

# Cargo-Cult Science versus

*Traditional educational research does not yield reliable generalizations; instead...*

By E. D. Hirsch, Jr.

“We really ought to look into theories that don’t work, and science that isn’t science. I think the educational . . . studies I mentioned are examples of what I would like to call cargo cult science. In the South Seas, there is a cargo cult of people. During the war, they saw airplanes with lots of good materials, and they want the same thing to happen now. So they’ve arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head for headphones and bars of bamboo sticking out like antennas — he’s the controller — and they wait for the airplanes to land. They’re doing everything right. The form is perfect. It looks exactly the way it looked before. But it doesn’t work. No airplanes land. So I call these things cargo cult science, because they follow all the apparent precepts and forms of scientific investigation, but they’re missing something essential, because the planes don’t land.”

**Richard P. Feynman, “Cargo Cult Science,” *Surely You’re Joking, Mr. Feynman!***

After many years of educational research, it is disconcerting — and also deeply significant — that we have little dependable research guidance for school policy. Richard Feynman identifies part of the reason for this shortcoming — that while educational research sometimes adopts the outward form of science, it does not burrow to its essence.

For Feynman, a Nobel prize winner in physics, the essence of good science is doing whatever is necessary to get to reliable and usable knowledge — a goal not necessarily achieved by merely following the external forms of a “method.”

The statistical methods of educational research have become highly-sophisticated. But the quality of the statistical analysis is much higher than its practical utility.

Despite the high claims being made for statistical techniques like regression analysis, or experimental techniques like random assignment of students into experimental and control groups, classroom-based research (as contrasted with laboratory research) has not been able to rid itself of uncontrolled influences called “noise” that have made it impossible to separate out the relative contribution of the various factors that have led to “statistically-significant” results.

This is a chief reason for the unreliability and fruitlessness of current classroom research. An uncertainty principle subsists at its heart.

As a consequence, every partisan in the education wars is able to utter the words “research has shown” in support of almost any position. Thus, “research” is invoked as a rhetorical weapon — its main current use.

If we take an example of the best educational research — say the Tennessee class-size experiment (STAR) — and ponder why it fails to serve policymakers well, some very basic reasons present themselves.

The STAR researchers dutifully gathered complex data, and used control groups and experimental groups, and applied sophisticated statistical techniques. They even followed the rigorous practice of purely random assignment of students to the experimental and control groups.

The researchers found that reducing class sizes from 24 to 15 enhanced equity and achievement in early grades. Yet when California legislators spent \$5 billion to reduce class size, the predicted effect did not result. The planes didn’t land.

The process of generalizing from classroom research is inherently unreliable. There are fundamental difficulties standing in the way of such direct generalization.

Young children learn slowly. The cumulative effects of interventions are gradual, extending over years. Yet most educational research is conducted over spans measured in months, rather than years, ensuring that effect sizes will tend to be small.

These effects may be rendered almost invisible by another difficulty — the fact that the process of schooling is exceedingly context-dependent. Detailed analyses of the contextual factors that influence learning are greatly to be desired, but progress in understanding those contextual factors is unlikely to result from coarse-grained classroom studies. Progress is more likely to result from highly-controlled “artificial” experiments that reveal the fine-grained underlying causes.

It used to be thought that damp, low-lying air caused “swamp fever.” (The other term for swamp fever, “malaria,” means “bad air.”) That theory of the cause of the disease was accepted by medical science as long as researchers stuck to coarse-grained observations which indicated that if you live in a swamp you are likely to get swamp fever. It was not until the disease was put under the microscope that progress began to be made in determining the true causes and vectors of malaria.

Medical science continues to advance as it becomes allied with ever more refined laboratory understandings. Its most striking and reliable advances have occurred since medicine became closely tied to biochemistry at a still more fine-grained level — the molecular. By analogy, it is plausible to think that progress in educational research will follow this sort of pattern.

# Reliable General Principles

*... we should put our faith in more general principles gleaned from cognitive science.*

By E. D. Hirsch, Jr.

To lend some credence to the proposition that general cognitive principles tend to be more dependable than conclusions drawn from direct classroom research, I shall now outline some issues in cognitive science about which a degree of consensus has been reached.

Shrewd application of these consensus principles would almost certainly enhance classroom learning, and ought also to encourage a shift in the way policymakers use educational data and research.

## **Prior knowledge is a prerequisite to effective learning**

The research that offers the most dramatic evidence that relevant prior knowledge is critical to thinking skill is the area of expert-novice studies. The expert learns more from a given experience than a novice does, even though the novice has much more still to learn.

That's because being presented with too many not-yet-interpreted items overloads and confuses the mind, whereas prior knowledge makes experience salient and meaningful.

## **Meaningfulness**

In the expert-novice experiments, it is thought that prior knowledge enables the expert not only to connect the elements of an experience, but also to pick out what is meaningful and salient in it.

For example, the connection between the sound and the sense of many words is entirely arbitrary. But, once the arbitrary sound-sense connection is learned, the meaningfulness of those words ensures that they will be remembered.

## **The right mix of generalization and example**

The optimal mode for learning most subjects is through a carefully-devised combination of the general concept and well-selected examples. The idea of teaching by both precept and example is so old that its confirmation in experiment is no surprise.

One famous experiment showed that the concept "bird" is stored (by North Americans) as something about the size of a robin, not the size of a hummingbird or ostrich. Concept and example are deeply connected with one another in how we think and remember, as well as how we learn.

## **Attention determines learning**

Although "motivation" and "interest" are perennial themes of education, it is sobering to discover from cognitive science that motivation is only an indirect and dispensable aid to learning. Some things that we involuntarily pay attention to are learned and remembered better than things we are trying to learn.

Attention is an aspect of our "working memory," a function that lasts just a few seconds. Out of the whirl of perceptual features that impinge on working memory every instant, we attend only to a salient few. That few is very, very limited in number, even for the most brilliant minds.

An expert with prior knowledge will be able to attend to many more things than a novice, not because of greater mental capacity, but because of "chunking." For an expert, noticing one thing is automatically to notice a myriad of things implied by it and known to be chunked with it, whereas the novice would need to get through dozens of connections, which, because of the limitations of working memory, is impossible.

If the attended-to things are given meaning by being connected with what we already know, we will learn (remember) them. If we do not attend to them and do not accommodate them to some known structure, we will usually not learn them.

This finding is a sobering reminder that we should not be overly distracted by the vast and unreliable literature on what will or will not properly motivate students — a debate that seems baffling to many teachers, since what motivates some students does not motivate others.

## **Repetition is usually necessary for retention**

How long something will be remembered is typically determined by how often it has been attended to. Repetition has the double purpose of retention and making meaningful connections between experiences. There is evidence that the need for repetition has a physical basis in the neuron structure of the brain.

## **Automaticity (through rehearsal) is essential to higher skills**

Rehearsal serves other purposes beyond long-term retention and the constructing of meaningful connections. It also makes certain operations non-conscious and automatic.

An obvious example comes from sports: the more one has to think about all the motions required for hitting a tennis ball well, the less one is likely to do so. In sports, no one doubts the need to gain automaticity.

In the academic context, automaticity frees up the working memory and allows it to concentrate on higher-order thinking.

## **Implicit instruction of beginners is usually less effective**

Cognitive science tells us that teachers should use both implicit and explicit teaching. It seems likely that explicit learning is best for a limited number of foundational elements, while implicit learning is best for advancing slowly on a broad front.

Cognitive science, in contrast to school-based research, gathers data from many sources and explains why they converge on a consensus interpretation. The difference in the two fields is that, whereas classroom research rarely converges on a consensus view, cognitive science has recently begun to do so.

*(Adapted with permission from "Classroom Research and Cargo Cults" in the Hoover Institution's Policy Review, No. 115, [www.policyreview.org](http://www.policyreview.org). Dr. Hirsch is the author of The Schools We Need.)*