

May 22nd, 9:00 AM - May 25th, 5:00 PM

## Commentary on: Ilan Goldberg, Justine Kingsbury and Tracy Bowell's "Measuring critical thinking about deeply held beliefs"

Robert H. Ennis

*University of Illinois, Professor Emeritus of Philosophy of Education*

Follow this and additional works at: <https://scholar.uwindsor.ca/ossaarchive>



Part of the [Philosophy Commons](#)

---

Ennis, Robert H., "Commentary on: Ilan Goldberg, Justine Kingsbury and Tracy Bowell's "Measuring critical thinking about deeply held beliefs"" (2013). *OSSA Conference Archive*. 59.

<https://scholar.uwindsor.ca/ossaarchive/OSSA10/papersandcommentaries/59>

This Commentary is brought to you for free and open access by the Department of Philosophy at Scholarship at UWindsor. It has been accepted for inclusion in OSSA Conference Archive by an authorized conference organizer of Scholarship at UWindsor. For more information, please contact [scholarship@uwindsor.ca](mailto:scholarship@uwindsor.ca).

# Commentary on: Ilan Goldberg, Justine Kingsbury and Tracy Bowell's "Measuring critical thinking about deeply held beliefs"

ROBERT H. ENNIS

*Emeritus Professor of Philosophy of Education*

*University of Illinois*

*3904 Trentwood Place*

*Sarasota, FL 34243*

*USA*

[rhennis@illinois.edu](mailto:rhennis@illinois.edu)

The authors, all critical thinking teachers, have presented the results of a comparison of five possible ways to measure critical thinking, the fifth of which served as the criterion variable for judging the others. The ultimate goal is to have a valid critical thinking test to check the effectiveness of different approaches to teaching critical thinking.

My comments are based on the slides provided me in advance, the full paper as submitted to OSSA, and what I can remember of the actual presentation.

As far as I have been able to determine it, their approach was to compare five different possible ways of assessing critical thinking, the CCTDI, the RPDS, each of the first two questions of an Interview, and the third Interview question. The response to the third Interview question (henceforth the "Fifth Test") was assumed to be the criterion measure used to compare the abilities of the other four (CCTDI, RPDS, and each of the first two questions) to measure critical thinking about deeply held beliefs. In their plan, the one that correlated the highest with the results of the third Interview question (the Fifth Test) would be likely to be the most valid test of the four.

*The Four Tests Being Evaluated for Their Ability to Assess Critical Thinking.*

One wonders why they chose that particular set of four possible tests. Perhaps one reason is that each is what is generally called an objective test, Questions 1 and 2 being semi-objective tests. That is, none of the four tests being evaluated required human graders to judge the degree of critical thinking of the students being tested. The Fifth Test, the one serving as the criterion test, did require human judgments of the degree of critical thinking of the students.

It is not clear why the CCTDI and the RPBS tests were chosen from the many publicly available objective critical thinking tests. Although the dispositions supposedly assessed by the CCTDI are very important, there is more to critical thinking than dispositions. Furthermore it is not clear why the CCTDI was deemed by the authors to be "internationally recognized". The only published discussion of it that I have seen (other than by people with a conflict of interest) does not recommend it (Possin, 2008, p. 209). Furthermore, unless it is administered

anonymously, I believe it can be gamed, and it is based on factor analysis, a controversial statistical procedure.

Regarding the RPBS, a person's opinions about so-called paranormal topics need more investigation than provided by the RPBS. All it does is determine the self-declared strength of a student's belief in a set of paranormal claims, assuming that the weaker the belief, the better and that strong disagreement is the weakest belief. It needs, for example, the kind of exploration of a person's position that the authors performed in administering the Fifth Test. Without exploration by human investigators, a student's position on "paranormal" propositions can be misleading. Varying backgrounds of students, including different amounts of deep experience with a particular controversial issue, might result in different amounts of strong disagreement with the beliefs.

Questions 1 calls for the graders (the instructors and authors of this presentation) on the basis of conversations with the students to judge (on a scale from one to seven) the extent to which the students thought that the authors' courses had a positive effect on the quality of the students' critical thinking. Question 2 calls for the graders to decide the extent to which students felt that the courses had a positive effect on the quality of the students' thinking about their deeply held beliefs. A high score on each question meant that the instructors felt that the students felt that the critical thinking course had a positive effect on the quality of the students' thinking. This is sort of an opinion poll of the students' beliefs about the quality of the instructors' courses (possibly marred by fear that the honest expression of opinion by students might adversely affect the instructors' opinions of the students). It is interesting that these opinion polls, as we shall see, did about as well as the CCTDI and the RPBS in correlating with the criterion test scores obtained from the Fifth Test).

Does this mean that asking students how well they did in our critical thinking courses is as good a way as at least the other two to find out whether the course actually was a good one? This would be an interesting conclusion that, if generalized, would make it not worth the trouble to make critical thinking tests at all. Perhaps just asking the students would be enough. This possible implication was not discussed in the presentation. Instead we were overwhelmed with uninterpreted numbers.

### *Results and Related Issues*

1. None of the correlations of the first four tests with the fifth test is very impressive (.27, .36, .38, and .29, the last two being the relationships between Questions 1 and Question 2 and the Fifth Test, as I read their extensive tables. However, having unimpressive concurrent evidence of validity is a common occurrence for correlations between test results and a criterion variable. Robert Linn (1982) investigated this matter for a number of popular tests and found such validity indices commonly to run between .2 and .4. This fact is unfortunately not highly advertised by test makers, who prefer to advertise the high numbers they get for "test reliability", by which they mean consistency, not validity.

2. The correlation for the CCDTI was a bit higher than that for the RPBS, contrary to the authors' expectations (.36 vs. .27).

3. Given the current implementation and grading of the Fifth Test, and assuming that we are limited to the tests considered, it appears that Questions 1 and 2 in the Interview would be as good as any choice from these four as a measure of critical thinking about deeply held beliefs, especially since Questions 1 and 2 are free, amenable to revision (as also is RPBS); and are simple, seem to have performed about as well as the others (.38 and .29), and focus directly on the ultimate objective, having a good critical thinking course. If we limit ourselves to the four choices the authors have made available to themselves, Questions 1 and 2 are the tests I would choose.

4. There is understandable disagreement among the three authors about the contents of, and how exactly to grade, the Fifth Test (e.g., inclusion of religion as a paranormal belief). Because it might have potential, I urge them to continue their exploration and try to develop useful rubrics. If the fifth test is offered as a comprehensive test of critical thinking courses, it should include critical thinking abilities as well as dispositions. Furthermore some obvious controversial issues need to be resolved. Its development into a good critical thinking test will be a difficult task, but if developed well, it might serve as a criterion variable for judging critical thinking courses and tests.

5. Lastly I wonder why "weak sense critical thinking" as they define it, is critical thinking at all. In their third slide, they define "weak sense critical thinking" as "critical thinking skills employed solely for the purpose of defending one's own beliefs and attacking competing beliefs." Such thinking seems to violate the basic critical thinking dispositions to seek the truth and avoid bias, made worse by the employment of the weapons of the fallacy labels and other distinctions and concepts we use in critical thinking. This is not just weak sense critical thinking. It is not critical thinking at all.

I thank the authors for sharing the results of their interesting study.

#### REFERENCES

- Goldberg, I., Kingsbury, J., & Bowell, T. (2013). Measuring critical thinking about deeply held beliefs. In D. Mohammed & M. Lewiński (Eds.), *Virtues of Argumentation. Proceedings of the 10th International Conference of the Ontario Society for the Study of Argumentation (OSSA), 22-26 May 2013* (pp. 1-13). Windsor, ON: OSSA.
- Linn, R. L. (1982). Ability testing: Individual differences, prediction and differential prediction. In A. K. Wigdor & W. R. Garner (Eds.), *Ability testing: Uses, consequences, and controversies* (Part II: Documentation section, pp. 335-388). Washington D.C.: National Academy Press.
- Possin, K. (2008). A guide to critical-thinking assessment. *Teaching Philosophy, 31* (3), 201-228.