The Mirage of Mount Stupid: Diving and the Dunning-Kruger Effect

Unskilled and unaware of it – references to the Dunning-Kruger effect are popular in the dive community. However, the findings of the original research are commonly misunderstood or misrepresented. On top of that, the effect itself may not even be real. A (somewhat) serious deep dive by *Tim Blömeke.*

Florida Man. The Darwin Awards. Some nitwit watched a video on rocketry and blew his garage to smithereens in an attempt to reach escape velocity. Or tried to make his own whisky, with identical results. The stories are interchangeable, but the moral is always the same: You don't know what you don't know, and a little knowledge is a dangerous thing. Ka-boom.

The world of diving has its own trove of stories of this type, many of them revolving around bad things happening in caves to divers who were only trained to dive in the ocean. Some are embellished, fictional accounts, but there is no shortage of true ones. Accident reports from the early days of cave exploration, or nonfiction books like Robert Kurson's *Shadow Divers*, make for some educational (and gruesome) reading.

In 1999, a pair of researchers named David Dunning and Justin Kruger published a paper[1] that provided these anecdotes with a scientific backdrop. The researchers conducted tests in which they asked participants of various ability levels to predict their performance in a series of academic exams. In a nutshell, Dunning and Kruger concluded that people of low ability systematically overestimate their performance, while those of higher ability tend to predict it more accurately and even slightly underestimate it.

The Internet picked up on their study and processed it into memes, going so far as to give names to specific terrain features in a fancifully shaped curve purporting to show the rise, fall, and rise-again of one's self-confidence along the course of one's learning journey: Mount Stupid, the Valley of Despair, the Slope of Enlightenment, the Plateau of Sustainability.

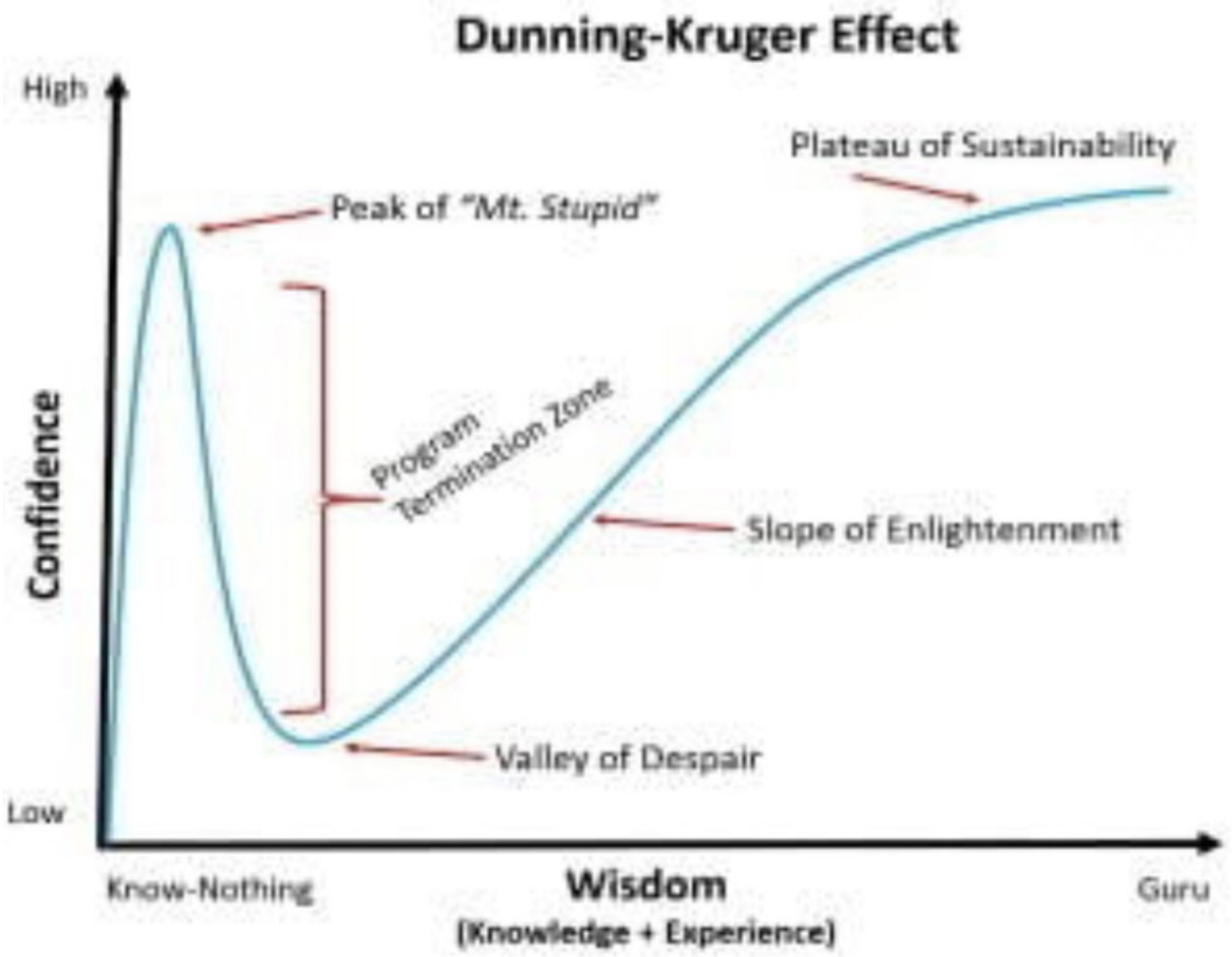

Fig. 1: You have probably seen something like this. It's not an accurate representation of what Dunning and Kruger said.

The memes were so successful that they even found their way into corporate consulting and management training programs. As a result, there are quite a few people with expensive educations who accept them as an accurate illustration of a real-world phenomenon. Many of us will have seen similar memes in the context of diving, at one point or another.

Memes spread not because they are true, but because they are appealing, and the one above does a great job of that by reminding us of the juicy anecdotes we all like to hear. Everybody has a story of "that guy" (and let's face it, it's mostly males who wind up as the protagonists in tales of grand stupidity). However, a little bit of applied skepticism quickly reveals why we absolutely shouldn't trust in personal experience when it comes to evaluating empirical claims.

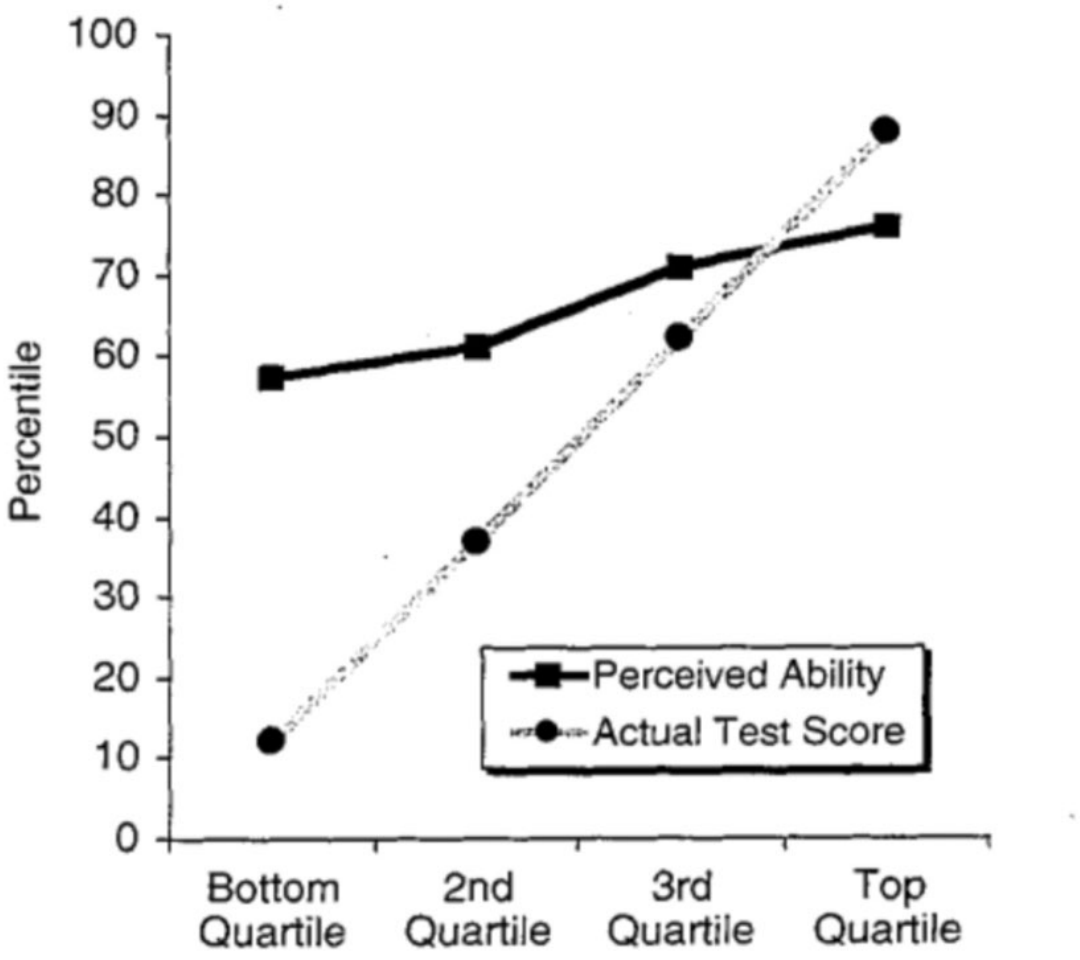

Fig. 2: This graph shows Dunning and Kruger's original findings. Mount Stupid is notably absent.

One big problem with relying on experience is that our information input is skewed. Overconfidence can produce spectacularly memorable outcomes, while lack of confidence rarely generates attention of any kind. Everybody's heard of Bob, the open water diver who entered a cave with a converted fire extinguisher for a scuba tank. He made international headlines when his body was found. Nobody ever heard of his classmate Alice, who underestimated her abilities to the point of (sadly) never diving again. The Bobs of this world become part of what we call our experience, while the Alices are quickly forgotten.

As someone who has harbored these thoughts for a while, I was excited to learn that not only the vulgar understanding of Dunning-Kruger but also their core claim has come under considerable fire within the scientific community,[2] culminating in a March 2022 cover feature in The Psychologist, with a rebuttal by David Dunning in the following issue. Criticisms center around the thesis that the effect Dunning and Kruger found is not a feature of human psychology, but instead a statistical artifact inadvertently created by the way the researchers set up their experiment and evaluated their data.

A relatively accessible (and elegant) version of this critique was published by Canadian economist Blair Fix in a blog post titled "The Dunning-Kruger Effect is Autocorrelation" (April 2022).

> "The Dunning-Kruger effect also emerges from data in which it shouldn't. For instance, if you carefully craft random data so that it does not contain a Dunning-Kruger effect, you will still find the effect. The reason turns out to be embarrassingly simple: the Dunning-Kruger effect has nothing to do with human

psychology. It is a statistical artifact — a stunning example of autocorrelation.

[…] The line labeled 'actual test score' plots the average percentile of each quartile's test score (a mouthful, I know). Things seem fine, until we realize that Dunning and Kruger are essentially plotting test score (x) against itself."

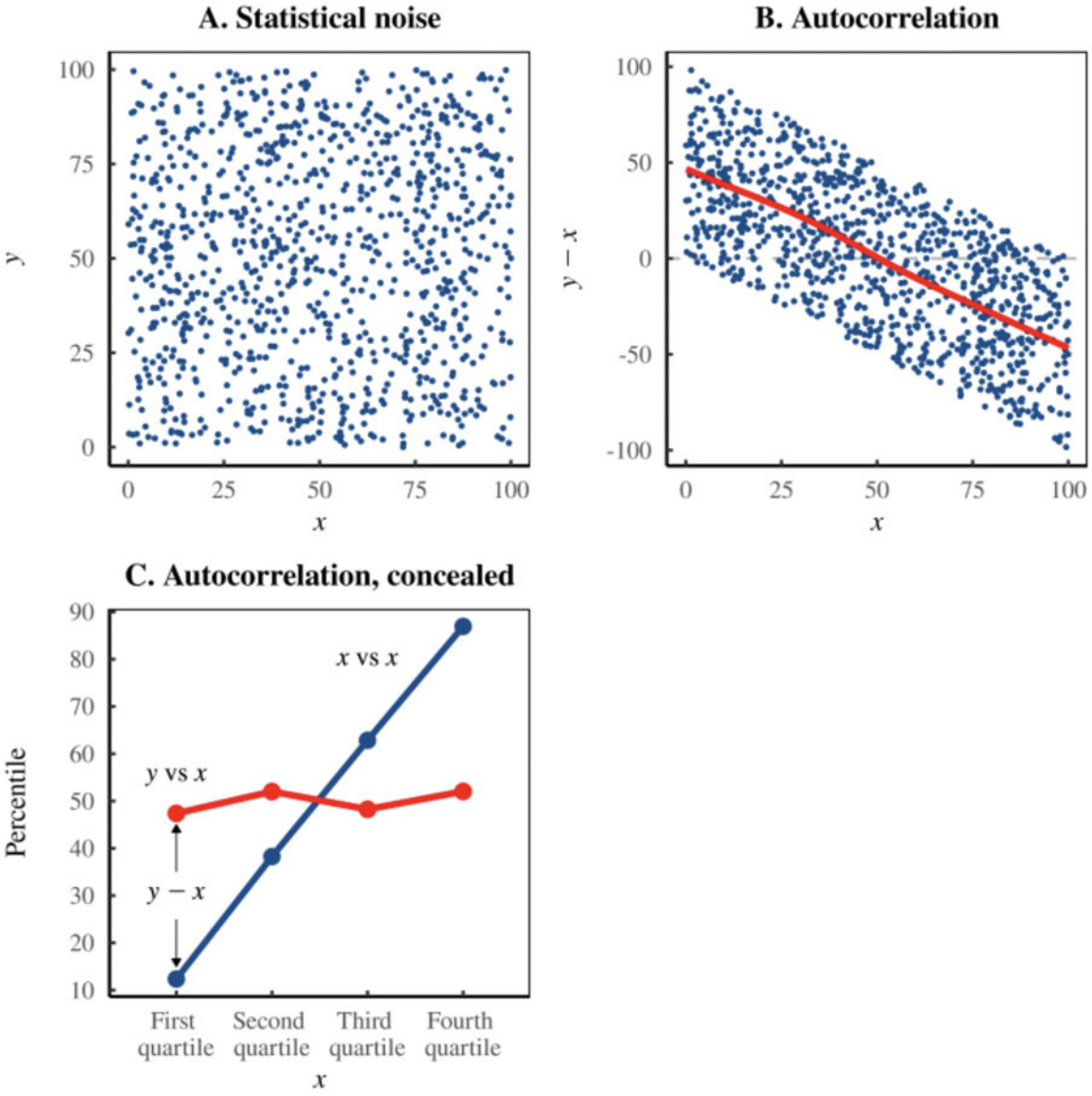

Fig. 3: A cloud of random data, and the same data after applying autocorrelation.
Source: "The Dunning-Kruger Effect is Autocorrelation"

My initial excitement was promptly followed by a sobering realization: I understand enough mathematics to find the argument compelling, but not enough to verify it. The fact that Fix's and others' criticisms confirmed something that I *wanted to believe anyway* didn't help: It could be just the same kind of cognitive trap that leads people into accepting the distorted, meme-ified version of Dunning-Kruger's claims. What if, by becoming more confident that the Dunning-Kruger effect isn't a real thing, I was propelling my ignorant self right onto the summit of Mount Stupid?

I reached out for help to an expert, Dr. Stephan Boes, a senior official at the Statistical Office of North Rhine-Westphalia in Germany[3], who validated Fix's critique: "The autocorrelation is definitely there. I can't say exactly how much of the effect it explains without reviewing the data, but it looks pronounced to me.

However, there is another problem further up the line: Participants in the experiment weren't really asked how competent they think they are. They were asked to predict how well they would do *in comparison to other participants*. There are two issues with this: One is that in order to make this prediction, participants would need to know the ability level of others in the test. Another is that a competitive ranking isn't very suitable for describing the distribution of outcomes for most real-world tasks, where you typically have a few people who consistently do badly, a few people who consistently excel, and a majority with middling outcomes who might do better than their peers in one test and worse in another. The way Dunning and Kruger present their data doesn't take this into account at all."

This connects to another criticism of the Dunning-Kruger effect: In a study published in 2020, authors Gilles E. Gignac and Marcin Zajenkowski find that the *above-average syndrome* (also known as *illusory superiority*) provides a better explanation for the discrepancies between predicted and actual performance relative to others that Dunning and Kruger found. Illusory superiority describes the observation that a majority of people consider themselves smarter, more competent, better drivers etc. than the average person (which is impossible; 50% are below average by definition).

In light of this information, the idea of applying Dunning-Kruger in the context of diving seems questionable. For starters, describing the ability of divers in terms of a competitive ranking is unhelpful. It doesn't matter if you were in the top or bottom quartile of your Advanced Open Water class. What matters is that your skills are adequate for the dives that you do – absolute ability, not relative. And even if we were to ignore all that and take Dunning-Kruger at face value, there are other human factors that come into play: In a PADI seminar on risk management I once attended, the lecturer emphasized that the majority of dive accidents during training don't happen under fresh-off-the-boat instructors who think they know everything. Accidents are more frequent with experienced instructors who become complacent.

Having absorbed all this, what are we to do when our instructor or buddy casually drops a reference to Dunning-Kruger or Mount Stupid in the classroom or over a beer? We could jump to our feet and launch into a maniacal rant about how the Dunning-Kruger effect isn't what they think it is, and how we read in *Alert Diver* that the effect maybe doesn't even exist, and even if it did, how it probably wouldn't apply to diving.

However, unless you're determined to spend the rest of the evening debating the methodology of quantitative psychological studies, regression to the mean, and statistical artifacts created by graphing x vs. $(x-y)$ when x and y have the same bounded value range, a better alternative would be to interpret mentions of Dunning-Kruger not literally but figuratively: as a cultural code, the short-hand version of a cautionary tale to warn us of underestimating the difficulty of a task we are about to attempt. Even if the Dunning-Kruger effect isn't real, overconfidence certainly is, in diving and elsewhere, and it's usually more dangerous than its opposite. We should always keep that in mind.

May the slopes of your learning curve be smooth and filled with joy.

Footnotes:

[1] Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

[2] Nuhfer, Edward, Christopher Cogan, Steven Fleisher, Eric Gaze, and Karl Wirth. "Random Number

Simulations Reveal How Random Noise Affects the Measurements and Graphical Portrayals of Self-Assessed Competency." *Numeracy* 9, Iss. 1 (2016): Article 4. DOI: http://dx.doi.org/10.5038/1936-4660.9.1.4

Gilles E. Gignac, Marcin Zajenkowski, "The Dunning-Kruger effect is (mostly) a statistical artifact: Valid approaches to testing the hypothesis with individual differences data." *Intelligence*, Volume 80, 2020, 101449, ISSN 0160-2896, https://doi.org/10.1016/j.intell.2020.101449.

Robert D. McIntosh and Sergio Della Sala, "The persistent irony of the Dunning-Kruger Effect." *The Psychologist,* Journal of the British Psychological Society, vol. 35, March 2020, https://www.bps.org.uk/volume-35/march-2022/persistent-irony-dunning-kruger-effect

David Dunning, "The Dunning-Kruger effect and its discontents." *The Psychologist,* Journal of the British Psychological Society, vol. 35, March 2020, https://www.bps.org.uk/psychologist/dunning-kruger-effect-and-its-discontents

[3] Views expressed here are personal and do not represent the opinion of Dr. Boes' employer.

About the author

Tim Blömeke teaches technical and recreational diving in Taiwan and the Philippines. He is also a freelance writer and translator, as well as a member of the editorial team of Alert Diver. For questions, comments, and inquiries, you can contact him via his blog page or on Instagram.