



Description of the Dunning-Kruger effect in general surgery residents during laparoscopic cholecystectomy: a blinded prospective study

Breno Cordeiro Porto¹ , Rodrigo Marcus Cunha Frati^{1,2} , Rafael Guisalberte Maltez² , Amanda Ferreira da Silva Lima² , Tatiane Alves Ferreira² , Larissa Cunha Baron² , Carlo Camargo Passerotti^{1,3} , Everson Luiz Artifon¹ , José Pinhata Otoch¹ , Jose Arnaldo Shiomi da Cruz^{1,2,3*}

1. Universidade de São Paulo – Faculdade de Medicina – Surgical Technique & Experimental Surgery – São Paulo (SP), Brazil.
2. Universidade Nove de Julho – School of Medicine – São Paulo (SP), Brazil.
3. Hospital Alemão Oswaldo Cruz – Urology Department – São Paulo (SP), Brazil.

ABSTRACT

Purpose: The purpose of this study is to assess whether the Dunning-Kruger effect occurs in surgical residents when performing laparoscopic cholecystectomy in a porcine model. **Methods:** Prospective blinded study, which counted with forty PGY-1 general surgery residents who agreed to participate in the study were blindly recruited to perform a laparoscopic cholecystectomy in a porcine model. At the end of the procedure, the participants assigned a score of 0-10 for their own performance and the video of the operation was independently assessed by 2 experienced laparoscopic surgeons using a validated tool. **Results:** Participants were divided into groups of 10 individuals according to objective performance and compared. The group with the worst objective result was inferior to the group with the best objective result (3.77 ± 0.44 vs. 8.1 ± 0.44 , $p < 0.001$), but they were similar in self-perception of performance (5.11 ± 1.69 vs. 6.1 ± 1.79 , $p = 0.999$). **Conclusion:** In the studied sample, it was possible to demonstrate the presence of the Dunning-Kruger effect.

Key words: Cholecystectomy, Laparoscopic. Motor Skills. Education.

Introduction

When individuals are not proficient in the strategies, they use to achieve success in a task, they face a dual challenge: they not only reach inaccurate conclusions but also, their ineptitude hinders their ability to recognize those misjudgments. As Charles Darwin insightfully noted over a century ago, “Ignorance more frequently begets confidence than does knowledge”. In essence, the very skills that lead to competence in a domain are the same ones required to evaluate competence in that domain. Consequently, underperformers often lack what cognitive psychologists term metacognition¹ or metacomprehension².

Research across various disciplines supports the idea that underachieving individuals often lack the necessary metacognitive skills for accurate self-assessment. In the realm of physics, novices are less adept than experts at gauging the complexity of problems. In tennis, beginners are less capable than seasoned players of accurately evaluating the success of their plays or their overall performance^{3,4}.

Initially, Kruger and Dunning conducted four parallel experiments across different fields of knowledge. The central hypothesis across these studies remained consistent: a test was administered to participants with varying expertise in the

*Corresponding author: arnaldoshiomi@yahoo.com.br

Received: Nov 17, 2023 | Accepted: Apr 04, 2024

Section editor: Marcus Vinicius Brito

Research performed at the University of São Paulo Faculty of Medicine, Surgical Technique & Experimental Surgery Department, Sao Paulo, Brazil.



subject, and while their scores were documented, they were also asked to estimate their performance. For instance, in one of the experiments, 65 college students were tasked with assessing the humor in jokes from renowned comedians. The students' evaluations were then compared with those of eight professional comedians. It was found that the lowest-scoring participants overestimated their ability to discern humor. Similar results were observed in subsequent experiments across various knowledge domains. This phenomenon was coined the "Dunning-Kruger effect"⁵.

A review of prominent databases, including PubMed, Embase, and Web of Science, in October 2023, identified only three studies addressing the Dunning-Kruger effect within the context of surgery. While all these studies discussed the potential risk of bias arising from the effect, none conclusively demonstrated its existence. Theoretically, given their extensive education and training, surgeons should be less susceptible to this bias⁶⁻⁸.

Furthermore, it has been suggested by Kruger and Dunning's subsequent research that extensive study and knowledge tend to diminish the Dunning-Kruger effect (DK effect). Performing surgeries entails considerable responsibility and risk⁵. It is deeply concerning to think that a surgeon might operate under the misconception of competence while lacking actual proficiency. Yet, no study was published so far investigating the presence of the DK effect among practicing surgeons or surgical residents. Thus, the purpose of this study is to assess the existence of the DK effect among first-year general surgery residents as they perform laparoscopic cholecystectomies on porcine models.

■ Methods

Type of study and research participants

This prospective, quantitative analytical study was conducted between 2020 and 2023, and involved 40 first-year General Surgery residents from a university hospital. Participants, aged 25 to 35, had minimal prior video laparoscopic surgical experience. Prior to the study, all residents completed an Informed Consent Form, though they remained "blinded" to the specific evaluations during the procedures. This study received prior approval from the Institutional Review Board.

Participants were selected to perform laparoscopic cholecystectomy on a porcine model under the guidance of an experienced laparoscopic surgeon. The entirety of these procedures was recorded. Once completed, each participant rated its own performance on a scale from 0 to 10. For self-assessment, a numeric scale from 0-10 was used. Zero was the worst possible performance and 10 was the best performance possible, for those who accomplished the procedure with proficiency. Following this, two laparoscopy-experienced surgeons reviewed and scored the recorded procedures using a validated tool⁹. The same numeric scale the participants used for self-assessment were used by the experienced surgeons assessing the videos. As this is not a validated scale, GOALS was also used to assess participants' performance. All videos were assessed by both surgeons, and the objective score was the simple mean of both surgeons scores. Both raters are experienced laparoscopic surgeons from university hospitals engaged in residents' training. Interrater reliability was not calculated between the assessments. The final performance evaluation was derived from the average of these two assessments.

Porcine model

The procedures utilized a porcine model. The animals, weighing between 10 and 15 kg, were sourced from a farm specializing in animals for experimentation. Each animal was anesthetized using a combination of telazol (4.4-6.6 mg/kg intramuscularly), xylazine (1.1-2.2 mg/kg intramuscularly), and atropine (0.04 mg/kg intramuscularly), with anesthesia maintained using isoflurane (1-4%) mixed with 1 to 2 L of oxygen. Throughout the procedure, the animals remained intubated and on mechanical ventilation. Upon procedure completion, the animals were euthanized via an intravenous potassium chloride (KCl) injection and later incinerated as biological material in line with national animal experimentation guidelines¹⁰.

Laparoscopic cholecystectomy

Laparoscopic cholecystectomy was performed with the animal in the supine position under general anesthesia. A 4-port technique was used. The experienced surgeon wielded the camera and the gallbladder with a grasper through the more lateral 5-mm trocar. Initially, the participant should dissect the pedicle of the gallbladder without separating cystic artery from the cystic duct, since these structures are closely related and technically difficult to separate in the porcine model, when compared to the human model.

After isolation of the pedicle of the gallbladder, the participant should clip the pedicle with 2 clips proximally and 1 distally to the point at which section of the pedicle should be made. The next step was to make a section of the pedicle and to free the gallbladder from its bed. Once the gallbladder was loose, the procedure was considered concluded. The blood loss was then aspirated and quantified. The entire operation was meticulously documented, and upon completion, participants self-rated their performance.

Procedures assessment

All surgical performances were reviewed through the video recordings, focusing on both quantitative and qualitative metrics. Quantitative aspects included procedure timing for individual steps and the overall duration. Qualitative evaluation, devoid of surgeon identification, was conducted by two seasoned laparoscopic surgeons, utilizing the global operative performance scale by Vassiliou et al. This scale encompasses five facets of minimally invasive procedures and offers a score between 1 and 5 for each, leading to a potential total ranging between 5 and 25 points (Table 1)⁹. Table 1 is similar to that of Vassiliou et al.'s manuscript. Two and 4 are intermediate values and are not explicit in the table just as in the original GOALS manuscript.

Table 1 – Global assessment scale: criteria for qualitative assessment of surgical skill¹¹.

| Qualitative assessment: global assessment scale | |
|---|--|
| Sense of depth | Tissue handling |
| 1. Often overshoots target, wide swings, slow to correct 2. Sometimes overshoots or misses the target, but is quick to correct 3. Directly points the instrument at the correct plane of the target | 1. Sudden movements, damage to the tissue and neighboring structures, poor control of the grasper and its fixation 2. Handles tissue reasonably well, resulting in less injury 3. Handles tissue well, applies appropriate traction, does not cause damage to neighboring structures |
| Bimanual dexterity | Autonomy |
| 1. Uses only one hand, ignores non-dominant hand, poor coordination between both hands 2. Uses both hands, but does not optimize interaction between them 3. Expertly uses both hands to provide optimized exposure | 1. Unable to complete the task even with verbal instructions 2. Able to complete task safely with moderate instruction 3. Able to complete task independent of instructions |
| Efficiency | Total Score |
| 1. Inefficient efforts, too many attempts at movement, constantly changing approach or insisting in the same approach without making any progress 2. Slow, but planned movements are rationally organized 3. Efficient and secure, maintains focus until another approach provides better performance | 5 - 25 / 25 |

A comprehensive assessment of all surgical procedures was conducted by reviewing video recordings, taking into account both quantitative and qualitative measures. Thus, the communication between participant and surgeon could be listened and assessed. If the participant did not need verbal instructions during the procedure, his autonomy was 5, according to GOALS. If all the procedure had to be directed by the surgeon, his autonomy was 0.

Statistical analysis

Data derived from video analyses were compiled into spreadsheets. Qualitative variables were presented in percentages, and quantitative variables were described with mean and standard deviation. As in the initial experiments uncovering DK effect, participants performed cholecystectomies into groups of ten, according to their objective performance assessed by laparoscopy experts.

Differences in performance across these quartiles were assessed using the Kruskal-Wallis test. In instances of statistical significance, pairwise method comparisons between quartiles were employed to discern distinctive quartiles. Statistical examinations were carried out with a 5% significance threshold using IBM SPSS Statistics for Windows (Version 29.0. Armonk, NY: IBM Corp.).

Results

The overall performance of the participants is shown in Table 2. When comparing interquartile performance, there was no difference between dominant hand ($p = 0.799$), age ($p = 0.585$), self-perception score ($p = 0.394$), time to ligate the gallbladder pedicle ($p = 0.240$) and time to section the gallbladder pedicle ($p = 0.883$). On the other hand, the objective score was clearly different ($p < 0.001$). This outcome was anticipated since grouping was based on this metric. In the quantitative assessment, the time for pedicle dissection ($p = 0.023$), the time for gallbladder removal ($p = 0.02$) and the total operative time ($p = 0.021$) showed statistically significant differences. In the qualitative analysis, all variables showed a statistically significant difference ($p < 0.001$) (Table 3).

Table 2 – Overall performance. Data presented as percentage and mean (\pm standard deviation).

| Overall performance | n (%) |
|--------------------------|------------------------|
| Right-handed/left-handed | 35 (87.5%) / 5 (12.5%) |
| Age (years) | 25.2 (\pm 1.43) |
| Self-perception score | 5.35 (\pm 1.67) |
| Objective score | 6 (\pm 1.88) |
| Quantitative assessment | Time (s) |
| Dissection | 410.9 (\pm 316) |
| Pedicle ligature | 375.2 (\pm 203) |
| Pedicle section | 64.13 (\pm 50.2) |
| Gallbladder removal | 746.18 (\pm 470) |
| Total operative | 1596.5 (\pm 621) |
| Qualitative assessment | n (%) |
| Depth perception | 2.59 (\pm 0.72) |
| Bimanual dexterity | 2.97 (\pm 0.95) |
| Efficiency | 2.75 (\pm 0.89) |
| Tissue management | 2.97 (\pm 1.14) |
| Autonomy | 4.43 (\pm 0.76) |
| Overall score | 15.72 (\pm 3.77) |

Table 3 – Performance by quartile. Data were presented as mean (\pm standard deviation).

| | Quartile 1 (worst result) | Quartile 2 | Quartile 3 | Quartile 4 (best result) | p-value |
|-----------------------------------|------------------------------|-----------------------|-----------------------|-----------------------------|-------------------|
| Number of participants | 10 | 10 | 10 | 10 | |
| Right-handed/left-handed | 10/0 | 8/2 | 10/0 | 10/0 | 0.799 |
| Age (years) | 25 (\pm 0.7) | 24.7 (\pm 0.86) | 25.3 (\pm 2) | 25.7 (\pm 1.76) | 0.585 |
| Self-perception score | 5.11 (\pm 1.69) | 4.88 (\pm 1.45) | 5.22 (\pm 1.71) | 6.1 (\pm 1.79) | 0.394 |
| Objective score | 3.77 (\pm 0.44) | 5.22 (\pm 0.44) | 6.66 (\pm 0.43) | 8.1 (\pm 0.45) | < 0.001 |
| Quantitative assessment (seconds) | | | | | |
| Dissection Time | 662.3 (\pm 314.8) | 400.7 (\pm 246.26) | 303.3 (\pm 342.4) | 290.7 (\pm 2 53.69) | 0.023 |
| Pedicle ligature time | 480.7 (\pm 231.5) | 338 (\pm 175.05) | 399.8 (\pm 187.8) | 291.8 (\pm 197) | 0.240 |
| Pedicle section time | 64.1 (\pm 48.9) | 56.2 (\pm 48.3) | 75.4 (\pm 58.6) | 61.1 (\pm 51.1) | 0.883 |
| Gallbladder removal time | 597.7 (\pm 273.4) | 926.2 (\pm 281) | 1070.7 (\pm 725.8) | 425.6 (\pm 78.36) | 0.002 |
| Total operative time | 1805 (\pm 465) | 1721.2 (\pm 388.4) | 1849 (\pm 807.8) | 1069.2 (\pm 451.4) | 0.021 |
| Qualitative assessment | | | | | |
| Depth | 2 (\pm 0.5) | 2.1 (\pm 0.6) | 2.88 (\pm 0.3) | 3.3 (\pm 0.48) | < 0.001 |
| Bimanual dexterity | 1.88 (\pm 0.33) | 2.55 (\pm 0.72) | 3.33 (\pm 0.70) | 4 (\pm 0) | < 0.001 |
| Efficiency | 1.66 (\pm 0.5) | 2.55 (\pm 0.52) | 2.88 (\pm 0.33) | 3.8 (\pm 0.42) | < 0.001 |
| Tissue management | 1.88 (\pm 0.92) | 2.44 (\pm 1.0) | 3.44 (\pm 0.52) | 4 (\pm 0.66) | < 0.001 |
| Autonomy | 3.44 (\pm 0.52) | 4.44 (\pm 0.72) | 4.77 (\pm 0.44) | 5 (\pm 0) | < 0.001 |
| Overall score | 10.8 (\pm 1.69) | 14.1 (\pm 1.83) | 17.3 (\pm 1.11) | 20 (\pm 0.73) | < 0.001 |

Comparison between quartiles

In line with the Dunning-Kruger effect study our primary focus was the comparison between quartiles 1 and 4. The Pairwise quartile method was applied in this case. Given multiple quartile comparisons, Bonferroni correction was applied to adjust the test significance value, thus reducing type I error likelihood.

Comparing quartiles 1 and 4 yielded results (Table 1) similar to those of Kruger and Dunning. Notably, while self-perception showed no differences between these quartiles, objective performance and the majority of the other studied variables did (Fig. 1).

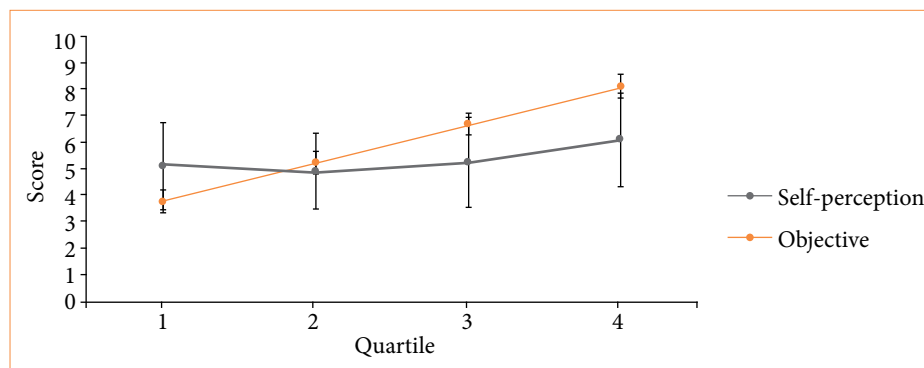


Figure 1 – Comparison of the behavior of objective and self-perceived scores between quartiles. Quartile one overestimated its performance, while the other quartiles underestimated its performance. The better the performance, the more it was underestimated.

■ Discussion

This experiment obtained very similar findings to that of Kruger et. al.⁵, which allows us to infer that, in fact, the DK effect exists among residents during their first laparoscopic procedures. At our institution, University of Sao Paulo Faculty of Medicine, there are 24 vacancies annually for general surgery residency, with a competition of around 10 to 15 candidates per vacancy, therefore, only the best 5 to -10% of candidates are admitted to the residency program. The selection process is expected to admit individuals with considerable minimum knowledge and some degree of metacognition. Despite this, the curves obtained in the experiments are similar to the initial manuscript by Kruger et. al.¹¹.

The research by Kruger et. al.¹¹ raises two questions. First, is there a DK effect? Second, why does it happen? Most studies recognize that there is a DK effect and provide a psychological explanation, either agreeing or disagreeing with the metacognitive explanations of Kruger and Dunning¹²⁻¹⁷. Theories try to explain the DK effect:

- **Metacognitive deficits:** The metacognitive hypothesis suggests that low performers don't have the required skills to recognize their inadequacy. Supporting this idea, studies have shown that training and education can reduce the DK effect by equipping individuals with better metacognitive tools¹⁸.
- **Anchoring and adjustment:** Another theory posits that individuals anchor their self-assessments based on their own performance and following inadequately adjust this self-evaluation when exposed to the performance of others. Meeran et al. (2020) provided evidence for this, showing that the DK effect can be understood as an anchoring-and-adjustment phenomenon¹⁹.
- **Regression to the mean:** A more statistically oriented explanation, the "regression to the mean" theory suggests that the DK effect is merely a statistical artifact. Since most people believe they are above average (a mathematical impossibility), this false self-assessment leads to an apparent overestimation of abilities among low performers and an underestimation among high performers²⁰.
- **Motivated reasoning:** DK effect arises from an innate human desire to view oneself in a positive light. This "self-serving bias" could lead to overestimations of one's abilities to maintain self-esteem²¹.

While the existence of the Dunning-Kruger Effect is broadly supported by empirical evidence, the underpinnings of its cause remain multifaceted, ranging from metacognitive shortcomings to statistical phenomena. Regardless of the reason behind this, the DK effect offers a cautionary reminder of the limits of self-awareness and the potential pitfalls of overconfidence²².

Our experiment certainly has limitations: the reduced number of participants could be a limiting factor, as well as the lack of similar studies in surgery to estimate the sample size. Nevertheless, by performing the post-hoc calculation of statistical power, we obtained a power above 80%, which is very positive, despite having all the inherent limitations of the post-hoc calculation of statistical power. The inputs to carry out this study are costly, both financially and ethically. As the work was performed in a public educational institution, resources were managed in the most optimized way.

Further studies should assess different groups of surgeons, such as senior residents and experienced surgeons. Furthermore, the impact that training will have on these surgeons must be assessed with the aim of optimizing their metacognition and mitigating the DK effect as much as possible in order to obtain the best results for patients.

■ Conclusion

Our study effectively depicted the DK effect among general surgery residents performing laparoscopic cholecystectomies on a porcine model. Further research is essential to validate our results and develop strategies to mitigate the DK effect's implications.

■ Conflict of interest

Nothing to declare.

■ Author's contributions

Substantive scientific and intellectual contributions to the study: Porto BC, Frati RM, Maltez RG; **Conception and design:** Cruz JAS, Porto BC, Frati RM; **Acquisition of data:** Lima AFS, Ferreira TA; **Analysis and interpretation of data:** Lima AFS, Ferreira TA, Baron LC; **Technical procedures:** Porto BC, Lima AFS, Ferreira TA, Baron LC; **Statistics analysis:** Lima AFS, Baron LC, Artifon EL; **Manuscript preparation:** Porto BC, Lima AFS, Ferreira TA, Maltez RG; **Manuscript writing:** Artifon EL; **Critical revision:** Cruz JAS, Passerotti CC, Otoch JP.

■ Data availability statement

Data available upon request.

■ Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

■ About the authors

Porto BC, Lima AFS, Ferreira TA, Baron LC are MSs;

Frati RMC, Maltez RG are MDs;

Passerotti CC, Artifon EL, Otoch JP, Cruz JAS are MDs, PhDs.

■ Acknowledgements

Not applicable.

■ References

1. Everson HT, Tobias S. The ability to estimate knowledge and performance in college: A metacognitive analysis. *Instructional Science*. 1998;26:65–79. <https://doi.org/10.1023/A:1003040130125>
2. Maki RH, Jonas D, Kallod M. The relationship between comprehension and metacomprehension ability. *Psychon Bull Rev*. 1994;1(1):126-9. <https://doi.org/10.3758/BF03200769>.
3. McPherson SL, Thomas JR. Relation of knowledge and performance in boys' tennis: age and expertise. *J Exp Child Psychol*. 1989;48(2):190-211. [https://doi.org/10.1016/0022-0965\(89\)90002-7](https://doi.org/10.1016/0022-0965(89)90002-7)
4. Moreland R, Miller J, Laucka F. Academic achievement and self-evaluations of academic performance. *Journal of Educational Psychology*. 1981;73(3):335-344. <https://doi.org/10.1037/0022-0663.73.3.335>
5. Kruger J, Dunning D. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*. 1999;77(6):1121–1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
6. Wang PZT, Xie WY, Nair S, Dave S, Shatzer J, Chahine S. A Comparison of Guided Video Reflection versus Self-Regulated Learning to Teach Knot Tying to Medical Students: A Pilot Randomized Controlled Trial. *J Surg Educ*. 2020;77(4):805-816. <https://doi.org/10.1016/j.jsurg.2020.02.014>

7. Miura N, Matsubara T, Takamori S, Akamine T, Haratake N, Shikada Y, Yamaguchi M, Okamoto T, Takenoyama M. Mutual checking system for assessing trainee skills of thoracic surgery. *Asian Cardiovasc Thorac Ann.* 2021;29(8):792-797. <https://doi.org/10.1177/02184923211025439>
8. Ahmed O, Walsh TN. Surgical Trainee Experience with Open Cholecystectomy and the Dunning-Kruger Effect. *J Surg Educ.* 2020;77(5):1076-1081. <https://doi.org/10.1016/j.jsurg.2020.03.025>
9. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D, Fried GM. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg.* 2005;190(1):107-13. <https://doi.org/10.1016/j.amjsurg.2005.04.004>
10. da Cruz JAS, Dos Reis ST, Cunha Frati RM, Duarte RJ, Nguyen H, Srougi M, Passerotti CC. Does Warm-Up Training in a Virtual Reality Simulator Improve Surgical Performance? A Prospective Randomized Analysis. *J Surg Educ.* 2016;73(6):974-978. <https://doi.org/10.1016/j.jsurg.2016.04.020>
11. Dunning, D. The Dunning-Kruger Effect: On Being Ignorant of One's Own Ignorance. In J. M. Olson, & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 44, pp. 247-296). Cambridge, MA: Academic Press. 2011; <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
12. Schlösser T, Dunning D, Johnson KL, Kruger J. 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.* 2013;85-100.
13. Williams EF, Dunning D, Kruger J. The hobgoblin of consistency: algorithmic judgment strategies underlie inflated self-assessments of performance. *J Pers Soc Psychol.* 2013;104(6):976-94. <https://doi.org/10.1037/a0032416>
14. Sullivan PJ., Ragogna M, Dithurbide L. An investigation into the Dunning–Kruger effect in sport coaching. *International Journal of Sport and Exercise Psychology .* 2029;17(6), 591–599. <https://doi.org/10.1080/1612197X.2018.1444079>
15. Magnus JR, Peresetsky AA. A Statistical Explanation of the Dunning–Kruger Effect. *Front. Psychol.* 2022;13:840180. <https://doi.org/10.3389/fpsyg.2022.840180>
16. Gabbard T, Romanelli F. The Accuracy of Health Professions Students' Self-Assessments Compared to Objective Measures of Competence. *Am J Pharm Educ.* 2021;85(4):8405. <https://doi.org/10.5688/ajpe8405>
17. Coutinho MVC, Thomas J, Alsuwaidi ASM, Couchman JJ. Dunning-Kruger Effect: Intuitive Errors Predict Overconfidence on the Cognitive Reflection Test. *Front. Psychol.* 2021;12:603225. <https://doi.org/10.3389/fpsyg.2021.603225>
18. Ehrlinger J, Johnson K, Banner M, Dunning D, Kruger J. Why the Unskilled Are Unaware: Further Explorations of (Absent) Self-Insight Among the Incompetent. *Organ Behav Hum Decis Process.* 2008;105(1):98-121. <https://doi.org/10.1016/j.obhdp.2007.05.002>
19. Meeran S, Harte D. Anchoring and the Dunning-Kruger effect. *Thinking & Reasoning.* 2020;26(3), 291-314.
20. Meeran S, Goodwin P, Yalabik B. A parsimonious explanation of observed biases when forecasting one's own performance. *International Journal of Forecasting.* 2016;32(1), 112-120. <https://doi.org/10.1016/j.ijforecast.2015.05.001>
21. Jansen RA, Rafferty AN, Griffiths TL. A rational model of the Dunning-Kruger effect supports insensitivity to evidence in low performers. *Nat Hum Behav.* 2021;5(6):756-763. <https://doi.org/10.1038/s41562-021-01057-0>
22. Gabbard T, Romanelli F. The Accuracy of Health Professions Students' Self-Assessments Compared to Objective Measures of Competence. *Am J Pharm Educ.* 2021;85(4):8405. <https://doi.org/10.5688/ajpe8405>