GL  Germaine load</td>
<td>How easy or difficult was it to understand the simulation?</td>
</tr>
<tr>
<td>OL  Overall load</td>
<td>Indicate on the scale the amount of effort you had to invest to follow the last simulation.</td>
</tr>
</tbody>
</table>

The learning environments automatically registered user actions. User actions that were logged included measures like user path through the learning environment (which parts of the learning environment were opened, when, for how long, and in what sequence) and the number and nature of manipulations carried out in the simulations (how many experiments were carried out and what were the exact input values of each experiment). The reason that the “apply”-button was added to the simulations, was to demarcate the manipulations by the user. Each time the user pressed this button, it was registered into the log file 1) that the button was pressed, 2) at which time that happened, plus 3) the exact settings of the input variables in the simulation at the very moment the button was pressed.

Procedure

At the beginning of the session the research was introduced briefly and globally. Then the course of the session was described. Participants were told that the session would consist of three parts (pre-test, learning phase, and post-test) and that they were allowed to work on their own pace through the three parts. Participants needed not to wait for others before they could continue with each next part. Despite the possibility to follow a non-linear path through the learning environment, participants were advised to keep to the order of sections because they build upon each other. The duration of the session was limited to three hours including a 15 minute break of the experiment. After this introduction, participants received a randomly handed-out log-in form. The indications on the form led the participants to one of the three conditions.

After logging on the pre-test, the participants were first presented with some background information with regard to the
experiment (general purpose of the research, the domain of interest, learning goals, etcetera). Furthermore, it was announced that the post-test would contain more and more difficult items than the pre-test but that the pre-test items nonetheless would give an indication of what kind of items to expect on the post-test. Before the pre-test was presented, the participants answered some questions with regard to their personal characteristics (e.g. age, gender, latest math grade, computer experience, spatial ability, visual/verbal preferences, and so on). At the end of the pre-test the participants received information about how to enter the learning environment. After finishing the last simulation in the learning environment, a message appeared on the screen, inviting the learners to close the learning environment and to ask the experimenter for further instructions. They then received log-on instructions for the post-test environment.

At the end of the post-test the participants were thanked for their cooperation and were allowed to leave the classroom.

Data preparation

The developers of the pre-test and post-test supplied a protocol for scoring the items as well. Items were scored on basis of the scoring rules in this protocol. For each open question, the protocol offers criteria and rules for determining the correctness of responses. Also for each item some examples of correct and incorrect responses are provided. In line with the protocol items were scored 0 if answered incorrectly, and 1 if answered correctly. Partial scores like 0.5 were not allowed. When doubting the (in)correctness of the response, the response was scored 0.

Results

Pre-Test

The pre-test overall scores are summarized in Table 5. In order to determine whether differences with regard to prior knowledge existed between conditions, a one-way analysis of variance (ANOVA) was conducted on the overall pre-test scores. No differences between conditions could be established, $F(2,55) = 0.062, p = .940$. Therefore, it is assumed that conditions did not differ with regard to level of prior knowledge.

Table 5

<table>
<thead>
<tr>
<th>Condition</th>
<th>$n$</th>
<th>$M$</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pictorial</td>
<td>21</td>
<td>5.33</td>
<td>1.978</td>
</tr>
<tr>
<td>Arithmetical</td>
<td>19</td>
<td>5.37</td>
<td>1.339</td>
</tr>
<tr>
<td>Textual</td>
<td>18</td>
<td>5.17</td>
<td>1.826</td>
</tr>
</tbody>
</table>
Post-Test

The post-test results are displayed in Table 6. All post-test measures were analyzed by one-way ANOVAs with condition as factor.

Table 6

Scores on post-test

<table>
<thead>
<tr>
<th>(Sub)type</th>
<th>Number of items</th>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Pictorial</td>
<td>Arithmetical</td>
<td>Textual</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>n</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Conceptual knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1 mc</td>
<td>4</td>
<td>2.26</td>
<td>1.28</td>
<td>19</td>
<td>2.57</td>
<td>0.81</td>
</tr>
<tr>
<td>C1i mc</td>
<td>13</td>
<td>9.11</td>
<td>2.77</td>
<td>19</td>
<td>10.43</td>
<td>2.40</td>
</tr>
<tr>
<td>C2 mc</td>
<td>4</td>
<td>2.42</td>
<td>1.12</td>
<td>19</td>
<td>2.62</td>
<td>0.81</td>
</tr>
<tr>
<td>C3 open</td>
<td>4</td>
<td>0.84</td>
<td>0.69</td>
<td>19</td>
<td>1.52</td>
<td>1.03</td>
</tr>
<tr>
<td>Procedural knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1 open</td>
<td>2</td>
<td>0.47</td>
<td>0.61</td>
<td>19</td>
<td>0.62</td>
<td>0.50</td>
</tr>
<tr>
<td>P2-near open</td>
<td>8</td>
<td>0.74</td>
<td>1.73</td>
<td>19</td>
<td>2.24</td>
<td>1.67</td>
</tr>
<tr>
<td>P2-far open</td>
<td>4</td>
<td>0.26</td>
<td>0.45</td>
<td>19</td>
<td>0.67</td>
<td>0.73</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>1.47</td>
<td>2.09</td>
<td>19</td>
<td>3.52</td>
<td>2.27</td>
</tr>
<tr>
<td>Situational knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>1.68</td>
<td>1.25</td>
<td>19</td>
<td>2.14</td>
<td>1.39</td>
</tr>
<tr>
<td>Overall</td>
<td>44</td>
<td>17.79</td>
<td>5.53</td>
<td>19</td>
<td>22.81</td>
<td>4.50</td>
</tr>
</tbody>
</table>

With regard to overall conceptual knowledge no main effect of condition, \(F(2,55) = 2.440\), \(p = .096\) was found. With regard to the subtypes of conceptual knowledge, one-way ANOVAs did not show any significant effects of condition on knowledge of relations between problem categories (C1) \([F(2,55) = 0.395, p = .676]\), on intuitive knowledge of relations (C1i) \([F(2,55) = 1.864, p = .165]\), on knowledge of relations within problem categories (C2) \([F(2,55) = 0.270, p = .764]\), or on knowledge of principles underlying calculations (understanding why) (C3) \([F(2,55) = 3.052, p = .055]\).

Analysis of the overall procedural knowledge revealed a significant main effect of condition, \(F(2,55) = 4.796\), \(p = .012\). A post-hoc LSD analysis showed that the arithmetical condition outperformed the pictorial condition \((p < .01)\). The textual condition did not differ significantly from both other conditions. On the subtype level of procedural knowledge the following results were obtained from one-way ANOVAs. Knowledge of calculations (knowing how) (P1), no significant effect of condition was found, \(F(2,55) = 0.320\), \(p = .727\). Analysis of near transfer (P2n) revealed a significant effect of condition, \(F(2,55) = 4.404\), \(p = .017\). Post-hoc LSD analysis showed the arithmetical condition to be superior to the pictorial condition \((p < .01)\). The textual condition did not differ from either the arithmetical condition \((p = .071)\) or the pictorial condition \((p = .316)\). On far
transfer (p2f) no significant effect of condition was found, \( F(2, 55) = 2.984, p = .059. \)

With regard to overall situational knowledge items no significant main effect of condition, \( F(2, 55) = 0.786, p = .461 \) was found.

Analysis of post-test overall scores found significant effects of condition, \( F(2, 55) = 5.130, p < .01 \), and a post-hoc LSD analysis revealed that both the arithmetical and the textual condition outperformed the pictorial condition (\( p = .003 \) and \( p < .05 \) respectively). The arithmetical and textual condition did not differ significantly from each other.

Cognitive load

All cognitive load measures were analyzed by one-way ANOVAs with condition as factor. Table 7 displays cognitive load measures that differ significantly from each other.

Table 7

<table>
<thead>
<tr>
<th>Cognitive load measures</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pictorial</td>
</tr>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Extraneous Load (EL3)</td>
<td>5.45</td>
</tr>
<tr>
<td>Germane Load (GL)</td>
<td>6.00</td>
</tr>
<tr>
<td>Overall Load (OL)</td>
<td>5.23</td>
</tr>
<tr>
<td>Externouse Load (EL1)</td>
<td>5.50</td>
</tr>
<tr>
<td>Germane Load (GL)</td>
<td>5.67</td>
</tr>
<tr>
<td>Overall Load (OL)</td>
<td>5.58</td>
</tr>
</tbody>
</table>

In simulation "11. Runners" the learners explored the relations between variables in the category "no replacement, order important". No effect of condition on any of the cognitive load measures was observed.

The same held true for the case of simulation "12. PIN-code" in which learners inquired the relations between variables in the category "replacement, order important".

In the case of simulation "13. Cell phones" where learners explored the category "no replacement, order unimportant", effects of condition were observed for extraneous load caused by accessibility of information (\( F(2, 34) = 4.126, p < .05 \)), germane load (\( F(2, 34) = 5.583, p < .01 \)), and overall load (\( F(2, 34) = 3.400, p < .05 \)). Post-hoc Bonferroni analysis revealed that with regard to extraneous load caused by accessibility of information (EL3), participants in the pictorial condition reported more load compared to participants in the arithmetical condition (\( p < .05 \)). Reports of participants in the textual condition did not differ from either the pictorial or the arithmetical condition. Second, with regard to germane load, higher load was reported by participants in the pictorial condition compared
to both the arithmetical condition \( (p = .01) \) and the textual condition \( (p < .05) \). The latter two conditions did not differ from each other.

Analysis of cognitive load measures with regard to simulation “14. Action figures”, in which learners inquired the category “replacement, order unimportant”, revealed an effect of condition on EL1, extraneous load caused by navigation \( (F(2,32) = 3.853, p < .05) \) and on overall load \( (F(2,31) = 4.302, p < .05) \). Post-hoc Bonferroni analysis showed that participants in the pictorial condition reported more extraneous load caused by navigation compared to participants in the arithmetical condition \( (p < .05) \). Participants in the textual condition did not differ from either the pictorial condition or the arithmetical condition. With regard to the mental effort invested in the task (overall load), the pictorial condition showed higher levels of effort compared to the arithmetical condition \( (p < .05) \). The textual condition did not differ from the other conditions.

Finally, in simulation “15. Mountain bikes”, the learners explored the relations between problem categories. Here, an effect of condition was found on intrinsic load \( (F(2,33) = 3.833, p < .05) \), germane load \( (F(2,33) = 3.772, p < .05) \), and overall load \( (F(2,32) = 3.659, p < .05) \). Compared to the textual condition, the participants in the pictorial condition experienced more intrinsic load \( (p < .05) \). The arithmetical condition did not differ from either the textual condition or the pictorial condition. More germane load was experienced by the pictorial condition compared to the textual condition \( (p < .05) \). The arithmetical condition did not differ from the textual or pictorial condition. The highest levels of invested effort (overall load) were again reported by the pictorial condition compared to the textual condition \( (p < .05) \). The arithmetical condition did not differ from the textual condition or the pictorial condition.

Interactivity

A summary of process measures ‘Time spent in the learning environment’ and ‘Number of experiments carried out in the simulations’ is displayed in Table 8. One-way ANOVAs found neither significant effects of conditions on time spent in the learning environment, \( F(2,55) = .392, p = .677 \), nor on the number of experiments carried out in the simulations, \( F(2,55) = .085, p = .919 \).

Table 8

<table>
<thead>
<tr>
<th>Time in Learning Environment and Number of Experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time spent in learning environment (in absolute minutes)</td>
</tr>
<tr>
<td>Condition</td>
</tr>
<tr>
<td>Pictorial</td>
</tr>
<tr>
<td>Arithmetical</td>
</tr>
<tr>
<td>Textual</td>
</tr>
</tbody>
</table>
Table 9 displays the average time spent in each of the simulations. Results were analyzed by a two-way ANOVA with condition and simulation as between-subject factors. The analysis revealed that condition had no effect on time spent in each simulation ($F(2,385) = .563, p = .570, \eta^2 = .003$). Simulation had effect on the average time learners spent in it ($F(6,385) = 4.876, p < .001, \eta^2 = .071$). However, the low $\eta$-square value indicates that this relation was only very weak. In all conditions learners spent most time in the first simulation (06. Replacement). No interaction between condition and simulation was found ($F(12,385) = 1.227, p = .262, \eta^2 = .037$). Figure 7 depicts the average time learners in each condition spent in each of the simulations. This graph shows that the average time spent in simulations drops sharply after the first simulation.

Table 9

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Pictorial</th>
<th>Arithmetical</th>
<th>Textual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$n$</td>
</tr>
<tr>
<td>06. Replacement</td>
<td>2.66</td>
<td>1.29</td>
<td>19</td>
</tr>
<tr>
<td>08. Order</td>
<td>1.54</td>
<td>1.13</td>
<td>19</td>
</tr>
<tr>
<td>11. No repl./order</td>
<td>1.48</td>
<td>1.00</td>
<td>19</td>
</tr>
<tr>
<td>12. Repl./order</td>
<td>1.58</td>
<td>1.24</td>
<td>19</td>
</tr>
<tr>
<td>13. No repl./no order</td>
<td>2.28</td>
<td>2.60</td>
<td>19</td>
</tr>
<tr>
<td>14. Repl./no order</td>
<td>1.49</td>
<td>1.45</td>
<td>19</td>
</tr>
<tr>
<td>15. Comparing categ.</td>
<td>2.15</td>
<td>2.29</td>
<td>19</td>
</tr>
</tbody>
</table>
Table 10 displays the number of manipulations carried out by the learner in each of the simulations. Results were analyzed by a two-way ANOVA with condition and simulation as between-subject factors.

Table 10

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Pictorial</th>
<th>Arithmetical</th>
<th>Textual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M  SD n</td>
<td>M  SD n</td>
<td>M  SD n</td>
</tr>
<tr>
<td>06. Replacement</td>
<td>7.56  4.91 18</td>
<td>6.95  4.72 20</td>
<td>8.59  8.80 17</td>
</tr>
<tr>
<td>08. Order</td>
<td>7.87  9.49 15</td>
<td>7.15  8.24 20</td>
<td>7.12 10.05 17</td>
</tr>
<tr>
<td>11. No repl./order</td>
<td>4.94  3.83 17</td>
<td>4.70  5.81 20</td>
<td>6.65 13.48 17</td>
</tr>
<tr>
<td>12. Repl./order</td>
<td>3.73  2.28 15</td>
<td>2.90  3.51 20</td>
<td>3.06  5.05 18</td>
</tr>
<tr>
<td>13. No repl./no order</td>
<td>4.00  3.12 16</td>
<td>3.19  3.97 21</td>
<td>2.75  2.67 16</td>
</tr>
<tr>
<td>14. Repl./no order</td>
<td>2.06  1.35 18</td>
<td>2.47  2.76 19</td>
<td>3.40  4.05 15</td>
</tr>
<tr>
<td>15. Comparing categ.</td>
<td>5.50  6.15 18</td>
<td>2.95  2.37 20</td>
<td>3.53  3.91 15</td>
</tr>
</tbody>
</table>

No effect of condition on the number of manipulations ($F(2,351) = .634$, $p = .531$, $\eta^2 = .004$). Simulation had effect on the average number of manipulations performed in it ($F(6,351) = 6.171$, $p < .001$, $\eta^2 = .095$). However, this relation was weak. In all conditions learners carried out most manipulations in the first simulation (section 06) and in the second simulation (08.). No interaction between condition and number of manipulations was found ($F(12,351) = .289$, $p = .991$, $\eta^2 = .010$).
Figure 8. Average number of manipulations carried out in each of the simulations

Figure 8 depicts the average number of manipulations learners carried out in each of the simulations. This graph shows that the number of manipulations drops sharply after the first two simulations (simulation 06 and 08).

Conclusions and discussion

Below, the effects of representational format and the cognitive processes underlying these effects will be discussed on basis of the criteria that were formulated earlier, namely: facilitation of problem solving, transfer, cognitive load, and interactivity.

Criteria

Facilitation of problem solving

This criterion refers to the extent to which problem solving is made easier or more difficult when using different representations. We argued that it is necessary to consider situational, conceptual, and procedural knowledge in order to assess the extent to which problem solving is easier or more difficult with different representations. With regard to conceptual knowledge, no differences between conditions were found. Therefore, the expected superiority of pictorial representations with regard to conceptual domain knowledge was not confirmed by our data. For procedural knowledge a main effect of condition was found in favor of the arithmetical condition. This was in line with the expectations. However, it should be noted that despite the main effect of condition, the scores on procedural knowledge were
relatively poor. With regard to situational knowledge no effects of condition were found. Textual representations were expected to enhance situational knowledge compared to pictorial and arithmetical representations, but our data do not confirm this expectation. Despite the slight advantage of arithmetical representations in problem solving, it must be concluded that problem solving as a whole, that is taken conceptual, procedural, and situational knowledge measures together, was relatively poor. On average, learners in all conditions hardly succeeded at solving correctly half of the 44 items in the post-test. Moreover, there was not much difference between conditions with regard to their overall post-test scores. Perhaps, the representational format does not affect problem solving much after all. Another explanation might be that the post-test items were too difficult, which may have caused floor effects. This explanation gains support from results obtained in other studies that used the same post-test. Post-test results obtained in those studies, were also relatively poor.

Transfer

The results with regard to transfer indicate that conditions differ with regard to near transfer. Participants in the arithmetical condition were better at solving near transfer problems compared to participants in the pictorial condition. This finding was in line with the expectations. With regard to far transfer no differences between conditions were found. So, the expected superiority of pictorial conditions was not confirmed by the data. On both near and far transfer the results were very poor. Still, the participants in the arithmetical condition were a little more able to apply their inferences to similar problems (near transfer).

Cognitive load

We expected pictorial representations to induce the least cognitive load and arithmetical representations the highest cognitive load. Of concern are extraneous cognitive load, germane cognitive load, and overall cognitive load (amount of mental effort invested).

With regard to extraneous cognitive load (or ineffective load), participants in the pictorial condition found it harder to collect information in simulation 13 (no replacement, order unimportant) compared to participants in the arithmetical condition. In simulation 14 ("replacement, order unimportant"), participants in the pictorial condition found working with the learning environment more difficult than participants in the arithmetical condition. This finding is striking, because working with simulation 14 did not deviate from working with the other simulations.

Concerning germane cognitive load (effective load), participants in the pictorial condition found simulation 13 more difficult to understand compared to participants from both the arithmetical condition and the textual condition. In simulation 15 (relations between problem categories), participants in the pictorial condition found the simulation more difficult to understand than participants in the textual condition.

With regard to invested mental effort (overall load), participants in the pictorial condition rated the amount of effort they invested in following simulation 14 higher than participants in the arithmetical condition. In simulation 15, participants in the pictorial
condition rated the amount of effort they had to invest in following
the simulation higher than participants in the textual condition.

We can conclude that the expected computational efficiency of
pictorial representations was not confirmed by our data. On the
contrary, in all cases participants in the pictorial condition reported
the highest cognitive load instead of the lowest. It is striking that
the differences in cognitive load did not seemingly affect problem
solving performance, as reflected by the performance on the post-test.
The possible floor effects of the post-test discussed before, may cover
real differences between different representational formats. On the
other hand, it may also be possible that higher levels of cognitive
load do not really affect problem solving performance. Discussions with
participants after the experiment, when they were allowed to see the
other conditions, revealed that almost all participants found the
pictorial representations most attractive. The general attitude of the
participants towards textual and in particular arithmetical
representations was that they found them "boring". So, despite the
higher cognitive load in the pictorial condition, the participants had
a positive attitude towards these representations. This attitude may
have compensated for the higher cognitive load, encouraging them to
persist in attempting to understand the meaning of the pictures and
preventing them from dropping out of the learning process.

Interactivity

No effects of conditions were found on time spent in the learning
environment, nor on the number of experiments carried out in the
simulations. Most time was spent on the first simulation, simulation
06. Most manipulations were carried out in simulation 06 and 08, the
first two simulations. Learners spent considerably less time on other
simulations and the number of manipulations also dropped significantly
after completing simulation 06 and 08. These effects are stable over
conditions. What accounts for these differences? Simulation 06 and 08
differ in two respects from the other simulations. First, in 06 and 08
learners explore the concepts of "replacement" and "order"
respectively. The simulations 11-15 treat combinations of these
congcepts (e.g. "no replacement/order important") and the effects of
these combinations on other variables in the domain. These simulations
are therefore conceptually more complex than simulations 06 and 08.
Second, the assignments that accompanied simulations 06 and 08
presented learners with three hypotheses. By carrying out manipulations
in the simulations and by analyzing the results of these manipulations,
the learners could determine which of the three hypotheses was (most)
correct. The other simulations were each time accompanied by an
assignment that presented learners with a short cover story that was
typical for the principle expressed in the accompanying simulation and
then the learners were asked to inquire the relations between the
variables in the simulation and to explain these relations (in fact,
they were asked to inquire and describe the principle underlying the
simulation). These relations are nonlinear and therefore quite complex.
In order to obtain an even modest level of understanding of these
relations, substantial amounts of manipulations need to be performed.
However, the log files revealed that the average number of
manipulations carried out by the learners, were by far not enough to
even discover the relations, let alone to understand them. In these
simulations learners carried out even fewer manipulations and spent
less time compared to the much more simple simulations 06 and 08. One
possible explanation would be that learners were affected by fatigue by the time they entered the simulations 11-15. However, when learners become tired, it is plausible that they will experience higher levels of cognitive load. The cognitive load measures that were collected over time in the learning environment do not show increased levels of cognitive load measures over time, therefore the fatigue-explanation does not seem plausible. Other possible explanations come from the records of learner answers to the assignments. What is striking in these answers is that most, if not all, learners stayed very close to the cover story that accompanied the simulation. Most of their responses consist of mere reproductions of the cover stories. If relations are described at all, the descriptions are very brief, consisting of only a few sentences at the most, and mostly in qualitative terms like “more than” and “less than” only. Clearly, learners do not feel called to consider relations on more abstract levels. Possibly, they implicitly assume the relations to be linear. This would explain why learners carry out only few experiments. If the relations were indeed linear, only a few manipulations would suffice to discover the nature of the relations. Additional research would be required to determine whether epistemological beliefs indeed affect the behavior of learners in simulations. This is beyond the scope of the current study however. On basis of the data it can be concluded that representational format does not affect interactivity, but it seems that the context of the simulation, for example the structure and content of the accompanying assignment, does influence the behavior of learners in manipulating simulations.
Overall conclusions

The current study centered around the question: What is the effect of representational format on knowledge construction processes in learning with interactive representations? The answer to this question is that in the domain of combinatorics and probability theory, arithmetical representations lead to enhanced levels of knowledge, in particular with regard to procedural knowledge and transfer, compared to textual and pictorial representations. Effects of pictorial representations fall short of the effects of textual and particularly arithmetical representations. However, the differences between the effects representational formats are only modest. In all cases, pictorial representations induced the highest levels of cognitive load. With regard to interactivity, no effects of representational format were found.

Future directions

Most pregnant issue for further research is to establish whether the post-test indeed suffers from floor effects. Floor effects may hide more outspoken differences between conditions. Furthermore, collecting thinking-aloud protocols may shine some light on differences in learner’s responses to different representational formats. For example, it may be interesting to find out why the higher levels of cognitive load in the pictorial condition did not seem to be very detrimental to performance. Also, the effects of structure of assignments on the manipulations performed by the learners, need further research. It is a common problem in inquiry learning with simulations that learners do not fully explore simulations. Finally, it would be interesting to contrast the instructional approach used in this study with other instructional approaches in order to determine whether representational format interacts with instructional approach.
REFERENCES


