

A Psychometric Evaluation of the Psychological Critical Thinking Examination for Course and Program Assessment

Nicole C. Rushing and Katherine L. Van Allen

Department of Psychology, Campbell University

A growing interest in higher education is measuring subject-specific critical thinking (CT), both for its relevance to course and program objectives as well as for examining the effectiveness of targeted instructional interventions. We report psychometric properties of the revised Psychological Critical Thinking Exam (PCTE), a subject-specific CT test recommended in the *American Psychological Association's (2013) Guidelines for the undergraduate psychology major*. Among 71 psychology students at a small, private university, PCTE items demonstrated moderate to high difficulty and mostly acceptable discrimination. We also found the PCTE to have excellent interrater reliability and good internal consistency and construct validity. This study adds to emerging evidence suggesting the revised PCTE is one useful measure of domain-specific CT skills.

Keywords: critical thinking, assessment, program evaluation

Critical thinking (CT) skills help us evaluate claims, solve problems, and make better decisions in an increasingly connected and information-rich world. Research reveals that CT skills are a significant predictor of positive life events across several domains including academics, employment, and relationships (Butler et al., 2017). In higher education, CT has been a focus for quite some time, having been shown to relate to important student outcomes such as academic motivation (Rugutt & Chemosit, 2009) and achievement (Gadzella et al., 1997; McCutcheon et al., 1992; Williams et al., 2003). Today, CT is frequently included in program and course learning objectives as well as university strategic and quality enhancement plans. However, some alarming evidence suggests that

college students are completing degree requirements while still lacking in CT skills. According to the results of the *National Association of Colleges and Employers' (NACE, 2018) Job outlook 2019* survey, employers indicated that the “critical thinking/problem-solving” career readiness competency was the most essential of those measured; yet, while 100% of employers rated this competency as “absolutely essential” or “essential,” they rated only 56.8% of college graduates entering the workforce as “proficient” in this category. Moreover, *PayScale's (2016)* study of over 60,000 managers in the United States revealed that 60% of supervisors believe CT skills are inadequate among recent graduates entering the workforce. These findings suggest that although employers place a high priority on CT skills, new employees are perceived to be wanting in that area. As the need for CT skills training becomes clearer, a first step of initiatives involves assessment of students' CT skills using valid and reliable measures.

Assessment of CT in higher education is somewhat difficult in part because there is no one agreed upon definition of CT. However, experts generally concur that CT is a multidimensional construct (Bensley & Murtagh, 2012), consisting of skills that allow individuals to use relevant information to reach and evaluate conclusions as

This article was published Online First February 24, 2022.
Nicole C. Rushing  <https://orcid.org/0000-0002-5534-4922>

Katherine L. Van Allen  <https://orcid.org/0000-0002-6842-9300>

The preliminary results from this research were presented as a poster at the 2020 National Institute on the Teaching of Psychology, St. Pete Beach, Florida, United States.

Correspondence concerning this article should be addressed to Nicole C. Rushing, Department of Psychology, Campbell University, P.O. Box 369, Buies Creek, NC 27506, United States. Email: rushing@campbell.edu

well as a disposition of inquiry characterized by dedication to reason and eagerness for reliable information (Facione, 1990). Although the exact nature of the relationship between CT skills and dispositions is still being debated, we chose to evaluate a test of CT skills in the present study because some research has demonstrated that CT skills may be more amenable to change than dispositions through educational experiences and interventions (Facione & Facione, 1997; Kalelioğlu & Gülbahar, 2014).

An additional distinction has been made between general versus domain-specific CT skills. Psychology educators and scholars Halonen and Gray (2000, p. 1) define general CT as a group of “thinking skills that promote conscious, purposeful, and active involvement . . . with new ideas.” Proponents of measuring general CT argue that a set of CT skills exist and transfer across domains, such that the same skills can be applied to problems in different subject areas.

On the other hand, an increasing number of scholars argue CT assessment should be domain specific for at least two reasons. First, researchers have demonstrated there are particular CT skills relevant to each discipline (Cheung et al., 2002) that are not always detected in general CT measures. Second, measuring accurate and practical application of CT skills in a subject is dependent on unique content knowledge (McPeck, 1990). For example, to explain why individuals with greater content knowledge perform better on domain-specific scenarios that require CT, Dwyer et al. (2015) suggest it is because they are better able to reason and evaluate the strengths and weaknesses of a given perspective within their discipline.

Educators from disciplines such as physics (Tiruneh et al., 2017), nursing (Daly, 2001), business (Anderson & Reid, 2013; Dwyer et al., 2015), and psychology (Renaud & Murray, 2008) have attempted to measure domain-specific CT. In psychology, subject-specific CT has been defined as “the ability to evaluate claims in a way that explicitly incorporates basic principles of psychological science” (Lawson, 1999, p. 207) and differs from CT skills emphasized in other disciplines (Bensley & Murtagh, 2012; Cheung et al., 2002). There is evidence that subject-specific CT in psychology demonstrates stronger growth in a single course (Stark, 2012) and is a better predictor of achievement in a course than general CT (Williams et al., 2004). The American Psychological Association (APA, 2013)

continues to include CT as one of the five goals for the undergraduate psychology major and outlines that undergraduate students in psychology should acquire domain-specific skills for “interpreting behavior, studying research, and applying research design principles to drawing conclusions about psychological phenomena” (p. 15). Therefore, it is important to measure psychological CT as it pertains to program evaluation and planning. Subject-specific CT tests may be useful in understanding the extent to which individual courses within a program influence CT, as well as for documenting the effectiveness of targeted CT interventions.

Despite the APA’s stated interest in psychological CT, we found only four psychology-specific CT tests, including the one evaluated in the present study. Bensley and Baxter’s (2006) unpublished *Critical Thinking in Psychology Test* is multidimensional in that it contains argument analysis, methodological reasoning, and causal reasoning subtests; however, psychometric information is not currently reported. Ricker (n.d.-a) developed the *Scottsdale Test of Critical Thinking* for internal assessment purposes in lower division courses at a community college. The 30-item multiple-choice test can be objectively and efficiently scored, and item-total correlations are mostly good. Internal consistency and test-retest reliability are reported as low to acceptable (Ricker, 2005, n.d.-a, n.d.-b). Although the *Scottsdale Test* was initially created with a focus on psychology, Ricker states that after revisions to content, the test may be better classified as a general CT test. Most recently, Cascio’s (2020) *Methodological Thinking Drill* (MTD) test asks students to respond to and self-score 10 short-answer questions about blog articles describing research in psychology. The instrument has utility as both a class activity and CT test, keeping in mind it was designed for introductory psychology courses and, thus, may not be applicable to upper-level students. Furthermore, although some of the MTD items do seem to tap CT (e.g., “Identify and describe one potential alternative explanation for the results”), other items capture basic knowledge about information literacy (e.g., “Is the article a primary source or a secondary source?”). Therefore, the MTD total score may not be a unitary measure of CT skills.

For the present study, we examined Lawson et al. (2015) revised *Psychological Critical Thinking Exam* (PCTE) because it is recommended by

the *APA guidelines for the undergraduate psychology major* (2013). The PCTE is a short-answer test that asks students to describe flaws in conclusions based on evidence. The original (1999) version contained 14 items scored as correct or incorrect and demonstrated good interrater reliability and construct validity (Haw, 2011; Lawson, 1999; Penningroth et al., 2007; Williams et al., 2003, 2004). The current 14-item version retained seven items from the original version and modified scoring such that items are scored 0–3. Lawson et al. (2015) reported that among students at one university in the Midwest U.S., the revised PCTE has good interrater agreement as well as excellent split-half reliability ($r = .88$) and 2-week test–retest reliability ($r = .90$). In support of the PCTE’s construct validity, advanced psychology students outperform students from other disciplines as well as less advanced psychology students, and PCTE scores correlate moderately with CT dispositions (Lawson et al., 2015).

We found only four additional published studies (Branch & Dubow, 2021; Jern, 2018; Rodefer et al., 2020; Tolbert, 2020) that utilized the revised PCTE. In these studies, investigators used change in PCTE score as an outcome in assessing effectiveness of instructional interventions in various psychology courses. Two 7-item pre- and posttest versions were created such that each version contained one question from each of the seven CT concepts included in the test. Similar to reports by Lawson et al. (2015) on the full 14-item version, Jern (2018) found that the number of psychology courses previously taken correlates positively with the 7-item PCTE score. Branch and Dubow (2021) reported good interrater reliability, with an intraclass correlation coefficient (ICC) ranging from .788 to .947. Furthermore, the PCTE detected changes in CT skills over a single course (Branch & Dubow, 2021; Rodefer et al., 2020). These findings suggest the PCTE may be one useful measure of subject-specific CT skills at the end of an academic program, as well as within a single course for the purpose of tracking growth in CT skills over time.

Modified versions of the PCTE appear to have good reliability and validity based on this small number of studies. However, no information on reliability and validity of the full 14-item current version beyond the test developers’ original report (Lawson et al., 2015) is available. Therefore, the purpose of the present study was to

address calls for more information on the psychometric properties of the instrument (Bensley & Murtagh, 2012; Lawson et al., 2015) by investigating item functioning, reliability, and validity of the full PCTE among psychology students at various stages in their academic program.

Method

Participants

Undergraduates at a small, private, religiously affiliated university in a rural county in the south-east U.S. participated in this research. Participants were 36 psychology majors in their final year, who were verified as having graduated from 2018 to 2020, and 35 students enrolled in either a 200-level social sciences statistics class or a 400-level abnormal psychology class in Fall 2019. The majority of students in the sampled classes were psychology majors (69%), with exercise science being the next most popular major (14.3%).

Measures

Psychological Critical Thinking Skills

The PCTE (Lawson et al., 2015) is a short-answer, subject-specific CT test designed for undergraduate psychology students. The PCTE contains two items for each of the seven CT principles represented. The item stems consist of a conclusion related to psychological phenomena and a few sentences of research or anecdotal evidence on which the conclusion is based. Students indicate there is no problem with the conclusion or identify and explain the problem. Scores of 3 indicate correct identification of the most relevant concept. Scores of 2 indicate the student identified the most relevant concept and at least one other less relevant concept. Scores of 1 indicate the student identified a problem unrelated to the most relevant concept, and scores of 0 indicate no problem was identified. If a student answered at least 12 of the 14 questions, we computed their total score and scored any skipped items as 0.

Content Knowledge in Psychology

The *Major Field Test* (MFT; Educational Testing Service, 2014) in Psychology consists of 140 multiple-choice questions and measures

graduating psychology students' mastery of concepts, principles, and knowledge.

CT Disposition

We administered the widely used 18-item short form *Need for Cognition Scale* (NFC; Cacioppo et al., 1984) as a measure of CT disposition. Individuals who are high in need for cognition have stable traits consistent with CT, such that they enjoy complex problem-solving, focus on relevant content rather than extraneous cues, and derive their attitudes from the tendency to organize, evaluate, and analyze information (Cacioppo & Petty, 1982). Participants rated items using a 5-point Likert scale ranging from 1 = *extremely uncharacteristic of me* to 5 = *extremely characteristic of me*, which were then summed to compute the total score. Two participants skipped one item on the NFC, and their missing scores were imputed based on their average item score. Higher total scores indicate higher need for cognition.

Academic Performance

For students in the statistics and abnormal psychology courses, we obtained data from academic records including number of completed credit hours in psychology, total number of completed credit hours, major(s), overall grade point average (GPA), and psychology GPA.

Procedure

The Campbell University Institutional Review Board (IRB) approved this research. Students in the statistics and abnormal psychology classes provided informed consent to participate in this study and completed the PCTE and NFC online during one class period between the first and third week of the Fall 2019 semester. Students had the option to receive a minimal amount of extra credit for participation. The course instructor or trained research assistant proctored the testing, and a department assistant obtained the academic record information.

Graduating psychology majors completed the PCTE and MFT in April or May of the year in which they planned to graduate as part of "exit testing" routinely conducted by the department. A department assistant proctored the exit testing at a day and time of each student's choosing. Although graduating seniors are not required to

complete exit testing and receive no compensation for doing so, we have observed that most students do complete the testing. The 2018 cohort completed a paper-and-pencil version of the PCTE, while the 2019–2020 cohorts completed the PCTE online. If a graduating student had previously completed the PCTE in their statistics or abnormal psychology class (as was the case with five students in the 2020 cohort), we included data for their first attempt only.

After reviewing the PCTE scoring guide, obtaining scoring clarification from the PCTE developer, and discussing sample responses, two faculty raters independently and blindly coded all responses. When there were discrepancies in item scoring, we averaged the ratings.

Results

Participation Rate

Participation was high, with about 80% of students enrolled in the sampled classes providing informed consent for this research and 80% of graduating psychology seniors in 2018 completing the PCTE. Our program did not track participation among graduating seniors after 2018.

PCTE Total Scores

We display average PCTE score by academic year in Table 1. The median PCTE total score was 17.5, and there were no outliers using the 1.5 times interquartile range method. Among students who completed the PCTE online, the median completion time was 34 min. With completion time, outliers removed from the analysis, as may have been the case for students who left the test window open for an extended period,

Table 1
PCTE Scores by Academic Year

Student group	<i>n</i>	<i>M</i>	<i>SD</i>
Graduating psychology majors	36	18.81	7.11
All fourth-year students ^a	54	17.77	6.70
Psychology major	45	18.32	6.82
Nonpsychology major	9	15.00	5.57
Third-year students	12	16.29	6.64
First- and second-year students	5	15.30	7.26
All students	71	17.35	6.67

Note. PCTE = Psychological Critical Thinking Exam.

^aIncludes graduating and continuing fourth-year students.

completion time was unrelated to PCTE score, $r(43) = -.015, p = .920$.

Item Analysis

For item analyses, we considered item scores 2+ to indicate a “correct” response. As shown in Table 2, the PCTE items displayed moderate to high item difficulty. We calculated item discrimination indices using the standard cutoff of 27% ($n = 19$) to define the upper and lower groups. All but three items demonstrated at least acceptable discrimination.

Reliability

Total scores displayed excellent interrater reliability, $r(69) = .975, p < .001$. After creating two halves of the test, such that each half contained one item representing each of the seven CT concepts, the PCTE demonstrated good internal consistency using the split-half method corrected by the Spearman–Brown formula, $r(69) = .788, p < .001$.

Validity

Graduating psychology majors obtained MFT total scores in the average range using norms from 2014 to 2017 U.S. test takers. The PCTE demonstrated good construct validity as evidenced by significant correlations with content knowledge in psychology on the MFT and academic achievement using psychology and overall GPA (Table 3). The strength of the correlations obtained between

PCTE and overall GPA versus PCTE and psychology GPA was not significantly different from one another, $z = 1.765, p = .078$ (see Lee & Preacher, 2013).

To further examine the relationship between GPA and PCTE, we used a cut point of 3.0, which was slightly higher than the median psychology and overall GPA ($Mdn = 2.86$). Students earning 3.0 or higher psychology GPA scored significantly higher on the PCTE ($M = 18.25, SD = 6.11$) than students with a psychology GPA below 3.0 ($M = 14.09, SD = 4.73$), $F(1, 30) = 4.63, p = .04$. The same pattern was observed for overall GPA using a 3.0 cut point, $M = 19.17 (SD = 5.49)$ versus $M = 13.53 (SD = 4.74)$, $F(1, 30) = 9.73, p = .004$. Additionally, among students with a 3.0 or higher overall GPA, we observed a strong positive correlation between overall GPA and PCTE, $r(13) = .708, p = .003$, whereas no relationship was observed among students with less than a 3.0 GPA, $r(15) = -.385, p = .127$. These results suggest that especially among high-achieving students (i.e., those with overall GPA of 3.0 or higher), small GPA increases were positively associated with PCTE score.

Students who had made more progress in their academic program, as evidenced by more credit hours in psychology, obtained higher PCTE scores (Table 1). Even among fourth-year students, the number of psychology credits was positively associated with PCTE score, $r(16) = .521, p = .027$. Discriminant evidence revealed the PCTE was not significantly related to overall credit hours completed or CT disposition.

Table 2

PCTE Item Difficulty and Discrimination

Psychological CT concept	Item no.	% scored 2+ by either rater	Item discrimination index
Coincidence/chance event	1	8.5	0.16
	14	42.3	0.21
Multiple causes	2	43.7	0.58
	13	42.3	0.58
Need for control group/placebo effects	3	43.7	0.42
	12	18.3	0.32
Falsifiability	4	11.3	0.16
	11	12.7	0.05
Correlation \neq causation	5	29.6	0.63
	8	18.3	0.47
Leading question	6	28.2	0.74
	9	45.1	0.68
Unrepresentative sample	7	46.5	1.0
	10	38.0	0.42

Note. PCTE = Psychological Critical Thinking Exam; CT = critical thinking.

Table 3
Measures and Correlations With PCTE

Measure	<i>M (SD)</i>	Correlation
Major Field Test (MFT) in psychology	158.19 (13.01)	$r(24) = .486^*$
No. of PSYC credits	12.17 (7.6)	$r(33) = .360^*$
No. of overall credits	85.56 (33.39)	$r(33) = .021$
PSYC GPA	2.80 (0.94)	$r(30) = .366^*$
Overall GPA	2.96 (0.58)	$r(30) = .534^*$
Need for Cognition (NFC) scale	60.32 (9.01)	$r(32) = .094$

Note. PCTE = Psychological Critical Thinking Exam; PSYC = psychology;

GPA = grade point average.

* Significant at $p < .05$.

Discussion

In our sample, we found the PCTE to display moderate to high item difficulty and mainly acceptable to good item discrimination as well as strong interrater reliability, good internal consistency, and good construct validity. As we hypothesized, having completed more credit hours in psychology predicted higher PCTE scores (whereas overall credits did not), and students who were more advanced in their academic program had higher PCTE scores than beginning students. These results specifically suggest that more experience in psychology courses is positively associated with CT in psychology.

Furthermore, students with higher academic achievement demonstrated higher CT skills. These results are in line with other findings suggesting that not only is education a significant predictor of CT skills (Butler, 2012) but that CT skills can be learned (Bensley et al., 2010; Muehlenkamp et al., 2015; Penningroth et al., 2007). While an alternate explanation for our result is that students with stronger CT skills may be more likely to be retained in their academic program, when we looked at only fourth-year students in our sample, we still found a moderate positive correlation between psychology credits and PCTE score. Thus, this observation was in line with our general conclusions, and we find our results promising for educators who are seeking to implement instructional interventions to strengthen students' CT skills in psychology.

Looking at PCTE scores, we observed a somewhat lower mean in our sample as compared to the instrument's developers (see Lawson et al., 2015). One explanation for this finding is that Lawson et al.'s (2015) normative data were

obtained *after* senior psychology students received feedback on practice questions similar to those on the PCTE, whereas our students had no prior experience with such CT-related questions. Therefore, it is possible that PCTE scores are sensitive to changes based on CT instruction or practice. Our observed PCTE scores, however, were similar to those reported by other investigators on their 7-item pretest versions when using the corrected total scores (see Branch & Dubow, 2021; Rodefer et al., 2020; Tolbert, 2020). Finally, a limitation of the present study is that the sample was restricted to psychology undergraduates at one small, religiously affiliated university in the southeast U.S., and we cannot be certain about the degree to which results generalize to other student populations.

One might expect students with a higher need for cognition to demonstrate greater CT skills; however, this was not the case in our study. Our finding was in line with research showing weak correlations between CT dispositions and skills among college students, such that CT disposition explains as little as 4% of the variability in measured CT skills (Facione & Facione, 1997). One explanation for the weak relationship may be due to a discrepancy between individuals' judgment of their CT ability and their demonstrated CT skills. For example, according to NACE's (2019) 2019 student survey report, 79% of college seniors perceive themselves as "very/extremely proficient" in CT. However, as noted previously, many employers perceive that new hires are lacking in this area (NACE, 2018). Moreover, in another study of adults who had taken college-level courses and were asked to apply CT skills to online materials and evaluate sources of information, researchers found participants greatly overestimated their ability to think critically (MindEdge, 2019).

Program directors and instructors should weigh the pros and cons of various general and subject-specific CT tests and consider the purpose of assessment when selecting particular instruments. CT comprises the second of five inclusive goals for undergraduate psychology programs specified by APA, and the PCTE is currently the primary test used to evaluate CT learning outcomes among graduating seniors in our department. We have also started using the PCTE to evaluate specific courses in our program, as well as new instructional interventions in those courses, that we believe may most impact CT. In examining the studies that have used either the original or revised PCTE, interventions that led to the greatest increase in scores (e.g., Branch & Dubow, 2021; Rodefer et al., 2020) involved a direct infusion approach utilizing CT instruction, practice questions, and feedback. This is consistent with one recommended approach to CT instruction whereby general CT principles are made explicit (Abrami et al., 2008; Ennis, 1989).

However, one drawback of the PCTE is completion time, with many of our students taking 45 min in class for the paper-and-pencil format. If the purpose of assessment is to track changes in CT skills over time, then instructors should be aware that two full class periods may be required for test administration. In an effort to improve testing efficiency, beginning in 2019, we administered the PCTE online, with a median time to complete of 34 min. When applying the Satterthwaite approximation due to unequal group variances, there was no significant difference in PCTE scores between graduating seniors who completed testing in paper-and-pencil ($M = 17.91$, $SD = 5.37$) versus online format ($M = 19.53$, $SD = 8.32$), $t(33) = -0.706$, $p = .485$. Therefore, we believe a strength of the PCTE is that it can be administered online.

Another strength of the PCTE is that item-level results allow faculty to compare students' performance across specific CT concepts. For example, we found that "falsifiability" and "correlation does not equal causation" were the most difficult concepts for our students, while Lawson (1999) identified "understanding multiple causes" as the most missed concept in their sample. Such item-level results can help inform course and program improvements as educators seek to strengthen particular CT skills. For example, our program is using item-level feedback to adjust which topics and activities to emphasize moving forward in relevant courses.

In summary, we know that CT skills are valued in higher education as well as in the workforce, as they are believed to contribute to both student and employee success. Yet, college graduates continue to overestimate their ability to think critically. This overconfidence bias is unsurprising, given that the majority of postsecondary institutions are not measuring general CT skills using standardized tests (Association of American Colleges and Universities [AAC&U], 2011). In our view, more widespread assessment of CT skills could shine a bigger light on this discrepancy.

It is also reasonable to surmise that most psychology departments are not measuring psychology-specific CT skills. According to the APA (2013), psychological CT is an important skills-based goal for several reasons. One reason is that strong psychological CT skills help reduce students' susceptibility to common thinking fallacies, thereby making them better consumers of scientific information. Additionally, stronger psychological CT skills are associated with better reasoning and problem-solving skills, which allow individuals to make better decisions and draw sound conclusions about their own and others' behavior. For these reasons, we advocate for use of the PCTE as one tool to assess students' psychological CT skills. Currently, the PCTE is the only subject-specific CT test recommended in the APA guidelines for the undergraduate psychology major (2013), and in order to increase acceptability and use of this test, it would be helpful to establish norms using larger and more diverse samples. Future research should also continue to explore the usefulness of the PCTE in examining targeted instructional interventions aimed at enhancing psychology-based CT skills. Psychometrically sound tests can help document the state of students' critical thinking, which is an important educational objective identified by the U.S. Department of Education, the APA, and administrators at 95% of U.S. institutions of higher learning (AAC&U, 2011).

References

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, Dai. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research*, 78(4), 1102–1134. <https://doi.org/10.3102/0034654308326084>

- American Psychological Association. (2013). *APA guidelines for the undergraduate psychology major: Version 2.0*. <https://www.apa.org/ed/precollege/about/psymajor-guidelines.pdf>
- Anderson, P. R., & Reid, J. R. (2013). The effect of critical thinking instruction on graduates of a college of business administration. *Journal of Higher Education Theory and Practice*, 13(3–4), 149–167.
- Association of American Colleges and Universities. (2011). *The LEAP vision for learning: Outcomes, practices, impact, and employers' view*.
- Bensley, D. A., & Baxter, C. (2006). *The critical thinking in psychology test* [Unpublished manuscript]. Frostburg State University.
- Bensley, D. A., Crowe, D. S., Bernhardt, P., Buckner, C., & Allman, A. L. (2010). Teaching and assessing critical thinking skills for argument analysis in psychology. *Teaching of Psychology*, 37(2), 91–96. <https://doi.org/10.1080/00986281003626656>
- Bensley, D. A., & Murtagh, M. P. (2012). Guidelines for a scientific approach to critical thinking assessment. *Teaching of Psychology*, 39(1), 5–16. <https://doi.org/10.1177/0098628311430642>
- Branch, J. G., & Dubow, E. F. (2021). Research experience in recitation sections within introductory psychology courses: The effectiveness of a semester-long laboratory component. *Scholarship of Teaching and Learning in Psychology*, 7(2), 149–159. <https://doi.org/10.1037/stl0000187>
- Butler, H. A. (2012). Halpern Critical Thinking Assessment predicts real-world outcomes of critical thinking. *Applied Cognitive Psychology*, 26(5), 721–729. <https://doi.org/10.1002/acp.2851>
- Butler, H. A., Pentoney, C., & Bong, M. P. (2017). Predicting real-world outcomes: Critical thinking is a better predictor of life decisions than intelligence. *Thinking Skills and Creativity*, 25, 38–46. <https://doi.org/10.1016/j.tsc.2017.06.005>
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131. <https://doi.org/10.1037/0022-3514.42.1.116>
- Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13
- Cascio, T. V. (2020). Methodological thinking drills: A novel technique for boosting real-world methodological thinking ability. *Teaching of Psychology*, 47(1), 63–67. <https://doi.org/10.1177/0098628319888093>
- Cheung, C.-K., Rudowicz, E., Kwan, A. S. F., & Yue, X. D. (2002). Assessing university students' general and specific critical thinking. *College Student Journal*, 36(4), 504–525.
- Daly, W. M. (2001). The development of an alternative method in the assessment of critical thinking as an outcome of nursing education. *Journal of Advanced Nursing*, 36(1), 120–130. <https://doi.org/10.1046/j.1365-2648.2001.01949.x>
- Dwyer, C. P., Boswell, A., & Elliot, M. A. (2015). An evaluation of critical thinking competencies in business settings. *Journal of Education for Business*, 90(5), 260–269. <https://doi.org/10.1080/08832323.2015.1038978>
- Educational Testing Service. (2014). *Major Field Test in psychology*.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, 18(3), 4–10. <https://doi.org/10.3102/0013189X018003004>
- Facione, N. C., & Facione, P. A. (1997). *Critical thinking assessment in nursing education programs: An aggregate data analysis*. California Academic Press.
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations*. California Academic Press.
- Gadzella, B. M., Ginther, D. W., & Bryant, G. W. (1997). Prediction of performance in an academic course by scores on measures of learning style and critical thinking. *Psychological Reports*, 81(2), 595–602. <https://doi.org/10.2466/pr0.1997.81.2.595>
- Halonen, J., & Gray, C. (2000). *The critical thinking companion for introductory psychology* (2nd ed.). Worth.
- Haw, J. (2011). Improving psychological critical thinking in Australian university students. *Australian Journal of Psychology*, 63(3), 150–153. <https://doi.org/10.1111/j.1742-9536.2011.00018.x>
- Jern, A. (2018). A preliminary study of the educational benefits of conducting replications in the classroom. *Scholarship of Teaching and Learning in Psychology*, 4(1), 64–68. <https://doi.org/10.1037/stl0000104>
- Kalelioğlu, F., & Gülbahar, Y. (2014). The effect of instructional techniques on critical thinking and critical thinking dispositions in online discussion. *Journal of Educational Technology & Society*, 17(1), 248–258. <https://www.jstor.org/stable/jeductechsoci.17.1.248>
- Lawson, T. J. (1999). Assessing psychological critical thinking as a learning outcome for psychology majors. *Teaching of Psychology*, 26(3), 207–209. <https://doi.org/10.1207/S15328023TOP260311>
- Lawson, T. J., Jordan-Fleming, M. K., & Bodle, J. (2015). Measuring psychological critical thinking: An update. *Teaching of Psychology*, 42(3), 248–253. <https://doi.org/10.1177/0098628315587624>
- Lee, I. A., & Preacher, K. J. (2013, September). *Calculation for the test of the difference between two dependent correlations with one variable in common* [Computer software]. <http://quantpsy.org/corrttest/corrttest2.htm>
- McCutcheon, L. E., Apperson, J. M., Hanson, E., & Wynn, V. (1992). Relationships among critical

- thinking skills, academic achievement, and misconceptions about psychology. *Psychological Reports*, 71(2), 635–639. <https://doi.org/10.2466/pr0.1992.71.2.635>
- McPeck, J. E. (1990). Critical thinking and subject specificity: A reply to Ennis. *Educational Researcher*, 19(4), 10–12. <https://doi.org/10.3102/0013189X019004010>
- MindEdge. (2019). *The third annual state of critical thinking study*. <https://digdeeper.mindedgeonline.com/>
- Muehlenkamp, J. J., Weiss, N., & Hansen, M. (2015). Problem-based learning for introductory psychology: Preliminary supporting evidence. *Scholarship of Teaching and Learning in Psychology*, 1(2), 125–136. <https://doi.org/10.1037/stl0000027>
- National Association of Colleges and Employers. (2018). *Job outlook 2019*. <https://www.nacweb.org/store/2018/job-outlook-2019/>
- National Association of Colleges and Employers. (2019). *The 2019 student survey report: Attitudes and preferences of bachelor's degree students at four-year schools*. <https://www.nacweb.org/store/2019/2019-nace-student-survey-report-four-year-schools/>
- PayScale. (2016). *2016 workforce-skills preparedness report*. <https://www.payscale.com/data-packages/job-skills>
- Penningroth, S. L., Despain, L. H., & Gray, M. J. (2007). A course designed to improve psychological critical thinking. *Teaching of Psychology*, 34(3), 153–157. <https://doi.org/10.1080/00986280701498509>
- Renaud, R. G., & Murray, H. G. (2008). A comparison of a subject-specific and general measure of critical thinking. *Thinking Skills and Creativity*, 3(2), 85–93. <https://doi.org/10.1016/j.tsc.2008.03.005>
- Ricker, J. P. (2005). *Teaching and assessing skeptical and critical thinking in abnormal psychology* [Unpublished manuscript]. Scottsdale Community College. <https://doi.org/10.13140/RG.2.1.4914.9686>
- Ricker, J. (n.d.-a). *Assessment of critical thinking at Scottsdale Community College: A pilot study (2003–2004)* [Unpublished manuscript]. Scottsdale Community College.
- Ricker, J. P. (n.d.-b). *Use of the Scottsdale Test of Critical Thinking in assessing critical thinking in ENG 101 and ENG 102* [Unpublished manuscript]. Scottsdale Community College.
- Rodefer, J. S., Lanz, J. L., Synek, S., & Rokusek, B. (2020, January 3–6). *Assessing the implementation of a critical thinking intervention in traditional and online courses* [Poster presentation]. 42nd Annual National Institute on the Teaching of Psychology, St. Pete Beach, FL, United States.
- Rugutt, J., & Chemosit, C. C. (2009). What motivates students to learn? Contribution of student-to-student relations, student-faculty interaction and critical thinking skills. *Educational Research Quarterly*, 32(3), 16–28.
- Stark, E. (2012). Enhancing and assessing critical thinking in a psychological research methods course. *Teaching of Psychology*, 39(2), 107–112. <https://doi.org/10.1177/0098628312437725>
- Tiruneh, D. T., De Cock, M., Weldelessie, A. G., Elen, J., & Janssen, R. (2017). Measuring critical thinking in physics: Development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education*, 15(4), 663–683. <https://doi.org/10.1007/s10763-016-9723-0>
- Tolbert, M. W. (2020). *The relationship between the flipped classroom and critical thinking, academic performance, student perceptions, and student evaluations in an introductory psychology course* [Doctoral dissertation, University of Tennessee at Chattanooga]. UTC Scholar. <https://scholar.utc.edu/theses/634/>
- Williams, R. L., Oliver, R., Allin, J. L., Winn, B., & Booher, C. S. (2003). Psychological critical thinking as a course predictor and outcome variable. *Teaching of Psychology*, 30(3), 220–223. https://doi.org/10.1207/S15328023TOP3003_04
- Williams, R. L., Oliver, R., & Stockdale, S. (2004). Psychological versus generic critical thinking as predictors and outcome measures in a large undergraduate human development course. *The Journal of General Education*, 53(1), 37–58. <https://doi.org/10.1353/jge.2004.0022>

Received August 2, 2021

Revision received January 11, 2022

Accepted January 16, 2022 ■