

CONTENTS

Volume 9 Number 3 • December 2008

Learning Styles: Concepts and Evidence

Harold Pashler, Mark McDaniel, Doug Rohrer, and Robert Bjork

- 106 Introduction
- 106 An Overview of Learning Styles: Doctrines and Industry
- 107 How Did the Learning-Styles Approach Become So Widespread and Appealing?
 - Origin and Popularity
 - Interactions of Individual Differences and Instructional Methods
- 108 What Evidence Is Necessary to Validate Interventions Based on Learning Styles?
 - Existence of Study Preferences
 - The Learning-Styles Hypothesis
 - Interactions as the Key Test of the Learning-Styles Hypothesis
 - Primary Mental Abilities: Relation to Learning Styles
- 111 Evaluation of Learning-Styles Literature
 - Style-by-Treatment Interactions: The Core Evidence Is Missing
 - Learning-Styles Studies With Appropriate Methods and Negative Results
- 113 Related Literatures With Appropriate Methodologies
 - Aptitude-by-Treatment Interactions
 - Personality-by-Treatment Interactions
- 116 Conclusions and Recommendations
 - Points of Clarification
 - Costs and Benefits of Educational Interventions
 - Beliefs Versus Evidence as a Foundation for Educational Practices and Policies
 - Everybody's Potential to Learn

About the Authors

Harold Pashler is Professor of Psychology and a faculty member of the Cognitive Science Program at the University of California, San Diego. His main areas of interest are human learning and the psychology of attention. Pashler's learning research focuses on methods for optimizing acquisition and retention of knowledge and skills. In the field of attention, Pashler's work has illuminated basic attentional bottlenecks as well as the nature of visual awareness. Pashler is the author of *The Psychology of Attention* (MIT Press, 1998) and the editor of *Stevens' Handbook of Experimental Psychology* (Wiley, 2001). He received the Troland Prize from the National Academy of Sciences for his studies of human attention, and was elected to membership in the Society of Experimental Psychologists.

Mark McDaniel is Professor of Psychology at Washington University in St. Louis, with a joint appointment in Education. He received his PhD from the University of Colorado in 1980. His research is in the general area of human learning and memory, with an emphasis on prospective memory, encoding and retrieval processes in episodic memory, learning of complex concepts, and applications to educational contexts and to aging. His educationally relevant research includes work being conducted in actual college and middle-school classrooms. This research is being sponsored by the Institute of Educational Sciences and the James S. McDonnell Foundation, and his work is also supported by the National Institutes of Health. McDaniel has served as Associate Editor of the *Journal of Experimental Psychology: Learning, Memory, and Cognition* and *Cognitive Psychology* and as President of the Rocky Mountain Psychological Association, and he is a fellow of Divisions 3 and 20 of the American Psychological Association. He has published over 200 journal articles, book chapters, and edited books on human learning and memory, and is the coauthor, with Gilles Einstein, of two recent books: *Memory Fitness: A Guide for Successful Aging* (Yale University Press, 2004) and *Prospective Memory: An Overview and Synthesis of an Emerging Field* (Sage, 2007).

Doug Rohrer is Professor of Psychology at the University of South Florida. He received his doctoral degree in Psychology from the University of California, San Diego, and he was a faculty member at George Washington University before moving to the University of South Florida. Before attending graduate school, he taught high-school mathematics for several years. Most of his research concerns learning and memory, with a recent emphasis on learning strategies.

Robert A. Bjork is Distinguished Professor and Chair of Psychology at the University of California, Los Angeles. His research focuses on human learning and memory and on the implications of the science of learning for instruction and training. He has served as Editor of *Memory & Cognition* and *Psychological Review* (1995–2000), Coeditor of *Psychological Science in the Public Interest* (1998–2004), and Chair of the National Research Council's Committee on Techniques for the Enhancement of Human Performance. He is a past president or chair of the Association for Psychological Science (APS), the Western Psychological Association, the Psychonomic Society, the Society of Experimental Psychologists, the Council of Editors of the American Psychological Association (APA), and the Council of Graduate Departments of Psychology. He is a recipient of UCLA's Distinguished Teaching Award, the American Psychological Association's Distinguished Scientist Lecturer and Distinguished Service to Psychological Science Awards, and the American Physiological Society's Claude Bernard Distinguished Lecturership Award.

Learning Styles

Concepts and Evidence

Harold Pashler,¹ Mark McDaniel,² Doug Rohrer,³ and Robert Bjork⁴

¹University of California, San Diego, ²Washington University in St. Louis, ³University of South Florida, and ⁴University of California, Los Angeles

SUMMARY—The term “learning styles” refers to the concept that individuals differ in regard to what mode of instruction or study is most effective for them. Proponents of learning-style assessment contend that optimal instruction requires diagnosing individuals’ learning style and tailoring instruction accordingly. Assessments of learning style typically ask people to evaluate what sort of information presentation they prefer (e.g., words versus pictures versus speech) and/or what kind of mental activity they find most engaging or congenial (e.g., analysis versus listening), although assessment instruments are extremely diverse. The most common—but not the only—hypothesis about the instructional relevance of learning styles is the meshing hypothesis, according to which instruction is best provided in a format that matches the preferences of the learner (e.g., for a “visual learner,” emphasizing visual presentation of information).

The learning-styles view has acquired great influence within the education field, and is frequently encountered at levels ranging from kindergarten to graduate school. There is a thriving industry devoted to publishing learning-styles tests and guidebooks for teachers, and many organizations offer professional development workshops for teachers and educators built around the concept of learning styles.

The authors of the present review were charged with determining whether these practices are supported by scientific evidence. We concluded that any credible validation of learning-styles-based instruction requires robust documentation of a very particular type of experimental finding with several necessary criteria. First, students must be divided into groups on the basis of their learning styles, and then students from each group must be randomly assigned to receive one of multiple instructional methods. Next, students must then sit for a final test that is

the same for all students. Finally, in order to demonstrate that optimal learning requires that students receive instruction tailored to their putative learning style, the experiment must reveal a specific type of interaction between learning style and instructional method: Students with one learning style achieve the best educational outcome when given an instructional method that differs from the instructional method producing the best outcome for students with a different learning style. In other words, the instructional method that proves most effective for students with one learning style is not the most effective method for students with a different learning style.

Our review of the literature disclosed ample evidence that children and adults will, if asked, express preferences about how they prefer information to be presented to them. There is also plentiful evidence arguing that people differ in the degree to which they have some fairly specific aptitudes for different kinds of thinking and for processing different types of information. However, we found virtually no evidence for the interaction pattern mentioned above, which was judged to be a precondition for validating the educational applications of learning styles. Although the literature on learning styles is enormous, very few studies have even used an experimental methodology capable of testing the validity of learning styles applied to education. Moreover, of those that did use an appropriate method, several found results that flatly contradict the popular meshing hypothesis.

We conclude therefore, that at present, there is no adequate evidence base to justify incorporating learning-styles assessments into general educational practice. Thus, limited education resources would better be devoted to adopting other educational practices that have a strong evidence base, of which there are an increasing number. However, given the lack of methodologically sound studies of learning styles, it would be an error to conclude that all possible versions of learning styles have been tested and found wanting; many have simply not been tested at all.

Address correspondence to Harold Pashler, Department of Psychology 0109, University of California, San Diego, La Jolla, CA 92093; e-mail: hpashler@ucsd.edu.

Further research on the use of learning-styles assessment in instruction may in some cases be warranted, but such research needs to be performed appropriately.

INTRODUCTION

The term *learning styles* refers to the view that different people learn information in different ways. In recent decades, the concept of learning styles has steadily gained influence. In this article, we describe the intense interest and discussion that the concept of learning styles has elicited among professional educators at all levels of the educational system. Moreover, the learning-styles concept appears to have wide acceptance not only among educators but also among parents and the general public. This acceptance is perhaps not surprising because the learning-styles idea is actively promoted by vendors offering many different tests, assessment devices, and online technologies to help educators identify their students' learning styles and adapt their instructional approaches accordingly (examples are cited later).

We are cognitive psychologists with an interest both in the basic science of learning and memory and in the ways that science can be developed to be more helpful to teachers and students. We were commissioned by *Psychological Science in the Public Interest* to assess, as dispassionately as we could, the scientific evidence underlying practical application of learning-style assessment in school contexts. This task involved two steps: (a) analyzing the concept of learning styles to determine what forms of evidence would be needed to justify basing pedagogical choices on assessments of students' learning styles and (b) reviewing the literature to see whether this evidence exists. Our team began this undertaking with differing—but not passionately held—opinions on learning styles as well as a shared desire to let the empirical evidence lead us where it would.

We start by offering the reader a brief overview of the learning-styles concept, including some of the publications and entrepreneurial ventures that have been developed around the idea. Next, we analyze the learning-styles concept from a more abstract point of view. Here, we grapple with some potentially confusing issues of definition and logic that in our opinion require more careful consideration in connection with learning styles than they have so far received. We argue that this analysis is a useful, and essential, prerequisite to organizing and appraising the evidence on learning styles. Finally, we describe the results of our search of published literature, draw some conclusions, and suggest lines of future research. We should emphasize, however, that the present article is not a review of the literature of learning styles; indeed, several such reviews have appeared recently (e.g., Coffield, Moseley, Hall, & Ecclestone, 2004; Kozhevnikov, 2007; Sternberg, Grigorenko, & Zhang, 2008). In brief, we sought to determine what kinds of findings

would provide sufficient evidence for the learning-styles concept, as detailed in the following sections, and then we searched for evidence that satisfied this minimal criterion.

AN OVERVIEW OF LEARNING STYLES: DOCTRINES AND INDUSTRY

As described earlier, the concept of learning styles encompasses not only a large body of written materials but also what seems to be a thriving set of commercial activities. The writings that touch on the learning-styles concept in its broadest sense include several thousand articles and dozens of books. These figures may seem surprisingly large, but one should keep in mind the sheer number of different schemes or models of learning styles that have been proposed over the years. For example, in a relatively comprehensive review, Coffield et al. (2004) described 71 different schemes, and they did not claim that their list was exhaustive.

The commercial activity related to learning styles is largely centered around the publishing and selling of measurement devices to help teachers assess individual learning styles; typically, although not always, these devices classify the learner into different style categories. Testing has been recommended by organizations at all levels of education that might be presumed to base their recommendations on evidence. For example, the National Association of Secondary School principals commissioned the construction of a learning-styles test that it distributed widely (Keefe, 1988). Similarly, the Yale Graduate School of Arts and Sciences (2009) currently maintains a Web site that offers advice for Yale instructors; the site informs visitors that “college students enter our classrooms with a wide variety of learning styles.” The site goes on to recommend that instructors determine their own “modality of learning” as well as assess their students' learning styles and make their instructional choices accordingly.

Furthermore, the learning-styles concept is embraced in a number of current educational psychology textbooks. For instance, Omrod (2008) wrote, “Some cognitive styles and dispositions *do* seem to influence how and what students learn. . . . Some students seem to learn better when information is presented through words (*verbal* learners), whereas others seem to learn better when it's presented through pictures (*visual* learners)” (p. 160, italics in original). Thus, educational psychology students and aspiring teachers are being taught that students have particular learning styles and that these styles should be accommodated by instruction tailored to those learning styles.

Some of the most popular learning-style schemes include the Dunn and Dunn learning-styles model (e.g., Dunn, 1990), Kolb's (1984, 1985) Learning Styles Inventory, and Honey and Mumford's (1992) Learning Styles Questionnaire. The assessment devices that have been developed in relation to the model of Dunn and Dunn are particularly popular and extensive.

Customers visiting the Web site of the International Learning Styles Network (www.learningstyles.net) are advised that

Learning style is the way in which each learner begins to concentrate on, process, absorb, and retain new and difficult information (Dunn and Dunn, 1992; 1993; 1999). The interaction of these elements occurs differently in everyone. Therefore, it is necessary to determine what is most likely to trigger each student's concentration, how to maintain it, and how to respond to his or her natural processing style to produce long term memory and retention. To reveal these natural tendencies and styles, it is important to use a comprehensive model of learning style that identifies each individual's strengths and preferences across the full spectrum of physiological, sociological, psychological, emotional, and environmental elements. (International Learning Styles Network, 2008)

As of June 2008, the company sells five different assessment tools for different age groups—ranging from the Observational Primary Assessment of Learning Style (OPAL) for ages 3 to 6 to Building Excellence (BE) for ages 17 and older (at a cost of approximately \$5.00 per student for the classification instrument). The vendor claims these assessments “measure the patterns through which learning occurs in individual students; they summarize the environmental, emotional, sociological, physiological, and global/analytic processing preferences that a student has for learning” (International Learning Styles Network, 2008). A summer certification program is also offered in connection with this approach (the basic certification program costs \$1,225 per trainee, excluding meals and lodging, with a higher level certification for conducting research on learning styles also offered for an additional \$1,000). The Dunn and Dunn assessment instrument for adults asks respondents to indicate, for example, whether they learn best when they hear a person talk about something, whether their desk is typically disorganized and messy, whether they would say that they normally think in words as opposed to mental images, and whether they would characterize themselves as someone who thinks intuitively or objectively (Rundle & Dunn, 2007).

Kolb's (1984, 1985) Learning Styles Inventory is another very popular scheme, particularly within the United States. It conceives of individuals' learning processes as differing along two dimensions: preferred mode of perception (concrete to abstract) and preferred mode of processing (active experimentation to reflective observations). The Learning Styles Inventory classifies individuals into four types on the basis of their position along these two dimensions: *divergers* (concrete, reflective), *assimilators* (abstract, reflective), *convergers* (abstract, active), and *accommodators* (concrete, active). The self-assessment requires people to agree or disagree (on a 4-point scale) with, for example, the idea that they learn best when they listen and watch carefully, or that when they learn they like to analyze things and to break them down into parts.

The Learning Styles Inventory is distributed by the Hay Group (<http://www.haygroup.com>) and sold in packs of 10 booklets for

approximately \$100.00 (as of June 2008). The Hay Group also distributes an informational booklet called “One Style Doesn't Fit All: The Different Ways People Learn and Why It Matters” (Hay Group, n.d.). According to the booklet, the practical benefits of classifying individuals' learning styles include “placing them in learning and work situations with people whose learning strengths are different from their own,” “improving the fit between their learning style and the kind of learning experience they face,” and “practicing skills in areas that are the opposite of their present strengths” (Hay Group, n.d., p. 11).

These three examples are merely some of the more popular and well-advertised products within the learning-styles movement. Readers interested in a more comprehensive view should consult Coffield et al. (2004).

HOW DID THE LEARNING-STYLES APPROACH BECOME SO WIDESPREAD AND APPEALING?

Origin and Popularity

The popularity and prevalence of the learning-styles approach may, of course, be a product of its success in fostering learning and instruction. Assessing the extent to which there is evidence that the approach does indeed foster learning is the primary goal of this review. However, there are reasons to suspect that other factors—in addition to, or instead of, actual effectiveness—may play a role in the popularity of the learning-styles approach.

Most learning-styles taxonomies are “type” theories: That is, they classify people into supposedly distinct groups, rather than assigning people graded scores on different dimensions. One can trace the lineage of these theories back to the first modern typological theorizing in the personality field, which was undertaken by the psychiatrist and psychoanalyst C.G. Jung (1964). Jung's ideas were explicitly incorporated into a psychological test developed in the United States, the Myers–Briggs Type Indicator test. This test became very popular starting in the 1940s and remains widely used to this day. The Myers–Briggs categorizes people into a number of groups, providing information that is said to be helpful in making occupational decisions. The assumption that people actually cluster into distinct groups as measured by this test has received little support from objective studies (e.g., Druckman & Porter, 1991; Stricker & Ross, 1964), but this lack of support has done nothing to dampen its popularity. It seems that the idea of finding out “what type of person one is” has some eternal and deep appeal, and the success of the Myers–Briggs test promoted the development of type-based learning-style assessments.

Another, very understandable, part of the appeal of the learning-styles idea may reflect the fact that people are concerned that they, and their children, be seen and treated by educators as unique individuals. It is also natural and appealing to think that all people have the potential to learn effectively and easily if only instruction is tailored to their individual learning styles. Another related factor that may play a role in the popu-

larity of the learning-styles approach has to do with responsibility. If a person or a person's child is not succeeding or excelling in school, it may be more comfortable for the person to think that the educational system, not the person or the child himself or herself, is responsible. That is, rather than attribute one's lack of success to any lack of ability or effort on one's part, it may be more appealing to think that the fault lies with instruction being inadequately tailored to one's learning style. In that respect, there may be linkages to the self-esteem movement that became so influential, internationally, starting in the 1970s (Twenge, 2006).

Interactions of Individual Differences and Instructional Methods

As we argue in the next section, credible evidence in support of practices based on learning styles needs to document a specific type of interaction between instructional method and assessments of an individual's learning style. Basically, evidence for a learning-styles intervention needs to consist of finding that a given student's learning is enhanced by instruction that is tailored in some way to that student's learning style.

Naturally, it is undeniable that the optimal instructional method will often differ between individuals in some respects. In particular, differences in educational backgrounds can be a critical consideration in the optimization of instruction. New learning builds on old learning, for example, so an individual student's prior knowledge is bound to determine what level and type of instructional activities are optimal for that student. Many research studies (see, e.g., McNamara, Kintsch, Butler-Songer, & Kintsch, 1996) have demonstrated that the conditions of instruction that are optimal differ depending on students' prior knowledge. Later in this review, we summarize some of the evidence suggesting that aptitude measures can help predict what instructional methods are most effective.

WHAT EVIDENCE IS NECESSARY TO VALIDATE INTERVENTIONS BASED ON LEARNING STYLES?

We turn now to the core of the learning-styles idea: an assessment of the degree to which it has been validated.

Existence of Study Preferences

In reviewing the literature on learning styles and examining the different ways in which this term is frequently used, we make a basic distinction between what we call the *existence of study preferences* and what we call the *learning-styles hypothesis*. The existence of preferences, as we interpret it, amounts simply to the fact that people will, if asked, volunteer preferences about their preferred mode of taking in new information and studying. Given that learning-style questionnaires focusing on preferences have at least some psychometric reliability (i.e., a person's score on one day predicts their score on another day; e.g.,

Henson & Hwang, 2002; Veres, Sims, & Shake, 1987), the existence of preferences with some coherence and stability is not in dispute. A study by Massa and Mayer (2006), which is discussed in more detail later, provides further evidence on this point. Massa and Mayer developed three instruments to assess people's preferences for receiving instruction verbally versus accompanied by pictorial illustrations. Responses on these instruments were significantly correlated with the degree to which college students chose to receive verbal elaboration versus pictorial elaboration of technical terms in an electronics lesson. Massa and Mayer also found significant correlations between the instruments they used to assess people's preference for certain kinds of representations and the mode of elaboration people elected to receive in the electronics lesson. (As discussed at more length later, however, the preference for visual versus verbal information intake had little, if any, relationship to an individual's objectively measured specific-aptitude profile.)

Having noted the reality of these preferences, we emphasize that the implications of such preferences for educational practices and policies are minimal. The existence of preferences says nothing about what these preferences might mean or imply for anything else, much less whether it is sensible for educators to take account of these preferences. Most critically, the reality of these preferences does not demonstrate that assessing a student's learning style would be helpful in providing effective instruction for that student. That is, a particular student's having a particular preference does not, by itself, imply that optimal instruction for the student would need to take this preference into account. In brief, the existence of study preferences would not by itself suggest that buying and administering learning-styles tests would be a sensible use of educators' limited time and money.

The Learning-Styles Hypothesis

What, then, is the version of the learning-styles hypothesis that has practical implications for educational contexts? It is the claim that learning will be ineffective, or at least less efficient than it could be, if learners receive instruction that does not take account of their learning style, or conversely, it is the claim that individualizing instruction to the learner's style can allow people to achieve a better learning outcome.

It is important to note that there is a specific version of the learning-styles hypothesis that evidently looms largest both within the educational literature and within the minds of most people writing about learning styles: the idea that instruction should be provided in the mode that *matches* the learner's style. For example, if the learner is a "visual learner," information should, when possible, be presented visually. We refer to this specific instance of the learning-styles hypothesis as the *meshing hypothesis*—the claim that presentation should mesh with the learner's own proclivities.

Most proponents of the learning-styles idea subscribe to some form of the meshing hypothesis, and most accounts of how in-

struction should be optimized assume the meshing hypothesis: For example, they speak of (a) tailoring teaching to “the way in which each learner begins to concentrate on, process, absorb, and retain new and difficult information” (Dunn & Dunn’s framework; International Learning Styles Network, 2008), (b) the learner’s preferred modes of perception and processing (Kolb’s, 1984, 1985, framework), or (c) “the fit between [people’s] learning style and the kind of learning experience they face” (Hay Group, n.d., p. 11). Note that the learning-styles hypothesis, as defined here, could be true without the meshing hypothesis being true—if, for example, individuals classified as visual learners profited more from verbal instruction in some situations or if individuals classified as verbal learners profited more from visual instruction. In our review, we searched for evidence for both this broad version of the learning-styles hypothesis and the more specific meshing hypothesis.

Interactions as the Key Test of the Learning-Styles Hypothesis

To provide evidence for the learning-styles hypothesis—whether it incorporates the meshing hypothesis or not—a study must satisfy several criteria. First, on the basis of some measure or measures of learning style, learners must be divided into two or more groups (e.g., putative visual learners and auditory learners). Second, subjects within each learning-style group must be randomly assigned to one of at least two different learning methods (e.g., visual versus auditory presentation of some material). Third, all subjects must be given the same test of achievement (if the tests are different, no support can be provided for the learning-styles hypothesis). Fourth, the results need to show that the learning method that optimizes test performance of one learning-style group is different than the learning method that optimizes the test performance of a second learning-style group.

Thus, the learning-styles hypothesis (and particular instructional interventions based on learning styles) receives support if and only if an experiment reveals what is commonly known as a *crossover interaction* between learning style and method when learning style is plotted on the horizontal axis. Three such findings are illustrated in Figures 1A to 1C. For each of these types of findings, the method that proves more effective for Group A is not the same as the method that proves more effective for Group B. One important thing to notice about such a crossover interaction is that it can be obtained even if every subject within one learning-style group outperforms every subject within the other learning-style group (see Fig. 1B). Thus, it is possible to obtain strong evidence for the utility of learning-style assessments even if learning style is correlated with what might, for some purposes, be described as ability differences. Moreover, the necessary crossover interaction allows for the possibility that both learning-style groups could do equally well with one of the learning methods (see Fig. 1C).

Figures 1D to 1I show some hypothetical interactions that would not provide support for the learning-styles hypothesis because, in each case, the same learning method provides optimal learning for every learner. Note that these findings are insufficient even though it is assumed that every interaction in Figure 1 is statistically significant. It is interesting to note that the data shown in Figures 1D and 1G do produce a crossover interaction when the data are plotted so that the horizontal axis represents learning method, as shown in Figure 2, but this mere rearrangement of the data does not alter the fact that the same learning method maximizes performance of all subjects.¹ Thus, as noted earlier, a style-by-method crossover interaction constitutes sufficient evidence for the learning-styles hypothesis if and only if the horizontal axis represents learning style, as in Figures 1A to 1C.

To provide the most liberal criterion in our search for evidence supporting the learning-styles hypothesis, we cast the hypothesis so that it requires only the style-by-method crossover interaction described previously. It does not require that the optimal method for each group would somehow match or conform to each group’s learning style (the meshing hypothesis referred to earlier).

Primary Mental Abilities: Relation to Learning Styles

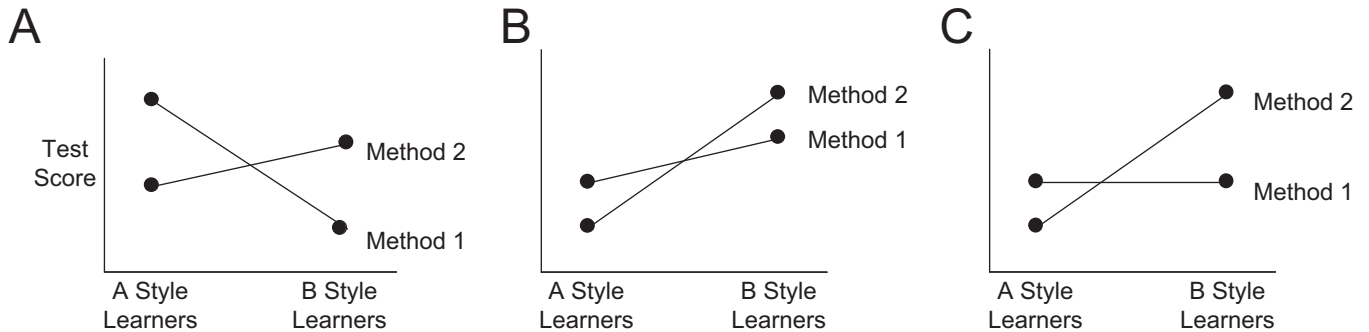
In our discussion of styles thus far, we have focused on *preferences* for how information would be presented to a person rather than on the notion of the person having different *ability* to process one kind of information or another. This focus is in conformity with the dictionary definition of *style* and matches at least the most typical usage of the term *learning style* within the education field. However, the notion of learning style as a set of preferences and the notion of learning style as a *specific aptitude* are very closely intertwined in many discussions of learning styles. Moreover, it is our impression that among the general public, the notion of learning styles and the notion of differential abilities are scarcely distinguished at all. There is, after all, a commonsense reason why the two concepts could be conflated: Namely, different modes of instruction might be optimal for different people because different modes of presentation exploit the specific perceptual and cognitive strengths of different individuals, as suggested by the meshing hypothesis.

Similar to the learning-styles hypothesis, the idea of specific abilities also implies a special form of crossover interaction. However, the interaction is different in kind from what was outlined earlier as the key test of the learning-styles hypothesis.

¹ A reviewer of an earlier version of this article noted that the interactions shown in Figures 1H and 1I might have potential practical importance, even in the absence of a true crossover. If one could sort people into two groups, one of which would benefit from an instructional manipulation and the other of which was completely unaffected by it, it might (on some assumptions) be worthwhile doing the sorting and selectively offering the manipulation. We agree. However, as we show later, the general conclusions reached here do not depend on this issue because we have not found any actual interactions of the types in Figures 1H and 1I in the learning-styles literature.

Acceptable Evidence

In examples A, B, and C, the learning method that optimized the mean test score of one kind of learner is *different* from the learning method that optimized the mean test score of the other kind of learner.



Unacceptable Evidence

In examples D through I, the *same* learning method optimized the mean test score of both kinds of learners, thereby precluding the need to customize instruction.

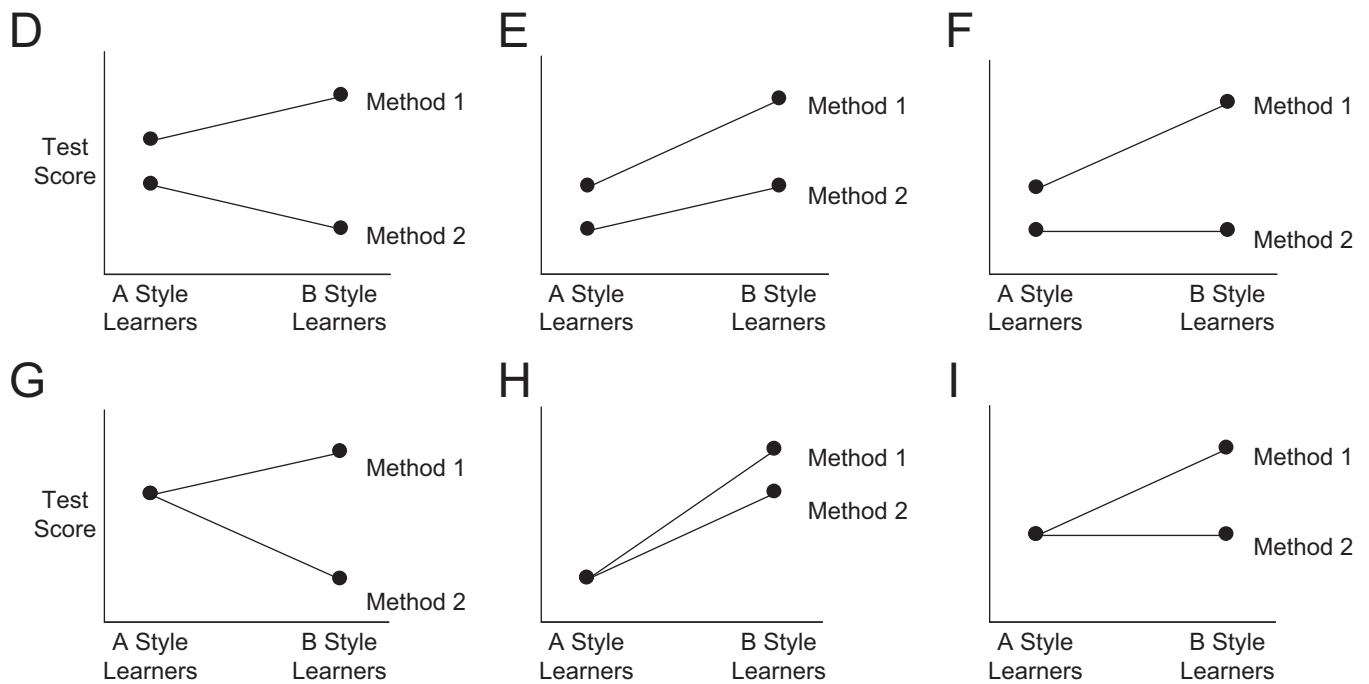


Fig. 1. Acceptable and unacceptable evidence for the learning-styles hypothesis. In each of the hypothetical experiments, subjects have been first classified as having Learning Style A or B and then randomly assigned to Learning Method 1 or 2. Later, all subjects have taken the same test. The learning-styles hypothesis is supported if and only if the learning method that optimized the mean test score of one group is different from the learning method that optimized the mean test score of the other group, as in A, B, and C. By contrast, if the same learning method optimized the mean test score of both groups, as in D through I, the result does not provide evidence. (Note that all nine interactions are assumed to be statistically significant.) In general, the learning-styles hypothesis is supported if and only if a study finds a crossover interaction between learning method and learning style, assuming that the horizontal axis represents the learning-style variable. See the text for more details.

If the notion of specific aptitudes or skills is valid, one ought to be able to divide subjects into two or more groups (e.g., Group A of learners with high auditory ability and Group B of learners with high visual ability). There should then be two tests such that Group A outscores Group B on one test, whereas Group B outscores Group A on the other test.

There is little doubt that specific-ability differences of this kind exist. The first psychologist to provide strong empirical evidence for the idea of specific-ability differences was Louis Thurstone (e.g., Thurstone, 1938). Thurstone proposed seven “primary mental abilities”: *verbal comprehension*, *word fluency*, *number facility*, *spatial visualization*, *associative memory*,

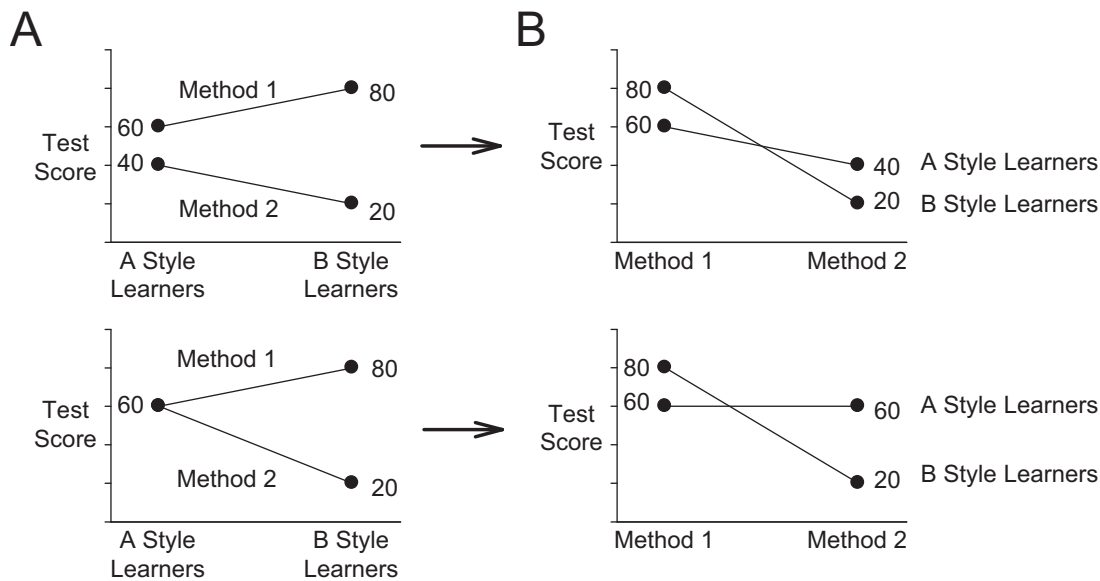


Fig. 2. Examples of crossover interactions that would not validate the learning-styles hypothesis. The two hypothetical outcomes in A are identical to the outcomes in B, and these examples demonstrate that the choice of variable for the horizontal axis can affect whether an interaction appears to “cross over.” Regardless of appearance, though, each of the graphs above demonstrates that the same learning method (Method 1) proved superior for all subjects. Thus, the data above do not provide evidence for the learning-styles hypothesis. However, if the horizontal axis depicts the learning-style variable, a crossover interaction is both sufficient and necessary to show evidence for the learning-styles hypothesis, as in Figures 1A–1C. Note that the above two results are identical to those in Figures 1D and 1G.

perceptual speed, and *reasoning*. Although these abilities are not completely uncorrelated (implying, to some, the idea of general mental ability or “g”; see Jensen, 1998; Spearman, 1927), they do show a moderate degree of independence (Thurstone, 1938). Although this provides evidence for specific aptitudes, it does not show that one needs to provide different groups with different forms of instruction to maximize their performance on any single outcome test. Thus, evidence for specific aptitudes does not, by itself, validate the learning-styles hypothesis.

There are few data on the relationship between preferences and specific aptitudes. However, one recent and well-executed study, which we discuss at more length later, discloses that preference for visual versus verbal information intake shows hardly any relationship to an individual’s objectively measured specific-aptitude profile (Massa & Mayer, 2006). Thus, the common assumption that preferences and abilities are closely tied is open to challenge. But as we have defined the learning-styles hypothesis, one could find evidence for the hypothesis regardless of whether the style measure involved a specific aptitude, a preference, or both.

EVALUATION OF LEARNING-STYLES LITERATURE

Style-by-Treatment Interactions: The Core Evidence Is Missing

For the reasons described earlier, it is our judgment that a validation of an intervention based on learning styles would need to

offer one kind of evidence, and one kind of evidence alone: a crossover interaction of the form illustrated in Figures 1A to 1C. On the basis of this analysis, we scoured the literature to identify studies that provided such evidence. Remarkably, despite the vast size of the literature on learning styles and classroom instruction, we found only one study that could be described as even potentially meeting the criteria described earlier,² and as we report in the following text, even that study provided less than compelling evidence.

The study in question was reported by Sternberg, Grigorenko, Ferrari, and Clinkenbeard (1999). In this study, 324 “gifted and talented” high school students were given the Sternberg Triarchic Abilities Test, which provided a rating of each student’s analytical, creative, and practical ability. On the basis of this test, the authors selected a subset of 112 subjects (35%) for whom one of these three abilities was much higher than the other two, and depending on their area of strength, these subjects were assigned to the high-analytical, high-creative, or high-practical groups. (Another 87 students were assigned to two additional groups not described here, and the remaining 125 students were excluded from the study.) The participating subjects enrolled in an introductory psychology summer course at Yale University, and each student was randomly assigned to class meetings that emphasized analytical instruction, creative instruction, practi-

²We also encountered one study in the domain of user information technology training that appears to offer one interaction of the form discussed here (see Bostrom, Olman, & Sein, 1990).

cal instruction, or memory instruction (a control condition). Their course performance was assessed by raters, and the ratings were “subjected to principal-component analyses” (Sternberg et al., 1999, p. 7). The authors reported several analyses, and, for the analysis of the interaction of interest, they compared the course performance of matched subjects (i.e., students who received instruction that matched their strongest ability) to mismatched subjects. The article states that after the data were “screened for deviant scores” (Sternberg et al., 1999, p. 10), matched subjects reliably outscored mismatched subjects on two of the three kinds of assessments.

Thus, the authors reported a style-by-treatment interaction. Although suggestive of an interaction of the type we have been looking for, the study has peculiar features that make us view it as providing only tenuous evidence. For one thing, the reported interaction was found only with highly derived measures (as noted above), and the untransformed outcome measures (e.g., the mean score on each final assessment) were not reported for the different conditions. Furthermore, and as noted previously, only about one third of the subjects were classified into the groups that produced the interaction. Finally, the interaction was achieved only after the outliers were excluded for unspecified reasons. In brief, although the article presents data that may be worth following up, it has serious methodological issues. Even for those who might disagree with this judgment, the potential support that this study could provide for any of the particular interventions based on learning styles that are being marketed at the present time is extremely limited because the instructional manipulation does not seem to correspond to any of the more widely promoted and used learning-styles interventions.

In summary, our efforts revealed at most one arguable piece of evidence for the learning-styles hypothesis in general. For the many specific assessment devices and interventions being actively marketed to teachers, as described earlier in this article, we were unable to find any evidence that would meet the key criteria discussed earlier (i.e., interactions of the form shown in Figs. 1A–1C). Moreover, we found a number of published studies that used what we have described as the appropriate research design for testing the learning-styles hypothesis and found results that contradict widely held versions of the learning-styles hypothesis; we turn to these studies now.

Learning-Styles Studies With Appropriate Methods and Negative Results

Massa and Mayer (2006) reported a particularly informative and well-designed study of learning styles with a set of three experiments. They constructed a reasonably realistic computer-based electronics lesson. Two different sorts of help screens were customized for verbal or visual learners, providing either supplementary printed text or carefully developed diagrams and illustrations, respectively. A wide variety of preference-based and ability-based individual-difference measures were admin-

istered to sort visual from verbal learners in various ways. In general, the results, which the researchers replicated, showed no tendency for better performance for those who received help screens matched to their preferences. Critically, Massa and Mayer found no support for any of these interactions despite exhaustive analysis of nearly 20 individual-difference measures that spanned their three proposed facets of verbalizer–visualizer learning styles. The authors concluded that their results provided no support for “the idea that different instructional methods should be used for visualizers and verbalizers” (Massa & Mayer, 2006, pp. 333–334).

Within a medical-education context, a recent study by Cook, Thompson, Thomas, and Thomas (2009) examined the hypothesis that learners with a “sensing learning style” would do better when given instruction in which the problem was presented prior to the content information used to solve the problem, whereas “intuitive learners” would do better with the reverse. The authors noted that this learning-styles taxonomy is similar to Kolb’s (1984, 1985) concrete–abstract dimension. Studying a sample of 123 internal medicine residents and presenting modules on four ambulatory medicine topics, they found no support for this prediction.

Another study reaching a similar conclusion, albeit using tasks with less direct correspondence to real educational activities, was reported by Constantinidou and Baker (2002). These investigators used a laboratory task to ask whether self-reported preferences in information uptake predicted ability to perceive and store information in different modalities. They examined the relationship between adults’ scores on the Visualizer–Verbalizer Questionnaire (VVQ; Richardson, 1977) to their verbal free-recall performance on a task that presented words through the auditory modality, the visual modality (as line drawings of the corresponding object), or both. The VVQ asks people a series of questions about their relative preference for taking in information through verbal versus visual means. VVQ scores were not related in any strong or clearly interpretable way to relative levels of free-recall performance for different input modalities. Visual presentations produced better free recall than did purely verbal presentations, and the authors reported finding “no relationship between a visual learning style and the actual learning of verbal items that are presented visually or auditorily” (Constantinidou & Baker, 2002, p. 306).

These studies, which we believe are methodologically strong, provide no support for the learning-styles hypothesis (or its popular specific version, the meshing hypothesis). As mentioned previously, however, it would clearly be a mistake to label these negative results as a conclusive refutation of the learning-styles hypothesis in general. Further research modeled on the work of Massa and Mayer (2006) may bring to light assessments paired with interventions that do meet our criteria. But at present, these negative results, in conjunction with the virtual absence of positive findings, lead us to conclude that any application of learning styles in classrooms is unwarranted.

RELATED LITERATURES WITH APPROPRIATE METHODOLOGIES

Aptitude-by-Treatment Interactions

Although the literature on learning styles per se has paid scarce attention to the need for group-by-treatment interactions, there has been a clear recognition of the importance of such interactions within an older educational psychology literature, going back to Cronbach's (1957) appeal for research to uncover interactions between *aptitude* and aspects of the instructional context (termed *treatments*). Although the validity of aptitude-by-treatment interactions (ATIs) is a separate issue from the validity of learning-style measures, which is the primary focus of the current article, we describe several ATIs so that the reader may gain an appreciation of a literature that recognizes the need to demonstrate the necessary interaction.

Initial attempts to demonstrate so-called ATIs were reviewed in a classic work by Cronbach and Snow (1977). According to Cronbach and Snow, these attempts were not highly successful because treatment durations were too brief, and aspects of the methodologies were inadequate. After that review, significant improvements were made in methodologies, with a number of studies examining treatments implemented in classroom settings for relatively long durations.

The kind of potential interaction that has received the most attention within the ATI tradition involves the degree to which the teaching approach provides ample structure or guidance for the learner. The primary hypothesis that has stimulated much of the work in this area is the idea that students with high ability tend to fare better in less structured learning environments than in highly structured learning environments. By contrast, students with low ability are hypothesized to fare better with instruction that is highly structured and provides explicit guidance than with instruction that is less structured and provides little guidance (see, e.g., Snow, 1977). A variant of this theme that also sparked interest is the idea that highly structured situations might reduce performance differences between students with high and low abilities (Freebody & Tirre, 1985). As detailed in the next paragraph, two key difficulties in evaluating this hypothesis are as follows: (a) The implementation of instructional methods that differ in structure (guidance) has been quite variable, and (b) the measures used to assess student abilities have varied considerably.

Freebody and Tirre (1985) reported an ATI in line with the above hypothesis that involved two competing reading-instruction approaches. One approach, the Matteson program (see Schlenker, 1978), provides a list of behavioral objectives in major reading-skill areas (e.g., word recognition, vocabulary development, literal and interpretative comprehension) combined with individualized learning packages that cover these areas, following precisely defined sequences. The other approach, the Scott Foresman (1972) program, is not strictly sequenced and monitored. Instead, the emphasis is on frequent

discussions focusing on the literal and inferential aspects of discourse. This approach is assumed to place a greater burden on the student for acquiring specific reading skills (see Freebody & Tirre, 1985).

All of the sixth-grade students in a large school district who had been in one of the two reading programs for 2 years or longer served as subjects ($N = 180$, nearly equally distributed across reading programs). Their aptitudes were assessed with a standardized test that included nonverbal and verbal measures of ability. The outcome measure was the reading test score achieved at the conclusion of the sixth-grade year. Multiple regression analyses produced a significant ATI. The interpretation of the interaction was based on *predicted* outcomes (from the regression equations) for particular low-ability values and particular high-ability values. These predicted outcomes indeed showed that students with low ability would generally perform better on the structured reading program (Matteson) than on the less structured reading program (Scott Foresman). The reverse would be predicted for the students with high ability: better performance on the less structured than on the more structured reading instruction method. Although suggestive, these data do not establish that students at a particular ability level (either low or high) fared significantly better (in terms of reading outcomes) as a function of the reading program in which the students were enrolled.

Additional direct support for the idea that learning outcomes for students with high and low abilities might reverse with a greater degree of structure embedded in instruction was reported in the domain of elementary school mathematics (Cramer, Post, & Behr, 1989). Fourth graders being taught fractions were given four lessons (in six 40-minute class periods) on completing rational numbers tasks that involved shading a particular fractional area (two thirds) of different kinds of visual figures (e.g., a rectangle divided into three columns). In the high-structured condition, instruction was teacher centered with little student choice. The teacher paced through each example in large-group lecture fashion. In the low-structured condition, the teacher provided an initial introduction to the problems and then students worked through examples at their own pace. The materials involved leading questions to guide the learner to discovery of the key concepts. Both instructional conditions used identical examples, and both contained a 10-minute practice phase that completed each 40-minute class period. Students in the higher and lower ranges of cognitive restructuring ability, as measured by the Group Embedded Figures test (see Witkin & Goodenough, 1981), were assigned to each instructional condition. At the conclusion of the lessons, the students completed a final test containing problems (rational numbers tasks) of the type taught in the lessons. For the more difficult problems—those requiring physical restructuring of the diagrams—a crossover interaction between ability and the degree of instructional structure emerged. The students with high ability performed better following low- than high-structured instruction; by contrast, students with low ability performed better following high- than

low-structured instruction. Particularly notable is that the students with low ability outscored the students with high ability (at least nominally) after both received the more highly structured instruction. This pattern thus provides evidence that learning is optimized when students with low ability are provided with structured instruction and students with high ability are provided with less structured instruction.

However, other studies that examined different content domains and used different assessment instruments did not always support the idea that high-ability students are better off with less structured instruction, whereas low-ability students profit more from higher-structured instruction. In Janicki and Peterson (1981), 117 grade school students completed a 2-week fractions unit in a “direct” instructional fashion (which involved homework assignments that students completed in class on their own) or in a less structured fashion (this involved mixed-ability four-student group seatwork with choice of homework or math games). Aptitude, as determined by a composite measure that included Ravens Progressive Matrices, did not interact with instructional method.

Greene (1980) similarly failed to find an interaction when fifth and sixth graders with high and low ability (as determined by Lorge-Thorndike verbal and nonverbal tests) were given high-structured instruction (specified sequence of workbook assignments and performance standards) or low-structured instruction (choice and pacing of which exercises to do in the workbook) on a letter-series task. The letter-series task was chosen to reflect general problem-solving goals in education. The basic result was that the students with higher ability performed better than the students with lower ability regardless of instruction.

In a well-conducted experiment, Peterson, Janicki, and Swing (1980, Study 2) manipulated instruction for a 2-week ninth-grade social studies unit across six classes (146 students). Two teachers taught each of three classes with one of three teaching methods. One teaching method was a standard lecture-recitation approach. In the second method, termed *inquiry*, students researched a historical question using primary sources. The third method, *public issues discussion*, required students to support a position on a current public issue using primary material. Aptitude was defined as verbal ability. The outcome measure was a test that included multiple-choice questions on historical facts and short essay questions requiring integration and evaluation of material. Critically, the test targeted readings and content common to all three instructional approaches. In line with the previous findings, there was no interaction between teaching method and ability for the essay performances, with students with higher ability performing better on the essay questions in general.

It is interesting to note that for the multiple-choice questions, there was a significant ATI such that students with high ability performed better with the lecture-recitation teaching method than with the inquiry or public issues discussion methods, whereas students with low ability performed better when receiving the inquiry or public issues methods than with the lecture-recitation method. This pattern would appear to counter the

main hypothesis being considered in reviewing this body of ATI work, because the inquiry and public discussion methods encouraged learner self-direction (less structure). However, Peterson et al. (1980) offered an interpretation based on the underlying cognitive demands placed on the students by the different instructional methods. They suggested that the lecture-recitation approach implemented in the study placed a heavier burden on students’ cognitive skills than did the other approaches. Specifically, students had to comprehend and attend to the lectures, take careful notes, and memorize target information. The idea is that students with high ability would have the requisite skills to accomplish these challenges. Of course, this interpretation does not clarify why the students with high ability would fare less well with the other instructional methods, relative to the lecture-recitation method.

One study activity that appears to be sensitive to individual ability differences is *concept mapping* (creation of diagrams that show the relationship among concepts), with students with low verbal ability profiting more from concept maps (in a chemistry learning activity) than students with high verbal ability (Stensvold & Wilson, 1990). Not surprisingly, in most studies the students with higher ability outperformed the students with lower ability in both instructional conditions.

However, complete crossovers have recently been reported with embedded-question techniques for learning from textbook chapters. In Callender and McDaniel (2007), the ability of interest was the degree to which learners can construct a coherent representation of presented content (either through text or lectures). Poor structure builders are assumed to perform relatively poorly at constructing a coherent representation of connected discourse that is either read or spoken (Gernsbacher, 1990). Such comprehenders appear to construct too many substructures to accommodate incoming information, rather than constructing a unified integrated representation of the target material. By contrast, good structure builders are able to extract coherent, well-organized mental representation of the text. Accordingly, Callender and McDaniel reasoned that embedding questions into a textbook chapter would orient poor structure builders to anchoring information around which to build a coherent representation and therefore improve learning for students at this level of comprehension ability. Embedded questions might be superfluous for good structure builders, however, because they are already able to construct coherent representations.

To test these predictions, Callender and McDaniel (2007) had college-age subjects read a chapter from an introductory psychology textbook with or without embedded questions. Afterward, the subjects were given a multiple-choice test consisting of questions targeting the information featured by the embedded questions and questions on information not targeted by the embedded questions. For poor structure builders, embedded questions significantly improved performance on target questions (relative to reading without embedded questions) but not performance on nontarget questions. Good structure builders

did not profit from embedded questions, and indeed their performance for nontarget information was better without embedded questions. Note that these patterns could be considered evidence for the general notion that more guided study activities are preferable for comprehenders of lower ability, whereas less guided presentations (no embedded questions) are preferred for comprehenders of higher ability. These patterns clearly require replication, as only one chapter was considered and the subjects were in a laboratory experiment and not an actual course. Yet, this finding illustrates the potential fruitfulness of attempting to link more specific cognitive processing abilities to instructional techniques designed to dovetail with those abilities.

In summary, ATIs evidently do occur, but it has not been easy to determine exactly when they occur. This diversity of outcomes is perhaps not surprising given that available studies vary on a number of potentially critical dimensions, including target content, particular implementations of variations in instructional structure, assessments used to index ability, and the kinds of criterial (outcome) tests used. In some studies, the ATIs can be reported for one type of criterial measure but not another (e.g., see Cramer et al., 1989; Peterson, 1979; Peterson et al., 1980). At best, then, the ATI literature provides a mixed picture. A few studies are consistent with the idea that structured instruction produces better learning outcomes for students of lower ability (relative to less structured instruction), whereas less structured instruction produces better learning outcomes for students with higher abilities (relative to structured instruction). But other studies either did not obtain significant ATIs involving general ability and the degree of structure in instruction or in some cases indicated that students with lower ability fared worse with structured instruction than with less structured instruction. The greater coherence of the literature assessing structure building suggests that a more fine-grained approach that focuses on individual differences in underlying cognitive processes, rather than general aptitudes, and implements instructional methods that target those processes may be more fruitful in producing robust interactions between learner ability and learner-directed activities.

Personality-by-Treatment Interactions

There are also some more fragmentary but methodologically sophisticated studies documenting personality-by-instructional treatment effects, though these findings, like the aptitude-treatment interactions described just above, do not speak to the validity of the learning-styles hypothesis. Several studies have looked at a personality measure called *locus of control*, which refers to an individual's belief about whether his or her successes or failures are a consequence of internal or external factors (Rotter, 1966). An internal locus of control indicates a belief that outcomes are a consequence of one's own actions. An external locus of control reflects the belief that outcomes are unrelated to one's own actions. One hypothesis that has received

consideration is that learners with an internal locus of control may fare better with less structured than with highly structured instruction, whereas learners with an external locus of control will achieve more with highly structured than with less structured instruction.

Several studies have examined this hypothesis in college mathematics classes for prospective elementary school teachers. Horak and Horak (1982) examined two instructional methods during a 2-week unit on transformational geometry, with each method randomly assigned to a particular class section (total number of students was 102). In the highly guided instruction ("deductive"), students were given rules or principles and then proceeded to apply the rules to examples. In the less guided instruction ("inductive"), students were given examples, with no rule or principle stated for the students or expected from them. The criterial test included questions designed to test lower levels of understanding (knowledge of terminology and reproduction of material presented) and higher levels of understanding (e.g., problem solving). Marginally significant support for the predicted interaction was found for the questions testing lower levels of understanding: Students with an external locus of control performed better after the highly guided instruction than after the less guided instruction. The reverse was observed for students with an internal locus of control, with performance after less guided instruction exceeding performance after highly guided instruction (this also occurred with the questions tapping higher levels of understanding).

Parallel findings of marginal magnitude were reported in similar mathematics classes for elementary school teachers with shorter treatment periods (McLeod & Adams, 1980/1981). Three experiments were conducted using somewhat different instantiations of amount of guidance given during instruction and somewhat different target content. In only one experiment was the interaction significant (although a second experiment showed the same pattern): In this experiment, all students spent 1 week learning about networks with an inductive set of materials (see earlier). The amount of guidance was manipulated by having students work individually on problems and encouraging help from the instructor (high guidance, here students asked many questions) or by having students work in groups of 4 (low guidance, very few questions were posed to the instructor). On an immediate but not a delayed (given several weeks after instruction) criterial test, students with an internal locus of control performed better with low guidance than with high guidance; the reverse was found for the students with an external locus of control. The absence of significant interactions in the other two experiments may have been a consequence of shorter treatments (75-min lesson in one experiment) or small sample size (just under 60 students in each experiment), as the authors suggested.

It is interesting to note that Janicki and Peterson's (1981) study that failed to find an interaction with general ability (reviewed in the preceding section) did observe a significant personality-by-treatment interaction with a composite factor of

locus of control and attitudes toward math (in teaching fractions to grade-school students). This composite factor of *accountability for learning* interacted with instructional method such that those students with higher accountability (more internal locus of control; 36% of the students) performed better on immediate and delayed computation and story-problem tests when in the less guided small group setting than when in the highly guided direct-instruction setting. Instructional setting did not produce differences for the students with lower accountability (external locus of control).

In summary, there is modest evidence for the idea that students with an internal locus of control benefit more from less guided or structured instruction than from more guided instruction, whereas students with an external locus of control might benefit more from guided (structured) instruction than from less guided (structured) instruction. Previous studies reinforce those reviewed herein with similar patterns (Daniels & Stevens, 1976; Horak & Slobodzian, 1980; Parent, Forward, Cantor, & Mohling, 1975; Yeany, Dost, & Matthews, 1980). The reliability and generalizability of these findings to other content areas and to longer instructional treatments remain to be demonstrated. A clear uncertainty is specifying the exact aspects of instruction (group vs. individual work; density of questions directed at the instructor; homework choice vs. no choice) that are interacting with locus of control.

CONCLUSIONS AND RECOMMENDATIONS

Our evaluation of the learning-styles concept led us to identify the form of evidence needed to validate the use of learning-style assessments in instructional settings (i.e., Figures 1A–1C). As described earlier, our search of the learning-styles literature has revealed only a few fragmentary and unconvincing pieces of evidence that meet this standard, and we therefore conclude that the literature fails to provide adequate support for applying learning-style assessments in school settings. Moreover, several studies that used appropriate research designs found evidence that contradicted the learning-styles hypothesis (Massa & Mayer, 2006; Constantinidou & Baker, 2002). Finally, even if a study of a particular learning-style classification and its corresponding instructional methods was to reveal the necessary evidence, such a finding would provide support for that particular learning-style classification only—and only then if its benefits surpass the high costs of student assessments and tailored instruction.

Our conclusions have particularly clear-cut implications for educational researchers, in our opinion. We urge investigators examining learning-styles concepts to embrace the factorial randomized research designs described in the earlier “Interactions as the Key Test of the Learning-Styles Hypothesis” section, because these alone have the potential to provide action-relevant conclusions. The kind of research that is needed must begin by classifying learners into categories based on

clearly specified measures and then randomize learners to receive one of several different instructional treatments. Equally crucial, the interventions must be followed by a common pre-specified learning assessment given to all the participants in the study. The paucity of studies using this methodology is the main factor that renders the learning-styles literature so weak and unconvincing, despite its large size.

Points of Clarification

Although we have argued that the extant data do not provide support for the learning-styles hypothesis, it should be emphasized that we do not claim that the same kind of instruction is most useful in all contexts and with all learners. An obvious point is that the optimal instructional method is likely to vary across disciplines. For instance, the optimal curriculum for a writing course probably includes a heavy verbal emphasis, whereas the most efficient and effective method of teaching geometry obviously requires visual–spatial materials. Of course, identifying the optimal approach for each discipline is an empirical question, and we espouse research using strong research methods to identify the optimal approach for each kind of subject matter.

Furthermore, it is undoubtedly the case that a particular student will sometimes benefit from having a particular kind of course content presented in one way versus another. One suspects that educators’ attraction to the idea of learning styles partly reflects their (correctly) noticing how often one student may achieve enlightenment from an approach that seems useless for another student. There is, however, a great gap from such heterogeneous responses to instructional manipulations—whose reality we do not dispute—to the notion that presently available taxonomies of student types offer any valid help in deciding what kind of instruction to offer each individual. Perhaps future research may demonstrate such linkages, but at present, we find no evidence for it.

Costs and Benefits of Educational Interventions

It should also be noted that even if the evidence had convincingly documented style-by-method interactions—which we have concluded is scarcely the case—the interactions would need to be large and robust, and not just statistically significant, before the concomitant educational interventions could be recommended as cost-effective. After all, there is no doubt that interventions built around learning styles will be costly. Students must be assessed and grouped by learning style and then given some sort of customized instruction, which, in turn, requires additional teacher training as well as the creation and validation of instructional activities for each learning style. Moreover, if one is to partition the children within a given classroom and teach each subset differently, this may require increasing the number of teachers. Ultimately, the practical question will be whether the benefits of learning-styles inter-

ventions exceed other ways of using the time and money needed to incorporate these interventions.

Beliefs Versus Evidence as a Foundation for Educational Practices and Policies

Basic research on human learning and memory, especially research on human metacognition, much of it carried out in the last 20 years or so, has demonstrated that our intuitions and beliefs about how we learn are often wrong in serious ways. We do not, apparently, gain an understanding of the complexities of human learning and memory from the trials and errors of everyday living and learning. Many demonstrations have shown that participants who are asked to predict their own future performance following conditions of instruction that researchers know to be ineffective will often predict better performance under poorer conditions of instruction than will participants provided with better conditions of instruction (for a review, see Schmidt & Bjork, 1992). Part of the problem is that conditions that make performance improve rapidly *during* instruction or training, such as blocking or temporal massing of practice, can fail to support long-term retention and transfer, whereas conditions that introduce difficulties for learners and appear to slow the learning process, such as interleaving different types of problems, or employing temporal spacing of practice on what is to be learned, often enhance long-term retention and transfer. As learners, we can also be fooled by subjective impressions, such as the ease or sense of familiarity we gain on reading expository text or how readily some information comes to mind, both of which can be products of factors unrelated to actual comprehension or understanding.

There is growing evidence that people hold beliefs about how they learn that are faulty in various ways, which frequently lead people to manage their own learning and teach others in non-optimal ways. This fact makes it clear that research—not intuition or standard practices—needs to be the foundation for upgrading teaching and learning. If education is to be transformed into an evidence-based field, it is important not only to identify teaching techniques that have experimental support but also to identify widely held beliefs that affect the choices made by educational practitioners but that lack empirical support. On the basis of our review, the belief that learning-style assessments are useful in educational contexts appears to be just that—a belief. Our conclusion reinforces other recent skeptical commentary on the topic (e.g., Coffield et al., 2004; Curry, 1990; Willingham, 2005, 2009). Future research may develop learning-style measures and targeted interventions that can be shown to work in combination, with the measures sorting individuals into groups for which genuine group-by-treatment interactions can be demonstrated. At present, however, such validation is lacking, and therefore, we feel that the widespread use of learning-style measures in educational settings is unwise and a wasteful use of limited resources.

Everybody's Potential to Learn

As a final comment, we feel the need to emphasize that all humans, short of being afflicted with certain types of organic damage, are born with an astounding capacity to learn, both in the amount that can be learned in one domain and in the variety and range of what can be learned. Children, unless stifled in some way, are usually virtuosos as learners.

As we asserted earlier, it is undeniable that the instruction that is optimal for a given student will often need to be guided by the aptitude, prior knowledge, and cultural assumptions that student brings to a learning task. However, assuming that people are enormously heterogeneous in their instructional needs may draw attention away from the body of basic and applied research on learning that provides a foundation of principles and practices that can upgrade everybody's learning. For example, the finding that learners' memory for information or procedures can be directly enhanced through testing (Roediger & Karpicke, 2006) is not something that applies to only a small subset of learners but (as far as can be told) applies to all. Although performance of a student on a test will typically depend on that student's existing knowledge, testing (when carried out appropriately, which sometimes requires providing feedback) appears to enhance learning at every level of prior knowledge.

Given the capacity of humans to learn, it seems especially important to keep all avenues, options, and aspirations open for our students, our children, and ourselves. Toward that end, we think the primary focus should be on identifying and introducing the experiences, activities, and challenges that enhance everybody's learning.

SUMMARY

Our review of the learning-styles literature led us to define a particular type of evidence that we see as a minimum precondition for validating the use of a learning-style assessment in an instructional setting. As described earlier, we have been unable to find any evidence that clearly meets this standard. Moreover, several studies that used the appropriate type of research design found results that contradict the most widely held version of the learning-styles hypothesis, namely, what we have referred to as the meshing hypothesis (Constantinidou & Baker, 2002; Massa & Mayer, 2006). The contrast between the enormous popularity of the learning-styles approach within education and the lack of credible evidence for its utility is, in our opinion, striking and disturbing. If classification of students' learning styles has practical utility, it remains to be demonstrated.

Acknowledgments—We are grateful to Maxwell Moholy for assistance with the literature review. We also thank Elena L. Grigorenko, Roddy Roediger, and two anonymous reviewers for their comments on an earlier version of this article.

REFERENCES

- Bostrom, R.P., Olfman, L., & Sein, M.K. (1990). The importance of learning style in end-user training. *MIS Quarterly*, 14, 101–119.
- Callender, A.A., & McDaniel, M.A. (2007). The benefits of embedded question adjuncts for low and high structure builders. *Journal of Educational Psychology*, 99, 339–348.
- Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning. A systematic and critical review*. London: Learning and Skills Research Centre.
- Constantinidou, F., & Baker, S. (2002). Stimulus modality and verbal learning performance in normal aging. *Brain and Language*, 82, 296–311.
- Cook, D.A., Thompson, W.G., Thomas, K.G., & Thomas, M.R. (2009). Lack of interaction between sensing-intuitive learning styles and problem-first versus information-first instruction: A randomized crossover trial. *Advances in Health Science Education*, 14, 79–90.
- Cramer, K.A., Post, T.R., & Behr, M.J. (1989). Cognitive restructuring ability, teacher guidance, and perpetual distracter tasks: An aptitude–treatment interaction study. *Journal for Research in Mathematics Education*, 20, 103–110.
- Cronbach, L. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671–684.
- Cronbach, L.J., & Snow, R.E. (1977). *Aptitude and instructional methods: A handbook for research on interaction*. New York: Irvington.
- Curry, L. (1990). One critique of the research on learning styles. *Educational Leadership*, 48, 50–56.
- Daniels, R.L., & Stevens, J.P. (1976). The interaction between the internal-external locus of control and two methods of college instruction. *American Educational Research Journal*, 13, 103–113.
- Druckman, D., & Porter, L.W. (1991). Developing careers. In D. Druckman & R.A. Bjork (Eds.), *In the mind's eye: Enhancing human performance* (pp. 80–103). Washington, DC: National Academy Press.
- Dunn, R. (1990). Rita Dunn answers questions on learning styles. *Educational Leadership*, 48, 15–19.
- Foresman, S. (1972). *Windows (Teacher's guide)*. Glenview, IL: Scott Foresman.
- Freebody, P., & Tirre, W.C. (1985). Achievement outcomes of two reading programmes: An instance of aptitude–treatment interaction. *British Journal of Educational Psychology*, 55, 53–60.
- Gernsbacher, M.A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.
- Greene, J.C. (1980). Individual and teacher/class effects in aptitude treatment studies. *American Educational Research Journal*, 17, 291–302.
- The Hay Group. (n.d.). One style doesn't fit all: The different ways people learn and why it matters. Retrieved September 28, 2009, from http://www.haygroup.com/tl/Downloads/Why_People_Learn.pdf
- Henson, R.K., & Hwang, D.-Y. (2002). Variability and prediction of measurement error in Kolb's Learning Style Inventory scores a reliability generalization study, 62, 712–727.
- Honey, P., & Mumford, A. (1992). *The manual of learning styles*. Berkshire, England: Honey, Ardingly House.
- Horak, V.M., & Horak, W.J. (1982). The influence of student locus of control and teaching method on mathematics achievement. *Journal of Experimental Education*, 51, 18–21.
- Horak, W.J., & Slobodzin, K.A. (1980). Influence on instructional structure and locus of control on achievement of preservice elementary science teachers. *Journal of Research in Science Teaching*, 17, 213–222.
- International Learning Styles Network. (2008). About learning styles. Retrieved August 6, 2009, from http://www.learningstyles.net/index.php?option=com_content&task=view&id=20&Itemid=70&lang=en
- Janicki, T.C., & Peterson, P.L. (1981). Aptitude–treatment interaction effects of variations in direct instruction. *American Educational Research Journal*, 18, 63–82.
- Jensen, A.R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Jung, C.G. (1964). *Psychological types: Or, the psychology of individuation* (H. Godwin Baynes, Trans.). New York: Pantheon Books.
- Keefe, J.W. (1988). *Profiling and utilizing learning style. NASSP learning style series*. Reston, VA: National Association of Secondary School Principals.
- Kolb, D.A. (1984). *Experiential learning*. Englewood Cliffs, NJ: Prentice-Hall.
- Kolb, D. (1985). *Learning style inventory*. Boston: McBer.
- Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: Toward an integrated framework of cognitive style. *Psychological Bulletin*, 133, 464–481.
- Massa, L.J., & Mayer, R.E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer-visualizer cognitive style? *Learning and Individual Differences*, 16, 321–336.
- McLeod, D.B., & Adams, V.M. (1980/1981). Locus of control and mathematics instruction: Three exploratory studies. *Journal of Experimental Education*, 49, 94–99.
- McNamara, D.S., Kintsch, E., Butler-Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1–43.
- Ormrod, J.E. (2008). *Educational psychology: Developing learners* (6th ed.). Upper Saddle River, NJ: Pearson.
- Parent, J., Forward, J., Cantor, R., & Mohling, J. (1975). Interactive effects of teaching style and personal locus of control on student performance and satisfaction. *Journal of Educational Psychology*, 67, 764–769.
- Peterson, P.L. (1979). Aptitude $\times$ Treatment interaction effects of teacher structuring and student participation in college instruction. *Journal of Educational Psychology*, 71, 521–533.
- Peterson, P.L., Janicki, T.C., & Swing, S.R. (1980). Aptitude–treatment interaction effects of three social studies teaching approaches. *American Educational Research Journal*, 17, 339–360.
- Richardson, A. (1977). Verbalizer–Visualizer: A cognitive style dimension. *Journal of Mental Imagery*, 1, 109–126.
- Roediger, H.L. III., & Karpicke, J.D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1, 181–210.
- Rotter, J.B. (1966). Generalized expectancies of internal versus external control of reinforcements. *Psychological Monographs*, 80, (Whole no. 609).
- Rundle, S., & Dunn, R. (2007). The Building Excellence Survey [self-directed learning tool]. Retrieved October 23, 2009, from http://www.learningstyles.net/index.php?option=com_content&task=view&id=25&Itemid=78&lang=en
- Snow, R.E. (1977). Individual differences and instructional theory. *Educational Researcher*, 6, 11–15.
- Schlenker, J. (1978). *Matteson Four-Dimensional Reading Project*. Matteson, IL: Reading Service Center.
- Schmidt, R.A., & Bjork, R.A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3, 207–217.
- Spearman, C. (1927). *The abilities of man*. London: Macmillan.

- Stensvold, M.S., & Wilson, J.T. (1990). The interaction of verbal ability with concept mapping in learning from a chemistry laboratory activity. *Science Education*, 74, 473–480.
- Sternberg, R.J., Grigorenko, E.L., Ferrari, M., & Clinkenbeard, P. (1999). A triarchic analysis of an aptitude–treatment interaction. *European Journal of Psychological Assessment*, 15, 1–11.
- Sternberg, R.J., Grigorenko, E.L., & Zhang, L. (2008). Styles of learning and thinking matter in instruction and assessment. *Perspectives on Psychological Science*, 3, 486–506.
- Stricker, L.J., & Ross, J. (1964). An assessment of some structural properties of the Jungian personality typology. *Journal of Abnormal and Social Psychology*, 68, 62–71.
- Thurstone, L.L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Twenge, J. (2006). *Generation me: Why today's young Americans are more confident, assertive, entitled—and more miserable—than ever before*. New York: Free Press.
- Veres, J.G., Sims, R.R., & Shake, L.G. (1987). The reliability and classification stability of the learning style inventory in corporate settings. *Educational and Psychological Measurement*, 4, 1127–1133.
- Willingham, D.T. (2005, Summer). Do visual, auditory, and kinesthetic learners need visual, auditory, and kinesthetic instruction? *American Educator*, 29(2), 31–35.
- Willingham, D.T. (2009). *Why don't students like school: A cognitive scientist answers questions about how the mind works and what it means for the classroom*. San Francisco: Jossey-Bass.
- Witkin, H.A., & Goodenough, D.R. (1981). *Cognitive styles: Essence and origins*. New York: International University Press.
- Yale University Graduate School of Arts and Sciences. (2009). *Graduate teaching center: Teaching students with different learning styles and levels of preparation*. Retrieved September 28, 2009, from http://www.yale.edu/graduateschool/teaching/learning_styles.html
- Yeany, R.H., Dost, R.J., & Matthews, R.W. (1980). The effects of diagnostic-prescriptive instruction and locus of control on the achievement and attitudes of university students. *Journal of Research in Science Teaching*, 17, 537–546.