It’s not unusual to read that a new study has failed to replicate – or has even reversed – the findings of an earlier study. The effect can be disconcerting, particularly when medical research announces that what was supposed to be good for us turns out to be dangerous, or vice versa.

Qualifications and reversals also show up in investigations of education and human behavior, but here an interesting pattern seems to emerge. At first a study seems to validate traditional practices, but then subsequent studies – those that follow subjects for longer periods of time or use more sophisticated outcome measures – call that result into question.

That’s not really surprising when you stop to think about it. Traditional practices (with respect to teaching students but also to raising children and managing employees) often consist of what might be called a “doing to” – as opposed to
a “working with” — approach, the point being to act on people to achieve a specific goal. These strategies sometimes succeed in producing an effect in the short term. Research — which itself is often limited in duration or design — may certify the effort as successful. But when you watch what happens later on, or you look more carefully at the impact of these interventions, the initial findings have a way of going up in smoke.

Consider three such traditional practices: using rewards to change people’s behavior, making students do additional academic assignments after they get home from school, and teaching by means of old-fashioned telling (sometimes known as direct instruction). What happens in each case when you look at short-term results, and what happens when you then extend the length of the study?

1. An ambitious investigation of various types of preschools looked specifically at children from low-income families in Illinois. The two educational approaches that produced the greatest impact on achievement in reading and arithmetic were both highly structured, one of them a behaviorist technique called Direct Instruction (or DISTAR) that emphasizes the use of scripted drill in academic skills and praise for correct responses.

Most studies would have left it at that, and the press doubtless would have published the findings, which suggests that this model is superior to more child-centered preschools. (Take that, you progressives!) Luckily, though, this particular group of researchers had both the funding and the interest to continue tracking the children long after they left preschool. And it turned out that with each year that went by, the advantage of two years of regimented reading-skills instruction evaporated, soon proving equivalent — in terms of effects on test scores — to “an intensive 1-hour reading readiness support program” that had been provided to another group. “This follow-up data lends
little support for the introduction of formal reading instruction during the preschool years for children from low-income homes,” the researchers wrote.

One difference did show up much later, however: Almost three quarters of the kids in play-oriented and Montessori preschools ended up graduating from high school, as compared to less than half of the direct-instruction kids, which was about the same rate for those who hadn’t attended preschool at all. (Other longitudinal studies of preschool have found similar results: The longer you track the kids, the more likely that a drill-and-skill approach will show no benefits and may even appear to be harmful.)[1]

2. For various reasons that I’ve reviewed elsewhere, there’s reason to doubt that requiring children to do homework has any meaningful academic benefit. In elementary school in particular, there isn’t even a correlation between doing homework (vs. doing none) or doing more homework (vs. doing less), on the one hand, and any measures of achievement—even such conventional (and, I believe, dubious) measures as grades or standardized test scores. But one prominent researcher—who does place stock in these measures—noticed something interesting when he reviewed 48 comparisons in 17 published reports of research projects that had lasted anywhere from two to thirty weeks: The longer the duration of the study, the less impact that homework had.[2]

This researcher speculated that less homework may have been assigned during any given week in the longer-lasting studies, but he offered no evidence that this was true. So here’s another theory: The studies finding the greatest effect were those that captured less of what goes on in the real world by virtue of being so brief. View a small, unrepresentative slice of a child’s life and it may appear that homework makes a contribution to school achievement; keep watching and that contribution is eventually revealed to be illusory.
3. Many people who are concerned with promoting healthy lifestyles assume that it makes sense to offer an incentive for losing weight, quitting smoking, or going to the gym. The only real question on this view is how to manage the details of the reward program. In an experiment published in 2008, people who received either of two types of incentives lost more weight after about four months than did those in the control group. (Unfortunately, there was no non-incentive weight-loss program; subjects got either money or no help at all.) At the seven-month mark, however, the effect melted away even if the pounds didn’t. There was no statistically significant weight difference between those in either of the incentive conditions and those who received nothing. This result, by the way, is typical of what just about all studies of weight loss and smoking cessation have found: The longer you look, the less chance that rewards will do any good — and they may actually do harm.[3]

4. Belief in the value of rewards is, if anything, even stronger in the corporate world, where it’s widely believed — indeed, taken on faith — that dangling financial incentives in front of employees will cause them to work harder. Conversely, if workers are provided with such an incentive and it’s then removed, their productivity would be expected to decline. An unusual occurrence in a manufacturing company provided a real-world opportunity to test this assumption: A new collective bargaining agreement for a group of welders resulted in the sudden elimination of a long-standing incentive plan. The immediate result was that production did indeed drop. But as with the preschool study, this researcher decided to continue tracking the company records — and discovered that, in the absence of rewards, the welders’ production soon began to rise and eventually reached a level as high or higher than it had been before.[4]

5. Sometimes a different result emerges when a new study is done better as opposed to merely lasting longer. The topic of
homework provides a striking example. One of the most frequently cited investigations in the field was published in the early 1980s by a researcher named Timothy Keith, who looked at survey results from tens of thousands of high school students and concluded that homework had a positive relationship to achievement, at least at that age. But ten years later, he and a colleague took a closer look – this time considering homework alongside other possible influences on learning such as quality of instruction, motivation, and which classes the students took. When all these variables were entered into the equation simultaneously, the result was “puzzling and surprising”: Homework no longer had any meaningful effect on achievement at all, even in high school.[5]

6. Finally, what happens when a second researcher comes along and does a study that’s both longer and better than the original? Consider a report published in 2004 that showed third and fourth graders who received “an extreme type of direct instruction [in a science unit] in which the goals, the materials, the examples, the explanations, and the pace of instruction [were] all teacher controlled” did better than their classmates who were allowed to design their own procedures. Frankly, the way they had set up the latter condition wasn’t representative of the strategies most experts recommend for promoting discovery and exploration. Nevertheless, the finding may have given pause to progressive educators – at least in the context of elementary school science teaching.

Or, rather, it may have given them pause for three years. That’s how much time passed before another study was published that investigated the same issue in the same discipline for kids of the same age. The two differences: the second study looked at the effects six months later instead of only one week later; and the second study used a more sophisticated type of assessment of the students’ learning.
Sure enough, it turned out that any advantage of direct instruction disappeared over time. And on one of the measures, pure exploration not only proved more impressive than direct instruction but also more impressive than a combination of the two — which suggests that direct instruction can be not merely ineffective but positively counterproductive.[6]

Despite their diversity, these six sets of studies hardly exhaust the universe of research that forces a reevaluation of what came before. Still, any observer willing to connect the dots may end up not only waiting for replications to be performed before accepting any preliminary conclusion — a reasonable posture in general — but more skeptical of studies that seem to support traditional practices in particular.

NOTES


