Scientific and Social Research, 2025, Volume 7, Issue 4
https://ojs.bbwpublisher.com/index.php/SSR
Online ISSN: 2981-9946
Print ISSN: 2661-4332

Cognitive Bias: The Dunning Kruger Effect — A Case Study of Postgraduate Students in Fujian Normal University

Shanshan Hong*

Fujian Normal University, Fuzhou 350000, Fujian, China

**Author to whom correspondence should be addressed.*

Abstract: The Dunning-Kruger effect is a purported bias of human metacognitive insight in which people who are incompetent in a given domain are unaware of their incompetence. In this exploratory study, a random sample of twenty-eight postgraduate students from the Fujian Normal University was selected to test this possibility. The results show that the grammar area basically follows the theory raised in the Dunning-Kruger effect — the top quartile students (those achieving above 75% in the test) showed a tendency to underpredict their test performance. Those who scored in the bottom 25% tended to overestimate their ability and test scores. Even when they are aware of what skilled performance would look like, individuals who are unskilled in a grammar domain might overestimate their performance.

Keywords: Dunning-Kruger effect; Cognitive bias; Overconfidence; Self-evaluation

Online publication: May 2, 2025

1. Introduction

A cognitive bias places disproportionate weight for or against an idea, leading to closed-minded or unfair behavior. Amos Tversky and Daniel Kahneman first introduced the concept in 1974, and since then, researchers have identified various biases affecting decision-making [1].

The Dunning-Kruger (D-K) effect is a type of cognitive bias where people believe they are smarter and more capable than they are. First defined by Cornell psychologists David Dunning and Justin Kruger, the phenomenon showed that low-ability individuals overestimate their skills because they lack the self-awareness needed to recognize their incompetence [2]. However, some critics argue that the D-K effect may be a statistical artifact rather than a true cognitive bias [3–4]. Mathematical studies have shown that experts and amateurs both tend to misjudge their abilities. Gignac and Zajenkowski suggest that the D-K effect may stem from flawed measurement rather than actual cognitive bias [5].

This study explores whether the D-K effect applies to postgraduate students at Fujian Normal University

(FJNU). In educational environments, people often see competent students underestimate their performance while less skilled students overestimate theirs. Simons argued that even individuals aware of their skill rank tend toward overconfidence, reinforcing the need to study this effect in academic settings [6–8]. Pennycook et al. further highlight that metacognitive failures in self-assessment are widespread across educational levels, not limited to novices [9]. Thus, this research examines whether the pattern holds true in a Chinese educational context.

2. Literature review

Kruger and Dunning set out to test their hypotheses on Cornell undergraduates in various psychology courses. In a series of studies, they examined the subjects' self-assessment of logical reasoning skills, grammatical skills, and humor. After being shown their test scores, the subjects were again asked to estimate their own rank, whereupon the competent group accurately estimated their rank, while the incompetent group still overestimated their own rank. According to Dunning and Kruger, across four studies, the authors found that participants scoring in the bottom quartile on tests of humor, grammar, and logic grossly overestimated their test performance and ability. Although test scores put them in the 12th percentile, they estimated themselves to be in the 62nd.

Meanwhile, people with true ability tended to underestimate their relative competence. Roughly, participants who found tasks to be relatively easy erroneously assumed, to some extent, that the tasks must also be easy for others. According to Sanchez and Dunning, beginners with limited exposure to a skill area are particularly prone to overconfidence due to their initial learning experiences creating an illusion of competence [10]. This pattern aligns with findings that metacognitive deficits exacerbate self-assessment errors [11]. Additionally, Burson et al. demonstrated that the D-K effect persists even when controlling for task difficulty, suggesting its robustness across domains [12].

A follow-up study, reported in the same paper, suggests that grossly incompetent students improved their ability to estimate their rank after minimal tutoring in the skills they had previously lacked — regardless of the negligible improvement in actual skills. In 2003, Dunning and Joyce Ehrlinger published a study showing that participants' self-assessment could be influenced by external cues, such as whether they were given positive or negative feedback [13]. Daniel Ames and Lara Kammrath extended this work to sensitivity to others and the subjects' perception of how sensitive they were [14]. Other research has suggested that the effect is not so obvious and may be due to noise and bias levels. Dunning, Kruger, and coauthors' latest paper on this subject comes to qualitatively similar conclusions to their original work after making some attempt to test alternative explanations. They conclude that the root cause is that, in contrast to high performers, "poor performers do not learn from feedback suggesting a need to improve."

The abundance of these studies provides additional empirical support for the D-K effect in self-evaluation and performance estimation.

3. Methods

This study was conducted in two phases. The first phase was based on the third domain of the original Kruger and Dunning study. The original Cornell study clearly showed the D-K effect present in the US collegiate elite. FJNU postgraduate students generally score above average on the postgraduate qualifying examination; however, due to the differences between Chinese and American culture, whether this phenomenon can be applied to Chinese universities is worth pondering. The purpose of the first phase is to see if the D-K effect was prevalent in the FJNU

postgraduate students using a grammar test.

The grammar test was constructed using ten items that were randomly selected from each of two sections in the Teacher Qualification Certificate Examination (NTCE). The first phase of the test consisted of ten questions. Each question had four options. The participants were to choose the grammatically correct answer from a list of four other choices.

After taking the grammar test, participants were asked to make three estimates about their ability and test performance. First, they rank their ability to recognize grammatically correct standard English on a percentile scale from one to one hundred, with one being poor and one hundred being excellent. Second, the participants compared their general ability with that of other students on the same percentile scale used before. Finally, they predicted the number of correct answers they would get out of the ten questions on the test. The participants then completed the first phase of the grammar test.

In the second phase of this study, the authors exchanged the tests completed by the top quartile and bottom quartile groups (the classification of the bottom-and top-quartile will be explained below in detail) in the first phase, and asked one group to assess how well the other group completed the test. The participants were asked to once again complete a questionnaire about the three estimates listed above: re-rate their own ability and performance on the test relative to their peers, re-estimate the number of test questions they answered correctly. After this, the participants completed the whole grammar exam research. In November of 2022, students from different majors in the Academic Writing course were asked to participate in the study and take the grammar exam. Participation was anonymous and voluntary, with 2RMB compensation each. Thirty students agreed to volunteer for the study, and twenty-eight students submitted a completed exam and accompanying predictions.

To visualize these results, this paper separated the participants into quartiles based on their actual test scores in the first phase, and the quartile score was converted to percentiles: those who performed in the bottom 25%, those who scored in the top 25%, and the two quartiles in the middle. For each quartile, the perceived ability and the perceived test score were plotted. Then, each quartile's perceived test was compared with the quartile's actual score using a one-sample *t*-test, paired *t*-test, and chi-square test. These methods are used to be generally consistent with the original Cornell study, while new methods were introduced to better ensure the accuracy of the sample data. These procedures align with those employed by Schlösser et al., who tested alternative explanations for the D-K effect through rigorous empirical validation [15].

4. Results and discussion

4.1. Results

4.1.1. Phase one: Grammar test

The phase one results of the FJNU postgraduate students were similar to the results of the Cornell students in the original study conducted by Kruger and Dunning (**Figure 1** and **Figure 2**).

In the Cornell study, the participants estimated their grammatical ability to be in the 71st percentile on average and their performance on the test to be in the 68th percentile on average, exceeding the actual mean of 50, $t(83) = 5.90$ and 5.13, respectively, $P < 0.0001$. The bottom quartile of students grossly overestimated their grammatical ability and perceived score on the grammar exam. Although the actual score of the Cornell students in the bottom quartile was in the 15th percentile on average, they estimated their grammatical ability and predicted scores on the exam to be in the 67th and 61st percentiles, respectively, $t(16) = 13.68$ and 15.75, $P < 0.0001$. The top quartile of

Cornell students underestimated their grammatical ability and perceived scores on the exam. Their actual exam score was in the 89th percentile, however, they estimated their grammatical ability and test performance to be in the 72nd and 70th percentile, $t(18) = -4.73$ and -5.08, respectively, $P < 0.0001$.

The FJNU postgraduate students followed a pattern similar to that of the Cornell students. The students estimated their grammatical ability at the 59th percentile on average, $t(98) = 9.350$, $P < 0.0001$, and their predicted scores in the 56th percentile on average, $t(98) = 10.80$, $P < 0.0001$. Overall, the students' perceived grammatical ability and predicted scores are only moderately correlated with the actual scores, $r = 0.301$, $P = 0.120$ and $r = 0.466$, $P = 0.013$, respectively. Therefore, the results obtained supported the predicted scores correlated with the actual scores, though there is no significantly stronger correlation between the students' perceived grammatical ability and the actual scores. Not surprisingly, the students' estimate of grammatical ability and predicted scores were strongly correlated, $r = 0.677$, $P < 0.001$.

The bottom quartile of participants scored in the 39th percentile on average. However, they predicted their grammatical ability at 56.4 on average. A paired *t*-test was used and support their predicted scores were almost the same as their actual scores ($t = 2.600$, $P = 0.035$). The top quartile of participants scored in the 66th percentile on average. They estimated their grammatical ability to be 80 on average and their predicted scores to be in the 75th percentile, $t(24) = 3.391$, $P = 0.002$.

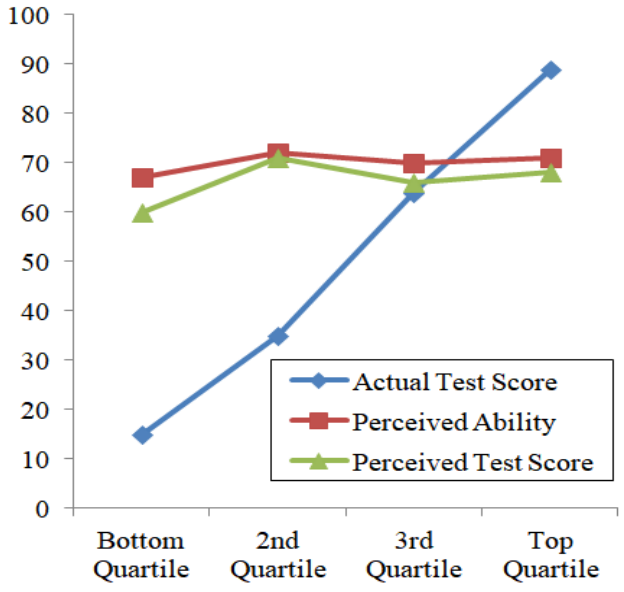

Figure 1. Perceived grammar ability and test performance as a function of actual test performance (Cornell University)

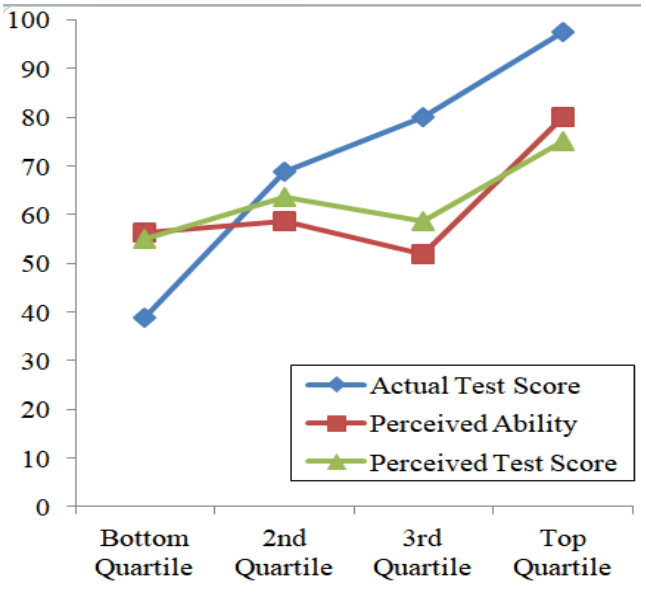

Figure 2. Perceived grammar ability and test performance as a function of actual test performance (Fujian Normal University)

4.1.2. Phase two: It takes one to know one

The phase two results of the FJNU postgraduate students were a little bit different from the results of the Cornell students in the original study conducted by Kruger and Dunning.

Both **Table 1** and **Table 2** display the self-assessments of bottom-and top-quartile performers before and after reviewing the answers of the test-takers shown during the grading task. As can be seen in **Table 1**, in the first phase, those in the bottom 25% thought they did much better than they did, and those in the top 25%

underestimated their performance. In phase two of the Cornell study, it looks like bottom-quartile participants tended to raise their already inflated self-estimates, although not to a significant degree, all *t*s(16) < 1.7. Top-quartile participants raised their estimates of their own general grammar ability after grading the test performance of five of their peers, $t(18) = 2.07$, $P = 0.05$, and their percentile ranking on the test, $t(18) = 3.61$, $P < 0.005$.

In Fujian Normal University, after taking the grammar exam, the students were far less confident in their original predicted ability. Both bottom quartile and top quartile groups lowered their prediction of how they would score on the exam after seeing the more competent choices of their peers. Bottom quartile performers struggled more (3.2%) than top quartile performers (1.2%) to readjust their own performance based on the answers of other participants. In other words, not only do they lack the ability to measure their own performance, but they also lack the ability to accurately measure performance in others.

On average, both two quartiles more accurately predicted their actual scores (Bottom: M percentile = 56.31, $t(24) = 4.766$, $P < 0.001$; Top: M percentile = 73, $t(75) = 0.581$, $P = 0.580$, respectively). The chi-square was used to analyze the difference in perceptions between the bottom-and top-quartile participants of their own answers, and the results showed a significant difference in the composition ratio between the bottom quartile and top quartile in terms of whether they underestimated or overestimated themselves ($x^2 = 24$, $P < 0.001$).

Table 1. Self-ratings (Percentile scales) of ability and performance on test before and after grading task for bottom- and top-quartile participants (Study 3, Phase 2, Cornell University)

	Participant quartile					
	Bottom			Top		
Rating	Percentile ability	Percentile test score	Raw test score	Percentile ability	Percentile test score	Raw test score
Before (Phase 1)	66.8	60.5	12.9	71.6	69.5	16.9
After (Phase 2)	63.2	65.4	13.7	77.2	79.7	16.6
Difference	-3.5	4.9	0.8	5.6	10.2	-0.3
Actual	10.1	10.1	9.2	88.7	88.7	16.4

Table 2. Self-ratings (Percentile Scales) of ability and performance on test before and after grading task for bottom- and top-quartile participants (Study 3, Phase 2, Fujian Normal University)

	Participant quartile					
	Bottom			Top		
Rating	Percentile ability	Percentile test score	Raw test score	Percentile ability	Percentile test score	Raw test score
Before (Phase 1)	57.6	58.6	55.0	80.0	66.0	75.0
After (Phase 2)	54.4	54.0	52.5	78.8	80.0	77.0
Difference	-3.2	-4.6	-2.5	-1.2	14.0	2.0
Actual	38.8	38.8		97.5	97.5	

4.2. Discussion

The figures and table shown above can be eye-openers and provide hands-on support for discussion and insights.

This study indicates that the D-K effect is mostly present in the FJNU postgraduate students. The results showed that overall, not only did most participants overestimate their performance, but the least competent participants — the ones that scored in the lower quartile — were likely to overestimate their performance. These students tend to have an inflated sense of how well they will do but do poorly. When taking a closer look at **Figure 2** above, it can be seen that the 3rd quartile actually made the worst estimate, which is different from the bottom performers in the Cornell study. The gap between their test score and their perceived ability is equal to 30 percentage points. When it finally comes to the top quartile, they predict that they will do well on the test and actually perform well on the test. However, it can be seen that the top quartile students underestimate themselves. The same pattern was evident in the original grammar study. Thus, they generally fit the D-K effect, while some differences still exist.

In addition, another point not shown in these figures is that the participants rate their performance higher than the performance of others. This shows the effect of illusory superiority, the tendency to overestimate their own qualities in relation to other people. Illusory superiority is also one of the biggest drives behind the D-K effect; that is, people are inclined to overestimate themselves and have trouble understanding what it means to be on target.

The results of the *t*-test analysis reveal that the unskilled students who are unaware of their own poor preparation for assessments are detrimentally affected. It may prevent those students from taking steps to improve, and it even compromises success for students who are putting a lot of effort into their studies. Competent students may fear to voice their opinions as they are unaware of their skilled nature. Thus, explicitly teaching metacognitive study methods will support students in overcoming the D-K effect by making their study more effective.

By the same token, being aware of the D-K effect can help teachers prepare teaching sessions. Teaching sessions focusing more on basic knowledge should be developed for students with lower skills and higher confidence. Top students with higher skills and lower confidence could be pushed and encouraged to learn more, thus expanding their knowledge. The difficulty would be in identifying in which quartile each student would be located. Teachers would need to take more time to more thoroughly evaluate each student, which in itself would be a good practice.

5. Limitations

The limitation of this research is that it only addresses twenty-eight FJNU postgraduate students who voluntarily took the grammar test. When comparing the perceived ability in the second phase of this study, the small sample size and large variations in the rankings make precise statistical analysis difficult. Therefore, more data would be needed to make a more detailed conclusive statement. Additionally, as Mahmood noted in a systematic review of D-K studies, self-assessments across various domains frequently suffer from measurement issues due to individual differences in interpreting self-evaluation criteria. Thus, further research should address these measurement challenges explicitly.

6. Conclusion

Based on this study, the D-K effect is evident among FJNU postgraduate students. The grammar test produced similar results to the Kruger and Dunning study at Cornell, reinforcing that low-performing students tend to overestimate their abilities while high-performing students underestimate theirs. This pattern extends beyond grammar to other areas like employability and interpersonal communication.

The key takeaway is that self-awareness is crucial for improvement. As Dr. Dunning notes, “The lesson of the effect was always about how we should remain humble and cautious about ourselves.” Acknowledging this bias

allows people to take steps to minimize its impact and improve learning outcomes.

Disclosure statement

The author declares no conflict of interest.

References

[1] Kahneman D, Tversky A, 1974, Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157): 1124–1131.

[2] 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.

[3] Huang S, 2013, When Peers are not Peers and Don't Know it: The Dunning-Kruger Effect and Self-fulfilling Prophecy in Peer-review. BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology, 35(5): 414–416.

[4] Nuhfer E, Fleisher S, Cogan C, 2017, How Random Noise and a Graphical Convention Subverted Behavioral Scientists' Explanations of Self-assessment Data: Numeracy Underlies Better Alternatives. Numeracy, 10(1): 1–31.

[5] Gignac GE, Zajenkowski M, 2020, The Dunning-Kruger Effect is (mostly) a Statistical Artefact: Valid Approaches to Testing the Hypothesis with Individual Differences Data. Intelligence, 2020(80): 101449.

[6] Simons DJ, 2013, Unskilled and Optimistic: Overconfident Predictions despite Calibrated Knowledge of Relative Skill. Psychonomic Bulletin & Review, 20(3): 601–607.

[7] Kruger J, 1999, Lake Wobegon be Gone! The "Below-average Effect" and the Egocentric Nature of Comparative Ability Judgments. Journal of Personality and Social Psychology, 77(2): 221–232.

[8] Mahmood K, 2016, Do People Overestimate their Information Literacy Skills? A Systematic Review of Empirical Evidence on the Dunning-Kruger Effect. Communications in Information Literacy, 10(2): 199–213.

[9] Pennycook G, Ross RM, Koehler DJ, et al., 2017, Dunning-Kruger Effects in Reasoning: Theoretical Implications of the Failure to Recognize Incompetence. Thinking & Reasoning, 24(2): 161–183.

[10] Sanchez C, Dunning D, 2018, Overconfidence among Beginners: Is a Little Learning a Dangerous Thing? Journal of Personality and Social Psychology, 114(1): 10–28.

[11] McIntosh RD, Fowler EA, Lyu T, et al., 2019, Wise Up: Clarifying the Role of Metacognition in the Dunning-Kruger Effect. Journal of Experimental Psychology: General, 148(11): 1882–1897.

[12] Burson KA, Larrick RP, 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.

[13] Ehrlinger J, Dunning D, 2003, How Chronic Self-Views Influence (and Potentially Mislead) Estimates of Performance. Journal of Personality and Social Psychology, 84(1): 5–17.

[14] Ames DR, Kammrath LK, 2004, Mind-Reading and Metacognition: Narcissism, not Actual Competence, Predicts Self-Estimated Ability. Journal of Nonverbal Behavior, 28(3): 187–209.

[15] Schlosser T, Dunning D, Johnson KL, et al., 2013, How Unaware are the Unskilled? Empirical Tests of the "Signal Extraction" Counter Explanation for the Dunning-Kruger Effect in Self-evaluation of Performance. Journal of Economic Psychology, 2013(39): 85–100.