

## Reply to Jonah Lehrer's "On Bad Reviews"

Christopher Chabris

5/13/12

I would like to thank Jonah Lehrer for commenting on my review of his book. Further discussion of our differences could be illuminating, so I am replying to him here. I will focus initially on Lehrer's comments on the central points in my critique, and return at the end to our disagreements over matters of fact.

First, a couple of preliminaries. I disagree with Lehrer's premise that my review is "a larger critique of [his] popular science writing." I have not read most of his writing, and in the review at issue, I addressed only the contents of *Imagine*. I also have no quarrel with the "story-study-lesson" format Lehrer adopts. I use it myself. As I said in my review, it is a "proven formula." This formula is so widely accepted as the best way to communicate science to the non-specialist reader that a book that doesn't follow it has to be a very special book to succeed (e.g., Daniel Kahneman's *Thinking, Fast and Slow* or Nassim Taleb's *The Black Swan*). A danger of story-study-lesson, however, is that the link between the story and the lesson might seem to be stronger than it is, merely because it involves some kind of scientific study. The study could be designed well or poorly; it could be relevant or irrelevant; it could be a robustly replicated or preliminary experiment; and its results could be interpreted reasonably or unreasonably. On such distinctions hangs the validity of the author's conclusions.

One could accuse me of harping on minutiae that are of merely academic interest in this discussion of *Imagine*—of missing the forest for the trees. But forests are made up of trees, and in the forest of science, the trees are the facts and inferences that we are justified in taking from studies and data. Lehrer starts his reply by bemoaning my failure to "engage with the ideas of *Imagine*." This was a deliberate choice on my part. Even if a book's ideas are interesting, when it purports to deliver the verdict of science, we should engage with those ideas only to the extent that they are supported by the evidence. That's the part of *Imagine* that I did engage with.

### Reaching Conclusions

The central part of my critique of *Imagine* concerned how it presents and interprets scientific evidence. Consider the Allen study of the relationship between workplace conversations and productivity, which I claim does not demonstrate that conversations *cause* productivity. Lehrer asks me to read Allen's book, which was his source for the original study. I had read the relevant parts of that book already, and I found nothing that tests causality or rules out the most plausible alternative explanation (that productivity leads to conversation). Despite its small sample size (16 members of 6 project teams), it seems like a valuable and well-conducted observational/correlational study—but this does not mean we should draw causal conclusions from it.

Lehrer appeals to converging evidence in support of the point about creativity that he uses the Allen study to make. However, the fact that other cleverly designed studies find similar correlations between communication and performance (traders) or between proximity and performance (scientists) does not show that these relationships (let alone the one in Allen's data) are causal. Aggregating findings of correlation does not alchemize them into evidence of causation. Regardless, in *Imagine*, Lehrer does appear to draw his conclusion that conversation causes productivity directly from the Allen study. The relevant sentence reads (p. 153): "According to Allen's data, office conversations are so powerful that simply increasing their quantity can dramatically increase creative production; people have more new ideas when they talk with more people." The *according to Allen's data* preamble unmistakably suggests that Lehrer has drawn a causal conclusion from a correlational study.

The situation is similar regarding Lehrer's point about fronto-temporal dementia and creativity. On p. 108 he writes, "Nevertheless, this awful affliction comes with an uplifting moral, which is that all of us contain a vast reservoir of untapped creativity." It's true that elsewhere in the same chapter Lehrer discusses TMS studies and other research on increasing creativity. But it is the dementia finding that he says provides the moral for "all of us" about tapping our potential. The fact that other lines of evidence support the notion of a creative reserve does not mean that the fronto-temporal dementia studies themselves support that notion.

I mentioned in my review that causal overreach is commonplace in *Imagine*, but did not have space to detail more than one example. Since I see it as a critical point, let me provide another one here. Remarking on a study that found higher levels of creative achievement among students with ADHD than those without ADHD, Lehrer says, "Their attention deficit turned out to be a creative blessing." (p. 35) Writing about the same finding in his *Wall Street Journal* column last year, Lehrer put it this way: "Their inability to focus turned out to be a creative advantage." The only way to conclude that ADHD is a blessing or advantage with respect to creativity is by assuming that the ADHD *caused* the increased creativity. In fact there can be no evidence for such a conclusion, because subjects cannot be randomly assigned to have ADHD or not. One could try to isolate the statistical effect of ADHD by controlling for other personality traits or cognitive abilities that might explain differences in creative achievement, but the study under discussion did not do that. And although Lehrer describes it as involving "a large sample of undergraduates" who completed tests of creativity, in fact only 60 students (30 in the ADHD group, and 30 in a control group) did the tests. The results of this study, provocative as they are, do not provide strong evidence that ADHD and creativity are even associated in the first place, and they come nowhere close to demonstrating a causal path from distractibility to creativity—the interpretation Lehrer gives them. Again, converging evidence from other similar studies cannot solve this problem.

The problems with Lehrer's presentation of these studies are symptomatic of a bigger issue: *Imagine* repeatedly takes research results at face value and fails to grapple with the limitations of the studies and their implications. Causality is hard to show and should not be assumed when the right experiments haven't been done. In social science particularly,

reverse causality (in the Allen study, the possibility that creativity caused increased conversations) or third-variable causality (the possibility that some third factor independently causes both of the things that we observe to be correlated) are almost always plausible alternatives to the more intuitive or preferred “forward” causal explanation. The possible confounds are many, and there is a human bias to interpret facts and evidence in the way that is most consistent with the story you want to tell. Science writers owe their readers a sophisticated approach to a question as central as what we can properly conclude—and what conclusions we should avoid jumping to—on the basis of scientific research. Lehrer seems to suggest that such discussion amounts to mere detail that obscures understanding. I think that because the discovery of valid, reproducible cause and effect relationships is at the heart of the scientific enterprise, discussing such issues can lead to the most important kind of understanding.

### **The Weight of Evidence**

I fear that Lehrer and I will have to agree to disagree regarding the “blue screen” study and the value of discussing “replication history.” First, even if this study were the last word on color priming and creativity, it is a leap from experiments that displayed questions on blue computer screens to the conclusion that “being surrounded by blue walls makes us more creative” (p. 51). In his response, Lehrer writes that “the scientists argue that, because blue is typically associated with the ocean and sky, it primes the mind to think in more open and expansive ways.” But a researcher’s explanation of how an experimental finding could come about does not constitute evidence that the finding is correct. Such claims of priming are controversial in their own right. They could well prove robust, but they shouldn’t make anyone call in the painters just yet.

More broadly, given the lack of incentives for researchers to conduct “mere” replications and the bias against publishing non-replications, the right question is not whether a study “has not been contradicted” but whether it *has been consistently replicated* by independent investigators (preferably multiple independent investigators). Calling such questions ones of “replication history” trivializes as a mere historical or technical dispute a matter that is—like the correct interpretation of observational studies—in fact a core issue in proper scientific thinking and good scientific communication. This applies most strongly to phenomena like color priming, where the findings are surprising and controversial in light of what is already known about the relevant cognitive mechanisms.

Lehrer claims that he gives frequent attention to replication, but in *Imagine* I could find only two uses of the word itself, and neither referred to efforts to reproduce scientific experiments and results. The clearest mention of the concept that I could find occurs when Lehrer critiques traditional brainstorming methods on pp. 158–159. He quotes the psychologist Keith Sawyer: “Decades of research have consistently shown that brainstorming groups think of far fewer ideas than the same number of people who work alone and later pool their ideas.” In this case, the lesson works well because the “study” part of the story-study-lesson frame is indeed based on decades of research. This may seem

like a minor point, but the the consistent conclusion of decades of research justifies a much stronger lesson than does the surprising conclusion of one recent article.

## Matters of Fact

Now back to the matters of fact. First, Lehrer says that when he described Bridgeport as wealthy (p. 189), he was referring to it not as a city, but as a “metropolitan region,” because the study he was discussing used such regions as its unit of analysis. Apparently this particular region includes Darien, the money management enclave of Greenwich, and other extremely wealthy parts of Connecticut, along with impoverished Bridgeport proper. It was not clear to me in the book’s text that Bridgeport referred not to the city itself but mostly to the surrounding counties and towns. However, Lehrer makes a fair enough point; I should have read this section more charitably before accusing him of getting it wrong.

Of the specific factual errors I pointed out in my review, Lehrer admits to two (regarding EEG and the anatomy of the visual system), but says that my “other ‘corrections’ are incorrect.” Besides Bridgeport, the only other correction he corrects concerns the remote associates test; more on that shortly. I must note, however, that Lehrer says nothing at all about the mistakes I flagged on COMT/dopamine and the Apple I’s memory capacity, so he leaves the impression that *Imagine* contains just two errors. Actually, there are more, but the space limitations of my review forced me to select a subset. So as not to leave the impression that the errors I mentioned there were the only ones I found, I will list a few more here:

- Functional magnetic resonance imaging (fMRI) is easily disturbed by head movement, but not because it involves “giant superconducting magnets” (p. 90)—the size or technology of the magnets has nothing to do with the danger of motion artifact.
- Lehrer says that activity in the medial prefrontal cortex indicates “self-expression” and “a kind of storytelling” (pp. 90–91), but such an inference is not consistent with the current state of knowledge about the function of such large cortical areas. Activity in most brain areas has been connected with multiple mental states, so picking just one is inappropriate.
- According to Lehrer, Benzedrine (W.H. Auden’s drug of choice for augmenting his creativity) affects behavior within five minutes of being swallowed because it penetrates the blood/brain barrier (p. 57). But the ability of a drug to penetrate the barrier is merely a prerequisite for *any* psychoactive effect. The fact that it penetrates the barrier does not determine the delay between the administration and the effect of the substance. (Prozac and similar drugs also penetrate the blood/brain barrier, but take weeks to affect depression symptoms; powder and crack cocaine both penetrate the blood/brain barrier, but their effects have different onset times.)
- Amphetamines increase working memory, says Lehrer (p. 62), but he gives no reference for this claim. The most recent review of this research that I could find states that the question is not yet settled (Smith & Farah, *Psychological Bulletin*, 2011).

- In the section on the creativity of cities, Lehrer says that if the number of patents generated by a city is a function of that city's population raised to the power of 1.15, then doubling the population will increase patents by 15%. This is incorrect; the correct value is approximately 11%.

Finally, is the Remote Associates Test (RAT) a measure of divergent thinking (as Lehrer says in *Imagine*) or convergent thinking (as I claim)? Perhaps we must agree to disagree on this as well. The relevant sentence in the book says that divergent thinking "is the kind of thinking that's essential when struggling with a remote associate problem." While the generation of multiple possible solutions may be a part of what's going on in the RAT, I stand by the claim that the test is best thought of as measuring convergent thinking. That is, if one had to use the RAT as an indicator of one or the other form of creative thinking, convergent would be the best choice. Many tests require the generation of multiple possible solutions; what's special about the RAT is that it requires a highly non-obvious solution that uniquely relates to all of the words provided to the subject—in other words, it requires a solution on which they all converge.

Perhaps Jonah Lehrer and I see the point of science writing differently. To my mind, properly interpreting and judiciously weighing the evidence of research studies is not just sweating the details—it is the least the writer can do. Helping his readers understand how to do this for themselves—how to appreciate the limits as well as the reach of science—should also be part of the mission. Lehrer has the ability to understand a broad range of complex scientific topics and the ability to communicate ideas in an engaging way. I hope he uses his skills to give his vast audience a deeper and more nuanced understanding of his future subjects than he did in *Imagine*.