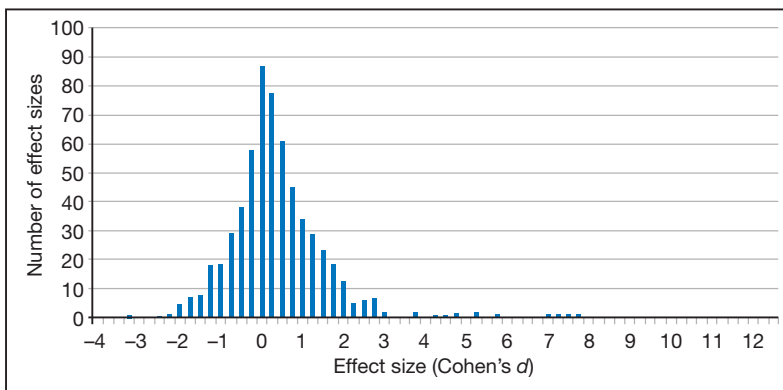


tion, and an effect size of  $-2$  would take an average person down to the bottom 2 percent. The effect sizes that Kluger and DeNisi found are shown in Figure 5.1.

Perhaps the most surprising thing about Figure 5.1 is that although there are some extreme values that can probably be discounted—it is hard to imagine what an effect size of 12 might mean, but it is almost certainly a spurious result—most of the values cluster quite closely around zero. The average effect size was 0.41, which equates to raising the performance of an average student up to the 66th percentile. Perhaps even more surprisingly, *38 percent of the effect sizes were negative*.

This is one of the most counterintuitive results in all of psychology. In each of the studies reviewed, feedback had been intended to improve performance, but in almost two out of every five experiments, feedback didn't just have no effect; it actually made things worse. The participants in the experiments would actually have done better if those providing the feedback had not given the feedback.

**Figure 5.1:** Bar Chart of Effect Sizes Found by Kluger and DeNisi (1996)



In a more recent study, focusing specifically on feedback in mathematics, science, and technology education, Ruiz-Primo and Li (2013) identify 9,000 potentially relevant studies, but after careful review, find that only 111 of the papers contain actual empirical evidence about the effects of feedback (with sixty, thirty-five, and sixteen focusing primarily on mathematics, science, and engineering/technology education, respectively). Like Kluger and DeNisi, Ruiz-Primo and Li find that feedback is generally effective, but the range is again wide. But more importantly, they conclude that hardly any of the studies they review are relevant to classroom practice. One “stunning” finding (their words, not ours) is that in over three-fourths of the studies, feedback is a single event, lasting minutes, with the whole experiment—teaching and testing the participants, giving feedback, testing the participants again—taking place in a single session of no more than three hours. There is not one single example of an experiment that looks at the effect of feedback on student achievement weeks, let alone months, later.