CHAPTER OUTLINE

Contiguity Principle 1: Place Printed Words Near Corresponding Graphics

Violations of Contiguity Principle 1
- Avoid Separation of Text and Graphics on Scrolling Screens
- Avoid Separation of Feedback from Questions or Responses
- Avoid Separating Lesson Screens with Linked Windows
- Avoid Presenting Exercise Directions Separate from the Exercise
- Avoid Displaying Captions at the Bottom of Screens
- Avoid Simultaneous Display of Animations and Related Text
- Avoid Using a Legend to Indicate the Parts of a Graphic

Contiguity Principle 2: Synchronize Spoken Words with Corresponding Graphics

Violations of Contiguity Principle 2
- Avoid Separation of Graphics and Narration Through Icons
- Avoid Separation of Graphics and Narration in a Continuous Presentation

Psychological Reasons for the Contiguity Principle

Evidence for Presenting Printed Words Near Corresponding Graphics

Evidence for Presenting Spoken Words at the Same Time as Corresponding Graphics
Applying the Contiguity Principle

WHAT’S NEW IN THIS CHAPTER?

SOMETIMES IN E-LEARNING that uses on-screen text to explain graphics, a scrolling screen reveals the text, followed by the graphic further down the screen. When you scroll down to the graphic, the corresponding text has scrolled out of the window from above; when you scroll up to see the text, the corresponding graphic has scrolled out of the window from below. The result is a physical separation of the text and the graphic. Alternatively, audio narration may be presented before or after the graphics it describes. When you click on a speaker icon, you can hear a brief narration, and when you click on a movie icon, you can see a brief animation, but the narration and animation are separated in time. In this chapter we summarize the empirical evidence for learning gains resulting from presenting text.
and graphics in an integrated fashion (that is, placing printed words placed next to the part of the graphic they describe or presenting spoken words at the same time as a corresponding graphic), rather than presenting the same information separately. The psychological advantage of integrating text and graphics results from a reduced need to search for which parts of a graphic correspond to which words, thereby allowing the user to devote limited cognitive resources to understanding the materials.

In this third edition, we present new evidence concerning the contiguity principle. The new evidence includes research on eye-tracking and pop-up windows. In this new edition, we also clarify some of the boundary conditions under which the contiguity principle applies most strongly.

**DESIGN DILEMMA: YOU DECIDE**

The e-learning design team is reviewing storyboards for their course on spreadsheets for small business owners. To accommodate different learning styles, they have decided to include both text and audio options in the lessons. To apply the multimedia principle discussed in Chapter 4, Ben has added some simple but relevant visuals to illustrate the concepts. For example, to show how to use the logic functions in spreadsheets, he gives an explanation in text and includes two small examples. As shown in Figure 5.1, he asks the learner to click on the small example screens to view the examples.

In reviewing the screens, Reshmi feels that the text explanations and the visual examples should be viewed together. “I recall reading research proving that it is better to allow the learner to view both text and visuals in close alignment.” “That’s a good idea in many situations,” Ben replies. “However, it would take too much screen real estate to include a large graphic and a coherent text explanation!” Based on your own experience or intuition, which of the following options is best:

A. Ben is right. To make sense, the visual examples must be displayed as small screens to be viewed after reading the text explanation.

B. Reshmi is right. Learning is more efficient when visuals and text are integrated. The text explanation should be integrated close to the visual examples.
Chapter 5: Applying the Contiguity Principle

Contiguity Principle 1: Place Printed Words Near Corresponding Graphics

The principle of contiguity involves the need to coordinate printed words and graphics. In this chapter, we focus on the idea that on-screen words should be placed near the parts of the on-screen graphics to which they refer. We recommend that corresponding graphics and printed words be placed near each other on the screen (that is, contiguous in space).

In designing or selecting e-learning courseware, consider how on-screen text is integrated with on-screen graphics. In particular, when printed words
refer to parts of on-screen graphics, make sure the printed words are placed next to the corresponding part of a graphic to which they refer. For example, when the graphic is a diagram showing the parts of an object, the printed names of the parts should be placed near the corresponding parts of the diagram, using a pointing line to connect the name to the part, rather than at the bottom of the graphic as a caption or legend. Similarly, when a lesson presents words that describe actions (or states) depicted in the series of still frames, make sure that text describing an action (or state) is placed near the corresponding part of the graphic, using a pointing line to connect the text with the graphic, rather than in a caption or in the main text.

When there is too much text to fit on the screen, the text describing each action or state can appear as a small pop-up message that appears when the mouse touches the corresponding portion of the graphic. This technique is called a mouse-over or rollover. For example, Figure 5.2 shows an application...
screen that uses the rollover technique. When learners place their cursors over different sections of the application screen, a text caption appears that explains that section. In Figure 5.2 the mouse has rolled over section 1 and the text window below it appears as long as the mouse remains in that area of the screen. One problem with rollovers is that they are transient. The text box disappears when the cursor moves to a different location on the screen. Thus, rollovers may not be appropriate for situations in which it's important for the learner to view more than one block of rollover text at a time or to take an action that relies on rollover text.

Violations of Contiguity Principle 1

Violations of the contiguity principle are all too common. The following list gives some of the most common violations (although there are more) of this principle that are frequently seen in e-learning courseware:

- In a scrolling window, graphics and corresponding printed text are separated, one before the other, and partially obscured because of scrolling screens.
- Feedback is displayed on a separate screen from the practice or question.
- Links leading to an on-screen reference appear in a second browser window that covers the related information on the initial screen (that is, printed text is in one window and graphics are in another).
- Directions to complete practice exercises are placed on a separate screen from the application screen in which the directions are to be followed.
- All text is placed at the bottom of the screen away from graphics.
- An animation plays on one half of the screen while text describing the animation is displayed simultaneously on the other half of the screen.
- Key elements in a graphic are numbered, and a legend at the bottom of the screen includes the name for each numbered element.
Avoid Separation of Text and Graphics on Scrolling Screens

Sometimes scrolling screens are poorly designed so that text is presented first and the visual illustration appears further down the screen, as illustrated in Figure 5.3. As the user scrolls down to view the graphic, the text is no longer visible and vice versa. This is a common problem we see in many courses that use scrolling screens to present instructional content. This particular problem can be remedied by integrating text and visuals on a scrolling screen, as shown in Figure 5.4. Another remedy to the scrolling screen problem is to use text boxes that pop up over graphics when the graphic is touched by the cursor (as shown in Figure 5.2). Alternatively, fixed screen displays can be used when it is important to see the text and graphic together. On a fixed screen, the graphic can fill the screen and text can be embedded within the graphic near the element being described.

Figure 5.3. Text and Graphic Separated on Scrolling Screen.
Avoid Separation of Feedback from Questions or Responses

Another common violation of the contiguity principle is when feedback is placed on a screen separate from the question or from the learner’s answers. This requires the learner to page back and forth between the question and the feedback, adding cognitive load to learning. For example, in Figure 5.5 from our pharmaceutical sales example lesson, a multiple-select question (not shown) requires the learner to select physicians whose practice would benefit from a new drug. When the learner clicks “done,” he or she is routed to a screen (Feedback A) that shows the correct answers. In order to compare their answers with the correct answers, the learners must page back to the question screen. A better solution is shown in the Feedback B screen. In this screen the learner’s answers (checks in boxes) have been carried over from the
Figure 5.5. Ineffective and Effective Placement of Feedback.

**Ineffective Feedback**

**Question:**
Which doctors are potential accounts for Lestrian?

When you have made your choices, click Ok.

- □ A) Dr. Jona - internist staff physician at community hospital
- □ B) Dr. Chi - private practice in low socio-economic area of city
- □ C) Dr. Marks - HMO GYN department head
- □ D) Dr. Zuri - family practice in small clinical rural setting
- □ E) Dr. Mettel - general practice in retirement community

The correct answers are shown. Click Back to compare your answers. Click Next to continue.

**Effective Feedback**

**Question:**
Which doctors are potential accounts for Lestrian?

When you have made your choices, click Ok.

- ✓ A) Dr. Jona - internist staff physician at community hospital
- ✓ B) Dr. Chi - private practice in low socio-economic area of city
- ✓ C) Dr. Marks - HMO GYN department head
- ✓ D) Dr. Zuri - family practice in small clinical rural setting
- ✓ E) Dr. Mettel - general practice in retirement community

Yes. Dr. Jona and Dr. Chi have current patients that would benefit from Lestrian. In the near future, Dr. Mettel may gain patients with a need when the new wing opens.

Click Next to continue.
question screen and placed next to the correct answer, allowing a quick and easy comparison without paging back.

**Avoid Separating Lesson Screens with Linked Windows**

The use of links that lead to adjunct information is common in e-learning. However, when the linked information covers related information on the primary screen, this practice can create a problem. For example, a link on an application screen leads to a window containing a job aid. Having access to reference material is a good idea for memory support. However, if the resulting window covers the graphic example that it describes, the contiguity principle is violated. A better solution is to link to a window that is small, can be moved around on the main screen, and/or can be printed.

**Avoid Presenting Exercise Directions Separate from the Exercise**

Another common violation of the contiguity principle is the practice of presenting exercise directions in text separated from the screens on which the actions are to be taken. For example, in Figure 5.6 we see textual directions for a case study from an Excel e-learning lesson. When moving to the spreadsheet on the next screen, the learner no longer has access to the directions. A better alternative is to put the step-by-step directions in a box that can be minimized or moved on the application screen.

**Avoid Displaying Captions at the Bottom of Screens**

For consistency, many e-learning designs place all text in a box at the bottom of the screen such as the frame shown in Figure 5.7A. The problem with this layout is that the learner needs to scan back and forth between the words at the bottom of the screen and the part of the graphic they describe. A better arrangement is to relocate the text closer to the visual as well as to insert lines to connect the text and visual, as shown in Figure 5.7B. Alternatively, the text can be broken into shorter segments, with each segment placed next to the part of the graphic it describes.
Figure 5.6. Separating Exercise Directions from Application Screen Adds Extraneous Cognitive Load.

Figure 5.7. Text Placed at Bottom of Screen Versus Next to Visual.

A. Text at Bottom of Screen  
B. Text Next to Visual
Avoid Simultaneous Display of Animations and Related Text

You may want to use an animation to depict movement such as to show how to perform a computer application or to illustrate how equipment works. If the animation is playing at the same time as the text is displayed, the learners can either view the animation or read the descriptive text. If they read the text, they miss much of the animation or if they watch the animation then they will read the text after the animation has run. A better solution is to present the text for reading and instruct the learner to press a play button to view the animation after reading, as shown in Figure 5.8. The text remains on the screen for review as desired while the learner watches the animation.

Avoid Using a Legend to Indicate the Parts of a Graphic

Suppose you wanted students to learn about the parts in a piece of equipment. You could show them an illustration in which each equipment part is numbered and a legend below the illustration describes each one. The
problem with this layout is that the learner must scan between the number and the legend, which creates wasted cognitive processing. A more efficient design would place the name and part description in a separate box near the corresponding part on the visual. The text could be placed in a rollover box or in a fixed display on the screen. If the learner will benefit from seeing several parts simultaneously, leaving them on the screen in a fixed display would be better than a rollover box that disappears when the cursor is moved.

Contiguity Principle 2: Synchronize Spoken Words with Corresponding Graphics

Another version of the contiguity principle deals with the need to coordinate spoken words and graphics. In this section we focus on the idea that spoken words (narration) that describe an event should play at the same time as the graphic (animation or video) depicting the event. In short, we recommend that corresponding graphics and spoken words be presented at the same time (that is, contiguous—next to each other—in time).

When e-learning courseware contains narration and corresponding graphics (animation or video), you should consider how spoken words are integrated with on-screen graphics. In particular, when spoken words describe actions that are depicted in the on-screen graphics, make sure the corresponding spoken words and graphics are presented at the same time. For example, when the graphic is an animation showing the steps in a process, the narration describing a particular step should be presented at the same time that the step is shown on the screen. When the graphic is a video showing how to perform a task, the narration describing each step should be presented at the same time as the action shown on the screen.

Violations of Contiguity Principle 2

Violations of the contiguity principle include the following:

- A link to audio is indicated by one icon and a link to video is indicated by another icon.

- A segment provides a narrated description followed by animation or video.
Avoid Separation of Graphics and Narration Through Icons

Suppose you click on “How the Heart Works” in an online encyclopedia, and two buttons appear—a speaker button indicating that you can listen to a short narration about the four steps in heart cycle and a movie button indicating that you can watch a short animation, as illustrated in Figure 5.9. You click on the speaker button and listen to a description of the four steps in the heart cycle. Then you click on the movie button and watch a narration showing the four steps in the heart cycle. You might think this is an excellent presentation because you can select which mode of presentation you prefer. You might like the idea that you listen to the explanation first and then watch, or vice versa, thereby giving you two complementary exposures to the same material.

What’s wrong with this situation? The problem is that, when a lesson separates corresponding words and graphics, learners experience a heavier load on working memory—leaving less capacity for deep learning. Consider the learner’s cognitive processing during learning when a narration is followed

Figure 5.9. Narration Is Presented Separately from Animation.
by an animation. After listening to the narration, the learner needs to hold all the relevant words in working memory and then match up each segment with the corresponding segment of the animation. However, having to hold so much information in working memory can be overwhelming, so the learner may not be able to engage in other cognitive processes needed for deep learning. This is the type of load we called extraneous processing in Chapter 2. Extraneous processing refers to mental load that does not contribute to learning. Therefore, we recommend that you avoid e-learning lessons that present narration and graphics separately.

Avoid Separation of Graphics and Narration in a Continuous Presentation

Even when a lesson presents graphics and narration as a continuous unit, a lesson may be designed so that an introduction is presented as a brief narration that is followed by graphics (such as an animation, video, or series of still frames depicting the same material). For example, consider a multimedia presentation on “How the Heart Works” that begins with a narrator describing the four steps in the heart cycle, followed by four still frames depicting the four steps in the heart cycle.

At first glance, you might like this arrangement because you get a general orientation in words before you inspect a graphic. Yet, like the previous scenario, this situation can create cognitive overload because the learner has to mentally hold the words in working memory until the graphic appears—thereby creating a form of extraneous cognitive processing. To overcome this problem, we recommend presenting the narration at the same time the static frames are presented. In this situation, the learner can more easily make mental connections between corresponding words and graphics.

Psychological Reasons for the Contiguity Principle

As we have reviewed in the examples shown in the previous sections, it is not unusual to see (1) corresponding printed text and graphics physically separated in e-lessons or (2) corresponding narration and graphics presented
at different times in e-lessons. The physical separation may occur because of vertical placement of printed text and graphics (one on top of the other), which separates them when the screen is scrolled, or by a text window covering a graphics window or vice versa, or by placing related information on separate fixed screen displays. The temporal separation may occur because a narrated introduction precedes a graphic, or graphics and narration are accessed through clicking on different icons.

Some designers separate words and pictures because they haven't stopped to think about whether it's an effective way to present information. Others reason that presenting the same material in two different places on the page or at two different times allows learners to choose the format that best suits their needs or even to experience the same information in two different ways. We recommend against separating words and pictures, even for environments with high traffic and low bandwidth, because it is not based on an accurate understanding of how people learn. Rather than being copy machines that record incoming information, humans are sense-makers who try to see the meaningful relations between words and pictures. When words and pictures are separated from one another on the screen or in time, people must use their scarce cognitive resources just to match them up. This creates what we call *extraneous processing*—cognitive processing that is unrelated to the instructional goal. When learners use their limited cognitive capacity for extraneous processing, they have less capacity to use to mentally organize and integrate the material.

In contrast, when words and pictures are integrated, people can hold them together in their working memories and therefore make meaningful connections between them. This act of mentally connecting corresponding words and pictures is an important part of the sense-making process that leads to meaningful learning. As we saw in Chapter 2, it is in working memory that the related incoming information is organized and integrated with existing knowledge in long-term memory. When the learner has to do the added work of coordinating corresponding words and visual components that are separated on the screen or in time, the limited capacity of working memory is taxed—leading to cognitive overload. Ayres and Sweller (2005) argue that putting corresponding words and pictures far apart from each other
(or presenting them at different times) creates what they call *split attention*, which forces the learner to use limited working memory capacity to coordinate the multiple sources of information. You should avoid instructional designs that cause split attention because they force the learner to waste precious cognitive processing on trying to coordinate two disparate sources of information.

**Evidence for Presenting Printed Words Near Corresponding Graphics**

Our first recommendation—presenting corresponding printed text and graphics near each other on the screen—is not only based on cognitive theory, but it is also based on several relevant research studies (Mayer, 1989b; Mayer, Steinhoff, Bower, & Mars, 1995; Moreno & Mayer, 1999a). In five different tests involving lessons on lightning formation and how cars’ braking systems work, learners received printed text and illustrations containing several frames (or on-screen text with animation). For one group of learners (integrated group), text was placed near the part of the illustration that it described, as you can see in Figure 5.10A. For another group (separated group), the same text was placed under the illustration as a caption, as you

---

**Figure 5.10. Screens from Lightning Lesson with Integrated Text and Graphics and Separated Text and Graphics.**

Adapted from Mayer, 2001, 2005c.
can see in Figure 5.10B. In five studies, the integrated group performed better on problem-solving transfer tests than the separated group. Overall, the integrated group produced between 43 and 89 percent more solutions than the separated group. The median gain across all the studies was 68 percent for an effect size of 1.12, which, as mentioned in Chapter 3, is a large effect. Figure 5.11 summarizes the results from one of the experiments.

Figure 5.11. Learning Is Better from Integrated Text and Graphics Than from Separated Text and Graphics.
Adapted from Mayer 2001a, 2005b.

Similar results have been found with training programs for technical tasks (Chandler & Sweller, 1991; Paas & Van Merrienboer, 1994; Sweller & Chandler, 1994; Sweller, Chandler, Tierney, & Cooper, 1990), practical training in physical therapy (Pociask & Morrison, 2008), and even with a single scientific illustration and explanatory text presented on a computer screen (Florax & Ploetzner, 2010). Erhel and Jamet (2006) found that people learned better from an online lesson on the human heart when pop-up windows containing text appeared next to the part of the graphic they described, rather than having the text at the bottom of the screen. In a systematic review of thirty-seven studies, Ginns (2006) found strong support for the benefits of spatial contiguity, with an average effect size of .72.

Additional evidence comes from eye-tracking studies involving text and corresponding diagrams. Successful learners tended to read a portion of the
text, then search the diagram for the object being described in the text, then read the next portion of text and search the diagram for the object being described, and so on (Hegarty, Carpenter, & Just, 1996; Schmidt-Weigand, Kohnert, & Glowalla, 2010). It seems reasonable that we can simplify this process for all learners by breaking text into chunks and by placing each chunk of text near the part of the graphic that it describes. For example, in a naturalistic eye-tracking study shown in Figure 5.12, newspaper readers were more likely to look back and forth between corresponding words and graphics (which contributes to meaningful learning) if the words were placed next to corresponding graphics on the newspaper page (Holsanova, Holmberg, & Holmqvist, 2009). Overall, there are numerous studies that support our recommendation.

Some possible boundary conditions are that the spatial contiguity recommendation may most strongly apply for low-knowledge learners (Mayer, Steinhoff, Bower, & Mars, 1995) and when the graphic and words are complex (Ayres & Sweller, 2005).
Evidence for Presenting Spoken Words at the Same Time as Corresponding Graphics

Our second recommendation—presenting corresponding speech and graphics at the same time—is also based on research evidence (Mayer & Anderson, 1991, 1992; Mayer, Moreno, Boire, & Vagge, 1999; Mayer & Sims, 1994). In one experiment, some students (integrated group) viewed a thirty-second narrated animation that explained how a bicycle tire pump works, in which the spoken words described the actions taking place on the screen. For example, when the narrator’s voice said, “the inlet valve opens,” the animation on the screen showed the inlet valve moving from the closed to the open position. Other students (separated group) listened to the entire narration and then watched the entire animation (or vice versa). On a subsequent transfer test the integrated group generated 50 percent more solutions than did the separated group, yielding an effect size greater than 1, which is considered large.

Overall, across eight different experimental comparisons involving pumps, brakes, lightning, and lungs, students who received integrated presentations generated 60 percent more solutions on a transfer test than did students who received separated presentations. The median effect size across all eight experiments was 1.30, which is considered a large effect in practical terms. Research by Baggett (1984) and Baggett and Ehrenfeucht (1983) shows that learners experience difficulty in learning from a narrated video even when corresponding words and graphics are separated by a few seconds. In a systematic review of thirteen studies, Ginns (2006) found strong evidence for temporal contiguity with an average effect size of .87. As you can see, when you have a narrated animation, narrated video, or even a narrated series of still frames, there is consistent evidence that people learn best when the words describing an element or event are spoken at the same time that the animation (or video or illustration) depicts the element or event on the screen. A possible boundary condition is that the temporal contiguity recommendation applies most strongly when the narration and animation segments are long and when students cannot control the order and pace of presentation (Mayer, Moreno, Boire, & Vagge, 1999; Micas & Berry, 2000).
What We Don’t Know About Contiguity

Overall, our goal is to reduce the need for learners to engage in extraneous processing by helping them see the connection between corresponding words and graphics. Two techniques we explored in this chapter are to present printed words near the part of the graphic they refer to and to present spoken text at the same time as the portion of graphic they refer to. Some unresolved issues concern:

1. How much detail should be in the graphics and in the words?
2. When is it better to use printed words and when is it better to use spoken words?
3. How does the conversational style of the words affect learning?
4. How do characteristics of the voice affect learning with spoken words?

DESIGN DILEMMA: RESOLVED

Ben and Reshmi are debating the best placement of text in the Excel lesson. Some alternatives raised were:

A. Ben is right. To make sense, the visual examples must be displayed as small screens to be viewed after reading the text explanation.

B. Reshmi is right. Learning is more efficient when visuals and text are integrated. The text explanation should be integrated close to the visual examples.

C. Both ideas could be accommodated by placing text directions in a rollover box on top of a large screen shot example.

D. Not sure which option is best.

We recommend Option B for most situations. We show one alternative display in Figure 5.13. Although rollovers can be a useful way to ensure contiguity between visuals and text, rollovers can be transient with the information disappearing when the cursor is moved. In the case of text that will be referred to over time, such as directions for an exercise, a more permanent display that integrates text and graphic will impose less mental load on learners.
Figure 5.13. This Alternative to Figure 5.1 Applies the Contiguity Principle.

### WHAT TO LOOK FOR IN e-LEARNING

- Screens that place printed text next to the portion of the graphic it describes
- Feedback that appears on the same screen as the question and responses
- Directions that appear on the same screen in which the steps are to be applied
- Linked information does not appear in windows that obscure related information on the primary screen
- Animations that can be played independently of text that describes the animation
- Text placed next to or within graphics rather than below them
- Legend callouts that are embedded within the graphic rather separated from it
- Narrated graphics in which corresponding words and graphics are presented at the same time
COMING NEXT

In this chapter, we have seen the importance of (1) the on-screen layout of printed text and graphics and (2) the coordination of corresponding narration and graphics. Next we will consider the benefits of presenting words in audio narration rather than in on-screen text. We know that audio adds considerably to file sizes and requires the use of sound cards and sometimes headsets. Does the use of audio add anything to learning? In the next chapter we examine the modality principle, which addresses this issue.

Suggested Readings


