of the Sternberg triarchic abilities test in three international samples: An empirical test of the triarchic theory of intelligence. *European Journal of Psychological Assessment, 17*(1), 1–16.
Sternberg, R. J., Grigorenko, E. L. (1995). Styles of thinking in school. *European Journal for High Ability, 6*(2), 201–219.
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), 1–11.
Sternberg, R. J., Grigorenko, E. L., & Zhang, L.-F. (2008). *Styles of learning and thinking matter in instruction and assessment*. Perspectives on Psychological Science, 3, 486–506.
Sternberg, R. J., & Kaufman, S. B. (Eds.)(2011). *Cambridge handbook of intelligence.* New York: Cambridge University Press.
Sternberg, R. J., & Wagner, R. K. (1992). *Thinking Styles Inventory.* Unpublished test, Yale University.
Sternberg, R. J., & Zhang, L. F. (Eds.). (2001). *Perspectives on thinking, learning and cognitive styles*. Mahwah, NJ: Erlbaum.
Witkin, H. A., Dyk, R. B., Fattuson, H. F., Goodenough, D. R., & Karp, S. A. (1962). *Psychological differentiation: Studies of development.* Oxford, England: Wiley.
Zhang, L. F. (2004). Revisiting the predictive power of thinking styles for academic performance. *Journal of Psychology, 138,* 351–370.
Zhang, L. F., & Sternberg, R. J. (2002). Thinking styles and teachers' characteristics. *International Journal of Psychology, 37*(1), 3–12.
Zhang, L. F., & Sternberg, R. J. (2006). *The nature of intellectual styles.* Mahwah, NJ: Erlbaum.
Zhang, L.-F., & Sternberg, R. J. (Eds.). (2009). *Perspectives on the nature of intellectual styles.* New York, NY: Springer.
Zhang, L.-F., Sternberg, R. J., & Rayner, S. (Eds.). (2012). *Handbook of intellectual styles.* New York, NY: Springer.

COMMENT ON STERNBERG'S REVIEW OF ZHANG

Robert Sternberg's review of L.-F. Zhang's *The Malleability of Intellectual Styles* devotes a surprising amount of attention to our critical review of learning styles (Pashler, McDaniel, Rohrer, & Bjork, 2008). That would be flattering—except that his comments do not represent accurately what we did and what we claimed.

Our article was commissioned by *Psychological Science in the Public Interest,* a journal whose purpose is to help the public and the media understand which behavior-related practices and conclusions rely on a strong evidence base and which do not. In that spirit, we took our charge to be helping teachers and other interested people determine whether there exists good evidence for the claim that students learn better when they are grouped by learning style and then provided with instruction that is tailored to their style.

Although Sternberg's review does not mention it, learning styles is not just an interesting academic topic; it is also an industry. There are many promoters who actively market tests, books, workshops, and so forth to school districts and to individual teachers. Our sense from talking to teachers is that although most of them do not personally administer learning styles tests to their students, many of them assume that they really *ought* to be doing that, and regret that they do not have the opportunity to do more of it.

As we saw it, our job was to figure out whether that regret was well founded, and our final conclusion was that it was not. In our article we warned that learning styles products and practices lacked the sort of validation evidence that, in our opinion, ought to be presented *before* an industry emerges, not after. In the pharmaceutical world, the Food and Drug Administration requires strong validation before allowing drugs onto the market. At no point in our article did we claim that evidence ruled out the possibility that learning style measures (existing ones or new ones not yet developed) might be validated in *future* research. Indeed, we called for such research and suggested a basic roadmap for how it would need to be done.

As Sternberg correctly notes, our analysis of the learning style literature rested on our argument that a certain minimal standard would have to be met for a study to provide convincing evidence for the educational usefulness of learning styles. This standard was neither exotic nor idiosyncratic. To the contrary, it is very much in line with widely accepted principles of intervention research.

In brief, we said that any research study validating the use of learning styles for instruction must, as a first step, classify learners using a test of some sort. Then it must show that this classification successfully predicts which of several different instructional procedures will produce the best learning outcomes (a particular type of interaction). These learning outcomes must be assessed with a common measure in order for comparisons to be meaningful.

To show that our reasoning is wrongheaded, it seems to us it would have been sufficient to describe a research design that is *missing* one or more of our key ingredients but nevertheless shows the efficacy of instructional tailoring. We find no such description in Sternberg's article. Instead, he makes a number of other arguments.

First, he contends that our insistence on randomized controlled trials (RCTs) is overly rigid and even possibly antiquated. In our view, Sternberg could hardly be more wrong here. Across a broad range of human activities and fields, recognition of the crucial value of RCTs has increased dramatically in the last 20 years. Partly this is because of a few famous debacles, such as hormone replacement therapy. Large-scale RCTs revealed that numerous nonrandomized studies performed over four decades had not even correctly identified the *direction* of hormone replacement therapy's effects on important outcomes such as heart disease. The results produced a dramatic change in prescription practices and a broad awareness in the biomedical world about how easily nonrandomized designs can mislead us (Krieger et al., 2005).

Since that time, researchers in many fields have come to realize that rather than being an isolated oddity, this pattern of results is quite common. Stanley Young and Karr (2011) examined 12 different observational results reported in the medical literature and contended that none of them were replicated when proper RCTs were conducted. Similar findings have been noted in intervention studies of welfare-to-work strategies, where even sophisticated observational data analysis designs mispredicted the results found with RCTs (Michalopoulos, Hill, & Lei, 2002). We think it is fair to say that respect for the need for RCTs in the intervention research community is greater than ever before.

This awareness of the necessity of RCTs is now spreading far beyond academia and medicine. Thanks to the growth of the Internet, corporate interest in RCTs is exploding (Glennerster & Takavarasha, 2013), and the number of behavioral RCTs conducted every year by the technology industry probably dwarfs the number of experiments done by research psychologists. In parallel with that development, forward-looking governments (e.g., in the United Kingdom and Washington State in the United States) are beginning to undertake field RCTs on a scale never before contemplated, evaluating causal impacts on behaviors ranging from retirement savings to organ donation. Thus, as we see it, the suggestion that randomized trials are a tired and fading old custom is not just wrong, it is dramatically wrong. There has never been a time when RCTs have inspired as much excitement as they are doing right now (see Manzi, 2012, for an excellent overview of this excitement).

Second, Sternberg implies that our article made the bald claim that learning styles do not even exist. We explicitly avoided making any such claim, acknowledging that people express fairly consistent beliefs about what kind of instructional materials work best for them. Based on that, we see no problem with someone saying, if so inclined, that someone has a certain learning style. Our point, though, was that the fact people can be sorted based on their self-reported impressions about what works for them does not make them right about what works for them. One can measure opinions and call them "styles" or anything else you want to call them; the important question is whether doing so provides any predictive leverage. One broad and important conclusion based on the last quarter century or so of research on metacognition and learning is that we, as learners, often have a faulty mental model of how we learn, making us susceptible to illusions of comprehension and to preferring poorer conditions of study or practice over better conditions (for a review, see Bjork, Dunlosky, & Kornell, 2013).

As we mentioned in our original review, the fact that people have strong and consistent preferences for how materials are presented is only one reason why the learning style approach is appealing to many people. To think that one has a unique style of learning is also appealing, as is being able to attribute our not learning as efficiently as we might like to teachers and trainers not presenting material in a way that meshes with our style. The "meshing hypothesis"—namely that material should be presented in a way the meshes with our style—seems logical, but we could find no serious support for that view.

That people have preferred styles has led some educators to be skeptical of our conclusions, as have other considerations, including that Person X might be high in math and low in verbal, whereas Person Y is the opposite, that different disciplines lend themselves to different optimal styles (e.g., geometry should be visual), and that the optimal approach may often involve multiple approaches (e.g., combining verbal and visual instruction). However, none of those considerations constitutes evidence that supports the use of learning style classifications in education.

In a related vein, Sternberg contends that our review offered no evidence to either support or rule out molding instruction to complement learning styles. In fact, we discussed at length the study by Massa and Mayer (2006), which struck us as both exemplary and illuminating.

Third, Sternberg says that we failed to notice that certain other traits, ones he assumes we take more seriously than we do learning styles (e.g., intelligence), are also subject to the very same criticisms we offered. He is right to suspect we take ability measures more seriously than we do learning style measures, but he is wrong that ability measures lack validation. There is a large literature going back many decades documenting the predictive validity of IQ measures for a wide array of real-world outcomes (e.g., Gottfredson, 2002). Research on ability is a mature field, one that reflects years of diligent work by psychometricians who have focused heavily on predictive validity.

On the other hand, we agree with Sternberg that aptitude–treatment interactions in education (i.e., the ability of mental ability measures to predict the optimal instructional procedure for an individual) have indeed been difficult to nail down firmly. If we had been charged with evaluating the utility of measuring mental ability for the optimal tailoring of instruction, we would have insisted on that kind of evidence and complained about its thinness.

We also agree with Sternberg that conducting RCTs in classroom settings is challenging. Indeed, each of us has some experience with long-term, classroom-based studies. On the positive side, however, the current funding environment and presence of education evaluation companies devoted to implementing such designs make the challenge easier to meet. Some of the hurdles Sternberg mentions are possible to surmount, such as randomization of instructors (see Zacharis, 2011, for a learning style quasiexperiment in which the same instructor implemented different instructional formats). Moreover, there is ample opportunity for researchers to follow the lead of Massa and Mayer (2006), whose laboratory study we touted in our article. These investigators took the most popular style dimension (visual vs. verbal) and conducted a very credible intervention study using precisely the design we advocated. The study was short term and lab based, but the tasks, the instructional interventions, and the tests were meaningful and well grounded in both cognitive theory and common sense. For example, the topic that was taught was one that could very reasonably be presented with either a heavy reliance on words or a heavy reliance on diagrams.

From what we have heard, our article is prompting many other investigators to adopt similar research designs looking at a range of other learning style dimensions, and we look forward to seeing the results. Our article may have displeased Sternberg, but we are hopeful that it will have encouraged the development of a literature that in 5 or 10 years will provide a much better guide for action than what exists at the moment.

Harold Pashler
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093
E-mail: hpashler@ucsd.edu

Robert Bjork
University of California, Los Angeles

Mark McDaniel
Washington University

Doug Rohrer
University of South Florida

REFERENCES

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64,* 417–444.

Glennerster, R., & Takavarasha, K. (2013). *Running randomized evaluations: A practical guide*. Princeton, NJ: Princeton University Press.

Gottfredson, L. S. (2002). Where and why *g* matters: Not a mystery. *Human Performance, 15*(1–2), 25–46.

Krieger, N., Löwy, I., Aronowitz, R., Bigby, J., Dickersin, K., Garner, E., . . . Weisz, G. (2005). Hormone replacement therapy, cancer, controversies, and women's health: Historical, epidemiological, biological, clinical, and advocacy perspectives. *Journal of Epidemiology and Community Health, 59*(9), 740–748.

Manzi, J. (2012). *Uncontrolled*. New York, NY: Basic Books.

Massa, L. J., & Mayer, R. E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer–visualizer cognitive style? *Learning and Individual Differences, 16*(4), 321–335.

Michalopoulos, C., Hill, C. J., & Lei, Y. (2002). *Can nonexperimental comparison group methods match the findings from a random assignment evaluation of mandatory welfare-to-work programs?* New York, NY: MDRC.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles concepts and evidence. *Psychological Science in the Public Interest, 9,* 105–119.

Stanley Young, S., & Karr, A. (2011). Deming, data and observational studies: A process out of control and needing fixing. *Significance, 8*(3), 116–120.

Zacharis, N. Z. (2011). The effect of learning style on preference for Web-based courses and learning outcomes. *British Journal of Educational Technology, 42*(5), 790–800.

RESPONSE TO PASHLER, BJORK, MCDANIEL, AND ROHRER

Pashler et al. make three specific points in their review. All three points are misrepresentations of what I wrote.

Point 1 is that "He [Sternberg] contends that our insistence on randomized controlled trials (RCTs) is overly rigid and even possibly antiquated." I did not contend this. Rather, the appropriate method depends on the substantive problem to which the method is addressed. For example, RCTs are useful for instructional efficacy studies; they are not useful for determining, say, the degree of (correlational) relationship between styles and, for example, ability, personality, or achievement measures. The tool should fit the problem, lest one become like the carpenter always looking to use his hammer.

As I stated in my original review, other methods can be useful in naturalistic contexts for ascertaining the educational usefulness of styles in instruction. Examples are structural and hierarchical regression. In the latter, one predicts educational outcomes by multiple kinds of measures (e.g., ability, styles), entering styles (or any crucial variable) *last*. Styles are useful if they add predictive variance to educational outcomes, over and above (holding constant) the effects of other potentially relevant kinds of measures.

Point 2 is that "Sternberg implies that our article made the bald claim that learning styles do not even exist." I did not imply this. On the contrary, I quoted them as saying that "there is no adequate evidence to justify incorporating learning-styles assessments into general educational practice" (p. 105). My objection was not to their article in general but rather to this sweeping and overly broad conclusion. The Zhang (2013) book; Kozhevnikov, Evans, and Kosslyn (2014, published in the same journal where Pashler et al. published their original article); and the numerous other sources I cite in my review show that there are many kinds of evidence for the usefulness of styles in general educational practice.

Point 3 states, "He is right to suspect we take ability measures more seriously than we do learning style measures, but he is wrong that ability measures lack validation." This is a misrepresentation of what I wrote. Of course ability measures have validation. I have been using them extensively in my own research, starting with my dissertation (see Sternberg, 1977). But their validity always depends on the purpose for which they are used, the population with which they are used, and the circumstances under which they are administered.

Pashler, McDaniel, M., Rohrer, D., and Bjork (2008) went way beyond their data in drawing a sweeping conclusion about the lack of utility for learning style assessments in educational practice. Had they limited their conclusion to the data they presented and the arguments they actually made, I would have largely agreed with them.

Robert J. Sternberg
Department of Human Development
B44 MVR
Cornell University
Ithaca, NY 14853
E-mail: robert.sternberg@cornell.edu

REFERENCES

Kozhevnikov, M., Evans, C., & Kosslyn, S. M. (2014). Cognitive style as environmentally sensitive individual differences in cognition: A modern synthesis and applications in education, business and management. *Psychological Science in the Public Interest, 15*(1), 3–33.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*(3), 105–119.

Sternberg, R. J. (1977). Component processes in analogical reasoning. *Psychological Review, 84,* 353–378.

Zhang, L.-F. (2013). *The malleability of intellectual styles.* New York, NY: Cambridge University Press.

HAPPINESS: A THEORY OF RELATIVITY

The Myths of Happiness: What Should Make You Happy, but Doesn't, What Shouldn't Make You Happy, but Does

By Sonja Lyubomirsky. New York, NY: Oxford University Press, 2014. Paperback, $27.95.

A little misery may enhance the quality of one's life. (Wedell & Parducci, 2000, p. 240)

Happiness is a basic human concern involving both desire and pursuit. Most people strive for satisfaction, contentedness, and meaning in life, yet contemporary culture suggests that happiness is difficult to attain and fleeting when it comes. In *The Myths of Happiness,* Sonja Lyubomirsky argues that lasting happiness may be found by those who "think instead of blink" (p. 251). Thoughtful people can gain insight into the limitations of many culturally prescribed notions about the attainment of happiness and thereby expand their freedom to pursue and achieve genuine life satisfaction.