7-2019

# Dunning-Kruger Effect: A Simple System-Based Explanation

Griselda Acosta
*University of Texas at El Paso*, gvacosta@miners.utep.edu

Eric Smith
*University of Texas at El Paso*, esmith2@utep.edu



Comments:
Technical Report: UTEP-CS-19-68

# Dunning-Kruger Effect: A Simple System-Based Explanation

Griselda Acosta[1], Eric Smith[2], and Vladik Kreinovich[3]
[1]Department of Electrical and Computer Engineering
[2]Department of Industrial, Manufacturing, and Systems Engineering
[3]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
gvacosta@miners.utep.edu, esmith2@utep.edu, vladik@utep.edu

**Abstract**

In their 1999 paper, psychologists David Dunning and Justin Kruger showed that, in general, experts not only provide better estimates of different situations, but they also provide a better estimates of the accuracy of their estimates. Which this phenomenon has been confirmed by many follow-up experiments, it remains largely unexplained. In this paper, we provide a simple system-based qualitative explanation for the Dunning-Kruger effect.

## 1 Formulation of the Problem

**Dunning-Kruger effect: a brief reminder.** In their 1999 paper [4], Justin Kruger and David Dunning from Cornell University showed that:

- not only experts have a better knowledge and produce more accurate estimates than novices,
- experts also assess their own accuracy with better accuracy than novices;

see also [2, 8]. This phenomenon became known as the *Dunning-Kruger effect*.

*Comment.* It should be mentioned that in their original paper [4], due to faulty statistical analysis, the above claim was more specific:

- that experts usually underestimate their abilities (in particular, overestimate the inaccuracy of their estimates),

- while novices usually overestimate their ability (in particular, underestimate the inaccuracy of their estimates).

A more accurate statistical analysis has shown that this specific claim is *not* supported by the evidence. However, the above claim – that experts estimate the accuracy of their own estimates better than novices – is definitely statistically valid; see, e.g., [5, 6].

**Challenge.** The effect is there, but how can we explain it?

**What we do in this paper.** In this paper, we provide a simple system-based qualitative explanation for the Dunning-Kruger effect.

**Remaining open question and future work.** We hope that future research will help transform our qualitative explanation into a more quantitative one.

## 2 Our Explanation

**Towards formulating the problem in precise terms.** What we want from experts is an estimate of the state of the corresponding system:

- this system is a patient if this expert is a medical doctor,
- this system is a complex machine if the expert is an engineer,
- this system is a mineral deposit if the expert is a geophysicist, etc.

Most information is usually described in terms of real numbers. In this sense, what we want from an expert is to provide some estimates of the numbers describing the corresponding system.

For example, based on his/her expertise (and on other available information), a geophysicist can provide us with:

- an estimate of the amount of oil in a given oil field,
- an estimate of the depth at which this oil will most probably be found, etc.

In all these cases, we want to predict the value of the desired quantity $y$ based on the available values $x_1, \ldots, x_n$ of related quantities. In other words, we need to have some prediction algorithm

$$y = f(x_1, \ldots, x_n).$$

Depending on the previously available data, we can have different prediction algorithms. For example, if we want to predict tomorrow's weather, then:

- in some areas (e.g., areas shielded by mountain ranges), it is sufficient to take into account today's weather patterns only in the nearby areas, while

- in other places (e.g., on the plains where cyclones and other weather phenomena can travel large distances fast), we need to take into account today's weather in a much wider area.

In general, instead of a *single* algorithm

$$y = f(x_1, \ldots, x_n),$$

we have a *family* of algorithms

$$y = f(x_1, \ldots, x_n, c_1, \ldots, c_m)$$

depending on some parameters $c_1, \ldots, c_m$. These parameters $c_1, \ldots, c_m$ need to be determined based on the available data.

- In statistics [7] and in machine learning (see, e.g., [1, 3]) the corresponding family of algorithms is usually given explicitly.
- In expert decision making, a lot of the reasoning is happening at the subconscious level. So we do not have explicit expressions for the corresponding prediction algorithm – but, in effect, experts make predictions – and thus, subconsciously use some algorithms to make these predictions.

**Resulting formulation of the problem.** We have a model

$$y = f(x_1, \ldots, x_n, c_1, \ldots, c_m)$$

that describes the dependence of the desired quantity $y$ on observable quantities $x_1, \ldots, x_n$. To make a prediction, we need to estimate the values of the parameters $c_1, \ldots, c_m$ that best describe the current situation.

In reality, measurements are approximate, and models are approximate. So, in the actual dependence of $y$ on $x_1, \ldots, x_n$, we do not have the exact equality, there is also some noise $N$:

$$y = f(x_1, \ldots, x_n, c_1, \ldots, c_m) + N.$$

To estimate the values of these parameters, we can use the results of previous observations. Let $K$ denote the number of such observations. For each observation $k = 1, \ldots, K$, we know the values $x_1^{(k)}, \ldots, x_n^{(k)}$ and $y^{(k)}$ of the corresponding quantities. We want to find the values of the parameters $c_1, \ldots, c_m$ for which

$$y^{(k)} \approx f\left(x_1^{(k)}, \ldots, x_n^{(k)}, c_1, \ldots, c_m\right)$$

for all $k$.

*Comment.* This problem is ubiquitous in applications of statistics, it is known as *regression*.

- The most widely used and the most well-known is the *linear* regression, when the dependence $f(x_1, \ldots)$ on $x_1, \ldots, x_n$ is linear.

- However, nonlinear regression is also actively used; see, e.g., [7].

**How the estimate's accuracy depends on the number of observations.** The main difference between an expert and a novice is that an expert is aware of a much larger number of previous observations: the value $K$ corresponding to the expert is much larger than the number of observations corresponding to the novice. Thus, to understand the difference between estimates by experts and estimates by novices, we need to analyze how the accuracy of an estimation depends on the available number of observations $K$.

According to statistics, if we have $K$ observations with standard deviation $\sigma$, then, in general, the accuracy with which we can estimate the values of the corresponding parameters is proportional to $\frac{\sigma}{\sqrt{K}}$; see, e.g., [7].

This formula is easy to derive in the simplest situation, when we simply observe the desired quantity $y$ several times, i.e., if we have $K$ results $y^{(1)}, \ldots, y^{(K)}$ of measuring this same quantity. In this case, a reasonable way to combine these results into a single more accurate estimate is to take the average

$$\overline{y} = \frac{y^{(1)} + \ldots + y^{(K)}}{K}. \tag{1}$$

The inaccuracy of each measurement is described by the difference $\Delta y^{(k)} \stackrel{\text{def}}{=} y^{(k)} - y$ between the measurement result $y^{(k)}$ and the actual (unknown) value $y$ of the corresponding quantity. From the formula (1), we conclude that the inaccuracy $\Delta y \stackrel{\text{def}}{=} \overline{y} - y$ of the arithmetic average is equal to

$$\Delta y = \frac{\Delta y^{(1)} + \ldots + \Delta y^{(K)}}{K}.$$

Inaccuracies $\Delta y^{(k)}$ corresponding to different measurements are usually independent. For independent random variables, the variance of the sum is equal to the sum of the variances. Since each of the inaccuracies $\Delta y^{(k)}$ has the variance $\sigma^2$, the variance of the sum

$$\Delta y^{(1)} + \ldots + \Delta y^{(K)}$$

is thus equal to $K \cdot \sigma^2$. Hence, the standard deviation of the sum is equal to the square root of this variance, i.e., to $\sqrt{K} \cdot \sigma$.

When we divide a random variable by a positive constant, its standard deviation divides by the same constant. Thus, the standard deviation of the inaccuracy $\Delta y$ – corresponding to the use of $K$ observations – is equal to

$$\frac{\sqrt{K} \cdot \sigma}{K} = \frac{\sigma}{\sqrt{K}}.$$

This derivation is only valid for the simplest case, but a similar asymptotics $\frac{\sigma}{\sqrt{K}}$ holds in the general situation as well [7]. This confirms the intuitive idea

that the more experience the expert, the more accurate are this expert's estimates.

**Towards understanding the Dunning-Kruger effect: how accurately can we estimate the accuracy of our own estimates?** Accuracy estimates presented in the previous subsection depend on the standard deviation $\sigma$ of the measurement error – i.e., on the standard deviation of the differences $\Delta y^{(k)}$. Usually, we do not know this value exactly, but we can estimate it based on the results of previous observations.

For example, if we simply observe the desired quantity $y$ several times, i.e., if we have $K$ results $y^{(1)}, \ldots, y^{(K)}$ of measuring this same quantity, then we can estimate this standard deviation $\sigma$ as

$$\sigma \approx \overline{\sigma} = \sqrt{\frac{1}{K-1} \cdot \sum_{k=1}^{K} \left(y^{(k)} - \overline{y}\right)^2}.$$

Similar formulas can be used in the general case: once we have estimates $\overline{c}_1, \ldots, \overline{c}_m$ for the parameters $c_1, \ldots, c_m$, we can then estimate $\sigma$ as

$$\sigma \approx \overline{\sigma} \sqrt{\frac{1}{K-1} \cdot \sum_{k=1}^{K} \left(y^{(k)} - f\left(x_1^{(k)}, \ldots, x_n^{(k)}, \overline{c}_1, \ldots, \overline{c}_m\right)\right)^2}.$$

It is known that the relative accuracy $\dfrac{\sigma - \overline{\sigma}}{\sigma}$ of this estimate decreases with $K$ as $\dfrac{1}{\sqrt{K}}$ – at the same rate as the accuracy itself. Thus, the more experienced the expert – i.e., the larger the corresponding $K$ – the more accurately this expert can estimate the accuracy of his/her estimates.

This is exactly what the Dunning-Kruger effect is about. Thus, we have indeed found a simple explanation for this effect.

## Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

## References

[1] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.

[2] D. Dunning, "The Dunning-Krueger effect: On being ignorant of one's own ignorance.", In: J. M. Olson and M. P. Zanna (Eds.), *Advances in Experimental Social Psychology*, Vol. 44, Academic Press, San Diego, California, 2011, pp. 247–296.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Leaning*, MIT Press, Cambridge, Massachusetts, 2016.

[4] J. Kruger and D. Dunning, "Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessment", *Journal of Personality and Social Psychology*, 1999, Vol. 6, pp. 1121–1134.

[5] E. Nuhfer, C. Cogan, S. Fleischer, E. Gaze, and K. Wirth, "Random number simulations reveal how random noise affects the measurements and graphical portrayals of self-assessed competency", *Numeracy*, 2016, Vol. 9, No. 1, Paper 4.

[6] E. Nuhfer, S. Fleischer, C. Cogan, K. Wirth, and E. Gaze, "How random noise and a graphical convention subverted behavioral scientists' explanations of self-assessment data: numeracy underlies better alternatives", *Numeracy*, 2017, Vol. 10, No. 1, Paper 4.

[7] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

[8] T. Schlosser, D. Dunning, K. L. Johnson, and J. Kruger, "How unaware are the unskilled? Empirical tests of the 'signal extraction' counterexplanation for the Dunning-Kruger effect in self-evaluation of performance", *Journal of Economic Psychology*, Vol. 39, pp. 85–100.