Chapter 1

Training Fads and Fiction

Do you talk on your cell phone while driving? If yes, you are among the 73 percent of drivers who do. But like seat belt use, your driving habits are about to change. Evidence shows that even hands-free cell phones are potentially lethal distractions putting you at four times greater risk of a crash!

In spite of the evidence, as of early 2009 only five states banned hand-held phones while driving. However, in 2009 the National Safety Council is calling on all states to outlaw use of mobile devices while driving. The journey from evidence to application of evidence is often slow and workforce learning is no exception.

We will see how evidence can save your organization time and money wasted on training fads that don't work, and at the same time can also guide you to invest your resources more productively in proven training methods.

Blood, Phlegm, Black and Yellow Bile

Our story starts in the early 1600s—the birth years of evidence-based practice. There are a number of examples from those years, but in line with the current popularity of vampires, I'll use blood. In the early seventeenth century, folks believed that blood was produced by the heart and the liver and was continuously used up by the body. In other words, there was no accurate conception of blood circulation. In 1628 William Harvey introduced the revolutionary idea that blood was not consumed by the body. Based on measures of blood volume and anatomical observations, he proposed that blood was pumped from the heart and circulated throughout the body returning again to the heart. Harvey, along with Galileo, Descartes, and others, turned the seventeenth century world upside down by advocating evidence and reason rather than traditional wisdom and faith as the basis for knowledge and decisions.

We've come a long way from the days when medical diagnosis and treatments were based on a balance of the four body humors of blood, phlegm, black and yellow bile. If you were lucky your treatment prescribed an amulet that at least did no harm. If you were not so lucky, you were subjected to bloodletting. Although great strides were made in medical science, over 400 years passed before health science professionals formally adopted evidence-based practice. Old habits die hard. Even though we've had evidence about the dangers of cell phones for more than five years, that data is just starting to be translated into policy changes regarding the use of mobile devices while driving.

What Do You Think?

Mark each statement below as Fad or Fact.

A. **Fad or Fact:** To accommodate different learning styles, it's best to explain a visual with words presented in text *and* in audio.
B. **Fad or Fact:** Instructor-led classroom training results in better learning than computer-delivered instruction.
C. **Fad or Fact:** Courses that get higher student ratings generally produce better learning outcomes.

D. **Fad or Fact:** Individuals with visual learning styles learn best from lessons with graphics.
E. **Fad or Fact:** Interesting stories and lesson themes promote learning.

What Is Evidence-Based Practice?

In the last part of the twentieth century, the medical profession was the first applied field to formally adopt the incorporation of evidence into clinical decisions. Sackett and colleagues (1996, 71-2) define evidence-based medicine as the "conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients." Only in the last 20 years has *evidence-based practice* migrated to the social sciences. In 1992 the American Psychological Association recommended that emerging clinical practice guidelines be based on both research data and clinical expertise. This initial statement evolved to the 2006 publication of "Evidence-Based Practice in Psychology" advocating a fundamental commitment to evidence-based psychological practice (APA Presidential Task Force 2006).

You might wonder whether evidence-based practice has any value for workforce learning. Unlike medical practitioners, most trainers do not see their work as having life or death consequences. Although there are some notable exceptions. Remember US Air 1549—the miracle of the Hudson? Here's what Captain Chesley "Sully" Sullenberger had to say: "One way of looking at this might be that, for 42 years, I've been making small regular deposits in this bank of experience: education and training. And on January 15, the balance was sufficient so that I could make a very large withdrawal" (CBS Evening News, February 10, 2009).

Even if you are not involved in safety-critical training, chances are your organization invests a great deal in workforce learning. In 2008 the United States alone allocated around $130 billion (Paradise 2009). What kind of return does your organization get on its training investment? Think of the last class that you developed or facilitated. To what extent did the content sequencing, training methods, and facilitation techniques

promote learning? Many common training practices are based more on fads and fables than on evidence of what works. Let's look at the facts behind four of my favorite training myths.

Training Myth 1: Learning Styles

Like the four body humors of blood, phlegm, black and yellow bile, I think *learning styles* represent one of the more wasteful and misleading pervasive learning myths of the past 20 years. From audio learners to visual learners or from "sensors" to "intuitive," learning styles come in many flavors. Corporations and universities alike frequently incorporate the concept of learning styles and sometimes even use learning style assessments as part of their instructor training. For some reason, the idea of a learning style has a kind of cosmic intuitive appeal that is very compelling. Ask almost anyone whether they are a visual learner or a verbal learner and you will get an immediate commitment to a specific learning style!

The learning style myth leads to some very unproductive training approaches that are counter to modern evidence of what works. For example, many trainers believe that visuals should be described by words in text format for visual learners and narration mode for auditory learners. To accommodate visual and auditory learners a visual on a slide is explained with text and audio narration of that text. As we will see in chapter 6, this practice has been proven to actually depress learning.

The time and energy spent perpetuating the various learning style myths can be more wisely invested in supporting individual differences that are proven to make a difference—namely, prior knowledge of the learner. If you make one change as a result of reading this book, *give up the learning style myth!*

Evidence about Learning Styles

Do we have any evidence about learning styles? Kratzig and Arbuthnott (2006) compared the relationship among three learning style indicators.

They asked a group of university students to do three things. First, each participant rated their own learning style as visual, auditory, or kinesthetic. Second, each individual took a learning style test that classified them as a visual, auditory, or kinesthetic learner. Finally, each person was given three tests to measure visual memory, auditory memory, and kinesthetic memory. If the learning style concept had substance, we would expect to find some positive relationships among these measures. For example, someone who considered themselves a visual learner would score higher on the visual part of a learning styles test and have better memory for content presented visually. However, when all of the measures were compared, there were absolutely no relationships! A person who rated themselves an auditory learner was just as likely to score higher on the kinesthetic scale of the learning style test and show best memory for visual data. The research team concluded that "in contrast to learning style theory, it appears that people are able to learn effectively using all three sensory modalities" (Kratzig & Arbuthnott 2006, 241).

Another research study focused on sensing versus intuitive learning styles. Cook et al. (2009) compared learning of medical residents who tested as having a sensing learning style with learning of medical residents who tested as having an intuitive learning style. Each resident completed four web-based training modules on ambulatory internal medicine. Half the lessons started with a clinical problem followed by traditional information. The other half reversed the sequence, starting with information and ending with a clinical problem. Sensing learners should learn better with a case-first approach while intuitive learners should learn better from a traditional rule-example approach. Knowledge tests were administered at the end of each module as well as several months later. As in the experiment described previously, there was no association between learning style and instructional method. The research team concluded that "it appears from the preponderance of evidence that sensing-intuitive styles have little impact, if any, on educational outcomes" (88).

The lack of evidence about learning styles is the basis for my first recommendation.

♦ ♦ ♦

Fads & Fiction Guideline 1:

Do not waste your training resources on any form of learning style–based efforts including instructor training, measurement of learning styles, or training methods that attempt to accommodate learning styles.

♦ ♦ ♦

Training Myth 2: Media Panaceas

Only a few years ago, computer-delivered instruction incited a revolution in training. Of course computers were not the first technology to cause a stir. Decades prior to computers, radio, film, and television were hailed as having high potential for educational revolution. The first widespread dissemination of computer-based training (CBT) was primarily delivered on mainframe computers. Soon, however, advances in digital memory, display hardware, programming software, and Internet distribution catalyzed the rapid evolution of CBT to recent technological panaceas including web-based training, social media, digital games, and virtual worlds, to name a few. With each new technology wave, enthusiasts ride the crest with claims that finally we have the tools to really revolutionize training. And yet, if you have been around for a few of these waves, those claims begin to sound a bit hollow. In just a few years, today's latest media hype will fade, yielding to the inexorable evolution of technology and a fresh spate of technological hyperbole.

What's wrong with a *technology-centric view* of instruction? Instructional scientists have learned a lot about how humans acquire new knowledge and skills. Like Harvey, who gave birth to the modern mental model of blood circulation, instructional psychology has revealed the strengths and limits of a human brain which is the product of thousands of years of evolution. When we plan instruction around the latest technology gismo, we ignore the psychology of human learning which, as we have learned again with cell phones and driving, has severe limits. In fact, technology

today can deliver far more information faster than the human brain can absorb it.

When we assume a technology-centric view, our focus is on all of the wrong things. Instead of designing training to support human learning processes, we get caught up in the latest technology trends without regard for how they can be most effectively used.

Evidence Against the Technology Panacea

For over 60 years, instructional scientists have attempted to prove the superiority of each new technology over old-fashioned classroom instruction. One of the first published media comparison studies appeared in the 1940s. The U.S. Army believed it could improve instructional quality and reliability by replacing many instructors with films. To its credit, before setting policy based on this idea, the Army tested it. It compared learning a simple procedure from a lesson delivered by film, by instructor, and by print. Each version used similar words and visuals. What do you think the studies found?

- ☐ A. Instructor-led training led to the best learning.
- ☐ B. Paper-based, the least expensive, led to the best learning.
- ☐ C. Films could replace instructors as they led to the best learning.
- ☐ D. Learning was the same with instructor, print, and film.

When the Army tested learners from each group, they discovered that participants from all three lesson versions learned the procedure equally well. In technical terms we say that there were "no significant differences in learning" among the three groups. Since that early experiment, hundreds of studies have compared learning from classroom instruction with the latest technology—the most recent being various forms of digital distance learning. In fact, so many media comparisons have been published, that a synthesis of all of the results (called a meta-analysis) found the same basic conclusion that the Army reported so many years ago: No major differences in learning from

classroom lessons compared with electronic distance learning lessons (Bernard et al. 2004). Therefore, Option D is correct.

But wait! There is an important caveat to this conclusion. The basic instructional modes and methods must be the same in all versions. In other words, if the classroom version includes graphics and practice exercises, the computer version must include similar graphics and practice opportunities. That's because what causes learning are the psychologically active ingredients of your lessons regardless of what media you are using. Rather than asking which technology is best for learning, you will find more fertile ground by considering how instructional modes such as graphics and instructional methods such as practice can best promote learning. More than 50 years of media comparison research is the basis for my second recommendation.

♦ ♦ ♦

Fads & Fiction Guideline 2:

Ignore panaceas in the guise of technology solutions in favor of applying proven practices on best use of instructional modes and methods to all media you use to deliver training.

♦ ♦ ♦

As a postscript to this media discussion, what were once considered distinct and separate delivery technologies are increasingly converging. For example, Kindle merges books and computers. Handheld mobile devices merge functionalities of computers, newspapers, telephones, cameras, radios, clocks, and navigational devices to name a few. Will a time come when the media distinctions blur to the extent that this discussion becomes obsolete? Time will tell.

Training Myth 3: The More They Like It, the More They Learn

Do you collect student ratings at the end of your courses? Your answer is probably yes. Over 90 percent of all organizations use end-of-training surveys to gather participant evaluation of the quality of the course, the

effectiveness of the instructor, how much was learned, and so on. These rating sheets are commonly called smile sheets, or *Level 1 evaluations*. If you are an instructor or a course designer, chances are you have reviewed ratings sheets from your classes. You might also have a sense of how much learning occurred in that class. Based on your own experience, what do you think is the relationship between participant ratings of a class and the actual learning that occurred?

- ☐ A. Classes that are higher rated also yield greater learning.
- ☐ B. Classes that are higher rated actually yield poorer learning.
- ☐ C. There is no relationship between class ratings and learning from that class.

To answer this question, researchers have collected student satisfaction ratings as well as lesson test scores that measure actual learning. They then looked to see what overall relationships surfaced. For example, did higher ratings correlate with more learning or less learning?

Evidence on Liking and Learning

A recent meta-analysis synthesized more than 1,400 student course ratings with student test data. Sitzmann et al. (2008) found there is in fact a positive relationship between ratings and learning. But the correlation was very small! In fact, it was too small to have any practical value. Specifically, the research team concludes that "Reactions have a predictive relationship with cognitive learning outcomes, but the relationship is not strong enough to suggest reactions should be used as an indicator of learning" (289).

What factors are associated with higher ratings? The two most important influencers of ratings are instructor style and human interaction. Instructors who are psychologically open and available—in other words who are personable—are associated with higher course ratings. In addition, the opportunity to socially interact during the learning event with the instructor as well as with other participants leads to higher ratings.

Evidence from comparisons of hundreds of student ratings and student learning outcomes is the basis for my third recommendation.

♦ ♦ ♦

Fads & Fiction Guideline 3:

Don't rely on course evaluations as indicators of learning. Use valid tests to assess the pedagogical effectiveness of any learning environment.

♦ ♦ ♦

Training Myth 4: Stories (Games or You-Name-It) Promote Learning

As I wrote this chapter, I happened to overhear a caller into a conservative talk radio show. "Those liberals," she claimed, "are just like teen agers—they act impulsively and don't think things out." I had to laugh at the blanket characterization of liberals, and it reminded me of many generalizations about various learning methods.

Have you heard that stories (or games or collaboration or you-name-it) are really good training techniques? Training lore is full of claims and recommendations about a variety of training methods like these. What's wrong with these kinds of recommendations?

First, we are using such broad terms for our techniques that statements about them are meaningless. Take games for instance. Do you mean puzzle games, adventure games, strategy games, or simulation games? Do you mean individual paper and pencil games, video games, or group participation games? As a category, games include so much diversity that it is just about impossible to make any generalizations about their instructional effectiveness. The same goes for many other instructional techniques such as graphics or stories.

Second, even if we narrow down to a fairly specific set of criteria for any given instructional method, its effectiveness will depend on the intended learning outcome and the learners. Is your goal to build awareness, to help learners memorize content, to teach procedural skills, to motivate, or to promote critical thinking? Regarding learner differences, prior knowledge (not learning styles!) is the most important factor that

moderates the learning effects of instructional methods. Techniques that help novice learners are not necessarily going to apply to a learner with more expertise.

The lack of universal effectiveness of most instructional techniques is the basis for what I call the "No Yellow Brick Road Effect." By that I mean that there are few best practices that will work for all learners and for all learning goals. The evidence that has accumulated over years of research on general categories like graphics and games is the basis for my fourth recommendation.

Fads & Fiction Guideline 4:

Be skeptical about claims for the universal effectiveness of any instructional technique. Always ask, How is the technique defined? For whom is it useful? For what kinds of learning outcomes does it work?

What Is Useful Evidence?

As an undergraduate science major I have always been drawn to experimental evidence to guide our instructional decisions. Useful experimental evidence requires several conditions. First, subjects are randomly assigned to at least one experimental treatment as well as a comparison treatment. For example, from a pool of 50 students, 25 are randomly assigned to a multimedia lesson with background music while another 25 are randomly assigned to the same multimedia lesson with no background music. Random assignment assures us that individual differences among the learners are evenly distributed. For example, if our pool of 50 students included some who routinely liked and listened to music and some who preferred to study in silence, these differences would be "neutralized" in that there is an equal probability that both types would be assigned to the two lesson versions.

Second, our two lesson versions are identical in every respect except for the one variable of interest. Each lesson in our experiment would have

exactly the same words and same graphics and would vary only in the presence or absence of music. By keeping everything consistent except for a single variable, we assure that differences in learning outcomes can be attributed to our experimental variable and not some other features of the lesson. If our experimental lessons had several differences other than background music, we would be comparing apples and oranges and could not really make any sense of the results.

After the learners complete the lesson, they are tested. The type of test is an important element of the experiment. Many educational experiments use a recall test in which learners are asked to write down what they learned. For workforce learning purposes, however, I prefer application types of tests that require the learners to use what they learned to solve a problem or complete a task. This is because in the world of workforce learning, we are most concerned with job-relevant application outcomes—not recall. Therefore in our music experiment, we would want to judge learning from a test that measured understanding and application of the lesson content—not just recall of content.

About Effect Sizes

I won't attempt a class on statistics here. But you need a couple of basic concepts to interpret the data I'll show you throughout the book. Let's start with *effect size*.

Suppose in our music experiment that on average the learners in the no-music group scored 80 percent while those in the music group scored 74 percent. On the surface, it seems as if learning is better without music. However, we need to do some statistical tests to determine whether a difference of six points is important enough to warrant a change in practice. Scientists report a statistic called effect size to help us make that determination. I'll be mentioning effect sizes throughout the book, so I'll give a brief explanation here.

Think of an effect size as a multiplier for the amount of variation (called a standard deviation) that occurs around an average set of scores. For example, suppose Sally scores an average of 60 percent with a standard

deviation of 10 on lessons with background music. That means most of her scores fall between 50 and 70 percent (60 percent, plus or minus 10). If the effect size from our experiment on music is .8, then we can expect her score on average to increase from 60 to 68 percent (10 times .8 equals 8 points) when we eliminate background music.

Instructional scientists give us some guidelines regarding effect sizes summarized in table 1.1. As you can see, if the effect size is .8 or above, the results of the experiment are probably worth implementing. In contrast, for effect sizes below .3, there may not be sufficient value to warrant recommendations for practitioners. As you review the research I've summarized in this book or in other sources, always look at the effect size as one indicator of the practical implications of the results.

By the way, if you are now curious about the value of background music during learning, you can read about the actual experiments and results in chapter 8.

Table 1.1. Effect Sizes and Practical Significance

Effect Size	Practical Significance	The instructional method results in	Example: Average score of 70% with standard deviation of 10 points would change to
Below .3	LOW	less than 3/10 of a standard deviation improvement in learning	Average of less than 73%
.3 to .8	MODERATE	a 1/3 to 4/5 of a standard deviation improvement in learning	Average of 73% to 78%
.8 to 1.0	HIGH	a 4/5 to 1 standard deviation improvement in learning	Average of 78% to 83%
Above 1.0	VERY HIGH	a greater than 1 standard deviation improvement in learning	Average of more than 83%

Pooling Multiple Experiments with Meta-Analysis

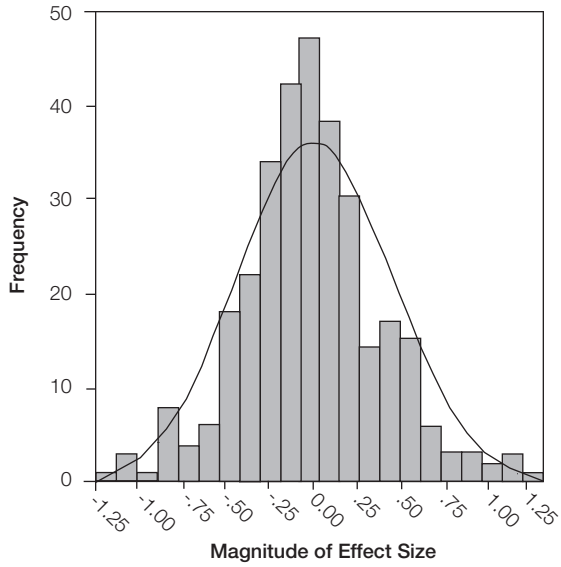
If a recommendation to eliminate background music from multimedia lessons is based on one or two experiments our confidence in that recommendation must be somewhat limited. However, imagine that a large number of experiments were published that compared learning with and without background music. These experiments may have used different lesson content, included short and long lessons, and involved learners of different ages and backgrounds. If we find consistent results from lots of experiments we can feel much more confident that the results will apply to our learners and our lessons. A relatively recent technique called a *meta-analysis* pools the effect sizes of many individual experiments and derives generalizations based on many results.

Earlier in this chapter I reported the results of a meta-analysis of media-comparison experiments. Figure 1.1 plots the values of the effect sizes in 318 comparisons of learning from instructor-led lessons with lessons delivered by some form of electronic distance learning technology. As you can see in the histogram, most of the effect sizes fall in the near-zero range. In other words, most research studies found little learning difference among lessons taught in an instructor-led or computer learning environment. Data like this gives me a high degree of confidence in my guideline to avoid a technology-centric focus.

Limits of Evidence-Based Training

Although I'm a passionate advocate of using evidence as the basis for best practices in training, there are definitely some constraints we practitioners must consider. First, real world decisions cannot be made on the basis of evidence alone. Factors such as budget, development or delivery time constraints, technology capabilities, and political pressures are just a few of the realities most trainers must juggle. Second, the number of evidence-based principles available today is somewhat limited. After 20 years of concentrated research, Richard Mayer (2009) concludes: "Research in multimedia learning is still in its early stages" (279). Not only is the number of research studies limited, but the experiments themselves have

Figure 1.1. A Histogram of 318 Effect Sizes from Learning in Classroom Versus Electronic Distance Learning Media.



From Bernard et al. 2004.

limits. Most of them use Western (U.S., Australian, and European) college students as experimental students, and the outcomes may not apply to other cultures or age groups. In most cases, the experimental lessons are relatively brief, lasting from just a few minutes to an hour in length. We do not know the extent to which the conclusions derived from these relatively short lessons will apply to longer instructional events.

But we do have enough of a start for some firm recommendations that I make in the chapters to follow. As evidence accumulates, I anticipate that the guidelines I offer will be refined and perhaps in some cases even superseded. Nevertheless, the research efforts of the last 20 years lay the foundation for a science of instruction—a science that can offer practitioners a basis for minimizing resources wasted on the myths in favor of practices proven to enhance learning.

The Bottom Line

Let's conclude by revisiting the questions I asked in the beginning of the chapter:

What Do You Think?

1. **Fad or Fact:** To accommodate different learning styles, it's best to explain a visual with words presented in text *and* in audio.
   FAD. The benefits of using text and audio to describe visuals is a common misconception among trainers. In chapter 6, we will examine the evidence and psychology of how to best use words to describe visuals.

2. **Fad or Fact:** Instructor-led classroom training results in better learning than computer-delivered instruction.
   FAD. We saw the evidence from hundreds of media comparison studies that learning effectiveness does not depend on the delivery medium but rather reflects the best use of basic communication modes and instructional methods. We will be reviewing the evidence for these modes and methods in most of the chapters in this book.

3. **Fad or Fact:** Courses that get higher student ratings generally produce better learning outcomes.
   FACT—but only marginally. There is a very small positive relationship between ratings and learning. However, it is too small to draw any conclusions about the learning value of a class from student ratings of that class. Don't rely on student rating to assess the learning value of any training event.

4. **Fad or Fact:** Individuals with visual learning styles learn best from lessons with graphics.
   FAD. There is no evidence for the prevalent myth of learning styles such as visual learners and auditory learners. Perpetuating this myth detracts resources from more productive proven training methods to be described throughout the book.

5. **Fad or Fact:** Interesting stories and lesson themes promote learning.

FACT—sometimes. Yes, sometimes stories can be useful. However their value depends on the type of story, the outcome goal, and the placement of stories during learning. Remember, there is no Yellow Brick Road. We will look at evidence showing how some stories can actually depress learning in chapter 8.

Applying Evidence-Based Practice

The evidence I will review in this book can guide your decisions regarding the best instructional modes and methods to use in your training. But more important, I will consider the book a success if you become a more critical consumer of the various training recommendations appearing in practitioner articles, webinars, and conferences. My hope is that the next time you hear or read some generalizations about the latest technology or hot training method you will ask:

- ☐ What exactly are the features of this method?
- ☐ What is the evidence for this method?
- ☐ How valid is the evidence to support the method?
- ☐ For whom is the method most appropriate?
- ☐ How does the method fit with our understanding of the limits and strengths of human learning?

Let's take a look in chapter 2 at what we know about human memory and learning.

For More Information

Clark, R. (2008). *Building Expertise,* 3d ed. San Francisco, CA: Pfeiffer. See ch. 3.