

A statistical explanation of the Dunning–Kruger effect

October 4, 2021

Jan R. Magnus

Department of Econometrics and Data Science, Vrije Universiteit Amsterdam
and Tinbergen Institute, Amsterdam, The Netherlands

Anatoly A. Peresetsky¹

National Research University ‘Higher School of Economics’, Moscow, Russia

Abstract: An explanation of the Dunning–Kruger effect is provided which does not require any psychological explanation, because it is derived as a statistical artefact. This is achieved by specifying a simple statistical model which explicitly takes the (random) boundary constraints into account. This model fits the data perfectly.

JEL Classification: A22; C24; C91; D84; D91; I21.

Keywords: Dunning–Kruger effect; Boundary conditions; Tobit model.

¹Corresponding author. Address: Faculty of Economic Sciences, Higher School of Economics, Pokrovsky Boulevard, 11, Moscow, Russia, 109028

E-mail addresses: jan@janmagnus.nl (Magnus), aperesetsky@hse.ru (Peresetsky).

1 Introduction

The Dunning–Kruger (DK) effect states that people with low ability tend to overestimate their ability. This hypothetical cognitive bias was first described in Kruger and Dunning (1999) and, if true, it is potentially important and dangerous, because it means that people of low ability not only perform tasks poorly but (even worse) that they think that they perform these tasks well. Dunning and Kruger claim that the reason for this bias is that people of low ability are not good in seeing and judging themselves (a deficit in metacognitive skills). A closely related effect, also important but arguably less dangerous, is that people of high ability tend to underestimate their ability. This second effect, although not discussed in Kruger and Dunning (1999), is often also associated with their names. The DK effect and Dunning and Kruger’s explanation of it has been discussed and challenged extensively.

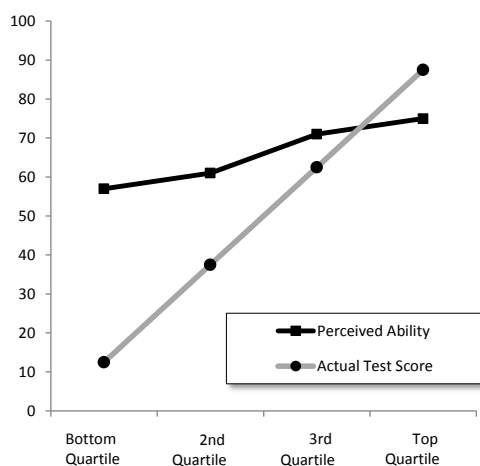


Figure 1: Perceived ability to recognize humor as a function of actual test performance (from Kruger and Dunning, 1999)

In their original paper, Kruger and Dunning (1999) tested undergraduate students enrolled in various psychology courses at Cornell University for their ability in humor, logical reasoning, and English grammar. After the test they asked the students to assess their performance in the test. The students were then split in four groups according to their actual test scores. Calculating the average perceived ability in each group, Dunning and Kruger obtained Figure 1. The accuracy of the prediction was high in the top group and low

in the bottom group, and the prediction in the bottom group was highly overestimated.

The Kruger–Dunning paper raises two questions. First, is there a DK effect? And second, is the explanation provided by Dunning and Kruger correct?

There has been both criticism and support. Most studies agree that there is a DK effect, although many disagree on Kruger and Dunning’s metacognitive explanations; see Ehrlinger et al. (2008), Schlosser et al. (2013), Williams et al. (2013), Sullivan et al. (2018), West and Eaton (2019), Gabbard and Romanelli (2021), and Mariana et al. (2021); and partial responses in Kruger and Dunning (2002), Dunning et al. (2003, 2004), and Dunning (2011).

But there has also been much criticism and this criticism typically relies on a statistical rather than a psychological explanation of the DK effect. The attack on Dunning and Kruger was initiated by Krueger and Mueller (2002), who suggested a regression better-than-average approach which is parsimonious and ‘does not require mediation by third variables, such as metacognitive insights into one’s own problem-solving abilities’.² Their approach is based on two empirical facts. First, it is well-known that people tend to overestimate their performance. Most people think they drive better than average (Svenson, 1981). In a survey of engineers, 42% thought their work ranked in the top 5% among their peers (Zenger, 1992); and in a survey of college professors, 94% thought they performed ‘above average’ (Cross, 1977). Second, the slope in the linear regression of estimated performance on actual performance is not equal but less than one. This phenomenon is called ‘regression to the mean’ and has been known since Galton (1886) studied the relationship between the height of sons and fathers. Combining these two facts leads to the regression better-than-average approach, and it explains the asymmetry of the DK effect: overestimation in the bottom quartile and underestimation in the upper quartile. A more precise formulation of the regression better-than-average approach was provided in the noise-plus-bias model (Burson et al., 2006).

Several other studies attempted to provide statistical explanations. Krajc and Ortmann (2008) assumed a nonsymmetric J -distribution for the talent of the undergraduates studied by Kruger and Dunning (1999), which leads to more students in the left tail of the students’ ability distribution, resulting in the DK effect. McIntosh et al. (2019) experimented with movement and memory tasks, and concluded that the DK effect exists as an empirical phenomenon. But they disagreed with the explanation that poor insight is

²To avoid confusion: Joachim Krueger of Brown University and Justin Kruger of Cornell University are two different people.

the reason for overestimation among the unskilled. Gignac and Zajenkowski (2020), using a sample of general community participants, tested the validity of the DK effect with the Glejser test of heteroskedasticity and by nonlinear (quadratic) regression, and found much less evidence in favor of the DK effect than Kruger and Dunning (1999). Jansen et al. (2021) replicated two of Dunning and Kruger’s studies using a sample of 4000 participants. Their model for the probability of a correct answer is an extension of the one-parameter item response theory (IRT) model, known as the Rasch model (Embretson and Reise, 2013). They showed that this model is consistent with the data.

Our explanation of the DK effect is based on the fact that the data are bounded. This feature of the data has not received much attention, with the exception of Burson et al. (2006), who concluded that the boundary restriction ‘is an important concern that should be addressed in future research’; and Krajc and Ortman (2008), who noted that students in the bottom quartile can only make optimistic errors placing themselves into a higher quartile, while students in the top quartile can only make pessimistic errors placing themselves in a lower quartile.

The remark by Krajc and Ortman provides the essence of our story. Consider a brilliant student who typically scores 95 or 99 points out of 100. Because of the bound at 100, there is not much room to predict higher than her ability but there is plenty of room to predict lower, so she would typically predict 85 or 90, thus underestimating her score. The same happens at the bottom end of the scale, where there is a bound of 0 and a student would typically overestimate. This simple observation is the basis of our model.

We shall employ data on 665 undergraduates at the International College of Economics and Finance of the Higher School of Economics in Moscow, who predict their grade on a 0–100 scale for a statistics exam. We use a simple statistical model which explicitly specifies the (random) boundary constraints. This model fits the data perfectly. There is thus no need for a psychological explanation of the DK effect: it is a statistical artefact.

The remainder of the paper is organized as follows. In Section 2 we present and discuss the data. In Section 3 we present a simple one-parameter model that accounts for censoring, and show that this model explains the DK effect, although not yet perfectly. In Section 4 we extend this simple model to a more realistic three-parameter model where the bounds are random rather than fixed, and present the results based on this extended model. The fit is now near-perfect. Section 5 concludes. A mathematical Appendix contains the statistical theory underlying the required conditional expectation functions.

2 The data

We shall study the DK effect by comparing exam results with predictions of these results, and in this section we describe the data in some detail.

The International College of Economics and Finance (ICEF) in Moscow was established in 1997 jointly by the London School of Economics and Political Science (LSE) in London and the Higher School of Economics (HSE) in Moscow. The college offers a four-year bachelor's program, which is considered to be one the top programs in economics in Russia. Each year about 200 students enter the program, typically immediately after high school. In their first year the students follow, among other subjects, a course called Statistics-1, and in their second year they follow Statistics-2. Both courses are compulsory. Our data are obtained from four cohorts of students following Statistics-2 in the period 2016–2019. In total, after removing students who took the course for a second time, 665 students remained who took this course and provided a prediction.

In Statistics-2 students take three exams every year, at the end of October (exam 1), the end of December (exam 2), and the end of March (exam 3). The exams are written exams, not multiple choice, and each exam consists of two parts (80 minutes each) with a ten minute break between the two parts. The level of the exam questions is the same in the two parts. To avoid cheating, students are not allowed to leave and come back during each part of the exam. At the end of part 1 and at the end of part 2 the examiner collects each student's work. Each part is graded out of 50 points.

At the end of the first part of each of the three exams each student is invited to predict (out of 100) their grade for this exam (the two parts together). When writing down the prediction, students know the questions and their answers in part 1, but not yet the questions of part 2. To encourage students to provide a prediction and try their best, a bonus is promised as follows. If the difference between the prediction and the grade is less than or equal to 3 in absolute value, then one bonus point is added to the grade. For example, if the prediction is 49 and the grade is 52, then the grade for this exam is marked up to 53. This procedure had to be and has been approved by the ICEF administration. As a result of the procedure and the possibility of a bonus, the response rate was extremely high (97%). The idea of giving each student an incentive to express their opinion was successfully used earlier in experiments by Blackwell (2010) and Magnus and Peresetsky (2018).

In the current study we take data only from the second exam in each year. This is the most representative of the three exams, because in the first exam students may not yet be familiar with the benefits of a careful prediction, and in the third exam there is the problem that smart (or risk averse) students

utilize the bonus to maximize the probability that their grade is ≥ 25 , which is a requirement for passing the course. The student's optimal strategy is then to choose their prediction between 21 and 27 in which case a grade of 24 would be marked up to 25. Many students actually use this strategy which leads to an overrepresentation of 24 and 25 in the sample of the third exam.

In each year t we thus have one grade and one prediction per student. Let us define

$$\begin{aligned} x_{it} &: \text{actual grade of student } i \text{ in year } t, \\ y_{it}^{raw} &: \text{raw (unadjusted) prediction of student } i \text{ in year } t, \\ d_{it} = y_{it}^{raw} - x_{it} &: \text{difference between raw prediction and actual grade.} \end{aligned}$$

In each year we can average over students and this gives

$$\bar{x}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} x_{it}, \quad \bar{y}_t^{raw} = \frac{1}{n_t} \sum_{i=1}^{n_t} y_{it}^{raw}, \quad \bar{d}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} d_{it},$$

where n_t denotes the number of students in year t .

We don't want to use the raw predictions directly, because of the variation in the student cohort's strengths and in the difficulty of the exam over the years. To filter out these variations we define an adjusted prediction

$$y_{it} = y_{it}^{raw} - \bar{d}_t \tag{1}$$

with the property that $\bar{y}_t = \bar{x}_t$, so that in each year the average prediction equals the average grade.

Year	# Students	Exam grade	Raw prediction	Difference	St. dev.
t	n_t	\bar{x}_t	\bar{y}_t^{raw}	\bar{d}_t	τ_t
2016	144	41.8	39.0	-2.76	12.1
2017	168	33.3	38.5	5.24	12.7
2018	185	41.2	37.5	-3.71	13.1
2019	168	43.0	39.3	-3.65	12.0
Total	665	39.8	38.6	-1.23	13.0

Table 1: Descriptive statistics (means) of the data

In Table 1 we present a summary of the data. Per year we provide the number of students n_t , the average exam grade \bar{x}_t , the average raw prediction \bar{y}_t^{raw} , and the difference \bar{d}_t between these two averages. The difference d_{it} varies a lot within each year, as shown by the standard deviation τ_t in the last column. The second exam in 2017 turned out particularly difficult (or the cohort was less motivated) leading to relatively low grades.

$$\sum_{i=1}^{n_t} d_{it}^2 = \sum_{i=1}^{n_t} (y_{it} - x_{it})^2 + n_t \bar{d}_t^2$$

3 Fixed bounds

As a first attempt to model the predictions we propose the following equation:

$$y_{it}^{raw} = \alpha_t + x_{it} + \zeta_{it}, \quad (2)$$

where the constant α_t may vary per year to adjust for (over)confidence and the difficulty of the exam, and the errors ζ_{it} are assumed to be independent and identically distributed (iid) as $N(0, \sigma_\zeta^2)$. Writing (2) in deviation form gives

$$y_{it}^{raw} - \bar{y}_t^{raw} = (x_{it} - \bar{x}_t) + (\zeta_{it} - \bar{\zeta}_t).$$

The adjusted prediction is given by

$$y_{it} = y_{it}^{raw} - \bar{d}_t = y_{it}^{raw} - \bar{y}_t^{raw} + \bar{x}_t,$$

which leads to the simple equation

$$y_{it} = x_{it} + \epsilon_{it}, \quad \epsilon_{it} = \zeta_{it} - \bar{\zeta}_t.$$

After adjustment, the year plays no longer any role, so we may simplify the notation and write

$$y_i = x_i + \epsilon_i. \quad (3)$$

The difference between the (adjusted) prediction y_i and the grade x_i is thus random noise, and the only thing to estimate is the variance of that noise.

This first attempt does not, however, take into account that the left-hand side of (3) is bounded by $0 \leq y_i \leq 100$, so that the right-hand side is similarly bounded. The right-hand side $x_i + \epsilon_i$ does not automatically fulfill this constraint; it has to be censored to do so. The basic censoring model in statistics and econometrics is the tobit model introduced by Tobin (1958). In the tobit model we introduce a latent (unobserved) random variable y_i^* defined as

$$y_i^* = x_i + \epsilon_i, \quad (4)$$

where the ϵ_i are independent and identically distributed (iid) as $N(0, \sigma_\epsilon^2)$. Then we model y_i as

$$y_i = \begin{cases} 0, & \text{if } y_i^* \leq 0, \\ y_i^*, & \text{if } 0 < y_i^* \leq 100, \\ 100, & \text{if } y_i^* > 100. \end{cases} \quad (5)$$

This is the standard tobit model, double-censored due to the fact that we have both a lower and an upper bound.

Model (5) is more realistic than model (3), and once we have estimated σ_ϵ^2 we can compute the expectation $h(x_i) = E(y_i)$ as given in (9) in the Appendix.

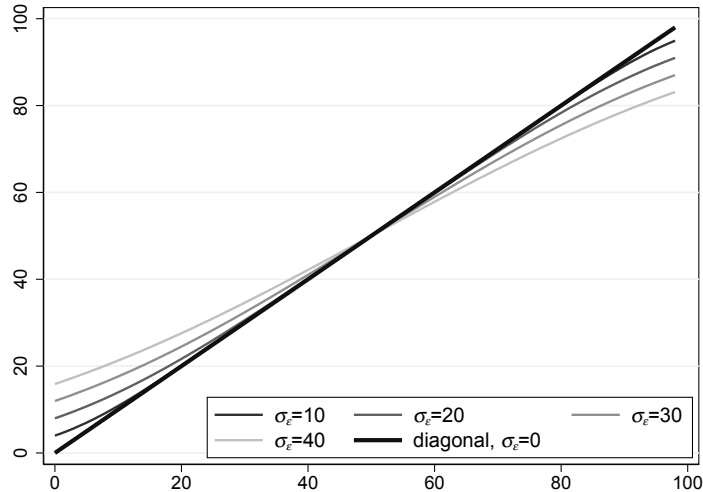


Figure 2: Expectation functions for the one-parameter censored tobit model for $\sigma_\epsilon = 0, 10, 20, 30, 40$

Estimating σ_ϵ by maximum likelihood (ML) gives $\hat{\sigma}_\epsilon = 12.5$ with standard error 0.35. In Figure 2 we plot the expectation $h(x_i) = E(y_i)$ for four values of σ_ϵ : 10, 20, 30, and 40 (and 0, which is the 45° line). Figure 2 already demonstrates how the two bounds force the expectation function in the direction of the DK effect in Figure 1.

But the fixed-bound model is not yet completely satisfactory due to the fact that it is not realistic to assume that $y_i = 0$ if $y_i^* \leq 0$ or that $y_i = 100$ if $y_i^* > 100$: no student predicts 0 (however poor) or 100 (however brilliant). This leads to the random-bounds model presented in the next section.

4 Random bounds

A more realistic model is given by

$$y_i = \begin{cases} |u_i|, & \text{if } y^* \leq 0, \\ y_i^*, & \text{if } 0 < y^* \leq 100, \\ 100 - |v_i|, & \text{if } y^* > 100, \end{cases} \quad (6)$$

where we assume that ϵ_i , u_i , and v_i are iid, independent of each other, and that all three are normally distributed as $N(0, \sigma_\epsilon^2)$, $N(0, \sigma_u^2)$, and $N(0, \sigma_v^2)$, respectively.

When applying (6) there is one further complication, namely that the lower bound must not only satisfy $|u_i| > 0$, but also $|u_i| < 100$, while the upper bound must not only satisfy $100 - |v_i| < 100$, but also $100 - |v_i| > 0$. Hence, we must require that $|u_i| < 100$ and $|v_i| < 100$. This will be ‘almost’ true in most applications. For example, we have $\Pr(|u_i| < 100) = 99.9$ and 95.5 for $\sigma_u = 30$ and 50 , respectively. We deal formally with this situation by considering the conditional expectation function

$$h(x_i) = E(y_i | 0 < y_i < 100). \quad (7)$$

We derive the mathematical expression for this conditional expectation function in the Appendix, resulting in (8).

Restriction	σ_ϵ	σ_u	σ_v	$\log L$
None	12.69 (0.40)	22.81 (2.81)	31.29 (3.38)	-2582.39
$\sigma_u = \sigma_v$	12.67 (0.39)	24.79 (2.30)	24.79 (2.30)	-2582.73

Table 2: Maximum likelihood estimates for the three-parameter model (standard errors in parentheses)

The ML estimates are presented in Table 2, first without restriction and then under the restriction that $\sigma_u = \sigma_v$. The estimates take on reasonable values and they are estimated rather precisely. The restriction $\sigma_u = \sigma_v$ is not rejected by a Wald (p -value 0.061) or likelihood ratio (p -value 0.412) test.

In Figure 3 we present the conditional expectation functions based on the ML estimates in Table 2. As expected, there is not much difference between the restricted ($\sigma_u = \sigma_v$) and the unrestricted plot, and both plots show clearly the DK effect based purely on the fact that the observations are bounded. The observed S -shape is very similar to the empirical plots reported in the literature: overestimation for the weak students (‘unskilled and unaware of it’ in the words of Dunning and Kruger) and underestimation for the strong students.

As a benchmark comparison we also provide a nonparametric plot based on locally-weighted scatterplot smoothing (lowess) with bandwidth 0.2. The ML plots are close to the nonparametric plot when x is small but not so close

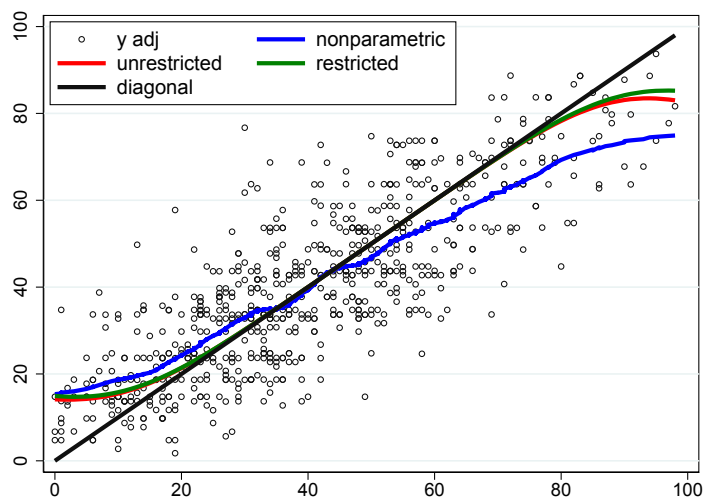


Figure 3: Conditional expectation functions for the three-parameter censored tobit model based on ML and nonparametric estimates

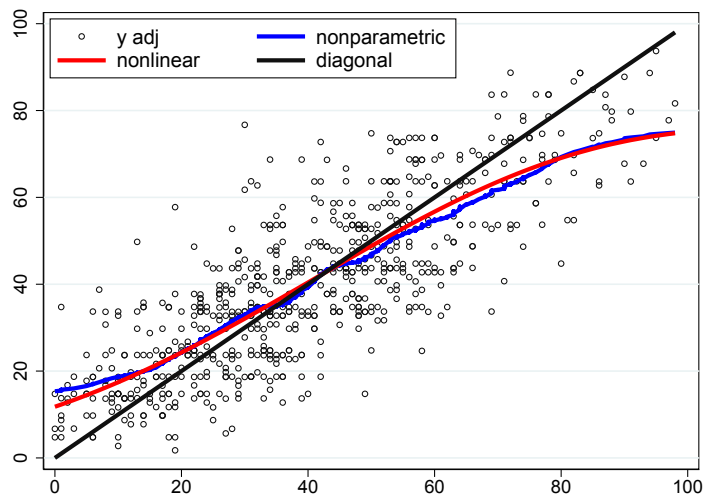


Figure 4: Conditional expectation functions for the three-parameter censored tobit model based on NLS and nonparametric estimates

when x is large. This is caused by the fact that only 14% of the observations fall in the interval $x > 60$.

We can try and fit our three parameters such that the conditional expectation function is ‘as close as possible’ to the nonparametric plot. If we employ nonlinear least-squares (NLS), then we find $\hat{\sigma}_\epsilon = 29.6 (0.70)$, $\hat{\sigma}_u = 0$,

and $\hat{\sigma}_v = 33.1$ (2.13). The NLS plot in Figure 4 is now very close to the non-parametric plot. In fact, the shape of the conditional expectation function is quite robust against changes in the three parameters. If we fix the parameters at $\sigma_\epsilon = 30$, $\sigma_u = 10$, and $\sigma_v = 35$; or at $\sigma_\epsilon = 25$, $\sigma_u = 15$, and $\sigma_v = 40$, then we obtain conditional expectations that are almost indistinguishable from Figure 4.

5 Conclusion

In this paper we have attempted to provide an explanation of the DK effect which does not require any psychological explanation. By specifying a simple statistical model which explicitly takes the (random) boundary constraints into account, we achieve a near-perfect fit, thus demonstrating that the DK effect is a statistical artefact. In other words: there is an effect, but it does not reflect human nature.

Mathematical appendix

Consider the censored model (6), where we write y^* , y , and x instead of y_i^* , y_i , and x_i for simplicity. We have

$$\begin{aligned} p_1 &= \Pr(y^* \leq 0) = \Phi\left(\frac{-x}{\sigma_\epsilon}\right), \\ p_2 &= \Pr(0 < y^* \leq 100) = 1 - p_1 - p_3, \\ p_3 &= \Pr(y^* > 100) = 1 - \Phi\left(\frac{100 - x}{\sigma_\epsilon}\right), \end{aligned}$$

and

$$\begin{aligned} q_1 &= \Pr(y \leq 0) = \Pr(y \leq 0 | y^* > 100) \cdot \Pr(y^* > 100) \\ &= \Pr(100 - \sigma_v |v| \leq 0) \cdot p_3 = 2p_3 \Phi\left(\frac{-100}{\sigma_v}\right), \\ q_2 &= \Pr(0 < y \leq 100) = 1 - q_1 - q_3, \\ q_3 &= \Pr(y > 100) = \Pr(y > 100 | y^* \leq 0) \cdot \Pr(y^* \leq 0) \\ &= \Pr(\sigma_u |u| > 100) \cdot p_1 = 2p_1 \Phi\left(\frac{-100}{\sigma_u}\right). \end{aligned}$$

Let $0 < t < 100$. Then,

$$\begin{aligned}
G_1(t) &= \Pr(0 < y < t | y^* \leq 0) = \Pr(0 < \sigma_u |u| < t) = 1 - 2\Phi\left(\frac{-t}{\sigma_u}\right), \\
G_2(t) &= \Pr(0 < y < t | 0 < y^* \leq 100) = \Pr(0 < y^* < t | 0 < y^* \leq 100) \\
&= \frac{\Pr(0 < y^* < t)}{\Pr(0 < y^* < 100)} = \frac{\Pr(0 < x + \sigma_\epsilon \epsilon < t)}{p_2} \\
&= \frac{1}{p_2} \left[\Phi\left(\frac{t-x}{\sigma_\epsilon}\right) - \Phi\left(\frac{-x}{\sigma_\epsilon}\right) \right], \\
G_3(t) &= \Pr(0 < y < t | y^* > 100) = \Pr(0 < 100 - \sigma_v |v| < t) \\
&= 2 \left[\Phi\left(\frac{100}{\sigma_v}\right) - \Phi\left(\frac{100-t}{\sigma_v}\right) \right]
\end{aligned}$$

with derivatives

$$\begin{aligned}
g_1(t) &= \frac{2}{\sigma_u} \phi\left(\frac{-t}{\sigma_u}\right), \\
g_2(t) &= \frac{1}{p_2 \sigma_\epsilon} \phi\left(\frac{t-x}{\sigma_\epsilon}\right), \\
g_3(t) &= \frac{2}{\sigma_v} \phi\left(\frac{100-t}{\sigma_v}\right).
\end{aligned}$$

This gives

$$\begin{aligned}
m_1 &= \int_0^{100} t g_1(t) dt = 2\sigma_u \int_0^{100/\sigma_u} s \phi(s) ds \\
&= -2\sigma_u \int_0^{100/\sigma_u} \phi'(s) ds = 2\sigma_u \left(\phi(0) - \phi\left(\frac{100}{\sigma_u}\right) \right), \\
m_2 &= \int_0^{100} t g_2(t) dt = \frac{x}{p_2} \int_{-x/\sigma_\epsilon}^{(100-x)/\sigma_\epsilon} \phi(s) ds + \frac{\sigma_\epsilon}{p_2} \int_{-x/\sigma_\epsilon}^{(100-x)/\sigma_\epsilon} s \phi(s) ds \\
&= \frac{x}{p_2} \left(\Phi\left(\frac{100-x}{\sigma_\epsilon}\right) - \Phi\left(\frac{-x}{\sigma_\epsilon}\right) \right) - \frac{\sigma_\epsilon}{p_2} \left(\phi\left(\frac{100-x}{\sigma_\epsilon}\right) - \phi\left(\frac{-x}{\sigma_\epsilon}\right) \right), \\
m_3 &= \int_0^{100} t g_3(t) dt = 200 \int_{-100/\sigma_v}^0 \phi(s) ds + 2\sigma_v \int_{-100/\sigma_v}^0 s \phi(s) ds \\
&= 200 \left(\Phi(0) - \Phi\left(\frac{-100}{\sigma_v}\right) \right) - 2\sigma_v \left(\phi(0) - \phi\left(\frac{-100}{\sigma_v}\right) \right).
\end{aligned}$$

Letting

$$F(t) = \Pr(0 < y < t | 0 < y < 100) = \frac{p_1 G_1(t) + p_2 G_2(t) + p_3 G_3(t)}{q_2}$$

with derivative

$$f(t) = \frac{p_1 g_1(t) + p_2 g_2(t) + p_3 g_3(t)}{q_2},$$

we thus obtain the expectation

$$\int_0^{100} t f(t) dt = \frac{p_1 m_1 + p_2 m_2 + p_3 m_3}{q_2}, \quad (8)$$

which provides the expectation in the general case.

In the special case of model (5) we have $\sigma_u = \sigma_v = 0$, and hence $m_1 = 0$, $m_3 = 100$, and $q_2 = 1$. This gives the expectation

$$\int_0^{100} t f(t) dt = p_2 m_2 + 100 p_3. \quad (9)$$

References

- Blackwell, C. (2010). Rational expectations in the classroom: A learning activity. *Journal for Economic Educators*, 10, 1–6.
- Burson, K. A., Larrick, R. P., and Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology*, 90(1), 60–77.
- Cross, P. (1977). Not can but will college teaching be improved? *New Directions for Higher Education*, 17, 1–15.
- Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one’s own ignorance. *Advances in Experimental Social Psychology*, 44, 247–296.
- Dunning, D., Heath, C., and Suls, J. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69–106.
- Dunning, D., Johnson, K., Ehrlinger, J., and Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12, 83–87;
- Ehrlinger J., Johnson K., Banner M., Dunning D., and Kruger J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, 105, 98–121.

- Embretson, S. E. and Reise, S. P. (2013). *Item Response Theory*. Psychology Press.
- Gabbard, T. and Romanelli, F. (2021). The accuracy of health professions students' self-assessments compared to objective measures of competence. *American Journal of Pharmaceutical Education*, 85(4), Article 8405.
- Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 264–263.
- Gignac, G. E. and Zajenkowski, M. (2020). The Dunning–Kruger effect is (mostly) a statistical artefact: Valid approaches to testing the hypothesis with individual differences data. *Intelligence*, 80, 101449.
- Jansen, R. A., Rafferty, A. N., and Griffiths, T. L. (2021). A rational model of the Dunning–Kruger effect supports insensitivity to evidence in low performers. *Nature Human Behaviour*, 5, 756–763.
- Krajc, M. and Ortman, A. (2008). Are the unskilled really that unaware? An alternative explanation. *Journal of Economic Psychology*, 29, 724–738
- Krueger, J. and Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology*, 82(2), 180–188.
- Kruger, J. and 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, 1121–1134. 77(6), 1121–1134.
- Kruger, J. and Dunning, D. (2002). Unskilled and unaware — but why? A reply to Krueger and Mueller (2002). *Journal of Personality and Social Psychology*, 82(2), 189–192.
- Magnus, J. R. and Peresetsky, A. A. (2018). Grade expectations: rationality and overconfidence. *Frontiers in Psychology — Quantitative Psychology and Measurement*, doi:10.3389/fpsyg.2017.02346.
- Mariana, V. C., Coutinho, J. T., Alsuwaidi, A. S. M., and Couchman, J. J. (2021). Dunning–Kruger effect: Intuitive errors predict overconfidence on the cognitive reflection test. *Frontiers in Psychology*, 12, 603225.

- McIntosh, R. D., Fowler, E. A., Lyu, T., and Sala, S. D. (2019). Wise up: Clarifying the role of metacognition in the Dunning–Kruger Effect. *Journal of Experimental Psychology: General*, 148(11), 1882–1897.
- Schlosser, T., Dunning, D., Johnson, K. L, and Kruger, J. (2013). 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*, 39, 85–100.
- Sullivan, P. J., Ragogna, M., and Dithurbide, L. (2018). An investigation into the Dunning–Kruger effect in sport coaching. *International Journal of Sport and Exercise Psychology*, 1–9.
- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47, 143–148.
- Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26(1), 24–36.
- West, K. and Eaton, A. A. (2019). Prejudiced and unaware of it: Evidence for the Dunning–Kruger model in the domains of racism and sexism. *Personality and Individual Differences*, 146, 111–119.
- Williams, E. F., Dunning, D., and Kruger, J. (2013). The hobgoblin of consistency: Algorithmic judgment strategies underlie inflated self-assessments of performance. *Journal of Personality and Social Psychology*, 104(6), 976–994.
- Zenger, T. R. (1992). Why do employers only reward extreme performance? Examining the relationships among performance, pay, and turnover. *Administrative Science Quarterly*, 37, 198–219.