

## When Can Knowing More Lead to Worse Decisions?

Michael Smithson, The Australian National University

Posted 30/1/2011 on

[http://www.bestthinking.com/thinkers/science/social\\_sciences/psychology/michael-smithson?tab=blog](http://www.bestthinking.com/thinkers/science/social_sciences/psychology/michael-smithson?tab=blog)

Among the sacred cows of Western culture is the notion that the more information and knowledge we have, the better decisions we'll make. I'm a subscriber to this notion too; after all, I'm in the education and research business! Most of us have been rewarded throughout our lives for knowing things, for being informed. Possessing more information understandably makes us more confident about our decisions and predictions. There also is good experimental evidence that we dislike having to make decisions in situations where we lack information, and we dislike it even more if we're up against an opponent who knows more.

Nevertheless, our intuitions about the disadvantages of ignorance can lead us astray in important ways. Not all information is worth having, and there are situations where "more is worse." I'm not going to bother with the obvious cases, such as information that paralyses us emotionally, disinformation, or sheer information overload. Instead, I'd like to stay with situations where there isn't excessive information, we're quite capable of processing it and acting on it, and its content is valid. Those are the conditions that can trip us up in sneaky ways.

An intriguing [2007](#) paper by Crystal Hall, Lynn Ariss and Alexander Todorov presented an experiment in which half their participants were given statistical information (halftime score and win-lose record) about opposing basketball teams in American NBA games and asked to predict the game outcomes. The other half were given the same statistical information plus the team names. Basketball fans' confidence in their predictions was higher when they had this additional knowledge. However, knowing the team names also caused them to undervalue the statistical information, resulting in less accurate predictions. In short, the additional information distracted them away from the useful information.

Many of us believe experts make better decisions than novices in their domain of expertise because their expertise enables them to take more information into account. However, closer examination of this intuition via studies comparing expert and novice decision making reveals a counterintuitive tendency for experts to actually use *less* information than novices, especially under time pressure or when there is a large amount of information to sift through. Mary Omodei and her colleagues' [chapter](#) in a 2005 [book](#) on professionals' decision making presented evidence on this issues. They concluded that experts know which information is important and which can be ignored, whereas novices try to take it all on board and get delayed or even swamped as a consequence.

Stephan Lewandowsky and his colleagues and students found evidence that even expert knowledge isn't always self-consistent. Again, seemingly relevant information is the culprit. Lewandowsky and Kirsner ([2000](#)) asked experienced wildfire commanders to predict the spread of simulated wildfires. There are two primary relevant variables: Wind velocity and the slope of the terrain. In general, fires tend to spread uphill and with the wind. Given a downhill wind, a sufficiently strong wind pushes the fire downhill with it, otherwise the fire spreads uphill against the wind.

But it turned out that the experts' predictions under these circumstances depended on an additional piece of information. If it was a wildfire to be brought under control, experts expected it to spread downhill with the wind. If an identical fire was presented as a back burn

(i.e., lit by the fire-fighters themselves) experts predicted the reverse, that the fire would spread uphill against the wind. Of course, this is ridiculous: The fire doesn't know who lit it. Lewandowsky's group reproduced this phenomenon in the lab and named it *knowledge partitioning*, whereby people learn two opposing conclusions from the same data, each triggered by an irrelevant contextual cue that they mistake for additional knowledge.

Still, knowing more increases the chances you'll make the right choices, right? About 15 years ago Peter Ayton, an English professor visiting a Turkish university, had the distinctly odd idea of getting Turkish students to predict 32 English FA cup third-round match winners. After all, the Turkish students knew very little about English soccer. To his surprise, not only did the Turkish students do better than chance (63%), they did about as well as a much better-informed sample of English students (66%).

How did the Turkish students manage it? They were using what has come to be called the *recognition heuristic*: If they recognized one team name or its city of origin but not the other, in 95% of the cases they predicted the recognized team would win. If they recognized both teams, some of them applied what they knew about the teams to decide between them. Otherwise, they guessed.

So, how could the recognition heuristic do better than chance? The teams or team cities that the Turkish students recognized were more likely than the other teams to appear in sporting news because they were the more successful teams. So the more successful the team, the more likely it would be one of those recognized by the Turkish students. In other words, the recognition cue was strongly correlated with the FA match outcome.

Many of the more knowledgeable English students, on the other hand, recognized all of the teams. They couldn't use a recognition cue but instead had to rely on knowledge cues, things they knew about the teams. How could the recognition cue do as well as the knowledge-based cues? An obvious possible explanation is that the recognition cue was more strongly correlated with the FA match outcomes than the knowledge cues were. This was the favored explanation for some time, and I'll return to it shortly.

In two classic papers ([1999](#) and [2002](#)) Dan Goldstein and Gerd Gigerenzer presented several empirical demonstrations like Ayton's. For instance, a sample of American undergraduates did about as well (71.4% average accuracy) at picking which of two German cities has the larger population as they did at choosing between pairs of American cities (71.1% average accuracy), despite knowing much more about the latter.

It gets worse. An earlier study by Hoffrage in his 1995 PhD dissertation had found that a sample of German students actually did better on this task with American than German cities. Goldstein and Gigerenzer also reported that about two thirds of an American sample responded correctly when asked which of two cities, San Diego or San Antonio, is the largest whereas 100% of a German sample got it right. Only about a third of the Germans recognized San Antonio. So not only is it possible for less knowledgeable people to do about as well as their more knowledgeable counterparts in decisions such as these, they may even do better. The phenomenon of more ignorant people outperforming more knowledgeable ones on decisions such as which of two cities is the more populous became known as the "less-is-more" effect.

And it can get even worse than that. A [2007](#) paper by Tim Pleskac produced simulation studies showing that it is possible for imperfect recognition to produce a less-is-more effect as well. So an ignoramus with fallible recognition memory could outperform a know-it-all with perfect memory.

For those of us who believe that more information is required for better decisions, the less-is-more effect is downright perturbing. Understandably, it has generated a cottage-industry of research and debate, mainly devoted to two questions: To what extent does it occur and under what conditions could it occur?

I became interested in the second question when I first read the Goldstein-Gigerenzer paper. One of their chief claims, backed by a mathematical proof by Goldstein, was that if the recognition cue is more strongly correlated than the knowledge cues with the outcome variable (e.g., population of a city) then the less-is-more effect will occur. This claim and the proof were prefaced with an assumption that the recognition cue correlation remains constant no matter how many cities are recognized.

What if this assumption is relaxed? My curiosity was piqued because I'd found that the assumption often was false (other researchers have confirmed this). When it was false, I could find examples of a less-is-more effect even when the recognition cue correlation was *less* than that of the knowledge cue. How could the recognition cue be outperforming the knowledge cue when it's a worse cue?

In August 2009 I was visiting Oxford to work with two colleagues there, and through their generosity I was housed at Corpus Christi College. During the quiet nights in my room I tunnelled my way toward understanding how the less-is-more effect works. In a nutshell, here's the main part of what I found (those who want the whole story in all its gory technicalities can find it [here](#)).

We'll compare an ignoramus (someone who recognizes only some of the cities) with a know-it-all who recognizes all of them. Let's assume both are using the same knowledge cues about the cities they recognize in common. There are three kinds of comparison pairs: Both cities are recognized by the ignoramus, only one is recognized, and neither is recognized.

In the first kind the ignoramus and know-it-all perform equally well because they're using the same knowledge cues. In the second kind the ignoramus uses the recognition cue whereas the know-it-all uses the knowledge cues. In the third kind the ignoramus flips a coin whereas the know-it-all uses the knowledge cues. Assuming that the knowledge-cue accuracy for these pairs is higher than coin-flipping, the know-it-all will outperform the ignoramus in comparisons between unrecognized cities. Therefore, the only kind of comparison where the ignoramus can outperform the know-it-all is recognized vs unrecognized cities. This is where the recognition cue has to beat the knowledge cues, and it has to do so by a margin sufficient to make up for the coin-flip-vs-knowledge cue deficit.

It turns out that, in principle, the recognition cue can be so much better than the knowledge cues in comparisons between recognized and unrecognized cities that we get a less-is-more effect even though, overall, the recognition cue is worse than the knowledge cues. But could this happen in real life? Or is it so rare that you'd be unlikely to ever encounter it? Well, my simulation studies suggest that it may not be rare, and at least one researcher has informally communicated empirical evidence of its occurrence.

Taking into account all of the evidence thus far (which is much more than I've covered here), the less-is-more effect can occur even when the recognition cue is not, on average, as good as knowledge cues. Mind you, the requisite conditions don't arise so often as to justify mass insurrection among students or object surrender by their teachers. Knowing more still is the best bet. Nevertheless, we have here some sobering lessons for those of us who think that more information or knowledge is an unalloyed good. It ain't necessarily so.

To close off, here's a teaser for you math freaks out there. One of the results in my paper is that the order in which we learn to recognize the elements in a finite set (be it soccer teams,

cities,...) influences how well the recognition cue will work. For every such set there is at least one ordering of the items that will maximize the average performance of this cue as we learn the items one by one. There may be an algorithm for finding this order, but so far I haven't figured it out. Any takers?