SCIENTIFIC STUDIES OF READING, *7*(2), 105–126

Understanding the Styles of Science in the Study of Reading

Keith E. Stanovich
University of Toronto

Remarkable progress has been made in the last 30 years toward understanding the basic psychological processes that underlie the act of reading. Dissemination of these results has sometimes been hampered because they are associated with strong stances on 5 different dimensions that represent styles of doing science. Research into the psychology of reading has been characterized by an emphasis on correspondence theories of truth rather than coherence, an emphasis on analytic reductionism rather than holism, an emphasis on probabilistic prediction (as opposed to a case-based approach), the search for robust-process explanations (rather than actual-sequence explanations), and a concern for consilience. These scientific styles have served the field well to this point, but that does not mean that we have calibrated their use in the optimal manner. Critiques by those with differing scientific styles may help the field to adjust its stance on these style dimensions in ways that foster scientific progress. This article ends with some thoughts on the difficulty of defining the scientific method.

The honor of delivering the Distinguished Scientific Contribution Award address to the Society for the Scientific Study of Reading (SSSR) allowed me to discuss some of the obstacles we face in disseminating the results of scientific investigations on the acquisition and teaching of reading. The work honored was made possible by collaborative research with a host of individuals. I have been unusually blessed in my colleagues. In Richard West and Anne Cunningham I have had two continuous collaborators and friends for more than 2 decades. A host of others (among them Linda Allen, Jim Cipielewski, Ruth Nathan, and Linda Siegel) have

---

Requests for reprints should be sent to Keith E. Stanovich, Department of Human Development and Applied Psychology, Ontario Institute for Studies in Education, University of Toronto, 252 Bloor Street West, Toronto, Ontario, Canada M5S IV6. E-mail: kstanovich@oise.utoronto.ca

been associated with me for a decade or more. In my recent book (Stanovich, 2000), I acknowledge more than 40 colleagues, and I thank them again here.

In that book I focused on this problem of scientific dissemination in our field and how it is linked to the peculiar politics of the field of reading. In approaching the dissemination problem this time, I focus on something slightly different—what I call the styles of science in the study of reading. My conclusion will be that there are certain reasoning styles characteristic of members of organizations such as the SSSR that amount to something that is virtually a worldview—a worldview that can seem strange to those not sharing those styles and that can impede communication and understanding. My point here will not be to convert anyone either to or from these styles. My goal is a more modest one of simply bringing to awareness—to facilitate communication—the differing styles of science within our field. The marvelous progress in understanding reading that has occurred over the past 30 years is obscured and misunderstood because much of the public—and some of the field itself—does not share the reasoning styles that have made the progress possible.

Understanding the styles of science that have generated this period of unusual scientific progress will help to contextualize why the current period may seem to us to be both the best of times and the worst of times. All SSSR members are proud o see the article in the March 2002 *Scientific American* authored by SSSR award winners, officers, and members (Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001) and its much more fleshed-out companion in the journal *Psychological Science in the Public Interest* (see Marks) as well as the influential National Reading Panel report (Ehri, Nunes, Stahl, & Willows, 2001; Ehri, Nunes, Willows, et al., 2001), in part authored by several SSSR members, as was its predecessor report from the Committee on the Prevention of Reading Difficulties in Young Children (Snow, Burns, & Griffin, 1998). At the same time that these wonderful consensus research conclusions have been emanating from the literature, major publishers are presenting to teachers advertisements that put the term *science* in scare quotes (e.g., an advertisement in the International Reading Association's *Reading Today,* March/April 2002). In other publications (e.g., Taylor, 1998) leading researchers in SSSR are called "spin doctors" rather than scientists. Some (e.g., Bialostok, 1997) see the research like that presented at the SSSR conference as merely an exercise in rhetorical posturing (see Kamil, 1995, and McKenna, Stahl, & Reinking, 1994, for contextualizations of these disputes in our literature).

It is indeed the worst of times just as it is, from another perspective, the best of times.

Members of the reading research community are often frustrated when what they view to be the empirical facts about the reading process are distorted or denied in communications directed to the public and to teachers. It seems as if some are denying the facts. But I submit that in many cases, the facts are secondary—what is being denied are the styles of reasoning that gave rise to the facts; what is being denied is closer to a worldview than an empirical finding. Many of these styles are implicit; we are not conscious of them as explicit rules of behavior. I discuss five of these implicit reasoning styles.

CORRESPONDENCE VERSUS COHERENCE

The first style concerns a bias in weighting correspondence versus coherence criteria when conducting and interpreting research. This style dimension maps on to debates in philosophy about theories of truth. The correspondence theory of truth is the default model of most working scientists, and a bias toward it—rather than coherence—is a characteristic style that is associated with views of scientific progress within a society such as SSSR. Correspondence theories (see Searle, 1995) are easy to understand because, at a superficial level, they represent a view that would be shared by the person in the street: Simply, there is a real world out there that exists independently of our beliefs about it, researchers form theories about this world, and the theories that track the world best are closer to the truth and are thus a better basis for action. This is why planes don't fall out of the sky, why bridges rarely collapse, and why my headache medication works more often than not.

In contrast to the correspondence view, which dominates the thinking of most reading researchers, many in the qualitative research communities emphasize constructivist principles that put stress on the requirement that beliefs fit together in a reasonably logical way—the so-called coherence theories of truth. Sometimes, the coherentist view is promulgated in the social sciences in such an extreme form that the *sole* criterion of truth becomes merely whether a set of beliefs makes sense to the person constructing the beliefs. For example, presenting such a view to the audience of the *Harvard Educational Review,* Mishler (1990) stated that it is irrelevant whether interpretations correspond to or mirror objective reality because

> a concern with distortion places the burden of validity claims on the wrong shoulders. … Instead of assuming a past reality as a criterion, a potential warrant for the validity of my interpretation is whether it makes sense to the respondent. (p. 427)

Variants of the coherence view are apparent in some fairly extreme forms throughout the behavioral sciences and humanities where we often hear that what is called knowledge is not tracking an independently existing world but is instead simply evolving, internally constructed coherence. Philosopher Michael Devitt (1991) discussed how much postmodern scholarship is the amalgamation of Kantian ideas with 20th-century epistemological relativism (see also Hacking, 1999; Ruse, 1999). To these scholars, an exclusive emphasis on correspondence leaves something crucial out of the equation. Coherence values underpin the opposition to scientific syntheses that derive from investigators agreeing on correspondence values (such as those in SSSR). And sometimes those coherence values can be hard for those with our worldview to appreciate. For example, Mishler (1990), again in the *Harvard Educational Review,* called the concepts of reliability and objectivity *shibboleths*—rhetorical strategies that "serve a deviance-sanctioning function" (p. 420).

Although such a view is indeed extreme, it is important to understand some of the unappealing aspects of the correspondence view. For example, an extreme adherence to a correspondence theory of truth often necessitates the frustration of the strong human need for narrative coherence in explanation. Theoretical abstractions that follow from correspondence-driven empirical investigations often disrupt folk conceptions that have existed for centuries because they were congruent with unaided perception. Coherence views that stress congruence with introspection resonate with more narrative clarity in such cases. Understanding a table as a solid unbroken surface coheres well with unaided perception. Understanding it as composed of atoms—most of which are empty space—threatens the narrative coherence of our view of the world, yet it is just the type of counterintuitive theoretical explanation we are constantly led to by the correspondence view (see Flanagan, 2002, on the difference between manifest and scientific images). This is an unpalatable aspect of the correspondence view, which those of us enamored with it tend to forget or simply do not to notice. It causes us to be unnecessarily puzzled when those not sharing the correspondence view reject the results that emanate from it.

Cognitive scientist and philosopher Patricia Churchland (Farber & Churchland, 1995) discussed how, as the concept of fire developed in explanatory power in science, its narrative coherence in folk discourse was challenged and disrupted. The classical concept was used

> to classify not only burning carbon stuffs, but also activity on the sun and various stars (actually nuclear fusion), lightening (actually electrically induced incandescence), the northern lights (actually spectral emission); and the flash of fireflies (actually phosphorescence). In our modern conceptual scheme, because none of these things involves oxidation, none belongs to the same class as wood fires. Moreover, some processes that belong to the oxidation class—rusting, tarnishing, and metabolism—were not originally considered to share anything with burning, because felt heat was taken to be an essential feature of this class. (p. 1296)

In short, the explanatory frameworks that are generated by scientists working in the correspondence framework may not seem plausible to those who value coherence more, particularly the type of coherence that resonates with the narratives inherent in folk psychology. The principle of oxidation uniting the phenomena of a campfire and rusting—and cleaving away the phenomenon of lightening—may be exciting to a scientific mind steeped in the correspondence view, but it will alienate those seeking coherent narratives that are more easily assimilated to preexisting world models.

In reading research, there have been many examples of how narratively coherent concepts became reformed and reshaped under the assault of empirical test. For example, the changing concept of context effects in reading displayed exactly the trend that Farber and Churchland (1995) described. In his seminal 1971 book, Frank Smith told a very coherent story about the role of context in reading. Empirical work by Chuck Perfetti, Susan Goldman, and colleagues (e.g., Perfetti, Goldman, & Hogaboam, 1979) and some of my own work with Richard West (e.g., Stanovich, West, & Feeman, 1981; West & Stanovich, 1978), however, served to theoretically fractionate the concept of context in reading—showing that effects at different levels of processing (e.g., word recognition vs. comprehension) had to be distinguished because individual difference relationships varied at the different levels. These findings and the associated theory disrupted greatly the narrative about context in Smith's book, but correspondence considerations dictated these complications. As I have written about extensively, however, we underestimated how compelling was the original narrative about context effects to those who valued coherence more than the correspondence criteria on which we, and other investigators at the time, were myopically focused. Thus, we underestimated how difficult this conceptual change would be for the field.

By starting with correspondence versus coherence as my first scientific style dimension, I do risk stacking the deck in favor of my own scientific style and worldview. An emphasis on correspondence is so critical to the scientific edifice (and science is so central to the modern world) that it is easy to caricature the postmodernists, relativists, and other schools of thought that define themselves in part by their denial of the primacy of correspondence criteria. As Dawkins (1995) said, "show me a cultural relativist at thirty thousand feet and I'll show you a hypocrite" (pp. 31–32). Or one might paraphrase Walker Percy (as cited in Lehman, 1991, p. 113) and say that a postmodernist is a person who claims that texts have an almost indeterminate number of interpretations and who then leaves a note on their refrigerator requesting a pepperoni pizza for supper!

More seriously, numerous authors have written about how the coherence doctrine, by linking itself with ecumenical notions such as tolerance and personal validation, obscures its uglier aspects. What has been obscured is how indiscriminate belief validation, with no check in external reality, creates a world that most of us would consider a nightmare. In this world, the witnesses and evidence in a jury trial are not sifted as to credibility because any piece of evidence put forward is equal to any other for the reason that all are valid by *someone's* perspective.

Writing on the implications of relativism, philosopher Daniel Dennett (1995, p. 154) attempted to break through the warm-fuzzy fog that has enshrouded the relativistic coherence theory by conjuring thought experiments that reveal its implications. He asked, for example, if you would want to be operated on by a surgeon who sometimes hears a voice in his head telling him to ignore his medical training—and when he hears this voice, he is prone to listen to it!

But as I mentioned earlier, I have somewhat stacked the deck here. I have started out with a scientific style that is hard to argue we have gotten wrong. If others find it hard to assimilate our findings because of a heavy reliance on correspondence criteria, the problem seems not to reside with reading researchers themselves. There is every reason to believe that we have calibrated this value correctly.

Not so with the remaining scientific styles I discuss—all of which are much harder to weight correctly. When others find it hard to assimilate our findings because of a heavy reliance on *these* values, our critics cannot be so easily dismissed.

ANALYTIC REDUCTIONISM VERSUS HOLISM

With a few exceptions, the research conducted by SSSR members has been heavily biased toward a style of analytic reductionism characteristic of the physical sciences and certain subfields of the biological sciences such as molecular biology. Our field has progressed by understanding the subprocesses of reading—by analytic fractionation that has progressed from the theoretical arguments of the 1970s in which one talked of models of reading as a global process to the 1980s and 1990s when word recognition as a crucial subprocess was understood, and its critical subprocesses of phonological and orthographic coding became the focus. The same trend toward analytic fractionation is apparent in the area of comprehension research. Of course, the current intense work on fractionating the theoretical concept of phonological awareness represents another classic example.

Here, however, the field does create problems of assimilation for outsiders by not being more fully cognizant of how unnatural the analytic style that we take for granted actually is—how readily it leaves us open to the charge that we are leaving out contextual factors that are of immense importance. Here, we might be better off admitting that the analytic stance is a calculated gamble and that we have made a bet based on how this gamble has worked out in other sciences. And, indeed, other sciences have grappled with the best stance to take on the analytic and holistic continuum—especially biology, which is just like the field of reading in having a long-standing debate between analytic and holistic approaches. For instance, biologist John Maynard Smith (2000) articulated the source of his disagreements with Evelyn Fox Keller about the mechanisms of gene action. Maynard Smith speculated that their differences are actually due to what I have been calling here different styles of science:

> Our disagreements … concern differences about the best strategy to pursue when faced by the complexities of living organisms. I know that the world is complicated, but always seek for simple explanations of the complexity. For Keller, living organisms only work because they are complex; to simplify them is to leave out their essence. (p. 43)

This quote could be about *reading,* and things would virtually play out the same. Many of us in this field know that reading is complicated but always seek for simple explanations of the complexity. For many of our constructivist and whole-language critics, to simplify reading is to leave out its essence. For those with an analytic bias, the proof is in the pudding—the gamble is that a gain in explanatory power will not result in a loss of contextualized understanding.

But perhaps we should make greater efforts to show our awareness of this caution and an appreciation for alternative scientific styles as did Maynard Smith (2000) in the following quote from the same review:

> I am stimulated by Keller's ideas, but I sometimes find myself disagreeing with her conclusions. Some of our disagreements may be settled by history. Essentially, they are disagreements about the kind of scientific explanation that we expect to be fruitful. I see myself as a reductionist, although my friend the biologist Lewis Wolpert sees me as a woolly-minded holist. I seek simple models of the world whose consequences I can work out. If I have to ignore some of the detail, too bad. My favorite guide to scientific theorizing is a remark by Pete Richerson and Robert Boyd: "To replace a world you do not understand by a model of the world you do not understand is no advance." Keller, in contrast, thinks that the behavior of a living organism depends on its complexity: if you ignore the complexity, you will not understand anything. If you start by ignoring what you take to be details, or brushing them under the carpet, you will end by throwing away the baby and keeping the bathwater. I suspect that science needs both types of approach. (p. 46)

Make no mistake, however, that even though I endorse Maynard Smith's ecumenical recommendation, my own personal bet in reading is on the analytic approach, for I think that in reading, excessive holism has obscured progress as it has in biology. Discussing the latter discipline, Dawkins (1998) noted that he grants the holistic point that

> at every level the units interact with each other. … This has all been said many times before, and is so obvious as to be almost platitudinous. But one sometimes has to repeat platitudes in order to prove that one's heart is in the right place! … Reductionism is a dirty word, and a kind of "holistier than thou" self-righteousness has become fashionable. … But there are times when holistic preaching becomes an easy substitute for thought. (p. 113)

And I do think holistier than though platitudes have been dominant in the reading field—and have contributed to driving many researchers into the analytic camp. Holistic claims at their best are additive—they contextualize the analytic explanation and thus add to it. However, holistic claims can sometimes be subtractive as well—when, for instance, they seek to negate analytic progress. In reading, we have been subjected to too much subtractive holistic advocacy, again perhaps serving to unduly undermine the credibility of the holistic style in the reading field. The following quote is entirely too typical: "The brain does not have separate compartments as the reading models suggest. The brain is highly integrated. … We must assume that the brain processes language in a holistic manner" (Grundin, 1994, p. 8). Ignoring the "separate compartments" caricature, it is clear that such a stance serves to functionally negate a vast amount of work in cognitive neuroscience and the cognitive science of reading. Much work in these subdisciplines does—contrary to the Grundin quote—support the notion of the separability of the cognitive and brain processes underlying reading (e.g., Byrnes, 2001; Carr, 1992; Posner & Carr, 1992; Pugh et al., 2000; Pugh et al., 1997; Pulvermuller, 1999; Shaywitz et al., 1998). Not only do separate brain architectures subserve word recognition and comprehension, but even within the word recognition module, subprocesses are subserved by separate and localizable brain areas.

Our situation here is not unique. Advocacy of a vague holism characterized biology as well in the era before the analytic successes of molecular biology. For example, in his book *Unweaving the Rainbow,* Richard Dawkins (1998) discussed the following quote in an old textbook by an author who, after acknowledging that atoms were, indeed, separate entities, argued:

> Not so the gene. It exists only as part of the chromosome, and the chromosome only as part of a cell. If I ask for a living chromosome, that is, for the only effective kind of chromosome, no one can give it to me except in its living surroundings any more than he can give me a living arm or leg. The doctrine of the relativity of functions is as true for the gene as it is for any of the organs of the body. They exist and function only in relation to other organs. Thus the last of the biological theories leaves us where we started, in the presence of a power called life or psyche which is not only of its own kind but unique in each and all of its exhibitions. (p. 90)

Dawkins (1998) told us bluntly that

> this is dramatically, profoundly, hugely wrong. And it really matters. Following Watson and Crick and the revolution they sparked, a gene can be isolated. It can be purified, bottled, crystallized, read as digitally coded information, printed on a page, fed into a computer, read out again into a test tube and reinserted into an organism where it works exactly as it did before. (p. 90)

The analogy here is direct. As the field of reading research marches ahead by analyzing and fractionating reading processes to a now almost exquisite degree (e.g., Berent & Perfetti, 1995; Booth, Perfetti, & MacWhinney, 1999; Frost, 1998; Perfetti & Tan, 1998; Rayner, 1998), holistic critiques must become more sophisticated. They must become additive rather than subtractive.

And such positive, additive holistic critiques *have* occasionally occurred in our field. Take, for example, the interest in whether there are associations between the orthographic transparency of writing systems and the ease of reading acquisition and incidence of reading difficulty. Sakamoto and Makita (1973; Makita, 1968) conducted research on the issue in the early 1970s, and Gibson and Levin devoted a section of their influential 1975 book to this topic. Stevenson conducted a series of much-publicized investigations on cross-linguistic differences in reading acquisition in the 1980s (e.g., Stevenson et al., 1985). But in the late 1990s this area of research rocketed forward with seminal work by Heinz Wimmer, Karin Landerl, Uta Frith, Usha Goswami, and many others (e.g., Frith, Wimmer, & Landerl, 1998; Geva, Wade-Woolley, & Shany, 1993; Goswami, 2000; Goswami, Gombert, & DeBarrera, 1998; Landerl, Wimmer, & Frith, 1997; Wimmer, 1996). It is not just the quantity of studies that have changed but the much more sophisticated way these investigations are now contextualized compared to the work in the 1970s and 1980s. For example, in just the way that holistic advocates have urged all along, these investigations now show a much more sophisticated appreciation of differences in instructional environments across orthographies. Gone are the days when such investigations were couched as if comparing a disembodied mind interacting with a disembodied orthography. Investigators in this area appreciate the necessity of adding the learning environment and instructional context as interacting factors in the model of orthography and achievement links that is being developed. This area displays an additive holism rather than the subtractive holism that has soured so many scientists on that end of the analytic and holistic continuum. As holistic critiques of the analytic bias in our work become more responsible—as in the area of cross-linguistic differences—it will not be so easy to determine a location on the style continuum that is optimal for the field. It will become clearer that Maynard Smith's recommendation should be followed: We need a mix of both styles.

PROBABLISTIC PREDICTION VERSUS A CASE-BASED APPROACH

A third dimension of scientific style that causes disagreement within the field and that is the source of no end of difficulties in communicating our findings to teachers and to the public is an emphasis on probabilistic prediction as opposed to a case-based approach. We in the behavioral science of the study of reading are so accustomed to probabilistic explanation and prediction that we are prone to forget how alien it seems to the layperson, to teachers, and indeed to some working in our own field. In some cases, we may fail to see that the notion of probabilistic prediction needs to be brought to people's attention.

Even more important, we often fail to explain how to interpret our probabilistic predictions when faced with an audience immersed in a case-based approach. Many of us have had the following experience. We are presenting to students or to a group of teachers the findings on the effects of phonological awareness training, and a questioner gets up from the audience to ask the inevitable question, "But does it work for every beginning reader encountering difficulty?—It doesn't work for everyone does it?" We of course are accurate and responsible in answering the questioner—we tell the teacher that no, the procedures do not work for every child, but researchers are working hard to find out why. But too often we stop right there, without a further explanation of what does and does not follow from the fact that the training does not work for every pupil. We assume that the teacher is going to take away the same inference that we—steeped in the probabilistic style—will take away. We take it for granted that the teacher knows that the approach we are recommending provides a bedrock of probabilistic success on which any other explanation will have to build; that probabilistic prediction is cumulative prediction—that variation not predicted by the first factor or method does not negate the success of that method. Probabilistic explanations do not negate but explain overlapping and nonoverlapping variance—and so on and so forth.

And the "so forths" are considerable—I could go on for quite some time with them. And that is the point. There are a number of background assumptions that contextualize our almost unconscious use of the logic of probabilistic prediction, and these assumptions are often completely opaque to our audience. Instead of making all of the assumptions I have outlined, including our implicit model of cumulative prediction, the teacher holds instead an alternative model (one more in tune with the case-based style). The alternative model is not the cumulative model of probabilistic inference, but a replacement model—a model that says that if certain cases cannot be explained and treated successfully with the method, then a new model must be found that *replaces* the first model.

We, in contrast, worry that the replacement model will be costly—it will throw away the hard-won first-pass variance that has been acquired by what I call the baseline model. In searching for a model to help the smaller number of children not helped by the baseline method, the replacement model will lose the plurality of children who *are* helped by the baseline method. But we really need to understand how the probabilistic style itself has unpalatable aspects and to understand what is attractive about the case-based approach.

The thought processes that the teacher is using are similar to those that we know a lot about because they have been studied by cognitive psychologists for more than 4 decades. The case-based thought processes of the teacher are captured in the phenomenon of so-called probability matching, which was studied by Bill Estes, Ward Edwards, and Amos Tversky in the 1960s and continues to be a standard paradigm (e.g., Estes, 1964; Fantino & Esfandiari, 2002; Gal & Baron, 1996; Tversky & Edwards, 1966).

In one version of it, the participant sits in front of two lights (one red and one blue) and is told that she or he is to predict which of the lights will be flashed on each trial and that there will be several dozen such trials (participants are often paid money for correct predictions). The experimenter has actually programmed the lights to flash randomly, with the provision that the red light will flash 70% of the time and the blue light will flash 30% of the time. Participants do quickly pick up the fact that the red light is flashing more, and they predict that it will flash on more trials than the blue light. Most often, they switch back and forth, predicting the red light roughly 70% of the time and the blue light roughly 30% of the time.

This strategy of probability matching is suboptimal because it ensures that, in this example, the participant will correctly predict only 58% of the time ($.7 \times .7 + .3 \times .3$) compared to the 70% hit rate that could be achieved by predicting the more likely color on each trial. In fact, much experimentation has indicated that people fail to maximize expected utility in these probabilistic learning experiments even when given financial incentives for correct predictions. We now know about some of the thought processes that sustain the suboptimal probability-matching behavior, and one in particular is relevant to the question raised by the teacher with the case-based approach mentioned previously.

One thing that participants find unpalatable about the utility-maximizing strategy of constantly predicting the more probable red light is that this strategy gives up on the possibility of getting *all* of the predictions correct. Only by predicting 70% red and 30% blue can anyone even sustain the hope of predicting correctly each time. The maximizing strategy locks the participants into a 70% payoff. This seems problematic, but probabilistic thinking dictates that this seemingly problematic aspect be ignored (note again the counterintuitive thinking that is required here). Going in search of the extra 30% hits (by predicting on the basis of a 70/30 split) will reduce the hit rate to 58% compared with the sure 70% of the maximizing strategy. However, unlike the maximizing strategy, the 58% hits of the probability-matching strategy will have a variance—some participants will do worse than this and some better—some so far above 58% that they are above 70% as well. Again, the advocate of the probabilistic mindset will remind us that they will be matched by an equal number equidistant below 58%—but no matter, says the case-based gambler; I at least have the hope of getting 100% correct.

Of course, the hope of the teacher is even more motivated than that of the gambler—the teacher wants to help every child. The teacher's replacement logic is appropriate when a case-based approach is actually called for—when the teacher is not making an aggregate decision. But often teachers (and the public) carry over the case-based style into situations where the decision is clearly an aggregate one—choosing group activities in a classroom, choosing instructional time allocation, choosing a school's or district's curriculum, or choosing the allocation of in-service training time. There, the probabilistic style is the appropriate one, and the replacement logic is inappropriate—there is no replacement program that could beat the number of children who are helped by the baseline treatment, just as there is no probabilistic choice strategy that could beat the 70% hit rate of the maximizing strategy in the example earlier. The potential mistake that worries me when I answer the questioner alluded to previously, is that they will adopt a case-based approach to an aggregate decision.

In many cases though, in the complicated field of reading—in its domains of application—it really is unclear whether we should be in probabilistic or case-based mode. Hence, one source of friction in our field follows from some quite legitimate differences over how to calibrate style of explanation to the particular situation.

ROBUST-PROCESS EXPLANATIONS VERSUS ACTUAL-SEQUENCE EXPLANATIONS

This scientific style is related to a fourth-style dimension, which is discussed quite a bit in the philosophy of biology literature and deserves more discussion in ours: the difference between a robust-process explanation and an actual-sequence explanation. Kim Sterelny (2001; Sterelny & Griffiths, 1999) is a philosopher of biology who has written about the importance of this distinction in all emerging sciences. To a philosopher, someone seeking a robust-process explanation is looking for a causal model that defines a class of possible worlds—the set in which a posited series of causal linkages holds. Actual-sequence explanations seek to identify a particular world within this class—a particular world we have designated for interest. Better to use an example, and I use two from Sterelny's (2001; Sterelny & Griffiths, 1999) work—the first referring to his native New Zealand. He asks how we might explain the fact that the birthrate in New Zealand dropped dramatically after World War I. He points out that one way to answer the question would be possible in principle (if not in practice): Go through all of the birth and death registries and trace the microhistories of all of the individual people and families of New Zealand during that period (a daunting task even for so small a country, but possible in principle). This would be seeking an actual-sequence explanation. An alternative would be to abstract oneself from the microhistories of individual families and invoke large-scale principles. For example, it is known that technological change during wartime spawns increasing urbanization, which in turn spawns lower birthrates. One could then show that these large-scale causal mechanisms were at work in New Zealand at the time.

A robust-process explanation in this case would identify New Zealand as a member of the class of countries in which the purported causal conditions were fulfilled. The term *robust* is apt because, in a full causal explanation of this sort, a change in the initial starting conditions and parameter settings of nonessential variables should not (if the theory is correct) jeopardize New Zealand's membership in the large class for which the causal relations hold. To have found that New Zealand fulfills the conditions is satisfying to a scientist focused on robust process but unfulfilling to someone on the other end of this continuum—it does not tell us the precise micromechanisms, the particular states that occurred in New Zealand.

Sterelny's second example again concerns World War I, in which he pointed out that the assassination of Archduke Franz Ferdinand of Austria would loom large in any actual-sequence explanation (e.g., as indeed would the trajectory of the bullet that struck him), but this event would take on much less importance in a robust-process explanation, which would of course stress instead the geopolitical instability of the period. From that perspective, the sequence that followed the assassination of the Archduke was just one of a class of conflicts that follow from similar levels of instability.

The general implication for the reading field should now be clear. Those of us studying the psychology of the reading process are often after robust-process explanations, whereas we often address audiences who are interested in and oriented toward actual-sequence explanations. Teachers often want to know how this *particular* child reached this level of school achievement or this level of reading difficulty, as the case may be. The explanations we often offer amount to robust-process explanations: This child performs as a member of a class of children whose behavior is accounted for by the following causal model, and so on. The recipient of this explanation often feels that something is being left out—an actual-sequence explanation of how this child arrived at where they are, which often the researcher is unable to give (for the same reason that searching in the New Zealand birth archives would be difficult). But I do not mean to stack the deck with this example. In reading, there are qualitative researchers looking for better ways to reveal actual-sequence explanations, and these endeavors could provide useful information because robust-process and actual-sequence explanations do not compete with each other (as is clear from work in allied disciplines; see Jackson & Pettit, 1992; Sterelny, 1996). By subdividing robust-process explanations we get closer to actual-sequence explanations, and by aggregating actual-sequence explanations we get closer to a robust-process explanation. The two work in concert.

CONSILIENCE VERSUS UNIQUENESS

The fifth-style dimension that characterizes the scientific study of reading—and the last I discuss—is the quest for consilience versus uniqueness. On this dimension, our field is strongly characterized by the search for consilience. I take the term *consilience* from E. O. Wilson's (1998) notable book of the same name. In my own short-methods textbook (Stanovich, 2001) I called this the principle of connectivity. (Because Wilson's book has sold a couple hundred thousand more than mine, I use his term!) Wilson defines *consilience* as the "unification of knowledge by the linking of facts and fact-based theory across disciplines to create a common groundwork of explanation" (p. 8). The reading field has been strongly characterized by the quest for consilience in that the best work in the field has been notable for seeking connections to cognitive neuroscience, to work in computer simulation, and to the latest work in linguistics. The quest for consilience is clearly apparent in the report of the Rayner group, published in *Psychological Science in the Public Interest* (Rayner et al., 2001), where a concern for consilience was clearly evident as they attempted to link current reading theory with facts about different orthographies, connectionist modeling, cognitive neuroscience, and classroom studies of effective practice. Consilience is demonstrated when there is a continuity of explanatory mechanisms across all of the domains in question—something that has at least been partially achieved in our now growing list of synthetic research reports: the Rayner group collaboration (Rayner et al., 2001), the National Reading Panel report (Ehri, Nunes, Stahl, & Willows, 2001), the Snow report (Snow et al., 1998), and of course Marilyn Adams's (1990) seminal volume.

Articles displaying a strong concern for consilience—knitting together theory and data from the many subfields that interface with the scientific study of reading—are common in our field. A typical example of such efforts is provided by a recent article in *Developmental Review* by Steffler (2001), who deftly wove together findings on spelling development with the theory from a very well developed area of experimental psychology—the study of implicit memory and implicit learning. Also employed in the same article are theoretical ideas from one of the most influential theories in cognitive developmental psychology—Annette Karmiloff-Smith's (1992) theory of representational redescription.

This concern for consilience contrasts with the faddish tendencies in the field of education to search for magic bullets and miracle cures deriving from theories that do not cohere with the knowledge being developed by allied disciplines. The quest for a magic bullet always tempts education to stray from valuing consilience. Magic bullets are by their very nature new and unique—a break with the past. Emphasis on the magic bullet aspects of a treatment inevitably results in the weakening of consilience considerations, because anything that is deeply enmeshed in what we already know cannot be that new or unique. The style of valuing consilience is a useful inoculation against fad educational treatments.

So these are the five scientific styles that I see as having driven the extraordinary advance of the science of reading in the last 30 years: an adherence to a correspondence theory of truth, an emphasis on analytic reductionism, an emphasis on aggregate and probabilistic prediction, an emphasis on robust-process explanations—and on consilience. I am not saying that we have the emphasis on these styles properly calibrated—only that they are characteristic of some of the best work produced in the field defined by the interests of the members of this society. Perhaps we would progress faster if we recalibrated some of them—moved a bit and lessened our bias toward one end of the continuum. My own feeling about these styles varies among them. I think we know how best to calibrate some of them, but we are less sure about others—and our critics may be helping us think about some that are less secure. Maynard Smith (2000) was surely right that we would never want the field to become monolithic in its stance on any style dimension.

How far might we move around on these five different style dimensions and still be doing science? Quite far, I believe. Figure 1 presents the five style dimensions as continua. Every scientist probably has a position on each of the continua and thus a spiked profile through this space of styles. Does every profile through the space depict a possible style conjunction that represents a valid scientific approach? Or, are some style profiles so deviant that they represent a scholarly approach that is best not viewed as science at all? My view is that most—but not all—profiles through the space would be stances that are validly termed scientific. Figure 2 presents my best guess at what the restrictions on the scientific styles are. There is a slash on each of the continua to indicate that I view any profile passing to the left of all the slashes as validly termed as science. As is clear from the figure, I view three of the dimensions (analytic reductionism versus holism, probabilistic prediction, and robust-process explanation) as completely open—any position on these dimensions may be taken and still represent a valid scientific approach. In contrast, I have slashed the mid-

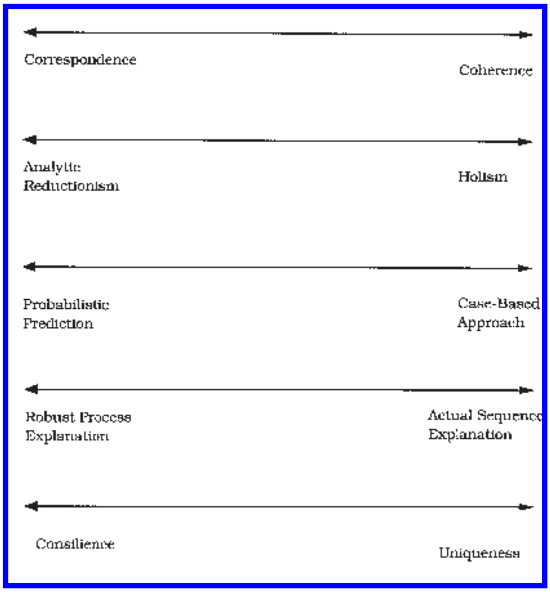

FIGURE 1 The five style dimensions represented as continua.

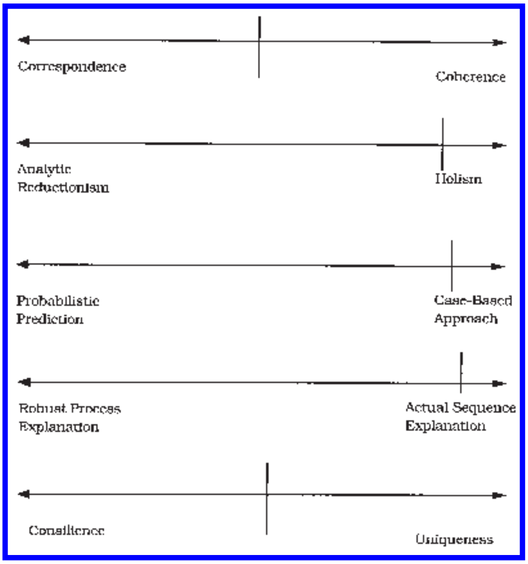

FIGURE 2 The five style dimensions with slashes indicating a conjecture about the limits of valid scientific approaches.

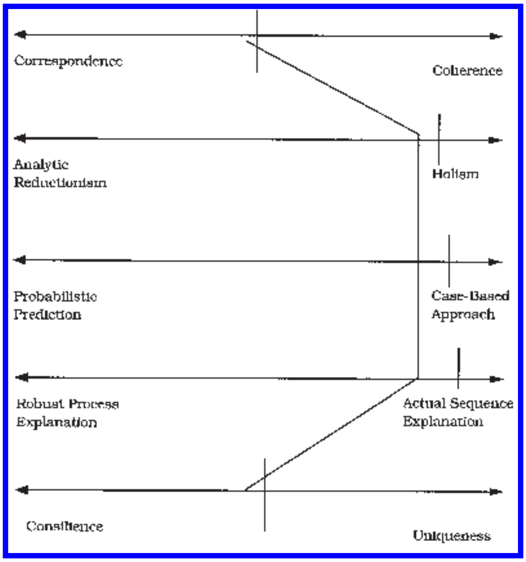

FIGURE 3 Profiles passing to the left of the demarcated area are conjectured to be valid scientific styles.

point of the correspondence versus coherence and the consilience versus uniqueness dimensions. This is meant to indicate that I view at least some minimal adherence to correspondence and consilience values as essential to the scientific attitude. One cannot completely reject correspondence or consilience criteria and still claim to be doing science.

Thus, my view of valid scientific styles is quite ecumenical (a vast number of style profiles would fulfill my criteria for valid scientific styles), but it is not completely unrestricted—two of the dimensions constrain the types of styles that can still be considered scientific. The area to the left of the quasi-polygon in Figure 3 identifies the areas of valid scientific styles in my view. Every profile that passes through this area defines a set of scientific styles that could, in theory, be encompassed within the community of science.

THE DIFFICULTY OF DEFINING THE SCIENTIFIC METHOD

I want to end this article with one warning and a suggestion—both contextualized by my discussion here of scientific styles. The warning concerns the current trend in legislation among educational policymakers that educational practice be based on valid research conclusions. Of course, it is heartening to finally see a concern in educational circles for research-based conclusions. Lately, however, we have been called on more and more to explain the mechanisms of science—for example, for definitions of reliable, replicable scientific research. We are being asked more and more to explain what the science of reading is. My only caution here is that this is going to be harder than some people think. For example, I have tried to respond as best I could when requested by various governmental bodies in the United States and Canada to provide inputs on the definition of scientific research.

I do fear that some would prefer rigid rules and bullet points on a PowerPoint presentation rather than the story of science in its full complexity—including the complexities of certain styles that are neither right nor wrong but represent continua (better viewed as parameters that we are constantly adjusting so as to facilitate the process). Science is a delicate epistemological game. There are no rigid rules for properly adjusting theory to evidence for example. And many of its modes of operation represent dispositions rather than rules—something I have stressed here by using the term *styles* of scientific thinking. Science's real uniqueness comes from its self-correcting nature. Its unique epistemic power comes from a very un-Promethean characteristic: its constant fiddling with things—with theory, experimental setups, techniques, and its styles of the type I have discussed. And we are not afraid to readjust—which itself, recursively, is one of science's characteristic features—we are not afraid to implicitly admit previous error when we make a readjustment.

I fear that this flexibility, this juggling, this self-corrective mindset will be lost if we too rigidly reduce science to a set of rules. Do not misunderstand me, though—I am wholeheartedly in favor of instructing teachers and other educational personnel in what science is and is not and what are the unique features that underlie its epistemic power. But these features should not become a prison. And we should not underestimate how difficult some of these styles and principles will be for nonscientists to assimilate. Many of the styles I have discussed are counterintuitive (e.g., probabilistic prediction and explanations in terms of robust process) as are many other principles of science such as falsifiability.

We are not the only field dealing with the issue of what constitutes scientific knowledge. Other fields grapple with it themselves, including such superordinate adjudicating bodies as the U.S. Supreme Court. In an effort to deal with the increasing prevalence of so-called junk science in the courtroom, the Supreme Court attempted to define what reliable, replicable scientific research was and marshaled considerable talent in attempting to develop a definition (see Foster & Huber, 1999).

Despite input from some of the best scientific minds and writers in the world (including the recently deceased Stephen J. Gould) some of the justices admitted to having difficulty with some of the concepts. One of the current justices of the Supreme Court, Chief Justice Rehnquist, said, for example

> I defer to no one in my confidence in federal judges; but I am at loss to know what is meant when it is said that the scientific status of a theory depends on it 'falsifiability' and I suspect some of them will be too. (Foster & Huber, 1999, p. 38)

But the difficulty of explaining these principles to the public and to teachers will not be aided by reducing them to bullet points. So my only caution is that we not succumb to the temptation to reduce the complex, evolving nature of the scientific process to a series of oversimplified rules in our admitted eagerness to infuse more of the scientific spirit into education—an eagerness that I share.

ADVERSARIAL COLLABORATION

Finally, I bring to the attention of reading researchers a procedure being used in some subareas of psychology to end scholarly disputes that had previously resisted resolution in the literature. I mentioned earlier that extreme differences in scientific style account for some of the internecine warfare that has plagued our field. How can these differing perspectives find a common ground? They can (see Kamil, 1995) if all parties are actually engaging in the scientific process (albeit from vastly different positions on the style dimensions I outlined). But how can we determine whether some disputants are engaged in science at all? I suggest that one criterion we can invoke is the willingness to engage in what has been termed adversarial collaboration.

A prominent example was recently published in the flagship journal of the American Psychological Society, the journal *Psychological Science,* by Mellers, Hertwig, and Kahneman (2001). The last two authors are the adversarial collaborators—they were engaged in a heated theoretical dispute that had gone on for more than a decade. The first author, Mellers, is the so-called arbiter. The theoretical dispute concerned whether the conjunction fallacy, a probabilistic reasoning error, is entirely eliminated when frequency formats are used, and a spirited set of articles had preceded the adversarial collaboration (see Hertwig & Gigerenzer, 1999; Kahneman & Tversky, 1996).

Mellers et al. (2001) described some of the steps in adversarial collaboration, and many are just what you would expect. The adversarial collaborators agree on an arbiter. The arbiter adjudicates agreements on how the experiments to resolve the dispute will be conducted. The arbiter has control of the data, the final decision on the outlet of publication, the final decision on the format of publication, and even the final authority to declare one of the collaborators uncooperative in print if that is warranted! If all goes well, the adversarial collaboration is published with all three parties as authors, usually with the arbiter first.

When we are dealing with the difficult task of delineating the difference between scientific research and political and rhetorical advocacy in education, the willingness to engage in adversarial collaboration might serve as a useful criterion of demarcation. In the middle of the heat of the internecine warfare that sometimes takes place in our field, it is important to be able to differentiate legitimate scientific criticism that derives from background assumptions from the opposite ends of the style dimensions I have discussed here—and criticisms deriving from critics who are not playing the game of science at all. Willingness to accept offers of adversarial collaboration might be a tool to use in distinguishing who is playing science from who is not. As 30 years of productive work in the study of the psychology of reading have proven, for those who wish to understand reading and to help children learn to read by using that understanding, science is a game eminently worth playing.

ACKNOWLEDGMENTS

This article is based on the Distinguished Scientific Contribution Award address presented at the meeting of the Society for the Scientific Study of Reading, Chicago, June 2002. Preparation of the article was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

I thank Frank Manis and Richard West for their helpful comments.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review, 102,* 146–184.

Bialostok, S. (1997, December). Offering the olive branch: The rhetoric of insincerity. *Language Arts, 74,* 618–629.

Booth, J. R., Perfetti, C. A., & MacWhinney, B. (1999). Quick, automatic, and general activation of orthographic and phonological representations in young readers. *Developmental Psychology, 35,* 3–19.

Byrnes, J. P. (2001). *Minds, brains, and learning: Understanding the psychological and educational relevance of neuroscientific research*. New York: Guilford.

Carr, T. H. (1992). Automaticity and cognitive anatomy: Is word recognition "automatic"? *American Journal of Psychology, 105,* 201–237.

Dawkins, R. (1995). *River out of Eden: A Darwinian view of life.* New York: Basic Books.

Dawkins, R. (1998). *Unweaving the rainbow.* Boston: Houghton Mifflin.

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meanings of life.* New York: Simon & Schuster.

Devitt, M. (1991). *Realism & truth* (2nd ed.). Cambridge, MA: Basil Blackwell.

Ehri, L. C., Nunes, S., Stahl, S., & Willows, D. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71,* 393–447.

Ehri, L. C., Nunes, S., Willows, D., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36,* 250–287.

Estes, W. B. (1964). Probability learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 89–128). New York: Academic.

Fantino, E., & Esfandiari, A. (2002). Probability matching: Encouraging optimal responding in humans. *Canadian Journal of Experimental Psychology, 56,* 58–63.

Farber, I. B., & Churchland, P. S. (1995). Consciousness and the neurosciences: Philosophical and theoretical issues. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1295–1306). Cambridge, MA: MIT Press.

Flanagan, O. (2002). *The problem of the soul.* New York: Basic Books.

Foster, K. R., & Huber, P. W. (1999). *Judging science: Scientific knowledge and the federal courts.* Cambridge, MA: MIT Press.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading, 2,* 31–54.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99.

Gal, I., & Baron, J. (1996). Understanding repeated simple choices. *Thinking and Reasoning, 2,* 81–98.

Geva, E., Wade-Woolley, L., & Shany, M. (1993). The concurrent development of spelling and decoding in two different orthographies. *Journal of Reading Behavior, 25,* 383–406.

Gibson, E. J., & Levin, H. (1975). *The psychology of reading.* Cambridge, MA: MIT Press.

Goswami, U. (2000). Phonological representations, reading development, and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia, 6,* 133–151.

Goswami, U., Gombert, J., & DeBarrera, F. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics, 19,* 19–52.

Grundin, H. (1994). Who's romancing reality? *The Reading Teacher, 48,* 8–10.

Hacking, I. (1999). *The social construction of what?* Cambridge, MA: Harvard University Press.

Hertwig, R., & Gigerenzer, G. (1999). The conjunction fallacy revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12,* 275–305.

Jackson, F., & Pettit, P. (1992). In defense of explanatory ecumenicalism. *Economics and Philosophy, 8,* 1–21.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103,* 582–591.

Kamil, M. L. (1995). Some alternatives to paradigm wars in literacy research. *Journal of Reading Behavior, 27,* 243–261.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, MA: MIT Press.

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition, 63,* 315–334.

Lehman, D. (1991). *Signs of the times: Deconstruction and the fall of Paul de Man.* New York: Poseidon.

Makita, K. (1968). The rarity of reading disability in Japanese children. *American Journal of Orthopsychiatry, 38,* 599–614.

Maynard Smith, J. (2000, December 21). [Review of the book *The Cheshire cat's DNA: The century of the geneI*]. *New York Review of Books, 47*(20), 43–46.

McKenna, M., Stahl, S., & Reinking, D. (1994). A critical commentary on research, politics, and whole language. *Journal of Reading Behavior, 26,* 211–233.

Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science, 12,* 269–275.

Mishler, E. (1990). Validation in inquiry-guided research: The role of exemplars in narrative studies. *Harvard Educational Review, 60,* 415–442.

Perfetti, C. A., Goldman, S., & Hogaboam, T. (1979). Reading skill and the identification of words in discourse context. *Memory & Cognition, 7,* 273–282.

Perfetti, C. A., & Tan, H. L. (1998). The time course of graphic, phonological and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning Memory and Cognition, 24,* 101–118.

Posner, M. I., & Carr, T. H. (1992). Lexical access and the brain: Anatomical constraints on cognitive models of word recognition. *American Journal of Psychology, 105,* 1–26.

Pugh, K. R., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Constable, R. T., et al. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. *Psychological Science, 11,* 51–56.

Pugh, K., Shaywitz, B. A., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Fletcher, J. M., et al. (1997). Predicting reading performance from neuroimaging profiles: The cerebral basis of phonological effects in printed word identification. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 299–318.

Pulvermuller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22,* 253–336.

Rayner, K. (1998). Eye movements in reading and information processing: 20 Years of research. *Psychological Bulletin, 124,* 372–422.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2,* 31–74.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2002, March). How should reading be taught? *Scientific American, 286*(3), 84–91.

Ruse, M. (1999). *Mystery of mysteries: Is evolution a social construction*? Cambridge, MA: Harvard University Press.

Sakamoto, T., & Makita, K. (1973). Japan. In J. Downing (Ed.), *Comparative reading* (pp. 440–465). New York: Macmillan.

Searle, J. R. (1995). *The construction of social reality.* New York: Free Press.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K., Fulbright, R., Constable, R., Mencl, W., et al. (1998). Functional disruption in the organization of the brain for reading and dyslexia. *Proceedings of the National Academy of Sciences USA, 95,* 2636–2641.

Smith, F. (1971). *Understanding reading.* New York: Holt, Rinehart and Winston.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers.* New York: Guilford.

Stanovich, K. E. (2001). *How to think straight about psychology* (6th ed.). Boston: Allyn & Bacon.

Stanovich, K. E., West, R. F., & Feeman, D. (1981). A longitudinal study of sentence context effects in second-grade children. *Journal of Experimental Child Psychology, 32,* 185–199.

Steffler, D. J. (2001). Implicit cognition and spelling development. *Developmental Review, 21,* 168–204.

Sterelny, K. (1996). Explanatory pluralism in evolutionary biology. *Biology and Philosophy, 11,* 193–214.

Sterelny, K. (2001). *The evolution of agency and other essays.* Cambridge, England: Cambridge University Press.

Sterelny, K., & Griffiths, P. E. (1999). *Sex and death: An introduction to philosophy of biology.* Chicago: University of Chicago Press.

Stevenson, H. W., Stigler, J. W., Lee, S. Y., Lucker, G. W., Kitamura, S., & Hsu, C. C. (1985). Cognitive performance and academic achievement of Japanese, Chinese, and American children. *Child Development, 56,* 718–734.

Taylor, D. (1998). *Beginning to read and the spin doctors of science.* Urbana, IL: National Council of Teachers of English.

Tversky, A., & Edwards, W. (1966). Information versus reward in binary choice. *Journal of Experimental Psychology, 71,* 680–683.

West, R. F., & Stanovich, K. E. (1978). Automatic contextual facilitation in readers of three ages. *Child Development, 49,* 717–727.

Wilson, E. O. (1998). *Consilience: The unity of knowledge.* New York: Knopf.

Wimmer, H. (1996). The nonword reading deficit in developmental dyslexia: Evidence from children learning to read German. *Journal of Experimental Child Psychology, 61,* 80–90.

Manuscript received September 4, 2002
Accepted September 25, 2002