

Confidence Can Be Measured and Calibrated

We read the article entitled “Medical Trainees and the Dunning–Kruger Effect: When They Don’t Know What They Don’t Know”¹ published in a recent issue of the *Journal of Graduate Medical Education*. We agree with the author’s identification of the problem of miscalibrated trainee confidence (of which the Dunning–Kruger Effect is a subset) and the serious shortcomings of self-assessment.

In addition to the suggestions by the author for improving calibration of trainee confidence, we would like to recommend quantitative methods for measuring confidence calibration at baseline and following specific interventions. The mathematical framework for these methods was first developed for improving the “diagnoses” of weather forecasters, but has been applied to physicians of all levels of training.^{2,3} We have summarized the medical application of these methods for quality improvement and education in 2 recent articles, with specific application to our field of practice (radiology).^{4,5}

The use of these methods requires: (1) a diagnostic question with a reliable reference standard; (2) trainees who are willing to quantify their confidence in their diagnoses in terms of percentages (eg, “I am only 25% sure that this patient has condition X.”); and (3) access to a web browser. We have created an online tool for generating the metrics, along with descriptions of the metrics, at <http://checkmyai.com>.

Behrang Amini, MD, PhD

Assistant Professor, Division of Diagnostic Imaging, Department of Diagnostic Radiology, and Program Director, Musculoskeletal Imaging Fellowship Program, University of Texas MD Anderson Cancer Center

Michael L. Richardson, MD

Professor of Radiology and Adjunct Professor of Orthopedic Surgery, University of Washington

References

1. Rahmani M. Medical trainees and the Dunning–Kruger Effect: when they don’t know what they don’t know. *J Grad Med Educ*. 2020;12(5):532–534. doi:10.4300/jgme-d-20-00134.1.
2. Poses RM, Cebul RD, Centor RM. Evaluating physicians’ probabilistic judgments. *Med Decis Making*. 1988;8(4):233–240. doi:10.1177/0272989X8800800403.
3. Poses RM, Smith WR, McClish DK, Huber EC, Clemons FL, Schmitt BP, et al. Physicians’ survival predictions for patients with acute congestive heart failure. *Arch Intern Med*. 1997;157(9):1001–1007.
4. Amini B, Bassett RL Jr, Haygood TM, McEnery KW, Richardson ML. Confidence calibration: an introduction with application to quality improvement. *J Am Coll Radiol*. 2020;17(5):620–628. doi:10.1016/j.jacr.2019.12.009.
5. Richardson ML, Amini B, Beckmann NM, Subhawong TK. Measuring and teaching confidence calibration among radiologists: a multi-institution study. *J Am Coll Radiol*. 2020;17(10):1314–1321. doi:10.1016/j.jacr.2020.06.035.