Self-evaluation bias: an alternative explanation to the Dunning-Kruger effect

Olivia Kim
Norwegian School of Economics

Cite as:
Kim Olivia (2025), Self-evaluation bias: an alternative explanation to the Dunning-Kruger effect. *Proceedings of the European Marketing Academy*, 54th, (123438)

Paper from the 54th Annual EMAC Conference, Madrid, Spain, May 25-30, 2025

Self-evaluation bias: an alternative explanation to the Dunning-Kruger effect

Abstract

In this paper, we explore the phenomenon similar to that of the Dunning-Kruger effect in non-cognitive traits, such as generosity. We hypothesized that individuals who are objectively more generous would perceive themselves as less generous than they are (H1), while those who are less generous would perceive themselves as more generous than they are (H2). We give scenarios where each participant rates the likelihood of their behavior (objective measure), and we ask them to estimate the percentage of people that are more likely than themselves to take the action (subjective measure). The pretest results supported H1 and H2, showing that the more objectively generous an individual was, the less generous they evaluated themselves in specific scenarios. These findings suggest that the Dunning-Kruger effect might not fully apply to non-cognitive traits, such as generosity, and the mechanisms behind self-assessment in these areas require further exploration.

Key words: self-evaluation, bias, personality

Track: Consumer Behavior

Introduction

An accurate self-assessment is an important skill in life, as it allows the individual to make effective life-decisions (career, lifestyle, life partner, etc.), which will lead to greater fulfillment and happiness in their life (Dunning et al., 2004; León et al., 2023; Strube et al., 1986). One of the ways people gain insight into their own competence or trait is by comparing themselves with others (Festinger, 1954; Gilbert et al., 1995). Assessing oneself accurately, however, is not a simple process. First, as demonstrated by Dunning and Kruger (Kruger & Dunning, 1999), they must be able to not only assess their own ability or trait accurately, but they must also be able to calibrate their assessment in the general context, which require what is termed as *metacognition* (Everson & Tobias, 1998). In their seminal paper (cited 10,072 times according to google scholar), Dunning and Kruger claim that the reason that the incompetent group (the bottom quartile) overestimate their own ability stems from their lack of metacognition – in that not only do they lack the competence to complete certain tasks correctly (in the case of their paper, humor, logical reasoning, and grammar), but they also lack the ability to accurately assess where they lie on the scale of the said competence.

However, some recent assessment have brought their explanation into question as perhaps a statistical artefact (Gignac & Zajenkowski, 2020). Furthermore, for areas that are independent of cognition, such as personality traits like generosity or patience, the lack of metacognition doesn't necessarily explain the phenomenon (assuming that personality traits don't have systematic correlation with cognitive abilities). In this paper we explore whether the Dunning Kruger effect – the phenomenon of people in the bottom quartile overestimating themselves to a larger extent compared to the objective scale while the people in the top quartile overestimate themselves to a lesser extent (or sometimes even underestimating their own abilities) compared to the objective scale– also in areas not related to cognitive abilities and the possible mechanisms that drive this phenomenon.

People tend to assess themselves slightly above average on many of the 'positive traits' regardless of their objective performance on task, known as 'Better than Average Effect' or BTAE for short (Zell et al., 2019), AND people tend to use their own ability as a base-line standard against which to measure others (Ross et al., 1977). So, both the bottom quartile and the top quartile probably think that whatever their trait is, it is slightly above average, but this 'average' is guided by their own ability or trait, which is explained by the tendency for social observers to perceive a "false consensus" with respect to the relatively

commonness of their own responses, known as the "false consensus effect" (Ross et al., 1977).

Generosity or other traits do not have an obvious unit of measurement. To say that someone is generous in practice means the same as saying that they are generous compared to the average person. We operationalize this way of thinking about such traits in the following way: Consider a specific scenario that involves a given personality trait. An example is the following question, which relates to the trait *generosity*:

Imagine you are in line at a grocery store with a rather full shopping cart. You notice a young man behind you in line carrying only one item (for instance a toothbrush or a snack). How likely are you to offer him to go ahead of you in the line?

Answer alternatives are Likert scale from 1 (very unlikely) to 7 (very likely). We denote i's response to question k by L_i, where the i-index is for respondent i. A higher L corresponds to higher generosity. Our objective measure of respondent i's generosity in this scenario is then the percentage of respondents with a lower L than respondent. We denote this percentage by x_i, so that

$$x_i = 100 \cdot \frac{n_i}{N} = 100 \cdot F(L_i)$$

where n_i is the number of respondents with a lower (Likert scale) response than i, and N is the total number of respondents. The expression on the right expresses x_i in terms of the true cumulative distribution (c.d.f.) $F(\cdot)$ where $F(L)$ gives the probability that a randomly chosen person has a Likert score equal to or less than L in this scenario. We think of x_i as i's true value of generosity in this scenario. Note that if we observe the L of a large number of respondents, we effectively know $F(\cdot)$ and can therefore calculate x_i.

However, this true value is not directly accessible to the respondent, since he or she does not know the distribution of other people's responses. As a follow-up question, we ask:

In the scenario in Q1, what percentage of people in general do you think are <u>more likely</u> than yourself to offer the young man to go ahead of them in the line?

where the answers take the form of multiple choice with 11 alternatives: 0%, 10%, ..., up to 100%. We define y_i as 100 – (response to the multiple choice question). This is our measure of respondent i's subjective self assessment of generosity in scenario k.

The self-assessment y_i involves a subjective estimate of the c.d.f. $F(\cdot)$, since the distribution of Likert scores in the population is not observable to respondent i. We denote i's subjective estimate of the c.d.f. by $\hat{F}_i(\cdot)$. Respondent i's estimate of x_i is then

$$y_i = 100 \cdot \hat{F}_i(L_i).$$

Our research question focuses on the relationship between y_i and x_i. Assuming a linear relationship, we can write

$$y_i = \alpha + \beta x_i + e_i$$

where the subjective self-assessment value y_i tends to change by βunits per unit increase in the true value x_i. We can use this equation to set out some alternative relationships. The simplest case is one in which the respondent has an unbiased assessment, possibly subject to some noise, so that

$$y_i = x_i + e_i$$

where the error term is mean zero and uncorrelated with x. We can summarize this as follows:

Case I. Unbiased self-assessment: $\alpha = 0, \beta = 1$.

A second possibility is that assessment is subject to an illusory-superiority or better-than-average effect, so that self-assessment is shifted up by some amount relative to the true value:

$$y_{it} = \alpha + x_{it} + e_{it}$$

where α is the bias in self-assessment, i.e. the amount by which the self-assessed value exceeds the true value, on average. This can be summarised as

Case II. Superiority bias self-assessment: $\alpha > 0, \beta = 1$

when the bias is constant across generosity levels, and therefore corresponds to the constant term α.

In this paper we hypothesize that respondents use their own Likert score as a heuristic for the likely value of other people's Likert scores, and partly rely on this heuristic when forming the estimate $\hat{F}_i(\cdot)$. In this way, respondents who themselves are more generous, think that other people are more generous. Concretely, if $L_i > L_j$, (i.e. respondent i is more generous than respondent j) then for a given Likert score L, respondent i will estimate that fewer people have a score less than L, compared to respondent j's estimate:

$$\hat{F}_i(L) < \hat{F}_j(L)$$

We further assume that on average in the population, the estimates $\hat{F}_i(\cdot)$ are approximately correct, i.e.

$$\frac{1}{N}\sum_{i=1}^{N} \hat{F}_i(L) \approx F(L)$$

for each L. It then follows that:

- for i with higher than average L_i, we have $\hat{F}_i(L) < F(L)$, for all L, so that

$$y_i = 100 \cdot \hat{F}_i(L_i) < 100 \cdot F(L_i) = x_i$$

- for i with lower than average L_i, we have $\hat{F}_i(L) > F(L)$, for all L, so that

$$y_i = 100 \cdot \hat{F}_i(L_i) > 100 \cdot F(L_i) = x_i$$

The *self-evaluation bias* is given by the difference

$$bias = y_i - x_i$$

which, according to the story set out above, should be positive for low values of x_i and then decrease to become negative for high values of x_i. Using the expression for x_i and x_i in terms of the estimated and true c.d.f.s, we get

$$bias(L_i) = y_i - x_i = 100 \cdot [\hat{F}_i(L_i) - F(L_i)]$$

and our hypothesis is that $bias(L_i)$ is positive for low L and negative for high L. Put differently, as we move from lower to higher values of the true generosity level x_i, the self-

assessed generosity level y_i starts out higher than x_i, but then increases at a rate of less than 1. In extreme cases, the bias could be so strong that the slope is negative. In terms of the linear model

$$y_i = \alpha + \beta x_i + e_i$$

we can summarize this as:

Case III. Generosity bias in self-assessment bias: $\alpha > 0, \beta < 1$.

We propose to test the theory by subjecting a number N of respondents i to different scenarios, and for each scenario estimating the equation

$$y_i = \alpha + \beta x_i + e_i$$

by OLS.

Pretest

We came up with 10 scenarios (4 on generosity, 4 on patience, and 2 on embarrassment), and for each scenario asked 2 questions:

1) How likely are you to (do the action mentioned in the question)?

   Likert scale (1 very unlikely to 7 very likely)

2) What percentage of people in general do you think are <u>more likely</u> than yourself (do the action mentioned in the question)?

   Multiple choice (0%-100%)

Scenarios:

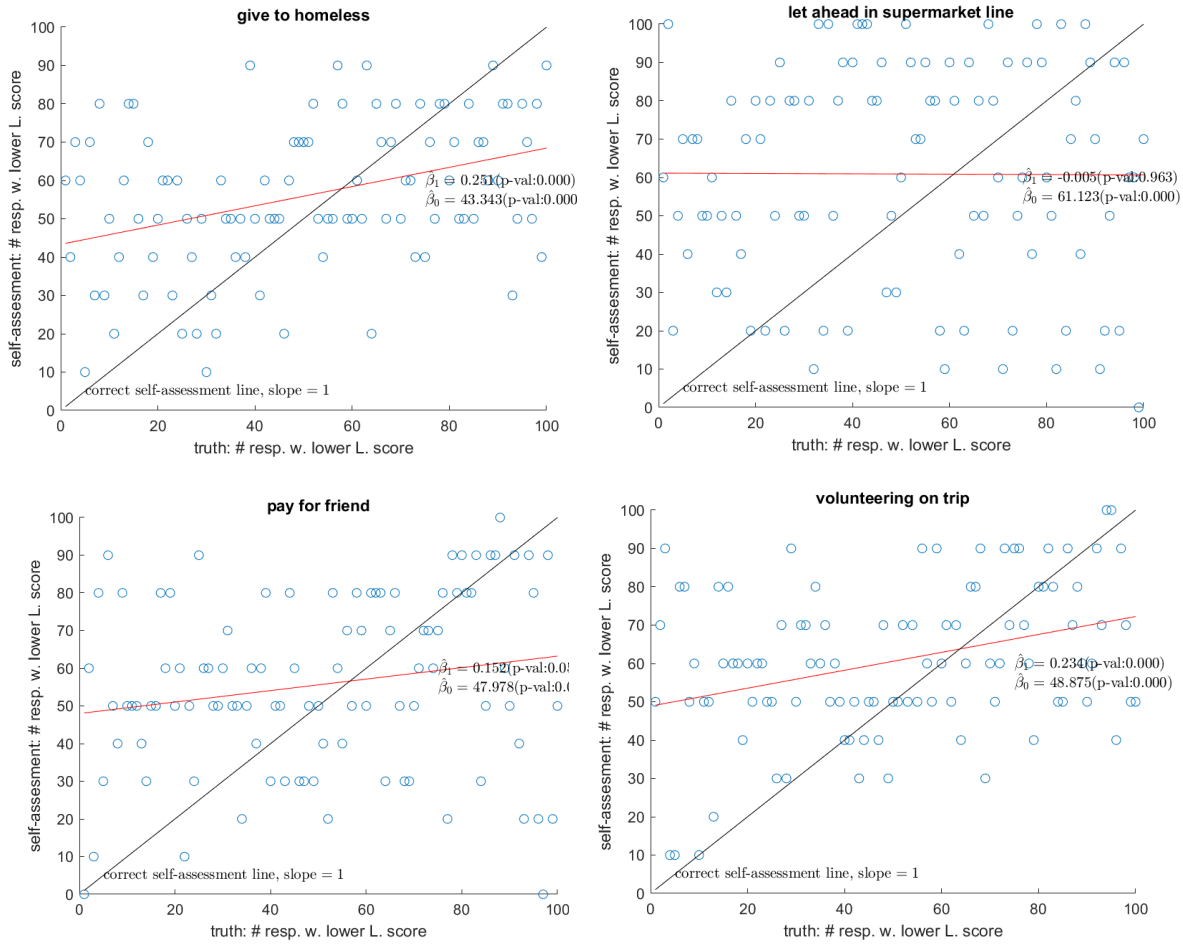
- Generosity:
  - Imagine that you are going on a work/school trip. The tour guide on this trip requires some volunteers to help out with the organization of activities. How likely are you to volunteer?

    - Imagine you are in line at a grocery store with 5-6 items. A young man behind you with just one item asks if he could go ahead of you because he needs to catch a bus. How likely are you let him go ahead of you?
    - Imagine that you are walking in the park, and you see a homeless person. How likely are you to give them some change?
    - Imagine that you are having dinner with a group of high school friends that you haven't seen in over a year. One of them forgot to bring their wallet. How likely are you to offer to pay for them and not ask to be paid back?
- Patience
    - Imagine that you are at a concession stand in the movie theater with just a few minutes before the movie starts, and someone cuts in front of you in line. How likely are you to get upset?
    - Imagine that you were up for a promotion at work, and that someone you are rivals with got the promotion, but you didn't. How likely are you to be upset?
    - Imagine that you are in a museum with a special exhibition of your favorite artist. You have been really looking forward to the exhibit. However, in the room that has the most famous piece from the artist are very loud groups of kids. How likely are you to be upset?
    - Imagine that you have to call a service center to get your computer fixed. You are put on hold for 30 minutes, and then the person on the other line transfer you to a different department and you have to explain your problem all over again. How likely are you get upset?
- Embarassed
    - Imagine that you are walking down the street, and someone waves at you, so you wave back even though you're not sure who it is. Then you find out that they were actually greeting someone who was right behind you. How likely are you to be embarrassed?
    - Imagine giving a presentation at school/work. You have been preparing for this presentation for a while now, and it is an important part of your grade/job.

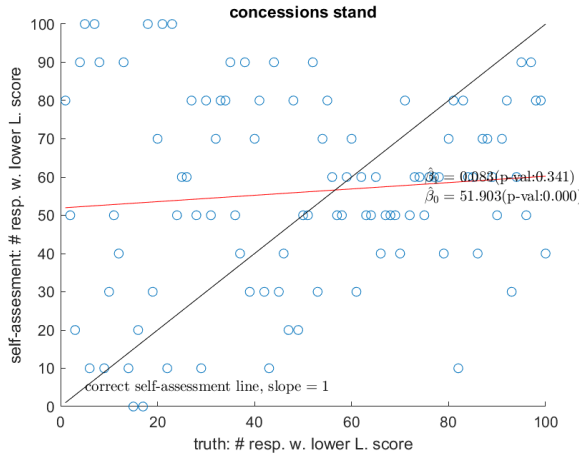
Unfortunately, you forget part of the presentation and mess it up. How likely are you to be embarrassed?

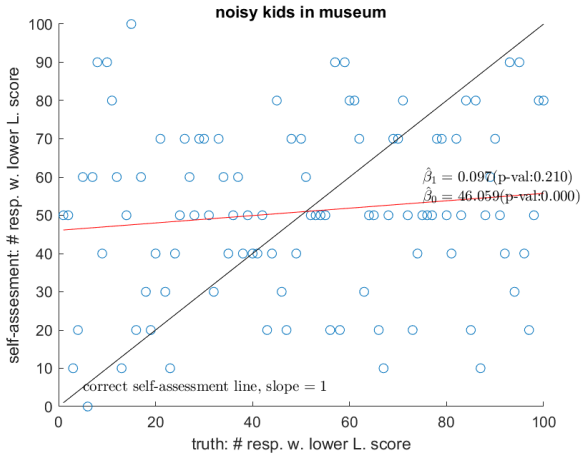
Pre-test Results:

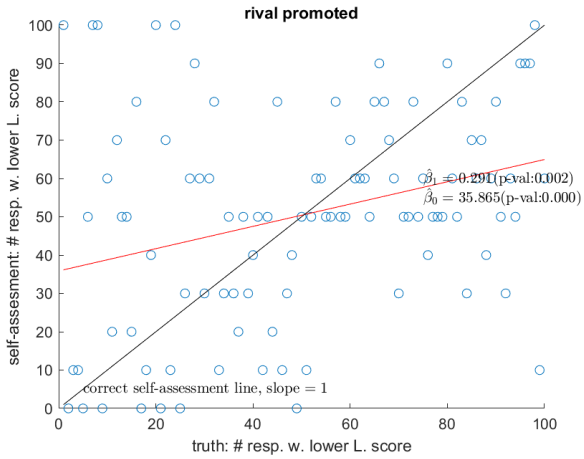
Generosity

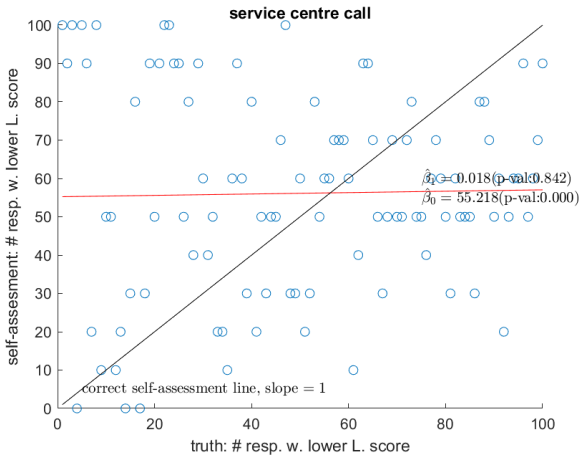


Patience









Embarrassment

References

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed Self-Assessment: Implications for Health, Education, and the Workplace. *Psychological Science in the Public Interest*, *5*(3), 69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Everson, H. T., & Tobias, S. (1998). [No title found]. *Instructional Science*, *26*(1/2), 65–79. https://doi.org/10.1023/A:1003040130125

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Gignac, G. E., & 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. https://doi.org/10.1016/j.intell.2020.101449

Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology*, *69*(2), 227–236. https://doi.org/10.1037/0022-3514.69.2.227

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

León, S. P., Panadero, E., & García-Martínez, I. (2023). How Accurate Are Our Students? A Meta-analytic Systematic Review on Self-assessment Scoring Accuracy. *Educational Psychology Review*, *35*(4), 106. https://doi.org/10.1007/s10648-023-09819-0

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279–301. https://doi.org/10.1016/0022-1031(77)90049-X

Smith, C. A., & Hill, J. P. (2009). *Toward the Measurement of Interpersonal Generosity (IG): An IG Scale Conceptualized, Tested, and Validated*. https://api.semanticscholar.org/CorpusID:146518178

Strube, M. J., Lott, C. L., Lê-Xuân-Hy, G. M., Oxenberg, J., & Deichmann, A. K. (1986). Self-evaluation of abilities: Accurate self-assessment versus biased self-enhancement. *Journal of Personality and Social Psychology*, *51*(1), 16–25. https://doi.org/10.1037/0022-3514.51.1.16