



# Leveraging multimedia for learning

## Use instructional methods proven to align with natural learning processes

Ruth Clark

If your organization is like many others, you're probably experiencing a demand for digitally delivered training. To save travel costs and instructional time, eLearning in both synchronous and asynchronous formats is increasingly supplementing or even replacing face-to-face classrooms.

However, all too often eLearning fails to live up to its potential, and as a result, learning suffers. Technostics ignore the unique instructional capabilities of eLearning by importing legacy materials from books or classroom manuals without employing engaging multimedia features. Figure 1 offers a typical example. This lesson is essentially a book transferred to a screen. There are no visuals, no audio, and no interactive elements to engage the reader.

At the other end of the spectrum, technophiles who are enamored with technological features use all of them at once. Take a look at the technophilic eLearning sample in Figure 2. Extraneous auditory and visual effects were added to the basic content to make it more interesting. The result is too much stimulus at once.

With these examples in mind, what works most effectively in reaching learners in digital environments? Fortunately, we can rely on over 20 years of research on the learning value of many of the multimedia features available to you today. With Richard Mayer, I have documented this research in our book *e-Learning and the Science of Instruction*. This white paper summarizes some of our most important evidence-based guidelines to help you get the most from your eLearning programs.

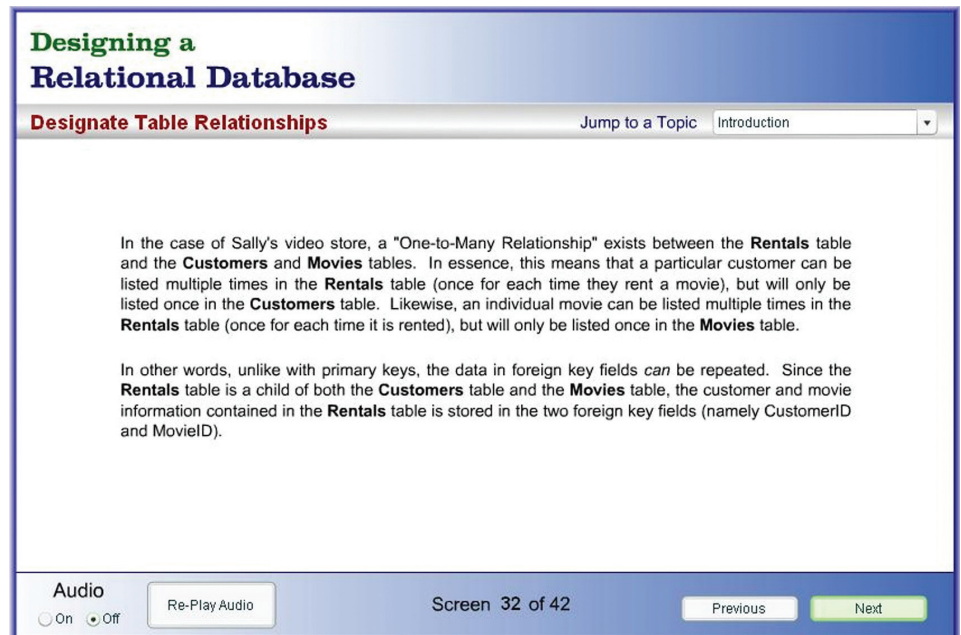


Figure 1. This technostic eLearning approach fails to leverage media features (from Clark & Mayer, 2007).

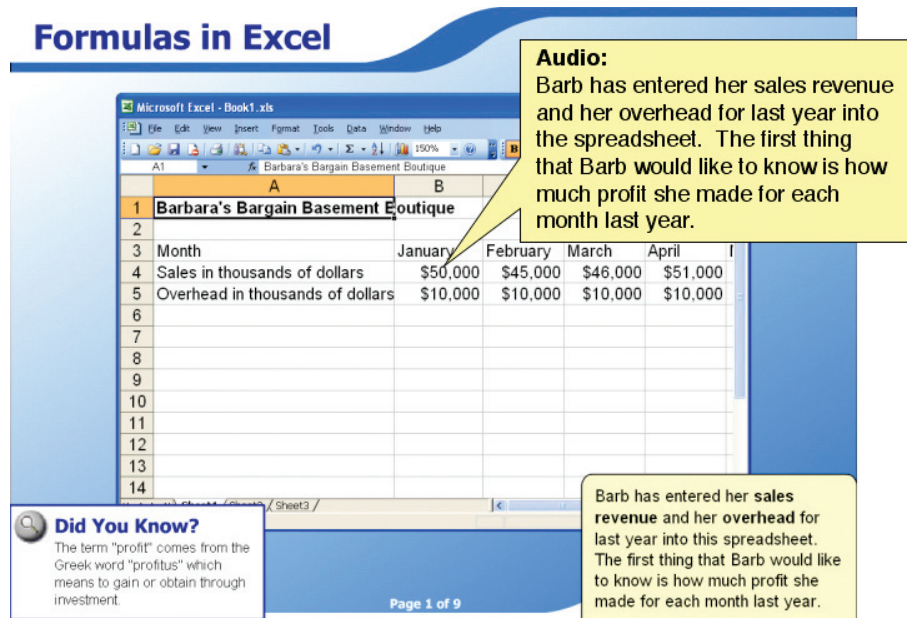


Figure 2. This technophilic approach to eLearning overuses technology features (from Clark, Nguyen, & Sweller, 2006).

### Face-to-face classroom learning versus eLearning

As new training technology emerges, we often see a barrage of research that compares learning from the new technology with learning from more traditional settings—usually classroom instruction. As these research studies are reported, they are brought together in a meta-analysis. In meta-analyses, the data from many individual research studies are integrated statistically. This allows us to make generalizations on the basis of any one research study.

For example, Figure 3 shows a histogram of the effect sizes from over 300 studies comparing learning from various forms of electronic distance technology to learning in face-to-face classrooms. As you can see, most of the effect sizes fall close to zero, indicating no practical learning differences between a digital and face-to-face delivery. However, in some cases, computer-delivered training resulted in more effective learning than classroom learning, and vice versa.

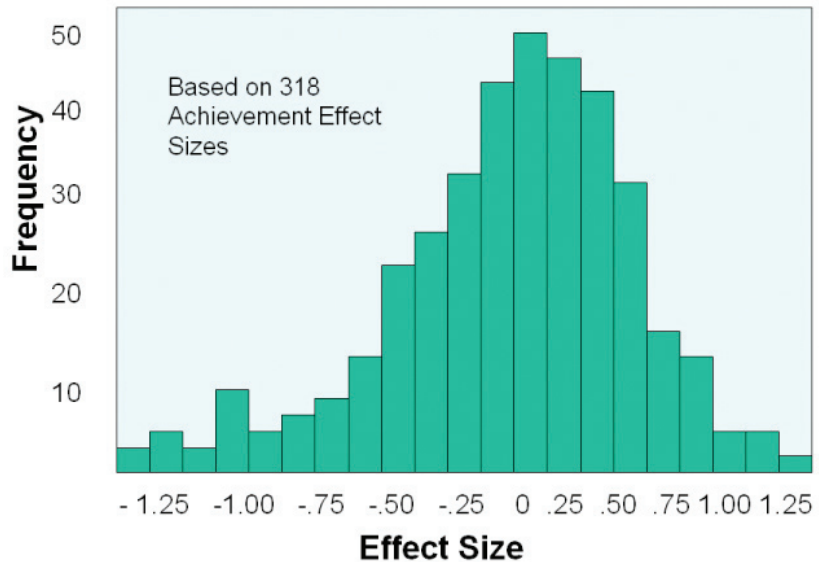


Figure 3. Most comparisons of learning from electronic distance media with classrooms show little difference (based on data from Bernard and others, 2004).

What causes these discrepancies? Have you ever attended an ineffective classroom course? Or experienced poor eLearning materials? The reason we see inconsistent learning outcomes is that the quality of a learning environment is not in the technology, but in how the technology is used.

After hundreds of media comparison studies, we've learned that it's not the delivery media that enables learning; it's how any given delivery technology supports human learning processes. If two lessons include all of the elements needed for learning, learning will occur whether the lesson is offered digitally or in a classroom. Alternatively, if a face-to-face classroom lesson is interactive, while a comparison digital lesson is not interactive, learning will be more easily achieved in the face-to-face version. And vice versa.

### How learning happens

No matter what mix of delivery media you use, it's imperative to accommodate the strengths and weaknesses of the human brain. Our learning is primarily shaped by two memory components: working and long-term memory, as shown in Figure 4. Working memory, as the name implies, is the active member of the pair. It is in working memory that ideas are generated and learning takes place. However, working memory has a very limited capacity. When working memory fills with even limited amounts of information, its processing power diminishes rapidly.

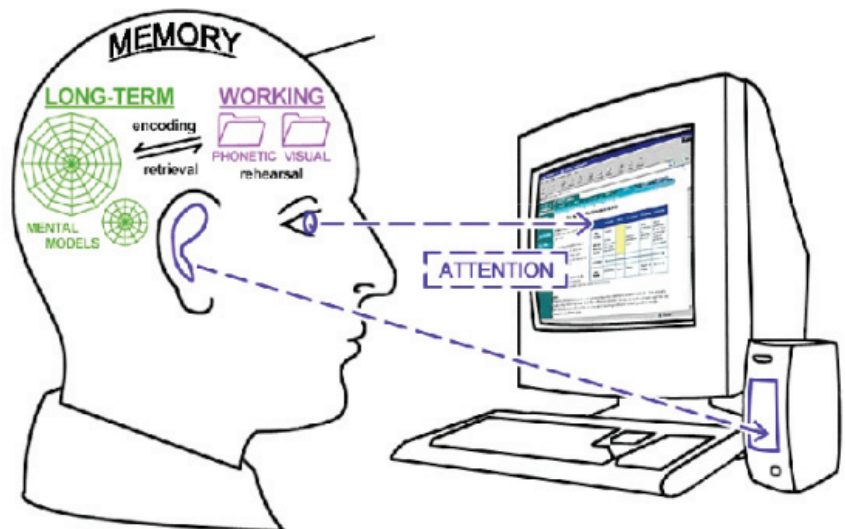


Figure 4. Effective eLearning supports critical psychological learning processes.

In contrast, long-term memory has a large capacity for information; it serves as your repository of knowledge and memories. However, long-term memory provides storage only—all the action takes place in working memory.

When it comes to learning, the goal is to create environments in which learners actively process new information in the working memory in ways that lead to storage in long-term memory. And when needed, can be retrieved back into working memory. Positive learning outcomes require instructional methods that accommodate the limits of working memory and encourage processing of new information for storage in long-term memory.

The main psychological processes you need to support include: attention, management of load in working memory, rehearsal of new information in working memory that results in encoding in long-term memory, and retrieval of new skills back into working memory when needed. Let's take a look at some of the basic instructional methods you can use to support these learning processes in digital learning environments.

**Principle 1: Visualization—use relevant graphics in eLearning**

Include explanatory visuals that illustrate relationships among the content in your lesson. Avoid decorative visuals that do not promote learning. Think visually when you design your instruction to take advantage of eLearning graphic features.

## Principle 1: Use relevant visuals to promote learning

Do learners prefer lessons that include visuals and text, or lessons using text alone? Does adding visuals to instructional materials improve learning? Are all visuals equally effective? Fortunately, we have research to guide us. Richard Mayer at the University of California compared learning from lessons that were all text with the same lessons that added relevant visuals. For example, he created two lessons on how a bicycle pump works. One version used text alone. A comparison version used the same text, but added simple visuals that showed how the pump works. You can see the differences in learning in Figure 5. In nine different experiments, Mayer found, on average, an 89% improvement in learning when a relevant visual was added to text.

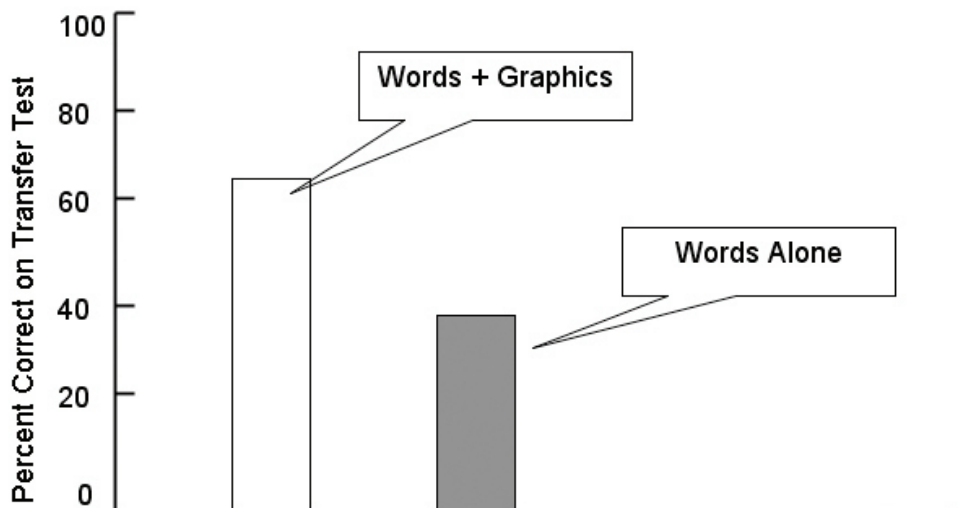


Figure 5. Learning is better using words and graphics than just using words (from Clark & Mayer, 2007).

In *Graphics for Learning*, I suggest different types of useful visuals based on the kind of content you are teaching and the background knowledge of your learners. In general, you should use visuals that illustrate relationships in your content rather than visuals that are decorative or simply reproduce content.

Now we know that visuals can improve learning. But are all visuals equivalent? What about the visual effects used in the lesson in Figure 2?

Mayer did experiments comparing learning with different types of visuals. For example, in a lesson on how lightning forms, he compared a basic version that used relevant visuals with a second version that included everything in the basic version plus additional visuals that illustrated interesting lightening facts. He added pictures of an airplane struck by lightning to illustrate how lightning affects airplanes. He also added visuals of a burned football uniform of a high school student struck by lightning.

Which lesson was more effective? The basic version or the enhanced version with the addition of interesting facts and related visuals?

You may be surprised to learn that the enhanced version depressed learning dramatically. The reason is that the additional information, while related to the topic of lightning, was not relevant to the instructional goal and distracted learners from the main ideas of the lesson. Mayer calls this a “coherence effect.” His research recommends avoiding decorative visuals that do not contribute to the instructional goals.

### Technostic and technophilic approaches to visuals

The technostic tends to ignore the potential for visuals in eLearning by relying primarily on text. In contrast, the technophile tries to supplement lessons with many animated visual effects. Both approaches ignore what research tells us about visuals and learning.

In eLearning, the main connection with the learner is the screen. While pages can deliver easy-to-read text, screens benefit from less text and more visuals. In screen-based media, such as computers and video, it’s much more important to visualize content than it is in paper-based media. However, it’s important to complement the text with relevant visuals rather than with graphics that detract from learning.

### Principle 2: Explain complex visuals with audio

Use audio narration, rather than text or text and audio, to explain a complex visual. Using audio and images manages a learner’s mental load during learning.

## Principle 2: Describe complex visuals with audio only

We’ve seen that relevant visuals can improve learning. However, what is the best way to explain a visual in eLearning? Some people think that to accommodate visual and verbal learning styles, words should be presented in both text and audio formats. Countless studies have touted the learning benefits of audio in explaining visuals. Research scientists refer to the benefits of audio as the “modality principle.” Let’s explore this idea.

Recall the capacity limitations of working memory mentioned previously. Within the framework of working memory there are two small areas for storage of limited information: one for auditory data and one for visual data. As shown in Figure 6, when you explain visuals with words in text, you overload the visual centers of working memory. When you explain visuals with words in audio, you balance your data between the two working memory sub-systems, thereby maximizing the limited capacity of working memory.

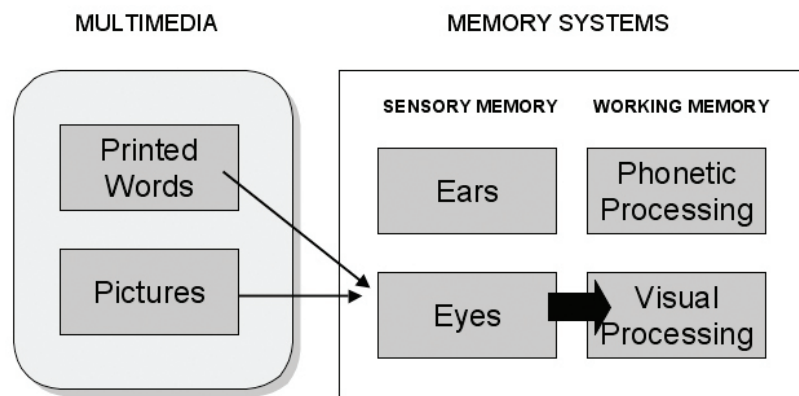


Figure 6. The visual channel is overloaded with the presentation of written text and graphics (from Clark & Mayer, 2007).

Figure 7 shows a good example of a complex visual. This visual is complex because it is animated and includes many details. Animated visuals show a great deal of visual information in a short time span and, therefore, are typically more complex than static visuals. To explain demonstrations like this, it is better to use audio narration rather than just text. However, because in some situations learners may not have access to sound, it is wise to make audio the default and provide a backup version that uses text. With Adobe® Captivate™ software, you can easily add closed captioning that learners can turn on and off depending on their needs.

Although it may seem like a good idea to explain a complex visual such as an animated demonstration with both text and audio, this combination actually results in poorer learning.

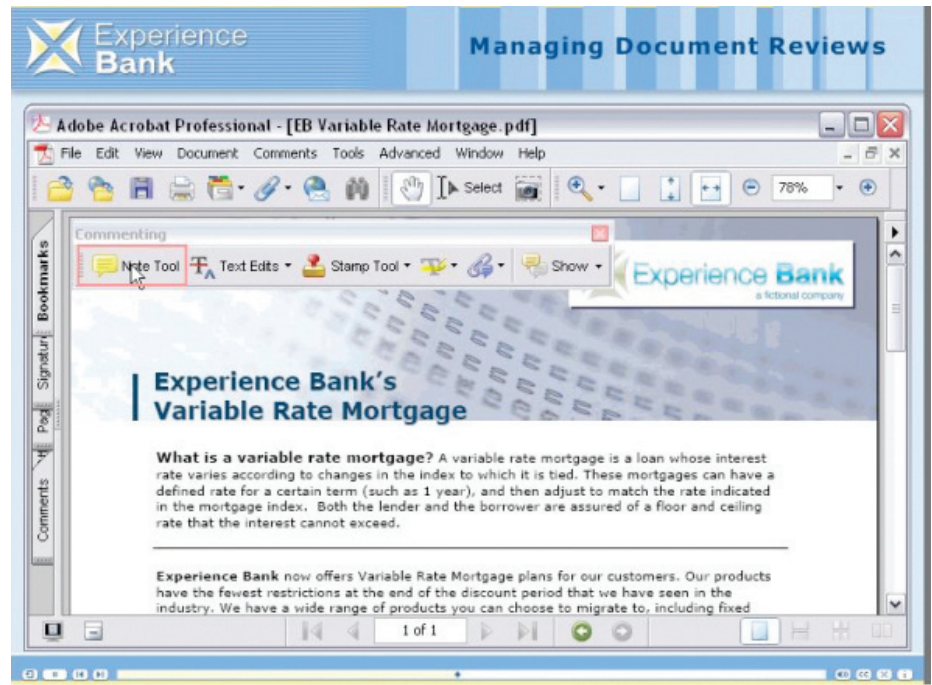


Figure 7. Explain complex visuals with audio narration. Project created with Adobe Captivate.

The depression of learning that occurs when you explain a visual with text, plus audio that repeats the text, is called the “redundancy effect.” The redundancy effect occurs when working memory becomes clogged with excess visual information from the text and graphics. Learners get overwhelmed trying to synchronize onscreen text with audio narration.

### Exceptions to the modality principle

There are times, however, that you may want to use text and not audio. Audio is transient. Once it plays, there is no record of the words. In any situation where learners need to refer to words over time, it’s better to use text. A good example is found in directions to exercises. When learners are responding to an exercise such as a simulation, it’s wise to put exercise directions and feedback in the text—not the audio. (Refer to Figure 8 to see an example.) That way, the learner can easily refer back to the directions as needed.

Another important exception to the modality principle is for learners who are studying in a second language. These learners benefit from taking the time they need to read onscreen text rather than listening to audio narration. It is recommended not to use text and audio at the same time.

## Principle 3: Use first and second person language and learning agents

### Principle 3. Make eLearning conversational

Use first and second person pronouns and learning agents to promote deeper learning.

Have you ever been in a conversation when someone asks you a question and you realize you have not heard much of what was said? All of us feel embarrassed when we are caught not attending to someone talking to us. This social convention is the basis for what Mayer calls “the personalization principle.” According to this idea, learning is better when participants in eLearning feel they are engaged in a conversation.

To engage your learners in a social experience, use informal writing that relies on first and second person language. Of course learners consciously know that they are working with a computer program and not a human partner. Nevertheless, Mayer found that just a few simple changes in language that involved adding “you” and “we” pronouns resulted in dramatic improvements in learning. The reason is that at an unconscious level, we tend to process more deeply when we are in a social-like setting.

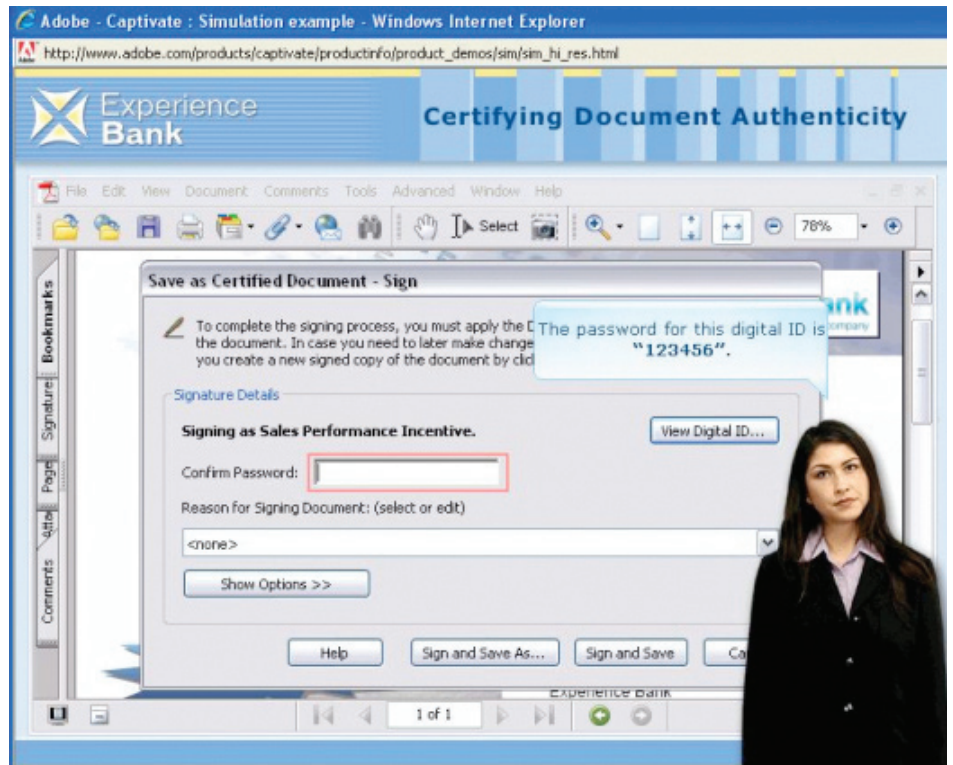


Figure 8. Use text when giving exercise directions as in this simulation exercise.

Another way to engage learners socially is to use onscreen characters, as shown in Figure 9. These characters are called learning agents. Comparisons of different types of images for agents have found that their appearance does not make much difference. You don't need to invest a great deal of effort to construct a highly realistic avatar. You can achieve the same learning effectiveness with simple images.

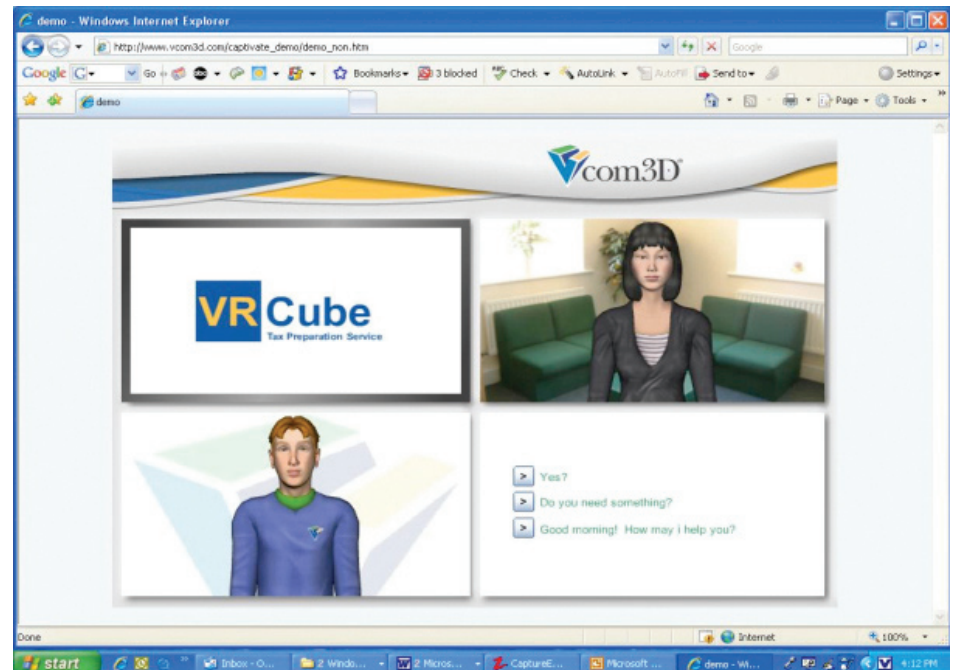


Figure 9. Simple images for learning agents are effective as long as the agent serves a valid instructional role. Created with Adobe Captivate and Vcom3D.

However, the agent’s voice is important. Agents that use conversational audio narration in a familiar human accent rather than onscreen text or machine-generated language have better learning results.

To get the most value from an agent, be sure that the agent serves some instructionally useful role. For example, in an experiment with “Herman the Bug,” an agent used in a botany learning game, Moreno (2004) found that the best learning occurred when Herman gave feedback to a learner’s actions in a friendly, informal tone. To an incorrect response, Herman might say, “Hmmm... Your choice of roots for this dry planet may not be the best to help the plant get the limited moisture available.” In Figure 9, the agent gives directions and informative feedback to the learner’s responses.

## Principle 4: Less is usually more

### Principle 4: Keep eLearning lean

Minimize the script to express essential content succinctly. Keep asynchronous eLearning lessons to 2–5 minutes. Segment complex content into multiple lessons.

In the classroom, it’s not uncommon for a lesson to last an hour or longer. Classroom manuals may contain lessons that include 20 or more pages of content. But what is the best length for an eLearning lesson? We can derive some guidelines based on comparing the learning achieved in longer versus shorter lessons.

There are several reasons you may be tempted to add words to an eLearning lesson. First, you may want to make a dull technical lesson more interesting by adding engaging information or stories. Second, you may want to be extra clear in an eLearning lesson because there is no instructor present to address questions. To do so, you may provide additional explanations of key points. Finally, subject matter experts often add what is commonly called “nice to know” information for technical depth. Based on research on each of these three forms of extraneous words, it is recommended to use only the essential words needed to support your instructional goals.

Mayer (2005) has compared lean lessons with lessons that included additional content to add interest, embellish key points, or provide technical details. For example, in one experiment, learning from a lightning lesson that included 600 words and five captioned illustrations was compared with learning from the same lesson that included 80 words and five illustrations. Learning was achieved much more readily from the lean version.

Often, however, you may be faced with having to communicate extensive technical information. In these situations, it is recommended to segment your lessons so that they can be completed in short chunks.

Although there is no research to prescribe an exact lesson length, it is smart to limit asynchronous eLearning lessons to 2–5 minutes and synchronous eLearning lessons to an hour. Recall that working memory has considerable restrictions on how much information can be held at one time. In addition, unlike a classroom setting, there are many competing activities that can easily distract learners from eLearning goals. Learners will not want to sit through lengthy audio segments. Attention will wander. Based on research and learner feedback, apply the “less is more” principle by designing brief lessons that use the least amount of words needed to achieve the instructional goal.

## Principle 5: Include frequent job-relevant interactions and feedback

### Principle 5. Engage learners meaningfully throughout your lesson

Include frequent job-relevant practice opportunities distributed throughout your lessons. Incorporate corrective feedback to learner responses.

All too often eLearning lessons, such as the one shown in Figure 1, fail to engage participants. For example, lessons may demonstrate how to use new software, but fail to use the simulation capabilities that involve learners in practice exercises. Perhaps the most important features of eLearning are the facilities used to interact with learners—those that ask learners to respond to job-realistic scenarios and get corrective feedback. Computers have a unique media capability to ask questions, judge responses, and provide feedback. Failure to use these features drastically undercuts the potential of eLearning.

Learning happens when new information is actively processed in working memory and results in new information stored in long-term memory. Some learners can passively view a lesson and actively process new information. Most, however, tune out. Or if they do process new information, they end up with misconceptions about it.



The surest road to learning is to engineer overt interactions. In *e-Learning and the Science of Instruction*, we summarize the research on type, frequency, and placement of practice opportunities in digital learning environments. The following is a brief synopsis of that research.

First, ensure that all of your interactions are job-relevant. Often we see questions that exercise memory level only, such as “fill in a blank” or “click on the guidelines for a good client response.” These questions merely ask learners to regurgitate information. This type of shallow processing fails to build the knowledge and skills that apply to the job. Rather than asking learners to “click on the guidelines for a good client response,” design a simulation in which they will respond to the client and see the client’s reactions, as shown in Figure 9.

Second, include frequent interactions dispersed throughout the lesson. Two lessons with the same number of practice items but grouped differently were compared. One lesson grouped the items into two batches, and the other distributed the items into four batches. The lesson with four batches showed much better long-term retention. Having frequent relevant interactions dispersed throughout your lesson will help sustain attention and promote the continuous processing of new content in working memory.

With this in mind, how much practice do learners need? We know that skill improvement can continue over many practice sessions—although with diminishing returns. The greatest amount of learning occurs in the first few practice sessions. How much practice to include depends on the criticality of the skills you are building, and on the extent to which performance can improve on the job. For some tasks, such as landing an airplane, it’s critical that the first performance is highly effective.

In other cases, learners can continue to practice and improve on the job. My recommendation is to adjust the amount of practice according to the criticality of the tasks and the cost benefits generated by additional practice opportunities.

It is also helpful to provide informative feedback with your practice exercises. Informative feedback tells respondents that they are correct or incorrect—and explains why. Moreno (2004) compared two versions of an eLearning game, one that gave explanatory feedback, and one that told learners only that they were right or wrong. Learning occurred more easily in the version with explanatory feedback. Presenting feedback in onscreen text rather than audio gives learners ample time to review it.

## Aligning eLearning with essential learning processes

eLearning offers great potential for performance improvement at low cost. However, for learners to achieve the best results, it is important to avoid a technophobic or a technophilic approach. Instead, evaluate the facilities available to you in your eLearning tool, and use them in ways proven to maximize learning.

In this white paper, we’ve looked at some of the most important guidelines for using visuals, audio, text, and interactivity in ways that promote human learning processes. To explore this information in greater depth, please refer to the books and articles cited.

## eLearning resources

Bernard, R. M., P. C. Abrami, Y. Lou, E. Borokhovski, A. Wade, L. Wozney, P. A. Wallet, M. Fixet, and B. Huang. 2004. "How does distance education compare with classroom instruction? A meta-analysis of the empirical literature." *Review of Educational Research* 74(3): 379–439.

Clark, R. C., and R. E. Mayer. 2007. *e-Learning and the Science of Instruction*. 2nd edition. San Francisco: Pfeiffer. (Available late 2007.)

Clark, R. C., and C. Lyons. 2004. *Graphics for Learning*. San Francisco: Pfeiffer.

Clark, R. C., F. Nguyen, and J. Sweller. 2006. *Efficiency in Learning*. San Francisco: Pfeiffer.

Mayer, R. E., ed. 2005. *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

Moreno, R. 2004. "Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia." *Instructional Science* 32: 99–113.

### For more information

For more details about Adobe Captivate and to view the interactive example, visit [www.adobe.com/products/captivate](http://www.adobe.com/products/captivate).



Adobe

**Adobe Systems Incorporated**  
345 Park Avenue  
San Jose, CA 95110-2704  
USA  
[www.adobe.com](http://www.adobe.com)

Adobe, the Adobe logo, and Captivate are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries. All other trademarks are the property of their respective owners. The names and logos referred to in the sample artwork are fictional and not intended to refer to any actual organization or products.

© 2007 Adobe Systems Incorporated. All rights reserved. Printed in the USA.  
95009711 6/07