How to Optimize Learning From Animated Models: A Review of Guidelines Based on Cognitive Load

Pieter Wouters, Fred Paas and Jeroen J. G. van Merriënboer

REVIEW OF EDUCATIONAL RESEARCH 2008 78: 645 originally published online 27 August 2008
DOI: 10.3102/0034654308320320

The online version of this article can be found at:
http://rer.sagepub.com/content/78/3/645

Published on behalf of

American Educational Research Association

and

SAGE
http://www.sagepublications.com

Additional services and information for Review of Educational Research can be found at:

Email Alerts: http://rer.aera.net/alerts
Subscriptions: http://rer.aera.net/subscriptions
Reprints: http://www.aera.net/reprints
Permissions: http://www.aera.net/permissions

>> Version of Record - Nov 20, 2008
OnlineFirst Version of Record - Aug 27, 2008
How to Optimize Learning From Animated Models: A Review of Guidelines Based on Cognitive Load

Pieter Wouters
Utrecht University, Netherlands

Fred Paas and Jeroen J. G. van Merriënboer
Open University of the Netherlands, Heerlen

Animated models explicate the procedure to solve a problem, as well as the rationale behind this procedure. For abstract cognitive processes, animations might be beneficial, especially when a supportive pedagogical agent provides explanations. This article argues that animated models can be an effective instructional method, provided that they are designed in such a way that cognitive capacity is optimally employed. This review proposes three sets of design guidelines based on cognitive load research: The first aims at managing the complexity of subject matter. The second focuses on preventing activities (attributed to poor design) that obstruct learning. The last incites learners to engage in the active and relevant processing of subject matter. Finally, an integrative framework is presented for designing effective animated models.

KEYWORDS: cognitive load theory, cognitive modeling, animations, instructional design/development, computers and learning, problem solving.

The current focus on lifelong learning and flexibility in task performance increasingly emphasizes the mastering of complex cognitive skills (Jonassen, 1999). Instructional methods such as modeling and vicarious learning, in which experts perform problem-solving tasks for learners while explaining the reasoning for their actions, fit this focus on complex learning. At the same time, rapid developments in computer and software technology in the last decades have enabled the use of animations to illustrate abstract cognitive processes and concepts (Casey, 1996; Chee, 1995; Collins, 1991) and programmable pedagogical agents to support learners.

We refer to the combined use of animations and pedagogical agents in modeling as animated models. Such models illustrate the solving of problems, such as scientific problems (e.g., problems about gravity), mathematical problems (e.g., probability calculation), and search problems (finding information on the Internet). The pedagogical agent functions as a social model and so guides the learner through
the animation—for example, by moving around the screen and guiding the learner’s attention, by addressing the learner in a personalized style, and by showing which errors typically occur and how the learner may avoid them. For instance, in solving a problem in the domain of probability calculation, it is important to know whether the answer involves a “drawing with or without replacement.” For novices, this concept may be rather abstract and difficult to understand. An animation can visualize the concept by showing what is happening in, say, a situation in which the probability has to be calculated so that a person guesses the correct pin code of a cash card in one trial. As depicted in Figure 1, the animated model may show someone standing before an ATM. For the first number of the pin code, this person may choose from the 10 figures located to the right of the ATM display.

For the second number of the pin code, the person again may choose from 10 figures. This is shown in Figure 2, in which the pedagogical agent guides the attention to the 10 figures (which have become highlighted) and explains that the figure used for the first number of the pin code can be used for the second number. This illustrative animated process and its explanation by the pedagogical agent continue until the problem is solved.

A potential danger of showing the performance of a complex task with visualizations and verbal explanations is that the limited cognitive capacity of learners might become overloaded. Cognitive load theory emphasizes this limitation as an important determinant for the effective use of instructional methods (Paas, Renkl,
In this article, we argue that animated models can be an effective instructional method, provided that they are designed in such a way that cognitive capacity is optimally employed. As such, we propose a set of design guidelines to accomplish this objective. For this purpose, we selected design guidelines from the review publications of leading researchers in the field (e.g., Mayer, 2001, 2005a, 2005b, 2005c; Mayer & Moreno, 2002, 2003; Moreno, 2005; Sweller, 1999, 2005, 2006; Sweller et al., 1998; van Merriënboer & Sweller, 2005). We also conducted a literature search, with these design guidelines as search terms, in the PsycINFO and EJS E-Journals databases. In addition, we searched these databases with broader descriptors, such as cognitive load, animations, dynamic visualizations, and multimedia. In cases in which too much output was generated, we limited the search results with terms such as learning, instruction, and training. Finally, the resulting scholarly output was then narrowed by selecting studies that were applicable to animated models.

In this article, we first give an outline of cognitive load theory. Second, we elaborate on the nature of animated models; that is, we pay attention to cognitive modeling, animations, and pedagogical agents. Third, we propose design guidelines that enable learners to engage in effective learning from animated models. In the last section, we draw some conclusions and provide directions for further research.
Cognitive Load Theory

Cognitive load theory tries to align the structure of information and the way that it is presented within human cognitive architecture. For learning to commence, people have to process information, and a qualifying factor is the degree in which the complexity of information varies. For the processing of information, two structures in human cognitive architecture are crucial: The first is working memory, where all conscious processing of information takes place. It has a limited processing capacity that is, by far, inadequate to meet the complexity of information that learners face in modern learning environments. The second structure, long-term memory, is a knowledge base with a virtually unlimited capacity that can serve as added processing capacity by means of schemas. Schemas comprise cognitive structures in which separate information elements are aggregated into one specialized element that can be processed by working memory as a single element (Paas et al., 2003). For example, for a complex skill such as driving a car—in particular, successfully changing gears—less experienced drivers need to bring the separate elements into working memory, one by one (e.g., declutching, shifting the gear, engaging the clutch). More experienced drivers, however, have aggregated these separate elements in one “changing gear” schema that can be processed by working memory as one element. The acquisition and automation of such schemas so that they can be processed unconsciously are important because it optimizes the processing capacity of working memory.

From the perspective of instructional design, information can impose a cognitive load in three ways: First, there is cognitive load, which depends on the element interactivity of the subject matter; that is, complex information consists of a multitude of elements that interact with one another. One can speak of understanding such complex information only when one processes the separate elements, as well as the way that they interact (Chandler & Sweller, 1994, 1996). For instance, in acquiring a foreign language, learning word pairs is associated with less element interactivity than that of understanding the grammar of a sentence. For learning word pairs, only two elements need to be active in working memory. However, for understanding a sentence, the words in the sentence have to be held in working memory, as do the grammatical relationships among these words. For example, to understand that a sentence such as “Two children are sitting on a couch” is correct but that the sentence “Two children is sitting on a couch” is not, the learner has to hold not only the separate words in working memory but also the grammatical parts, such as the subject and the verb and their relationship (i.e., the plural of the subjects has a consequence for the conjugation of the verb). In cognitive load theory, this is called intrinsic cognitive load, and it can be regarded as a necessary base load because it cannot be reduced without compromising full understanding. The more complex a skill, the higher the intrinsic cognitive load because of the higher element interactivity.

Second, the way that information is presented can impose a cognitive load. Extraneous, or ineffective, cognitive load is imposed on working memory because of poorly designed instructional material. Learners sometimes have to engage in cognitive activities that contribute not directly to learning but to overcoming the deficiencies of the design. One of the most investigated phenomena with respect to extraneous cognitive load is the split-attention effect (Kalyuga, Chandler, &
Design Guidelines for Animated Models

Sweller, 1999; Mayer & Moreno, 1998; Tarmizi & Sweller, 1988), which occurs when two or more sources of information must be simultaneously processed to derive meaning from subject matter. Take, for example, the situation of a diagram about assembling a machine for which explanatory text is presented on another page. The learner has to mentally search, match, and integrate both sources of information, which imposes a high extraneous cognitive load on working memory. This high load might interfere with learning.

Third, germane, or effective, cognitive load is imposed when information is presented in such a way that learning is enhanced—that is, when it facilitates the construction and/or automation of cognitive schemas. The assumption is that active processing will yield germane cognitive load. In this respect, the generation of self-explanations has proven to be an effective cognitive activity that enhances learning (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Renkl, 1997; Renkl & Atkinson, 2002). Consider, for example, an expert who has shown how to solve a problem in probability calculation. Novices might engage in relevant learning activities when they are incited to explain the observed problem-solving process and the resulting solution to themselves and in this way acquire or refine their cognitive schemas.

The three types of cognitive load are not isolated; rather, they act as additive components. The combined load of these components cannot exceed the available cognitive capacity—consequently, the high load of one component comes at the cost of another component. When intrinsic cognitive load is high, it becomes important to decrease extraneous cognitive load; otherwise, the combination of both might exceed the maximum cognitive capacity and thus prevent effective, or germane, activities to occur. From an instructional design point of view, extraneous cognitive load and germane cognitive load should be considered as communicating vessels because the reduction of extraneous cognitive load can free cognitive resources for an increase in germane cognitive load (Paas et al., 2003).

Modeling

The modeling and the vicarious learning literature emphasize that learning can be enhanced by observing experts or advanced novices display their performance of physical and/or cognitive skills (Bandura, 1976; Collins, Brown, & Newman, 1989; Cox, McKendree, Tobin, Lee, Mayes, 1999; van Merriënboer, 1997). Two arguments support this assertion: First, when observing an expert perform a complex task in which he or she integrates knowledge and skills, the learner can construct an adequate cognitive representation. This representation guides appropriate performance and so enables the learner to mentally or physically rehearse the task, which in turn refines the initial representation. Second, as compared with other instructional methods (e.g., worked-out solutions), learning by observation of a model might be beneficial because it shows not only what is happening but also why it is happening (Collins, 1991; van Gog, Paas, & van Merriënboer, 2004). Problem solving, for example, can be regarded as the application of several steps to solve a problem, but this approach does not take into account why some steps are chosen and others are not. In this way, learners might construct generalized schemas that they can apply in a variety of contexts and problem formats. Moreover, the expert might tell about false starts and dead ends and thus enable
the observer to learn what kind of response to avoid, without the need of making the error themselves (Bandura, 1976; Cox et al., 1999).

According to Collins et al. (1989), expert performance can be divided into the performance of physical skills and processes and the performance of cognitive skills and processes. On one hand, learning the appropriate muscle movements for learning to write, ski, play tennis, or throw darts typically involves behavioral modeling—that is, the expert shows the desired physical performance (Kitsantas, Zimmerman, & Cleary, 2000; Zimmerman & Kitsantas, 2002; for a review, see Wetzel, Radtke, & Stern, 1994). On the other hand, the modeling of cognitive skills and processes requires the explication of considerations, thoughts, and reasons that underlie the performance of actions or choices. Problem solving (Jonassen, 1999) and cognitive behavior modification (Meichenbaum, 1977) are examples of domains that essentially involve cognitive modeling.

Regarding skills and processes in the cognitive domain, it is a problem that they are not readily observable. When a novice observes an expert solve a problem, all the thoughts, considerations, and reasons might be traced back or concluded from the results, but the observer cannot actually perceive the cognitive performance. To overcome this problem, the cognitive skills and processes of the expert, which occur internally, have to be externalized. In a description of cognitive apprenticeship learning, Collins et al. (1989) discuss approaches in which the externalization of cognitive skills is practiced—namely, by having the models (i.e., the teachers) speak aloud their considerations with respect to heuristics (e.g., rules of thumb) and control processes in fields such as writing and mathematics.

When abstract concepts and processes are involved that have no physical counterpart, cognitive modeling might become difficult. For example, in debugging, which is an important aspect of computer programming, a novice programmer tries to find out what happens when an error occurs in the program code. Cognitive modeling could be used to show how an expert programmer finds out what cause-and-effect relations exist in the program code and which reasoning underlies these considerations. However, it is difficult to externalize the expert’s considerations about concepts such as readability and robustness and about processes such as inheritance of properties in such a way that it helps learners to construct a mental representation. In this respect, the use of dynamic visualizations such as animations might be helpful to illustrate these concepts and processes.

**Animations**

Mayer’s cognitive theory of multimedia learning (Mayer, 2001; Mayer & Moreno, 2003) is one of the most comprehensive theories about multimedia learning. It can be considered an extension of dual coding theory, from which it adopts the dual-channel assumption. This assumption implies that information is processed via two separate channels: a verbal system, which basically comprises written language and spoken language, and a nonverbal system, which processes pictorial materials (J. M. Clark & Paivio, 1991; Paivio, 1986). A fundamental prediction of dual coding theory is that both systems are additive and that people learn better when the presented information is encoded verbally and visually rather than in one system only. Information that has been encoded in two ways can be retrieved from memory more easily than information encoded in one way. Whereas the majority of dual coding theory research has been conducted with static visualizations, the
Design Guidelines for Animated Models

cognitive theory of multimedia learning has focused on dynamic visualizations. Animation is probably the most widespread type of dynamic visualization, and it can be regarded as the presentation of frames in such a way that each appears as an alteration of the previous one, with a speed that creates the illusion of apparent motion (Rieber & Kini, 1991). As such, the animation is often combined with explanatory verbal information. The cognitive theory of multimedia learning argues that different mental representations have to be constructed from verbal and pictorial information but that these representations have to be actively and simultaneously integrated for meaningful learning to commence (Mayer, 2001; Mayer & Moreno, 2003; Mayer & Sims, 1994).

Four characteristics of animations are relevant for cognitive modeling. The first is that animations can present information that changes with time, such as the working of a device or the explanation of a procedure, by movement of objects in the animation (Ainsworth & VanLabeke, 2004; Hegarty, 2004; Rieber, 1990; Tversky, Morrison, & Betrancourt, 2002; Weiss, Knowlton, & Morrison, 2002). Although the movement of objects is an important type of change within an animation, other types of changes can be distinguished. An interesting division is made by Lowe (1999, 2003), who conceives animations as consisting of one or more objects that can undergo several types of changes. First, there are transformations, which are changes in the properties of objects (e.g., color, shape, size). Subsequently, there are translations, which refer to movements of objects on the screen. Finally, there are transitions, which concern the appearance and/or the disappearance of objects. These three types of changes can occur in isolation (e.g., an object starts flashing), but they typically occur together (e.g., an object starts flashing and moving). In complex animations, more objects exist, and each can have its own regime of changes. For example, in a meteorological animation, several areas of high and low pressure may exist, which move into several directions (i.e., translations), expand or shrink (i.e., transformations), and arise or disintegrate (i.e., transitions). From a cognitive load perspective, the dispersion of information into parts that follow each other sequentially or are presented simultaneously might be problematic: Once a part of the information is missed or only partially processed, the remaining parts might become incomprehensible. To build a coherent representation, the learner has to hold and integrate information from these different parts in working memory and then store it in long-term memory; otherwise, he or she will not be able to retrieve it. If a subsequent part of an animation has to be processed before an earlier part is stored into long-term memory, this new information will interfere with the learner’s remembering the information of the earlier part. This phenomenon is called retroactive inhibition (Baddeley, 1997), and it may be reinforced by the fact that people have limited time to study each part of an animation because of its transient nature (Lowe, 1999, 2003). In the case of a poorly designed animation—that is, when extraneous cognitive load is involved—retroactive inhibition attributed to limited processing time uses up cognitive resources that could better be used for building the cognitive representation.

A second characteristic of animations is that they can be seen as depictive external representations (Schnotz, 2002). The depictive nature of animations enables the visualization of concrete concepts (e.g., the working of a bicycle pump, as depicted by the animation of its working) and abstract concepts represented by concrete events (e.g., the phrase “drawing without replacement” in probability calculation,
which can be depicted by drawing marbles from a vase without returning them). Moreover, an advantage of animations is that they can be shaped, distorted, and manipulated—for example, by showing an object from multiple perspectives and by making it larger or smaller (Hegarty, 2004; Schwan & Riempp, 2004).

The third characteristic is that an animation’s salient features, such as its motion and flashing, can focus the learner’s attention to relevant parts of the screen (Park & Hopkins, 1993; Wetzel et al., 1994). This aspect can be relevant for novices who might be overwhelmed by complex animations (Rieber, 1990). In the animation of a complex system, for example, a flashing arrow could highlight the critical features of the system.

A fourth characteristic reported by a number of researchers is that animations can motivate learners via their cosmetic appeal (Shah & Freedman, 2003; Weiss et al., 2002). For example, animated agents can be used to reduce potentially upsetting information, and popular cartoon figures can be used to engage young learners in learning (Wetzel et al., 1994).

These four characteristics make animations potentially useful in conjunction with cognitive modeling. Take, for example, an animation in which an expert is explaining how meteorological data have to be interpreted to give a sound weather report. The expert verbalizes that several low- and high-pressure areas exist, and he or she details how they interact and how they are geographically related. An animation could visualize this situation and thus make it easier for a novice to make a mental representation of it. Moreover, when the expert stresses the importance of a low-pressure area that is shrinking, the animation could focus the learner’s attention on this point by zooming in on the particular area.

Although dynamic visualizations may seem appealing, several studies and reviews have shown that dynamic visualizations are, at best, no more effective and occasionally even less effective than static visualizations. In an extensive review, Tversky et al. (2002) reported that, in general, dynamic visualizations were not more effective than static visualizations. In the cases in which they were more effective, their efficacy could be ascribed to information that was more detailed in the dynamic visualizations or to the benefits from study procedures (e.g., prediction) that were not available in the static visualizations. In the domain of mechanical systems, Hegarty, Kriz, and Cate (2003) compared learning from animated graphics with that from static diagrams and concluded that both types of visualizations resulted in learning but that the animated graphics did not lead to superior performance. In other cases, the use of dynamic visualizations led to more time spent on instruction without corresponding gains in learning outcomes (Koroghlanian & Klein, 2004). These studies make it clear that there is no strong evidence to ground any claim that dynamic visualizations are better than static visualizations. In their analysis, Tversky et al. (2002) formulated two principles that specify the conditions under which dynamic visualizations may be effective, although not necessarily more effective than static visualizations. First, they postulate the apprehension principle, stating that the structure and content of a dynamic visualization should be readily perceivable and comprehensible (e.g., a dynamic visualization should not go too fast). Second, the congruence principle explains that the structure of a dynamic visualization should correspond with the way that people conceive the processes and procedures that are visualized. For example, if operating
a machine is conceived as a sequence of discrete steps, a dynamic visualization should visualize it that way.

We concur with the notion that research should focus on identifying the conditions under which dynamic visualizations might indeed promote learning (Hegarty, 2004; Mayer & Moreno, 2002; Tversky et al., 2002). Furthermore, we contend that dynamic visualizations might become more effective when they are designed in such a way that cognitive capacity is optimally employed. In this respect, a series of four experiments conducted by Mayer, Hegarty, Mayer, and Campbell (2005) is of interest. Overall, the results of these experiments show that dynamic visualizations (narrated animations) result in poorer learning than that of static visualizations (illustrations on paper). Although Mayer et al. explained these findings in terms of the cognitive theory of multimedia learning, we contend that the dynamic visualizations in these experiments were not designed in such a way that learners’ cognitive capacity was optimally employed. We concur with the authors’ conclusion that the static visualizations are learner paced and segmented in meaningful units whereas the dynamic visualizations are computer paced and continuous. But as we argue later, we consider learner pacing and segmentation as design guidelines that can be used in combination with dynamic visualizations to decrease extraneous cognitive load and thus release cognitive capacity for genuine learning.

**Animated Pedagogical Agents**

Cognitive modeling involves complex skills that often have to be applied in specific contexts in which problems have to be examined from several perspectives. For novices, this itself can pose a problem; as such, support from a pedagogical agent (e.g., a tutor, a peer student, a software agent) might be helpful. Animated pedagogical agents are computerized characters that appear on-screen as a learner engages in a task. They provide support by guiding, coaching, and providing feedback via verbal communication (e.g., explanations) and nonverbal communication (e.g., gazing, gesturing; Atkinson, 2002; Clarebout, Elen, Johnson, & Shaw, 2002; Moreno, 2005). These animated pedagogical agents can be humanlike (e.g., “Herman the Bug,” an insect with some facial expression used in several research projects) or not (e.g., the well-known “Paperclip” of Microsoft Office).

Over the last 5 years, several reviews and studies have been published concerning the instructional value of animated pedagogical agents. An instructional advantage put forward by researchers involves the potential of animated pedagogical agents to motivate learners (Dehn & van Mulken, 2000; Moreno, Mayer, Spires, & Hiller, 2001; Moundridou & Virvou, 2002). For example, Moreno et al. (2001) found that learners in an environment with an animated pedagogical agent were more motivated and interested. Moreno and colleagues explained this motivation effect via social agency theory, which assumes that learners in a social-agent learning environment tend to work harder. Social agency theory was derived from the media equation hypothesis (Reeves & Nass, 1996), which claims that people view interaction with media, such as computers and software, as interaction with humans and that social rules that apply for human-to-human interaction therefore apply for human-to-media interaction. According to social agency theory, multimedia instruction can be regarded as information delivery or as a social event. When social cues are incorporated into the multimedia instruction, people interpret the interaction with the computer as a social event. The theory further argues
that these social cues prime conversation and so engage the learner in efforts to make sense of what the multimedia instruction is saying (Moreno et al., 2001). Furthermore, Moundridou and Virvou (2002) found that from the perception of the learners, an animated pedagogical agent makes learning in an environment with algebraic word problems easier and more pleasant.

Another didactical function is enabled by the current state of technology—namely, that animated pedagogical agents can be programmed to adapt to the characteristics of a specific learner or to the context in which a task is performed (Clarebout et al., 2002; R. E. Clark & Choi, 2005). For example, the agent could scaffold the amount of support and guidance by performing parts of the task that learners cannot perform on their own, by coaching, and by providing hints and feedback specific for a learner.

Reviews regarding the benefits of animated pedagogical agents have yielded mixed results: Some empirical studies found that animated pedagogical agents yield better learning (e.g., Moreno et al., 2001), whereas other studies did not find these benefits (R. E. Clark & Choi, 2005; Dehn & van Mulken, 2000). We contend that animated pedagogical agents must be applied carefully. To start with, it seems that the effect of these agents is domain specific. For example, in a review, Dehn and van Mulken (2000) concluded that the effect of an anthropomorphized agent on entertainment value is domain specific: In a technical system, an anthropomorphized animated pedagogical agent was more entertaining than an agent with a geometrical interface; however, in a system for introducing new employees in an organization, no difference in entertainment value was found. The same pattern was found for assessed task difficulty: When technical information was presented, lower task difficulty was reported when an animated agent referred to the information than when a pointing arrow did. Again, for the introduction of new employees, no difference in perceived task difficulty was found between the animated agent and the pointing arrow.

A second comment pertains to the preference of learners for animated pedagogical agents. Craig, Graesser, Sullins, and Gholson (2004), for example, investigated the relation between (a) different types of affect, such as boredom, flow, and confusion, and (b) learning in an environment about computer literacy that includes an animated conversational agent capable of synthesized speech, gestures, and facial expressions. While learners worked in the environment, their emotions were tracked and coded. Researchers found that affects such as confusion and flow correlate positively with learning gains, whereas an affect such as boredom correlates negatively with learning gains. Trying to ignore an animated pedagogical agent that does not motivate learners but bores and even annoys them imposes an ineffective cognitive load (i.e., extraneous load).

To conclude, we contend that the processing of a sophisticated animated pedagogical agent with many salient details might require so much cognitive capacity that little remains for processing the actual subject matter. We believe that animated pedagogical agents can be beneficial when they are designed according to the guidelines provided by cognitive load theory.

**Design Guidelines for Animated Models**

The purpose of the remainder of this article is to propose guidelines for decreasing extraneous cognitive load (and, if necessary, intrinsic cognitive load) and
increasing germane cognitive load. First, some guidelines are discussed to decrease
intrinsic cognitive load. Specifically, this category classifies guidelines that aim to
reduce or manage the element interactivity of subject matter. Second, guidelines
that decrease extraneous cognitive load are listed as those that reduce activities that
obstruct learning (e.g., visual search caused by split-attention effects). These
guidelines may help to free up processing resources that can be devoted to learn-
ing. Finally, design guidelines are presented to increase germane cognitive load.
These guidelines can make good use of the cognitive resources that have become
available through decreasing intrinsic and/or extraneous cognitive load. The crite-
rion for these guidelines is that they should prompt learners to engage in the active
processing of subject matter. It should be noted that some guidelines may have an
effect on more than one type of cognitive load. For example, element interactivity
may be reduced by providing a simple animated model. It is obvious that such a
model may involve less visual search and thus cause less extraneous cognitive
load. In these cases, the guidelines are classified under the category to which it
primarily aims. In addition, guidelines may be categorized differently in other
classifications. For example, Mayer (2005a) regards segmentation as a guideline
for decreasing intrinsic cognitive load, whereas in this review, we assume that it
decreases extraneous cognitive load. In these cases, we briefly discuss these
differences.

We provide an example for each guideline. For the sake of clarity, we apply
each guideline to the “pin code” animated model, described in the introduction. In
this model, a person has to guess the correct pin code in one trial. In the animated
model, a pedagogical agent explains the necessary steps to calculate the probabil-
ity that this person guesses the correct pin code.

**Guidelines for Decreasing Intrinsic Cognitive Load**

The first guideline to decrease intrinsic cognitive load focuses on scaffolding
learners when they perform so-called whole tasks. In such tasks, learners have to
coordinate and integrate multiple skills and knowledge so that they develop a holis-
tic view of the task (van Merriënboer, 1997). However, for novices, complex whole
tasks may be overwhelming and may thus impose a high level of intrinsic cognitive
load. Therefore, a sequence of simple-to-complex whole tasks is proposed, starting
with relatively simple whole tasks, which enable learners to construct and automate
schemas before they commence with more complex whole tasks (van Merriënboer,
Kirschner, & Kester, 2003; van Merriënboer & Sweller, 2005). A complex skill such
as “searching for literature,” for example, can be simplified by defining conditions
that make the task simple without compromising its whole-task nature, such as the
clarity of the concepts used, the number of articles in the domain, the number of
databases searched for relevant literature, the type of search, and the number of
search terms. The simplest whole task to which learners are initially confronted per-
tains to a domain with clearly defined concepts in which search terms or keywords
that are not connected by the operators and or or result in a limited number of arti-
cles originating from only one database. Whole tasks with increasing complexity can
be constructed by varying these conditions—for instance, learners have to conduct
a search in several databases with many search terms that have to be connected with
and and or operators (van Merriënboer et al., 2003).
With respect to the “pin code” animated model, an additional animated model with more complex conditions can be provided. One of the conditions that determine the complexity of probability calculation is that of the number of individual events that have to be considered. In the original “pin code” animated model, four such events had to be taken into account—that is, the four numbers of the pin code. As such, the additional animated model could be made more complex by increasing the number of individual events—for example, by the learner’s having to calculate the probability that someone correctly guesses a pin code consisting of six numbers.

A second guideline releases the whole-task approach. As such, it can be characterized as pretraining because it departs from the view that isolated components have to be instructed before learners are exposed to the interaction of these components (Mayer, 2005a; Mayer & Moreno, 2003). For example, in meteorology, learners are first instructed what high- and low-pressure systems are before they learn how these systems interact with one another and determine the weather in a region. Pollock, Chandler, and Sweller (2002, Experiments 1 and 3) reduced the complexity in the domain of testing electrical safety by presenting part of the subject matter in a pretraining. In the first phase, only the isolated components were presented, to enable novices to construct a schema of these components. In the second phase, all informational components were explained, including the interactions between them. Notably, the kind of reduction in the first phase may lead to an initial decrease in the learner’s understanding, which is compensated by an increase in the second phase. Pollock et al. found that a group of novices who were exposed to a similar two-phase instruction outperformed a group that was twice exposed to an instruction composed only of interacting components (there was no difference in instruction time). Similar results were found by Mayer, Mathias, and Wetzell (2002) with regard to learning how brakes and pumps work. Learners performed better on transfer when they first received a short training about the names and behaviors of the components, followed by a narrated animation about the way that these components interacted. Finally, in a cause-and-effect system (i.e., the origin of lightning), Mayer and Chandler (2001, Experiment 1) had novices first learn the isolated components, which enabled them to build a rudimentary schema; this was followed by an instructional phase in which the novices learned the causal relation between these components. The researchers found that learners who followed this treatment scored better on transfer than did (a) learners who twice received an instructional format in which the components and their causal relations were integrated or (b) a group who first received instruction about the causal relations between the components, followed by instruction of the components only. A problem with this study is that instruction time was not recorded; thus, it cannot be definitively concluded that the experimental treatment caused the effect or a prolonged instruction time.

In probability calculation problems such as the one used in the “pin code” animated model, the solution method depends on the interaction between “drawing with replacement or not” and “order relevant or not.” As such, in a pretraining strategy, definitions are introduced first—namely, those of relevant formulas and those of concepts such as “drawing with replacement or not” and “order relevant or not.” After which, the “pin code” animated model can be studied; that is, it can show how “drawing with replacement or not” and “order relevant or not” interact and thus determine which method can be used to solve this problem.
In conclusion, pretraining seems more appropriate for animated models in which causal relations prevail, such as cause-and-effect systems and the working of devices; however, it is less appropriate for animated models in which procedures are involved. Consequently, pretraining seems less suitable for the “pin code” animated model.

Guidelines for Decreasing Extraneous Cognitive Load

Pacing is the first guideline for decreasing extraneous cognitive load. It involves the control that the learner or the system (e.g., a computer) exerts over the continuation of the presentation of instructional material. Pacing might enable learners to adapt the presentation of instructional material to their cognitive needs (e.g., by pausing the instruction, by going backward in the material). In a study with a video about nautical knotting, Schwan and Riempp (2004) showed that pacing was heavily used (as instantiated by accelerating, decelerating, stopping, and repeating the video), especially as knot difficulty increased. Specifically, the more difficult that the knots became, the more that pacing was used, which resulted in a better understanding of the underlying processes—that is, less practice time needed to correctly reproduce the knots. Other studies have reported a learner pacing effect (Mayer & Chandler, 2001, Experiment 2; Mayer, Dow, & Mayer, 2003, Experiments 2a and 2b), but they did not take into account the effect of instruction time on transfer performance or the segmentation of instructional material. As such, it is not clear whether the positive effects can be ascribed to learner pacing, the prolonged instruction time, or the segmentation of instructional material.

However, studies in which pacing was a manipulated factor have reported mixed results. Moreno and Valdez (2005, Experiment 2) failed to find a learning advantage on transfer for learner pacing, as compared with system pacing. The learner-paced group even took less time than the system-paced group, which suggests that learners in the former condition might not have been motivated enough to work through the learning environment. Tabbers, Martens, and van Merriënboer (2004) further qualified the pacing effect. When learner pacing was involved, learners who received written explanations outperformed learners who received spoken explanations on transfer test performance. The absence of time pressure for learner pacing possibly enabled learners to process the written text strategically (i.e., scan the text, reread).

In the “pin code” animated model, a limited version of learner pacing may enable learners to pause and continue the animated model. In a more sophisticated version of learner pacing, learners might move forward and backward in the animated model via a slider bar (e.g., they can go forward by dragging the slider to the right).

To conclude, the findings suggest that considerations regarding pacing, such as when to apply pacing, might interrupt the processing of information and thus impose such an extraneous load on the cognitive system of novices that little resources remain for learning. Perhaps, pacing should be implemented in conjunction with other guidelines, such as segmentation.

Segmentation is a second guideline to apply in the decreasing of extraneous cognitive load. The segmentation of an event can be based on the model of event perception (Zacks & Tversky, 2001), which assumes that a continuous event is cognitively represented in a highly structured manner. According to the model of
event perception, events can be decomposed into segments that consist of activity steps. Zacks and Tversky (2001) provided evidence that activity steps with high informative value correctly characterize the segment of which they are a part whereas activity steps with low informative value fail to represent their segment. The places in the event where these highly informative activity steps occur are regarded as breakpoints. With respect to learning procedural tasks from videos (e.g., upgrading a computer), Schwan and Garsoffky (2004) found evidence that these breakpoints are important: The researchers observed that summaries of procedural tasks based on breakpoints are perceived as being equally comprehensible as the complete video but as more comprehensible than summaries based on nonbreakpoints.

Moreover, Schwan and Garsoffky (2004) found that the omission of breakpoints results in higher cognitive costs because the event structure is lost and learners have to use their cognitive resources to cope with this break in the coherence of the event. This observation is in line with that of Schwan, Garsoffky, and Hesse (2000), who argued that film cuts on places where breakpoints occur can facilitate the cognitive processing of breakpoints because such cuts make the breakpoints more salient. Because less effort is needed to search for breakpoints, more cognitive resources are available for building a cognitive representation. Furthermore, in a study of a cause-and-effect system (the origin of lightning), Mayer and Chandler (2001, Experiment 2) reported that learners who received an animation divided into segments scored better on transfer than did learners who receive a continuous animation. However, segmentation is effective only when the learner has completed the processing of one segment before the next segment is presented.

In the “pin code” animated model, the problem-solving process can be regarded as an event. As such, each segment could correspond with one important step in the problem-solving process. For example, Segment 1 could determine whether it is a “drawing with replacement or not,” and Segment 2 could determine whether the “order is relevant or not.” Based on this information, the problem-solving method is chosen in Segment 3. In Segment 4, the problem-solving method is applied, and finally, in Segment 5, the probability is calculated.

A final remark concerns the fact that other perspectives exist on the classification of the segmentation guideline. Whereas in this review, segmentation is regarded as a technique that may help learners avoid visual search, there is the point of view that regards segmentation as a technique for decreasing intrinsic load (Mayer, 2005a; Mayer & Moreno, 2003). In the present review, segmentation is considered in isolation, contrary to the other classification, in which segmentation is applied in conjunction with pacing. We agree that segmentation in combination with learner pacing may help learners to overcome the complexity of subject matter and in that respect decrease intrinsic cognitive load.

The application of the modality principle is the third design guideline. Modality refers to the sensory mode in which verbal code is presented—that is, in a written format or a spoken format (Penney, 1989). Research with respect to the modality of presentation has indicated that spoken verbal explanations are generally superior to written explanations, when used in combination with pictorial learning material (Mayer, 2005a; Moreno & Mayer, 1999; Mousavi, Low, & Sweller, 1995; Sweller et al., 1998; for a meta-analysis, see Ginns, 2005). This finding is ascribed to the modality principle: The combined use of (a) the visual channel for pictorial learning material and (b) the verbal channel for the explanation of this material...
increases effective working memory capacity and so facilitates learning. The modality principle has been recently qualified: To start with, Tabbers (2002) found that pacing is an important factor because written explanatory text is more effective than spoken explanatory text when learners have control over the pacing of the presentation. Perhaps, the absence of time pressure for learner-controlled pacing offers the learner the possibility to process the written text strategically. Second, Mayer, Sobko, and Mautone (2003, Experiment 1) observed that learners who heard a spoken explanation in a standard accent performed better on transfer than did learners who heard the same explanation with a foreign accent. Similarly, they found that a human voice resulted in better transfer than that of a synthesized voice (Experiment 2). Spoken explanations can be used if the “pin code” animated model requires so much visual search that little cognitive capacity remains for processing writing explanations.

Notably, Mayer (2001, 2005a; see also, Mayer & Moreno, 2003) considered the modality principle as a guideline to decrease intrinsic cognitive load. In this review, we regard the modality principle primarily as a guideline to decrease extraneous load by overcoming the split-attention effects typical for complex animated models. However, we concur with Mayer that using spoken explanations instead of written explanations implies that more information can be processed through the visual channel. In this way, more cognitive capacity becomes available for processing complex subject matter.

The use of the contiguity principle is the fourth guideline. This principle states that verbal explanations that accompany pictorial material should be presented contiguously in time or space to overcome the split-attention effect. The rationale underlying the spatial contiguity principle is that visual search between, for example, written explanations and pictorial information is reduced so that cognitive capacity is released for relevant learning activities (Mayer, 2005b). The rationale for temporal contiguity is that the explanation and the pictorial information are simultaneously held active in working memory, which is a condition for integrating both information sources (Mayer, 2005b). In a study of an animation about the formation of lightning, Moreno and Mayer (1999, Experiment 1) observed that learners who received written explanatory text that was close to the animation performed better on transfer than did those who received text that was far away from the animation. Mayer and Sims (1994) compared spoken explanatory text that accompanied an animation (i.e., text concurrently and successively delivered). As such, learners who received the concurrent narrated animation performed better on transfer than did learners who received the successive narrated animation. This result was confirmed by Mayer, Moreno, Boire, and Vagge (1999). The latter study also revealed that the temporal contiguity effect is eliminated when successive narration is broken up into small parts that last only a few seconds. Apparently, the fast alternation between narration and animation enabled the learners to make connections between the verbal information and the pictorial information without overloading the cognitive system.

In the “pin code” animated model, the expert points to the highlighted 10 figures from which the person can choose the second number of the pin code; at this time, the explanation must be spoken that the person can choose 1 of these 10 figures (or, when written, the text should appear close to the highlighted figures).
The application of signaling, or cueing, is the fifth and last guideline. According to Mayer and Moreno (2003), signaling is the provision of cues to the learner regarding how to select and organize the instructional material (see also Mayer, 2005b). In this respect, signaling covers a broad spectrum, stressing key words in a speech, organizing words in printed text by underlining them, presenting images (e.g., arrows) to focus the learner’s attention to a particular part of an animation, and so on. This review focuses on visual cues that are used to prevent visual search. As stated earlier, understanding commences only when the learner connects the verbal information with the pictorial information. With high visual search, learners unnecessarily use cognitive capacity for relating both information sources. Some studies have reported that visual cues fail in multimedia (Tabbers et al., 2004), whereas other studies have shown that cueing can be effective when the amount of necessary visual search is high, such as in complex animations (Jeung, Chandler, & Sweller, 1997). Mautone and Mayer (2001, Experiment 3) investigated the effect of signaling in a narrated animation and found that signaling is effective when both the animation and the narration are signaled (i.e., signaled words spoken with a slower, deeper intonation) but not when only the animation or the narration is signaled. In another study, Craig, Gholson, and Driscoll (2002) failed to prove the effectiveness of a pedagogical agent with a signaling function. In this study, cueing was implemented as global gestures in the direction of the screen to which the learner had to attend; as such, perhaps it was not directive enough to serve its purpose and so yield the desired effects.

To focus the learner’s attention to the parts of the “pin code” animated model that visualize it as a drawing with replacement, the pedagogical agent may first move to the ATM display and (a) point to the position for the second number of the pin code and then (b) point to the highlighted figures from which the second number can be chosen.

Guidelines for Increasing Germane Cognitive Load

Learners typically view animated models passively. With respect to modeling, Bandura (1976) observed that a stronger effect results when learners engage in active coding. Other researchers have advocated active learning (Chi et al., 1989; Mayer, 2001; Wittrock, 1974). Self-explanation has proven to be a successful approach to engaging in the active processing of learning material (Chi et al., 1989; Renkl, 1997; Renkl & Atkinson, 2002; Roy & Chi, 2005). By generating self-explanations, learners integrate newly learned information with prior knowledge, which yields a more integrated knowledge base with increased accessibility, better recall, and higher transfer of learning (Chi, de Leeuw, Chiu, & LaVancher, 1994). Moreover, self-explaining forces learners to explicate their understanding; it might also help them to discover what they do understand and what they do not (Renkl & Atkinson, 2002). The assumption is that active processing yields germane cognitive load. The following guidelines allow learners to engage in self-explanations when learning from animated models.

First, a broad group of guidelines can be summarized as expectancy-driven instructional methods, which enable learners to actively process instructional material by predicting the next step in a sequence (Renkl, 1997). The focus of these guidelines is to help learners to construct or refine an initial schema. Hegarty et al. (2003) reported that learners who were prompted by questions to predict an
outcome (i.e., how a device works) before continuing an animation comprehended the material (i.e., the working of the device) better than did learners who received no prompts. Also, Mayer, Dow, et al. (2003, Experiment 3) gave learners a question before showing an animation about the working of an electric motor and told them that they had to answer the question after the instruction. Learners who received prequestions scored better on transfer than did learners who did not receive prequestions. Furthermore, Renkl (1997) found that learners were successful in solving problems in the domain of probability calculation when they engaged in anticipative reasoning. In anticipative reasoning, learners first think about the next step in a task—for example, in the solution process of a problem—and then compare their understanding with the feedback provided by the learning environment before proceeding with the next step. To conclude, in a study in the field of biology, Moreno et al. (2001, Experiment 3) had learners design a plant—that is, determine the characteristics of its leaves, root, and so on—and relate these characteristics to environmental features, such as rainfall. Learners who participated in the design of a plant before they listened to a spoken instruction scored better on the more difficult transfer problems than did learners who only had to listen to the spoken instruction.

In conclusion, these studies indicate that inciting learners to actively anticipate the problem-solving process (by answering a prequestion, by predicting the next step in a process, etc.) is an effective instructional method that enables learners to engage in relevant learning activities.

The second guideline, subgoaling, seems especially useful for novices who can be easily overwhelmed by a complete solution process because they do not know which elements belong together. In subgoaling, learners are prompted to group coherent steps in a procedure into meaningful subgoals. As such, subgoals can facilitate learners’ solving novel problems by helping them to identify which parts of a previously learned solution procedure need to be modified to solve a novel problem (Catrambone, 1996, 1998). Cues such as labels and visual markers can support learners in creating subgoals and thus encourage a learner to self-explain the purpose of the steps. Take, for example, the situation in which a problem on gravity has to be solved. Without subgoals, learners might have to pass through several steps to solve the problem. However, the learners would not be encouraged to explain why some of these steps belong together. With subgoaling, the first subgoal could be a boldface text in which the learners are asked to identify the forces that act on an object. It is likely that learners would have to identify the forces that act on an object in other gravity problems as well, but the way to achieve this subgoal might be different. Subgoaling is closely related to segmentation. A segment indicates a coherent part within a process or event, and in this respect, it functions as a cue that might enable learners to create a subgoal for that segment.

Imagine that the “pin code” animated model is segmented. At the end of the segment that concludes whether the problem is a drawing with or without replacement, learners can be prompted to formulate a subgoal—namely, by asking them,
“Which factor(s) in this problem determine whether it is a drawing with or without replacement?” In this case, the learner has two cues for formulating the subgoal: the question and the segmentation.

The third guideline, imagination, is derived from studies involving motor skills—specifically, studies that have shown that imagining these skills before performing them leads to better results, when compared with not imagining them before performance (Cooper, Tindall-Ford, Chandler, & Sweller, 2001). Contrary to the expectancy-driven methods, which focus on the acquisition of an initial schema, the major effect of imagination is the facilitation of schema automation (although imagination may facilitate schema construction). Via imagination, an existing schema can be rehearsed and further automated. Because automated schemas can be performed without placing a load on working memory, imagination releases cognitive resources that can be used for other aspects of learning the task. Stimulating learners to imagine procedures and concepts can be an effective guideline for more advanced or proficient learners because imagination is possible only if a schema that can guide behavior has already been acquired. For example, to construct formulas in a spreadsheet, learners who have to imagine the procedure outperform learners who only have to study the procedure—both in the number of correct solutions and in the solution times (Cooper et al., 2001). Leahy and Sweller (2004) confirmed these findings with school teachers and young children in another domain (interpreting contour maps and graphs about weather). Furthermore, they found an interaction between imagination and split attention; that is, after reading a graph with integrated explanatory labels, learners who had to imagine performed better than did learners who had to read a graph with the explanatory labels on a separate page. Apparently, split-attention required so much cognitive resource that little capacity remained for performing the imagination technique.

As such, presume that more proficient learners first study the “pin code” animated model, which then disappears (or the learners turn away from the screen). Subsequently, learners have to imagine performing the problem-solving procedure shown in the animated model, and they have to try to understand it.

The fourth guideline, variability, focuses on the presentation of a sequence of tasks that differ in relevant features. The rationale behind the variability effect is that it encourages learners to identify and distinguish the relevant features from the irrelevant features and by doing so develop appropriate schemas. For example, Quilici and Mayer (1996) exposed one group to a set of statistical word problems that varied in their structural features (e.g., the mathematical procedure that was needed to solve the problem), whereas another group was exposed to a set that varied only in surface features (e.g., the story line of the problem). On a transfer test, the group exposed to variability in structural features outperformed the group that was exposed to surface features. Variability is closely connected to contextual interference—that is, the training conditions in which certain contextual factors prohibit a quick and smooth mastery of the skills being taught (van Merriënboer, Schuurman, de Croock, & Paas, 2002). High contextual interference may be realized by presenting problems in a random order so that each successive problem requires learners to apply different knowledge and skills. This practice schedule enables them to compare the solutions of the problems and thus construct more general applicable schemas that can then be used in larger classes of problems.
Although this might yield an increase in cognitive load and instruction time during the learning phase, it will generate higher transfer performance. For example, in the domain of computer-based, numerically controlled machinery programming, Paas and van Merriënboer (1994) investigated the effects of problem format and variability. They compared a low-variability conventional condition and a high-variability conventional condition in which practice problems had to be solved with worked-example conditions of low and high variability (i.e., worked examples had to be studied). The researchers found that learners who studied high-variable worked examples scored better on transfer than did learners who studied low-variable worked examples. Moreover, they found that high variability was effective (i.e., it imposed germane cognitive load) only in the worked-example condition, where the extraneous cognitive load was sufficiently low to allow learners to profit from increased variability. These findings have been confirmed in the domain of troubleshooting (de Croock, van Merriënboer, & Paas, 1998; van Merriënboer et al., 2002).

An example of variability in the “pin code” animated model involves the introduction of another “pin code” animated model, beside the original one, adapted in such a way that the pin code comprises four different numbers, which makes it a drawing without replacement. This adaptation varies the animated model in a structural feature because it changes the method that can be used to solve the problem. The original “pin code” animated model can be solved with a formula counting all possible combinations of four numbers—that is, $10^4$ (which results in 10,000 combinations). However, only one of these combinations is the correct one. In the adapted version, the drawing is now without replacement, and another formula has to be calculated (i.e., the permutation formula).

To conclude, the variability guideline is an effective instructional method, provided that the extraneous cognitive load is sufficiently low. Moreover, the variability guideline may take more cognitive load during training but yield higher posttest performance.

To conclude this classification, Table 1 provides an overview of the design guidelines, their descriptions, and examples. We should note that there is a close relationship between the different sources of cognitive load and the complexity of the animated models. For simple animated models, sufficient cognitive capacity is available for an increase of germane cognitive load. With animated models of intermediate complexity, germane cognitive load can be increased only when cognitive capacity is released by decreasing extraneous cognitive load. For highly complex animated models, both extraneous and intrinsic cognitive load should be decreased, and if possible, germane load should be increased.

Factors Mediating the Effect of Design Guidelines

Some caution should be taken when applying the guidelines. Several studies have revealed factors that mediate the instructional effects of the design guidelines. The prior knowledge of the learner is the first mediating factor. Recent research on cognitive load theory, for example, has proven that design guidelines that are beneficial for novice learners can be ineffective or even detrimental when applied to experts (Kalyuga, 2005; Kalyuga, Ayres, Chandler, & Sweller, 2003). Novices typically lack the cognitive schemas that release working memory resources and
<table>
<thead>
<tr>
<th>Guideline</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decrease intrinsic cognitive load</strong></td>
<td>Sequence of simple-to-complex whole tasks</td>
<td>Present animated models that require the integration of different skills and knowledge. Start with simple animated models with low element interactivity and gradually increase the complexity.</td>
</tr>
<tr>
<td></td>
<td>Pretraining</td>
<td>First present isolated components before the interaction between these components is instructed in the animated model.</td>
</tr>
<tr>
<td><strong>Decrease extraneous cognitive load</strong></td>
<td>Pacing</td>
<td>Allow learners to adapt the tempo of presentation of the animated model to their cognitive needs.</td>
</tr>
<tr>
<td></td>
<td>Segmentation</td>
<td>Divide animated models in several segments in which each segment corresponds with an important part of a procedure or process.</td>
</tr>
<tr>
<td></td>
<td>Modality principle</td>
<td>Present textual explanations in animated models in spoken format.</td>
</tr>
<tr>
<td></td>
<td>Contiguity principle</td>
<td>Present textual explanations in animated models contiguously in time or space.</td>
</tr>
<tr>
<td></td>
<td>Signaling, or cueing</td>
<td>Present cues to prevent visual search in animated models.</td>
</tr>
<tr>
<td><strong>Increase germane cognitive load</strong></td>
<td>Expectancy-driven methods</td>
<td>Present opportunities in animated models to predict the next step in a process.</td>
</tr>
<tr>
<td></td>
<td>Subgoaling</td>
<td>Prompt learners in animated models to group coherent steps of a procedure into a meaningful subgoal.</td>
</tr>
</tbody>
</table>
so enable the learner to process information effectively. In the case of novices, the application of the design guidelines can compensate for this lack of schemas. More experienced learners, however, already possess schemas to process information effectively, and the guidelines may yield instruction that is less effective for them. If the guidelines are nevertheless used by the designer of the instruction, more experienced learners will try connecting and integrating both schema information from their memory and the information based on the instruction. Given that this is redundant information that they cannot ignore, it can yield high cognitive load or even cognitive overload. As such, this moderating effect of one’s level of expertise is the expertise reversal effect (Kalyuga et al., 2003). For more experienced learners, the pretraining guideline would be less effective because they already possess the necessary schemas and are not confronted with a heavy intrinsic cognitive load. With respect to the imagination guideline, it is clear that this guideline is not appropriate for novice learners, because they do not have the necessary schemas (Cooper et al., 2001).

The spatial ability of learners is the second mediating factor. Mayer and Sims (1994) observed that learners with high spatial ability profit more from animations with concurrent narration than do learners with low spatial ability. The researchers concluded that the latter have to devote so much cognitive resources in constructing a mental visual representation that little resources remained for making connections between the visual and verbal representations. On the contrary, the learners with high spatial ability are able to build a visual representation with much less mental effort and can therefore devote more cognitive resources to the connection of visual and verbal representations.

The motivational aspects of learners is a third mediating factor. According to Fisher and Ford (1998), the allocation of effort toward learning activities is driven by individual motivational processes, such as personal goals and interests, incentives, individual personality differences, and metacognitive knowledge. In this respect, the pattern of cognitive load—that is, the specification of what is extraneous cognitive load and what is germane cognitive load—is not only affected by

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imagination</td>
<td>Stimulate learners to imagine procedures and concepts that are used in animated models.</td>
<td>Learners first study the &quot;pin code&quot; animated model and then have to imagine performing the problem-solving procedure just shown.</td>
</tr>
<tr>
<td>Variability</td>
<td>Present problems that vary in relevant features.</td>
<td>Adapt the animated model in such a way that each number of the pin code must be unique (which makes it a drawing without replacement). This has a consequence for the problem-solving method.</td>
</tr>
</tbody>
</table>

Note. In the Example column, the guidelines apply to the “pin code” animated model, in which a person has to guess the correct pin code of a cash card in one trial. A pedagogical agent in the animated model provides explanations.
instructional design but is mediated by the learners’ learning activities, which in turn depend on the personal goals and interests of the learner (Gerjets & Scheiter, 2003). Holladay and Quinones (2003) provided evidence for the mediating role of motivation. They found that self-efficacy generality—that is, efficacy beliefs related to a specific task that can be generalized to similar tasks—can be regarded as a mechanism to explain the relation between task variability and transfer performance. In a computer-based naval air defense simulation, the higher scores on far transfer could not be ascribed to the high-variability tasks that learners engaged in but to the higher self-efficacy generality resulting from the high-task variability. In other words, there is no direct relation between practice variability and far transfer when the effects of self-efficacy generality are taken away.

Age is the last mediating factor to be discussed. One of the central findings in cognitive aging research is that the efficiency of working memory deteriorates with aging. Several explanations have been proposed to account for this decline (Paas, van Gerven, & Tabbers, 2005). To start with, the view of reduced working memory suggests that elder people have reduced processing capacity, which becomes relevant with complex cognitive tasks (Gilinsky & Judd, 1994; Salthouse, Mitchell, Skovronek, & Babcock, 1989). When tasks become more complex, older adults tend to be slower than younger adults. A second view argues that reduced processing speed is a central mechanism in the explanation of age differences in performance (Fisk & Warr, 1996; Salthouse, 1996). A third view contends that older people cannot suppress irrelevant or extraneous information to the same extent that younger adults can. According to this view, the extraneous information imposes more load on the cognitive system of older adults. Finally, several studies report that older adults show reduced coordination and integration of information sources. For example, Light, Zelinski, and Moore (1982) observed, with regard to deductive reasoning, that older adults had difficulty integrating information across several premises, although they could recognize the separate premises perfectly. Because the total cognitive capacity of older adults is smaller than that of younger adults, the application of guidelines based on cognitive load theory and presented in this review might be proportionally more effective. Paas, Camp, and Rikers (2001) investigated the effects of goal specificity on younger and elder learners. In solving a problem with a specific goal, novices typically rely on weak problem-solving methods, such as working backward from the end goal (i.e., the solution) to subgoals, which imposes such a high level of extraneous cognitive load that little cognitive capacity remains for learning. When solving a problem without such a specific goal, they cannot use weak problem-solving methods and exert the cognitive capacity for learning. Paas et al. (2001) found that the absence of a specific goal has a larger beneficial effect for elderly learners than for younger learners on learning to solve maze problems. With respect to multimedia learning and age, little research has been conducted (for an overview see Paas et al., 2005). A study by van Gerven, Paas, van Merriënboer, Hendriks, and Schmidt (2003) revealed no proportional greater modality effect with older adults, although elder learners reported less cognitive load and needed less training time than younger learners did when they studied multimodal materials (visuals and spoken text) rather than unimodal materials (visuals and written text).
Conclusion and Discussion

In the previous sections, we argue that recent developments in information technology have enabled the application of animations and pedagogical agents with respect to cognitive modeling. Such modeling deals with cognitive processes that are not directly observable. To make this possible, the cognitive processes of the model have to be externalized. Because animations are transient, they fit the dynamic nature of modeling. Moreover, animations can facilitate the externalization of the cognitive processes, especially when they are difficult to describe in words. Animated pedagogical agents can give support to the learner and stimulate the learner to invest effort to understand the model performance depicted in the animation. We also argue that the application of animations can pose substantial extraneous cognitive load on the learner’s cognitive resources because information is dispersed sequentially and simultaneously. According to the current focus of cognitive load theory, extraneous cognitive load should always be minimized; if this is insufficient to prevent cognitive overload, intrinsic cognitive load may be decreased as well, at which time germane cognitive load is increased within the limits of totally available cognitive capacity. Therefore, we present three sets of guidelines. First, we discuss guidelines that can decrease intrinsic cognitive load, such as pretraining and the presentation of a range of tasks in a simple-to-complex sequence. Second, we propose guidelines that can decrease extraneous cognitive load, such as the implementation of pacing, segmentation, the modality effect, the contiguity effect, and signaling. Third, we discuss guidelines that stimulate germane cognitive load, such as the implementation of expectancy-driven instructional methods, subgoaling, imagination, and variability.

We also emphasize that the effectiveness of these guidelines depends on several mediating factors, such as prior knowledge, spatial ability, motivation, and age of the learner. Figure 3 shows an integrative framework for the design of animated models, based on a model by van Gerven et al. (2003). In the figure, the guidelines are located near the type of cognitive load that they influence. The rectangle labeled Available cognitive capacity is put in the center of the figure to indicate that optimizing animated models implies that the cognitive capacity has to be increased. To increase the available cognitive capacity, designers may decrease extraneous cognitive load (e.g., by using the contiguity guideline) and, if necessary, decrease intrinsic cognitive load (e.g., by using the sequence of simple-to-complex whole tasks). However, the available cognitive capacity can be increased or decreased—as depicted by the plus/minus symbol (+/−)—by moderating factors such as spatial ability. For example, when compared to learners with low spatial ability, learners with high spatial ability may have more available cognitive capacity when learning from animated models because they need less cognitive resources to construct a visual representation. The rectangles labeled Available cognitive capacity and Germane cognitive load overlap to indicate that sufficient cognitive capacity is only a requisite for germane cognitive load. It does not guarantee that learners will engage in relevant learning activities that impose germane cognitive load. In most cases, they have to be incited to do so via application of design guidelines such as expectancy-driven methods. Furthermore, Figure 3 makes it clear that an increase in germane cognitive load will increase performance in measures such as transfer and retention.
This review summarizes and evaluates guidelines that can be useful for animated models—that is, the use of animations in conjunction with a supportive pedagogical agent in modeling problem-solving processes. However, as stated here, animations are not always more effective than static visualizations. Furthermore, the creation of animations can be labor-intensive and expensive. Therefore, instructional designers should carefully consider whether animated models or static visualizations are most appropriate for the skills and knowledge that have to be learned.

The review also gives rise to some issues that justify a systematic research on animated models and the most appropriate guidelines to be used. First, researchers have focused on different kinds of domains. Most research is conducted with cause-and-effect systems, such as the origin of lightning (Mayer & Chandler, 2001) and the working of devices (Hegarty et al., 2003). But researchers have also studied the modeling of procedural tasks, such as performing a first-aid task (Michas & Berry, 2000) and tying nautical knots (Schwan & Riempp, 2004). Cause-and-effect systems involve knowledge about their components and knowledge about the behavior of these components. The two-phase approach is put forward as a method to decrease intrinsic cognitive load, and it is appropriate in this case—but less for sequentially oriented events with a strong procedural orientation, such as performing a first-aid task. In this latter case, a simple-to-complex whole-task sequencing approach seems more appropriate.

Second, this review shows that the design guidelines might interact—for example, between modality and pacing (Tabbers, 2002), problem format and variability (Paas & van Merriënboer, 1994), and imagination and spatial contiguity (Leahy & Sweller, 2004). This capability indicates that the design guidelines can be qualified when these interactions are taken into account.

Third, combinations of guidelines, such as the signaling and modality guidelines, might be particularly effective. The pedagogical agent in an animated model might be used to cue the learners’ attention to the relevant part in the animation and meanwhile provide auditory explanatory information.
Fourth, no research has been conducted on the relationships among the guidelines in large training programs. Especially for complex domains, such as learning to maintain computers, training programs are relevant. An interesting avenue for future research is to investigate whether the effect of guidelines in animated models within the context of such training programs is different from that of a single animated model. In a training program for computer maintenance, for example, segmentation and a sequence of simple-to-complex whole tasks can be applied. For computer maintenance, the simplest task might deal with only one obvious computer problem, whereas in the most complex task, several interrelated problems might occur. In the training program, each task might be first presented with an animated model, showing an expert explaining how a problem is solved, followed by a similar task that learners have to perform themselves with a real computer. Whereas a sequence of simple-to-complex whole tasks is used to decrease intrinsic cognitive load, the segmentation guideline for decreasing extraneous cognitive load can be applied for the separate animated models. It would be interesting to know whether segmentation in animated models that are part of a training program would yield effects different from those of segmentation applied to isolated animated models.

To design instructionally effective animated models and to develop a comprehensive design theory for learning from animated models, a thorough and systematic research program is required. In particular, this research program should investigate under which conditions animated models may be effective or not; that is, not only should it consider the guidelines and the mediating factors, but it should also take into account the four issues mentioned here. For example, this review proposes to apply the modality principle (i.e., use spoken explanations instead of written explanations) to decrease extraneous cognitive load. However, learner pacing seems to reverse the advantage of spoken over written explanations. In this case, the research program has to formulate clear research questions that unravel the conditions under which the modality principle is effective in animated models and the conditions under which it is not.

The application of animated models meets two focal points of contemporary educational theory. First, animated models that perform and show how they deal with real-life problems can enable the implementation of authentic learning in a meaningful context. Second, the modeling of cognitive processes with animated models is in line with the current focus on lifelong learning and problem-solving skills. From this perspective, animated models can be a promising instructional approach, provided that a balanced set of guidelines (based on the aforementioned comprehensive design theory) is applied to ensure an optimal use of cognitive resources.

Note

This research project is funded by the Netherlands Organization for Scientific Research (Project Number No. 411-02-163), The Hague.

References


Design Guidelines for Animated Models


Wouters et al.


674
Design Guidelines for Animated Models


Authors

PIETER WOUTERS is a researcher at the Department of Information and Computing Sciences, Utrecht University; e-mail: p.j.m.wouters@cs.uu.nl. He holds a doctorate in instructional design from the Open University of the Netherlands. His research focuses on cognitive and motivational processes in learning from “serious games.”

FRED PAAS is a full professor of educational psychology at the Educational Technology Expertise Center, Open University of the Netherlands, and at the Institute of Psychology, Erasmus University Rotterdam; e-mail: fred.paas@ou.nl. He holds a master’s degree in human movement science from Maastricht University and a doctorate in instructional technology from the University of Twente.

JEROEN J. G. VAN MERRIËNBOER is a full professor of educational technology and scientific director of the Netherlands Expertise Center for Lifelong Learning, Open University of the Netherlands; e-mail: jeroen.vanmerrienboer@ou.nl. His research interests include instructional design for complex cognitive tasks, the use of multimedia in education, and models for lifelong learning.