Cultivating Critical Thinking: Design and Evaluation of Training Modules for College Students

Yi Song[1], Brent Bridgeman[1], Zhitong Yang[2], & Chen Li[1]

[1] ETS Research Institute, Princeton, New Jersey, United States

[2] Graduate School of Education, University of Pennsylvania, United States

Abstract

Developing critical thinking skills is an important goal of higher education, as it can lead to better academic outcomes, improved problem-solving abilities, and greater success in the workforce. In this research project, we designed multiple critical thinking training modules to assess and support the development of critical thinking skills and collected empirical evidence from 84 college students with various backgrounds. Our pilot results showed that the tasks demonstrated acceptable reliability for assessing students' critical thinking skills, students' GPAs were significantly correlated with their critical thinking task scores, and students performed better on the second set of Research module tasks than on the first set. Students' survey responses highlighted overall positive perceptions on the critical thinking training modules. Students viewed the modules as a useful tool for developing critical thinking skills, recognizing their relevance to academic work and appreciating their engaging content and appropriate difficulty level. Students also valued the feedback provided within the modules, which they perceived as constructive and helpful for supporting their understanding and performance. This research has implications on designing tools that support the development and assessment of critical thinking skills in higher education.

Keywords: critical thinking, college students, training modules, assessment

Corresponding author: Yi Song, Email: ysong@ets.org

Acknowledgement

We thank Srinivasa Pillarisetti for programming the critical thinking training modules. We also want to thank Tenaha O'Reilly, Jesse R. Sparks, Jessica Andrews Todd, and Beata Beigman Klebanov for providing valuable feedback on a draft of this article.

Introduction

Multiple nationwide survey studies with employers and higher education institutions have pointed to the importance of critical thinking (Dondi et al., 2021; Finley, 2021). Critical thinking in the workplace is a "must-have" skill for solving problems, gathering multiple sources of evidence and perspectives, evaluating important information, and making well-informed decisions. According to a survey report, ninety-five percent of the 496 executives and hiring managers at U.S. companies rated critical thinking as either very or somewhat important (Finley, 2021). Furthermore, prior studies suggest that students with strong critical thinking skills tend to perform better academically (Halpern, 2003a; Hawkins, 2012; Liu et al., 2016), underscoring the importance of nurturing critical thinking skills for improved educational outcomes. The importance of critical thinking becomes even more pronounced in an age marked by the rapid proliferation of information. Today's youth and young adults are constantly bombarded with an overwhelming volume of information every day, since they are frequent users of social media. Thus, students must think critically about the information they are exposed to, demonstrating skills in evaluating the credibility of the sources, making sound inferences, and considering conflicting evidence, as without these skills, they may be misled by unreliable sources (McGrew et al., 2018). The need for critical thinking is further heightened in the era of AI, where students increasingly encounter AI-generated information, explanations, and arguments. Although the present study did not involve AI tools, research has shown that critical thinking is foundational for students' responsible engagement with AI-generated content (e.g., Melisa et al., 2025; Ng et al., 2021; Spector & Ma, 2019).

Despite its importance, there exist gaps between the skills that students mastered upon completion of their higher education and the skills that are required in the workplace (Dondi et al., 2021; Haviland et al., 2022). Many college students struggle to understand the concept of critical thinking, encounter difficulties in its application, and are unsure how they can develop critical thinking skills (Aston, 2024; Stupple et al., 2017). Students who lack critical thinking will be struggling with various tasks in their studies, such as analyzing and interpreting data, drawing conclusions, conducting literature reviews, writing papers, and making evidence-based decisions. However, efforts supporting college students to develop these critical competencies have been somewhat limited.

In this project, we aim to overcome the obstacle of constrained resources for students by designing online training modules and conducting a study to explore how college students of varying ethnic and academic backgrounds engaged with the modules. By analyzing their interactions and feedback, we sought to gather empirical evidence about the usefulness of our training modules. This research addresses the needs of developing students' critical thinking skills to enable them to demonstrate the kinds of thinking needed in accomplishing the academic goals and to be equipped with the higher-order thinking skills for future careers.

Literature Review and Rationale

Critical Thinking Theoretical Frameworks

There is no clear-cut definition for critical thinking because of its complex and multifaceted nature. In the academic context, many definitions for critical thinking stem from the well-known Delphi expert consensus where critical thinking was defined with six core dimensions: interpretation, analysis, evaluation, inference, explanation, and self-regulation (Facione, 1990). These include sub-skills such as categorization, clarifying meaning, examining ideas, analyzing arguments, assessing claims, drawing conclusions, stating results, justifying procedures, presenting arguments, self-examination, and self-correction. The experts also raised the importance of the disposition toward critical thinking, calling that "a critical spirit." Good critical thinkers are habitually disposed to engage in critical judgment (e.g., open-mindedness regarding divergent views, honesty in facing one's own biases and stereotypes, and willingness to reconsider and revise views).

Other influential research around critical thinking was conducted by Halpern (2003b), who described critical thinking as purposeful, reasoned, and goal-directed, playing a key role in decision-making, problem-solving, verbal reasoning, argument analysis, assessing likelihood and uncertainty, and hypothesis testing. She perceived critical thinking as a tool to facilitate decision-making and problem-solving. Additionally, Dwyer et al. (2014) developed an integrated critical thinking framework that identified and organized six key learning outcomes including memory, comprehension, analysis, evaluation, inference, and reflective judgment. They focused on the thinking component, describing how metacognition and self-regulation play an important role in understanding critical thinking in action. Further, Liu et al. (2014) published a critical thinking framework for higher education to develop operational measures of critical thinking, which put an emphasis on academic reading, writing, and reasoning skills. It consisted of two key

dimensions: (a) *analysis*, which includes evaluating evidence and its use and analyzing and evaluating arguments; and (b) *synthesis*, which includes understanding implications and consequences and developing sound and valid arguments. Relevant to both analysis and synthesis is the ability to understand causation and explanation.

At its fundamental essence, critical thinking involves the ability to analyze and evaluate information, make reasoned judgments and decisions, and solve problems effectively. These skills can be applied in a wide range of contexts, from academic disciplines to everyday life. Advocates of a generic approach, including Halpern (1998) and van Gelder (2005), believe critical thinking is intrinsically general in nature and can be taught outside a particular domain. Their stance underscores the malleability of critical thinking as a cognitive skill that can be honed independently of disciplinary boundaries. In contrast, proponents of domain specificity (Bailin, 2002; Garside, 1996; McPeck, 1990) perceive critical thinking skills as particular to each academic discipline and domain-specific knowledge as essential because what constitutes valid evidence, standards and arguments varies across disciplines. In this view, critical thinking intertwines with domain expertise, as the mastery of subject-specific conventions and contextual intricacies bolsters one's capacity to engage in rigorous critical analysis. Critical thinking is one of several generic attributes that are conceptualized and taught in different ways within each discipline.

In reality, the delineation between the two perspectives is often not starkly demarcated. Ennis (1989) and Facione (1990) proposed a middle ground, recognizing that while critical thinking possesses overarching principles, it also harbors elements that adapt to the particularities of different domains. This view harmonizes the recognition of general cognitive skills with the acknowledgment of the context-sensitive nature of critical thinking. A synthesis of these viewpoints, as exemplified by Davies (2013), underscores the existence of both general and domain-specific facets within the realm of critical thinking, wherein proficiency in one domain can bolster the ability to navigate others, albeit with a degree of contextual adaptation.

While the application of critical thinking often requires domain-specific knowledge, the overarching principles of analysis, evaluation, and problem-solving continue to resonate across various fields. The Interstate Passport program[1] administered by the Western Interstate Commissions for Higher Education highlights critical thinking as a cross-disciplinary process that involves inquiry, analysis, and problem solving as well as a habit of mind characterized by

thorough exploration of issues before making a decision or reaching a conclusion. The program suggests that critical thinking can be taught in a single course or multiple courses across lower-division general education curriculum (i.e., introductory courses designed for first-year and second-year students), such as science, sociology, history, and literature. This approach may help establish a strong foundation for overall academic success and intellectual growth. General critical thinking skills, such as problem-solving, logical reasoning, and information evaluation, serve as a universal toolkit that can be applied across academic disciplines and real-world contexts. By developing these skills, students can further develop the cognitive flexibility needed to adapt and excel in different domains. A general approach also helps students transfer their critical thinking skills to tackle complex, multifaceted challenges that go beyond specific subjects. Therefore, in this project, we took a general approach to design critical thinking training modules for college students from a variety of academic backgrounds.

Development of Critical Thinking Skills, Challenges, and Needs

According to a survey study conducted by the Hart Research Associates (2018), 84% of hiring managers and 78% of business executives identified critical thinking/analytical reasoning as a very important skill for college graduates. However, only 41% of hiring managers and 34% of business executives thought that recent college graduates were adequately prepared to apply critical thinking or analytical reasoning in the workplace (Hart Research Associates, 2018). Similar results emerged in a survey with 201 employers conducted by NACE. The results showed that even though 99.2% of U.S. employers considered that critical thinking and problem-solving skills crucial for job roles, only 55.8% of them believed that college graduates were proficient in these skills (NACE, 2017).

Several studies have examined the development of college students' critical thinking skills but reached mixed conclusions. Arum and Roksa (2011) found that a sample of approximately 2,300 students from 24 four-year institutions showed only a .18 standard deviation (SD) improvement on the Collegiate Learning Assessment (CLA), a standard measure of critical thinking in terms of analytic reasoning and problem solving, over three semesters. Additionally, 45% of the students displayed no improvement at all. In their subsequent study, Arum and Roksa (2014) reported a 4-year longitudinal gain of around .47 SDs in critical thinking. Another longitudinal study analyzed a sample of 2,200 students from 17 four-year institutions and revealed that the progress in critical thinking abilities, measured by the

Collegiate Assessment of Academic Proficiency (CAAP), was recorded at .44 SDs (Blaich & Wise, 2011). A meta-analysis conducted by Huber and Kuncel (2016) indicated that there is no clear pattern when the increasing gains in critical thinking of college students most likely take place. Possible explanations of these mixed results include: (1) different measures of critical thinking skills, and (2) samples varying by major, institution, and design (e.g., longitudinal vs. cross-sectional samples). Nonetheless, it appears that merely enrolling in a 4-year college program will not necessarily build a student's critical thinking skills to the level needed by the workforce (e.g., Hart, 2018), highlighting the necessity for more targeted training of these skills in postsecondary education.

Some researchers reported significant differences in critical thinking skills across various majors. For example, Liu et al. (2016) measured critical thinking using the ETS Proficiency Profile that targets skills of argumentation, recognizing hypotheses, drawing valid conclusions, etc. Their sample included cross-sectional freshmen and seniors in Education, Engineering, Health and Human Development, Business, and Social Sciences majors. They reported larger cross-sectional differences in students (between freshmen and seniors) in Education, Engineering, and Health and Human Development as compared to students in Business and Social Sciences. Further, Liu et al. (2022) investigated the developmental trends of 2,381 college students' critical thinking skills in their journey of pursuing a 4-year college degree. They found Natural Science majors achieved the highest overall critical thinking scores and had the largest cross-sectional difference from freshmen to senior year, measured by HEIghten Critical Thinking Assessment, and had one of the highest percentages of students at the Advanced level. These variations may be attributed to the distinct cognitive demands and problem-solving approaches inherent to each major.

The measures used to examine critical thinking development also vary in their underlying conceptualizations and assessment targets. In a review of 28 studies (conducted from 2013 to 2023) of performance-based critical thinking assessments in higher education, Hyytinen et al. (2025) found that most studies conceptualized critical thinking as a skill or set of skills and primarily used open-ended performance tasks. Evaluation and inference were the most frequently assessed skills, whereas dispositions, although included in some definitions of critical thinking, were not explicitly represented in assessment targets. The reviewed studies also lacked a

consistent theoretical framework, revealing potential misalignment between definitions of critical thinking and their operationalization in assessments.

A challenge faced by both students and educators is that the perceived complexity of critical thinking can hinder effective teaching and learning of these skills (e.g., Stupple et al., 2017). Students' attitudes and beliefs regarding critical thinking may influence their ability to demonstrate or further develop such skills. For example, Duro et al. (2013) conducted interviews and focus groups with psychology students and reported that some students perceived critical thinking as identifying limitations and emphasizing negative aspects, and some perceived themselves as lacking proficiency in critical thinking. Thus, encouraging students to recognize that critical thinking involves both identifying limitations and appreciating positive aspects can foster a more comprehensive and balanced perspective on this essential skill. Furthermore, addressing the self-perception of not being a good critical thinker may require targeted interventions and educational support to boost students' self-efficacy in critical thinking.

Some institutions offer courses in their general education curriculum to foster critical thinking, but a well-defined and structured learning and assessment approach may be needed to ensure their effectiveness. In many disciplines, critical thinking is often linked with reasoning, analysis, and evaluation of research evidence (e.g., Bensley, 1998). Consequently, when teaching critical thinking in academic contexts, the focus often lies on nurturing reasoning and problem-solving competencies (van Gelder, 2005). Despite recognizing the importance of critical thinking (DeAngelo et al., 2009), educators still need resources to incorporate it into their teaching (e.g., Halx & Reybold, 2006; Harris, 2021; Paul et al., 1997). For example, Silva et al. (2025) investigated faculty perspectives on integrating critical thinking into higher education courses and identified limited instructional time, insufficient training and assessment resources, and a lack of shared understanding of critical thinking as key barriers to its systematic integration. Liu et al. (2022) recommended that institutions develop plans and strategies to help students improve their critical thinking skills inside and outside classrooms, integrating critical thinking instruction into their courses and utilizing external resources. They also called for more research to identify high-impact programs and practices that promote critical thinking. Rothinam et al. (2025) synthesized 50 empirical studies and found that active learning, interdisciplinary instruction, and digitally enhanced learning environments can support critical thinking development. Their review also pointed out the influence of disciplinary and cultural contexts and recommended

student-centered instruction, blended assessment approaches, and institutional investment in relevant instructional resources. Together, these findings highlight the need for flexible, evidence-informed approaches that can provide students with structured opportunities to practice critical thinking across learning contexts. In response to these instructional and resource needs, our research team explored an online training approach to support critical thinking development. We started with designing three online training modules to cultivate critical thinking skills valued in postsecondary education.

Designing Online Critical Thinking Training Modules

Drawn on the existing critical thinking frameworks (e.g., Facione, 1990; Liu et al., 2014), our design blueprint focused on critical thinking skills as part of critical competencies valued in higher education and the workforce, such as evaluating arguments, identifying assumptions, making inferences, interpreting data, and drawing valid conclusions. The three modules were designed to progress from a general introduction to the application of these skills in argument analysis and research evaluation. We selected Argument and Research as the focal modules because analyzing arguments and critically evaluating research evidence and conclusions are central across critical thinking frameworks and commonly required in college coursework. To support the relevance and content representation of the targeted skills, we consulted 12 professors in various domains, including engineering, biochemistry, psychology, sociology, English, political science, and education, from nine universities. These content experts reviewed the module tasks, which we subsequently refined based on their feedback.

The critical thinking modules include both instructional and assessment components and can be delivered in a self-paced or instructor-led mode. To maximize learning opportunities, we built in immediate, explanatory feedback in assessment tasks. Immediate feedback can be viewed as a kind of scaffolding within a Vygotskian framework (Vygotsky, 1978), helping correct misconceptions and reinforce the rationales for critical thinking. Immediate feedback has been shown to enhance learning outcomes, helping students to achieve higher levels of academic success in some studies (Hattie & Timperley, 2007; van der Kleij et al., 2012).

When conceptualizing the instructional videos, our team acknowledged and embraced sociocultural diversity, recognizing its significance in creating a truly inclusive educational experience. For example, we created diverse representations of characters in the videos (see Figure 1) that are responsive to students' cultural identities, which has been recommended as one

design principle for culturally relevant classroom assessments (O'Dwyer et al., 2023). We incorporated real-world, authentic examples that reflect various cultures and perspectives. By doing so, we aimed to provide students with a relatable and enriching learning experience, allowing them to connect the content to their own lives and experiences. We also crafted passages with appropriate reading levels to ensure that students with varying reading skills can engage with the material effectively and independently. The length of the passages ranged from 61 to 243 words (M = 113.88, SD = 51.48), and the text difficulty measured by Flesch-Kincaid ranged from 6.2 to 13.4 (M = 9.74, SD = 1.96), with the majority reading level at the high school level.

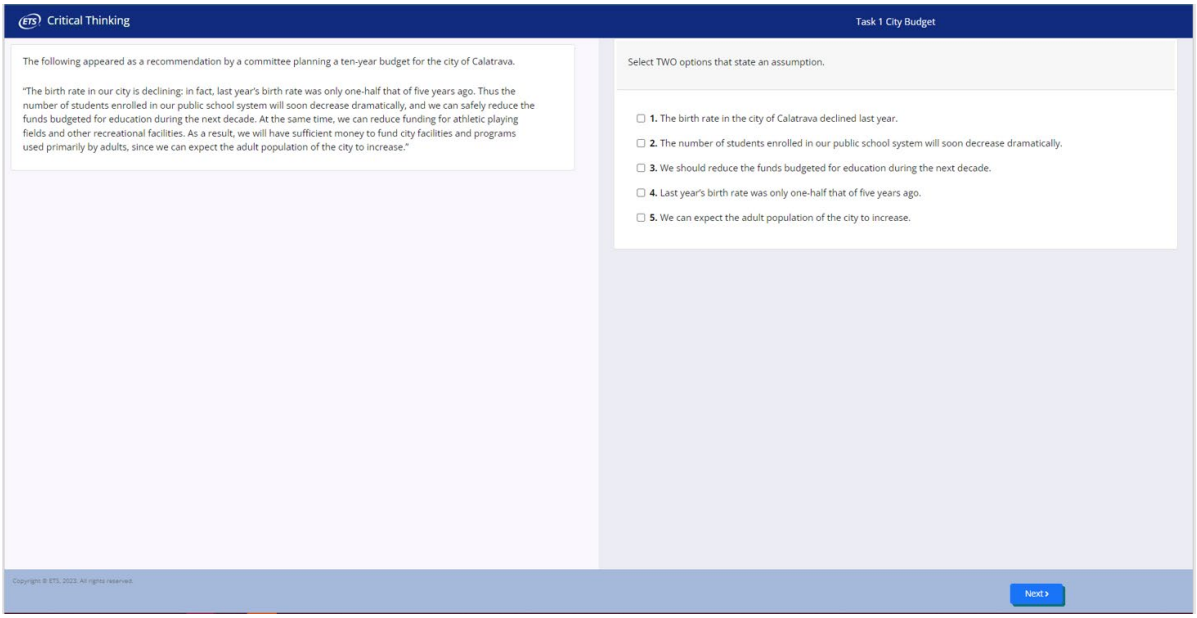
Figure 1. Screenshot of a Video in the Argument Module That Includes Characters From Different Ethnicities/Backgrounds



The first module (Overview) provides an overview of critical thinking that explains its importance and draws students' attention to its relevance to their study and everyday life. The second module (Argument) targets evaluating arguments and identifying assumptions. Students learn to distinguish valid arguments from invalid ones and identify underlying assumptions in the given arguments. It includes animated instructional videos about argument components, argumentation schemes, and critical questions, as well as practice tasks that students can

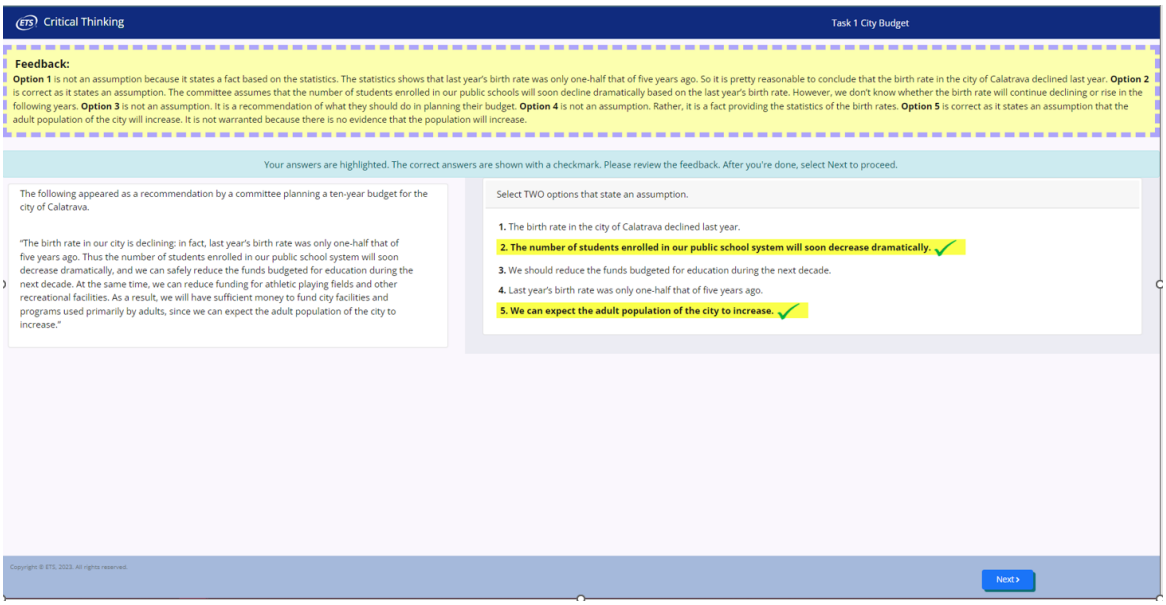
complete to apply these concepts and receive feedback. The last section is an assessment that consists of 14 selected-response items and two constructed-response tasks that measure students' critical thinking skills about evaluating arguments and identifying assumptions (see Figure 2 for a sample task).

Figure 2. Argument Module Assessment Task Example



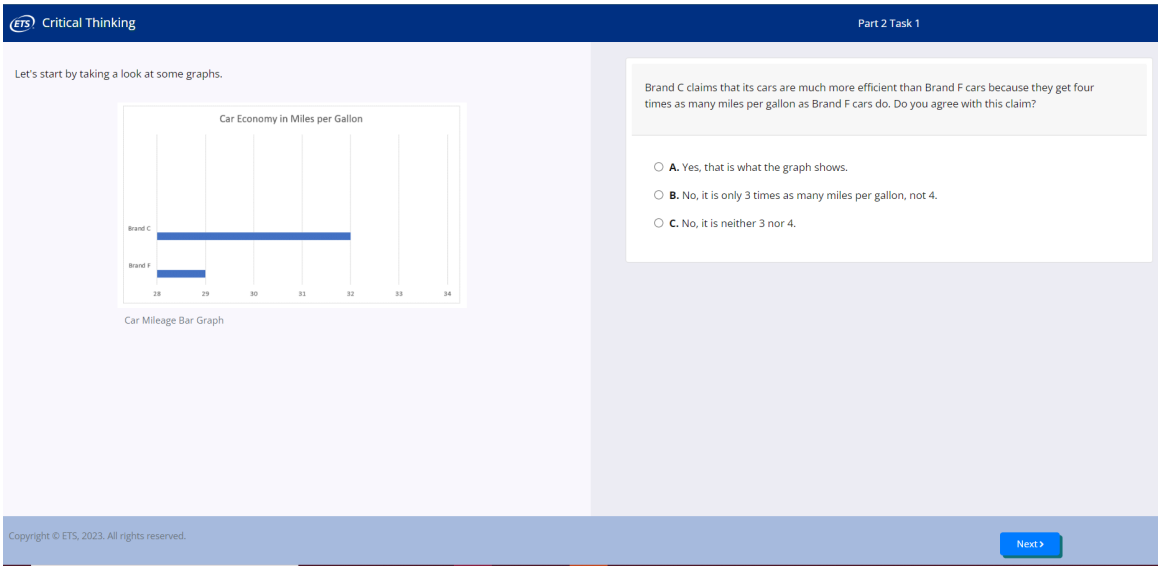
We designed immediate task feedback in this module. Such feedback not only shows the key, but also provides a content-based explanation to establish and reinforce the rationales that are critical for identifying and evaluating arguments, for the purpose of increasing the chance that struggling learners can learn and apply the target skills in subsequent tasks (see Figure 3 for a sample feedback screen).

Figure 3. Task Feedback Example



The third module (Research) focuses on thinking critically about research designs and conclusions. After the introductory video, students work on two sets of 10 selected-response items that were created based on the same blueprint, with the order of the sets randomly assigned (see Figure 4 for a sample task).

Figure 4. Research Module Assessment Task Example



Both sets have immediate task feedback that articulates the rationales for correct and incorrect task responses, which can be viewed as scaffolding. This design allows us to explore

whether students' performance improved from the first set to the second set after receiving feedback. In these modules, students not only watch animated instructional videos, but also have opportunities to work on a range of purposefully crafted tasks to practice the targeted skills.

We conducted a pilot study to evaluate how well the items function in the modules and whether students benefited from taking the training. We collected student feedback on their experience with the critical thinking training to inform potential modifications of the modules. Our research questions included the following:

- Do the items provide a reliable measurement of students' critical thinking skills?
- Do students show improved performance across two sets of Research module tasks?
- What are students' perceptions of the critical thinking training?

Method

Participants

Three U.S. institutions (a private research university, a state university, and a private historically Black college) agreed to participate and helped recruit their students. Among 96 students who participated, 84 completed all three training modules (Institution 1: $N = 26$; Institution 2: $N = 27$; Institution 3: $N = 31$) and were included in our data analyses. This sample represented a diverse group in terms of race/ethnicity: African American 35.7%; Asian 28.6%; Caucasian 19%; Hispanic 2.4%; Native American 1.2%; Mixed 6%; Unreported 7.1%. Slightly more than 51% of the students were male, and 45.2% were female (3.6% did not report gender). Roughly 57% of the participants were undergraduate students. Approximately 55% were STEM majors, including engineering, computer science, data science, etc.; 44% were non-STEM majors, such as business and sports management.

The three institutions had different participant profiles (see Table 1): Institution 1 participants consisted of 50% Caucasian students; 70% of the participants from Institution 2 were Asian, and many of them were international students in a graduate engineering program; participants from Institution 3 were undergraduates, and the majority were African American. This sample provided us with the opportunity to investigate how students from different backgrounds engaged with the critical thinking training modules.

Table 1. Participant Demographic Information by Institution

		Institution 1 (*N* = 26)	Institution 2 (*N* = 27)	Institution 3 (*N* = 31)
Gender	Male	46.2%	74.1%	35.5%
	Female	50.0%	22.2%	61.3%
Ethnicity	African American	7.7%	3.7%	87.1%
	Caucasian	50%	11.1%	0%
	Asian	19.2%	70.4%	0%
Year in College	Undergraduate	57.7%	11.1%	100%
	Graduate	42.3%	88.9%	0%
STEM	Yes	73.1%	100%	0%

Note. There were 1 or 2 missing cases from each institution, so the sum of the values did not equal 100%.

Measures

Training Modules

Students took three critical thinking training modules: overview, argument, and research. The Overview module does not have any measure of critical thinking. The Argument module has tasks with a total of 22 points, and the Research module has tasks with a total of 20 points. Selected-response items in the training modules were automatically scored. The two constructed-response tasks in the Argument module were human scored. Both constructed-response tasks were scored from 0 to 3 points. Each valid response earned 1 point. Students needed to provide three valid written responses in each task to get all 3 points. Two research assistants received training to apply the scoring rubric, and then scored students' responses independently. The interrater agreement was 77% for Task 1 (k = .64), and 88% for Task 2 (k = .76). The raters discussed and resolved all the disagreements in their initial scores.

Post-Training Survey

The post-training survey collected information about students' perceptions of the modules, including the difficulty level of the tasks, importance of various critical thinking skills, helpfulness of the feedback and module design, and engagement. Students selected a rating for the above aspects on a 5-point Likert-type scale. They were also given opportunities to provide open-ended written responses (e.g., if you think any aspects of the tasks are too difficult, please explain).

Student Background Information

We collected student background information, including demographics and overall GPA. Because GPA was optional, only 45 students provided this information.

Procedure

Each participant was assigned a unique ID for accessing the online critical thinking training modules. Using their own computers, students worked on the online training modules at their own pace in their spare time. No one reported any technical difficulties completing the modules. Seventy-two students (85.7%) completed the modules on the same day. After the training modules, they answered the online post-training survey. It took students up to two hours to complete this study.

Data Analysis

Data analysis first involved computing item-level descriptive statistics for the Argument module and the Research module (score means and standard deviations, Cronbach's alpha reliability, proportion correct, item-total correlations) and correlations between the Argument and Research module performance. Further, we conducted a paired-sample t-test to compare students' first-set and second-set scores in the Research module, followed by examining score improvements within each institution. For the survey data, we calculated frequencies and percentages to provide an overview of the students' ratings to the survey questions. Open-ended written responses were reviewed to provide illustrative examples and supplementary context for the survey findings (e.g., importance, task difficulty, feedback, engagement, helpfulness, improvement).

Results

Item Analysis and Task Performance

Table 2 presents the descriptive statistics of student performance in the Argument and Research modules. The mean score of the Argument module was 13.6 out of 22 points (SD = 4.10), ranging from 4 to 20 points. The mean score of the Research module was 12.9 out of 20 points (SD = 4.05), ranging from 4 to 20 points. The scores of these two modules had a relatively strong correlation, $r = .65$, $p < .01$. Overall Cronbach's alpha (internal consistency) reliabilities of the Argument and Research modules were acceptable ($\alpha = 0.76$; $\alpha = 0.78$, respectively). The

two Research task sets yielded slightly lower Cronbach's alpha coefficients (α = 0.66; α = 0.67), which may partly reflect the limited numbers in each set (10 items). Forty-five students (53.6%) provided their overall academic GPA (M = 3.68, SD = .38), and we found that their total critical thinking scores were moderately correlated with their GPA, r = .40, p < .01.

Table 2. Descriptives of Module Task Performance

Module	Institution	*N*	Mean	*SD*	Minimum	Maximum
Argument module	Institution 1	26	14.92	4.09	4	20
	Institution 2	27	14.93	3.45	6	20
	Institution 3	31	11.32	3.70	4	18
	Total	84	13.60	4.10	4	20
Research module	Institution 1	26	14.50	3.02	7	19
	Institution 2	27	15.63	3.10	6	20
	Institution 3	31	9.19	2.50	4	15
	Total	84	12.90	4.05	4	20

Further, item-level performance related to difficulty (p+) were calculated as mean score divided by the maximum score of each item. The Argument module item-level p+ values ranged from .36 to .88, indicating a reasonable variation of difficulty. All the Argument items had an adequate item-total correlation, ranging from .26 to .58. Similar results were found for the Research module. Its item-level p+ values ranged from .32 to .81, and all the items had an adequate item-total correlation, ranging from .28 to .62. Taken together, these results showed that the items possess the ability to offer a reliable measurement of students' critical thinking skills.

Learning Outcomes

As mentioned earlier, the Research module contained two parallel sets of items based on the same blueprint, with the order of the sets randomly assigned. There was no significant difference in mean performance between the two Research task sets; therefore, we conducted a paired-samples t-test to compare students' first-set and second-set scores in the Research module. Our analyses showed a significant improvement from the first set (M = 5.96, SD = 2.22) to the second set (M = 6.94, SD = 2.27) scores in the Research module, $t(1,83) = 4.62$, $p < .001$, d = .50. On average, the students increased nearly 1 point out of 10 from the first to second set. While only 13 students got 9 or 10 points in the first set, in the parallel second set 26 students

achieved 9 or 10 points. Despite the differences in student composition across the three institutions, students at each institution showed higher mean scores on the second task set than on the first. The within-institution effect sizes were positive: Institution 1, $d = .76$, Institution 2, $d = .42$; and Institution 3, $d = .52$. Because the institutional samples differed considerably in academic level, major, and demographic composition, we did not conduct between-institution comparisons. These descriptive patterns suggest that students in all three participating samples benefited from completing the modules.

Survey Results

Perceived Importance and Relevance of Critical Thinking

Among 84 participants, 82 completed the post-training survey. The majority of the students rated the critical thinking skills as important or very important: 89.3% for the skill of analyzing the relationship between claims and evidence, 85.7% for the skill of identifying assumptions and unsupported claims, 81% for the skill of evaluating research designs and conclusions, and 78.5% for the skill of distinguishing sound and unsound arguments. Some students provided examples of critical thinking skills required in their college study. A graduate student in computer science and engineering stated: "These skills are fundamental concepts in data science projects. Critical thinking while evaluating data for any data science project is very important to get an unbiased result from the code." A junior wrote: "In one of my engineering design courses, we analyzed a research paper and evaluated its arguments and conclusions." While acknowledging the connection to their study, many students believed that the training extends what they learned in their courses. For instance, a sophomore in finance commented: "The modules extend what I have already learned by going in-depth and giving me real-life experiences that help to stimulate my brain." A graduate student in mechanical engineering wrote: "I did not go through any course regarding critical thinking, most of the thinking I do is based on my previous experience and the mistakes I made by concluding something that did not have solid proof. A module like this will definitely help people in the research community differentiate and conclude arguments and hypotheses."

Task Difficulty and Feedback

More than half of the students perceived the difficulty level of the modules as being just right (61.9% for the Argument module; 56% for the Research module). Only one or two students

thought the modules very easy or very difficult. Most of the students (83.4%) thought the feedback in the modules was helpful or very helpful. A sophomore in finance wrote: "It was helpful because it let me know if my answer was correct, and even when I wasn't, it showed me the correct answers with explanation." A graduate student in engineering responded: "Feedback explained why choices were not correct and why one choice was clearly the correct answer. Even when getting the correct answer explanation is very helpful. This is because sometimes you can take an educated guess but your reasoning may not be concrete. The feedback solidifies your thinking approach." There were quite a few responses from other students showing similar understandings of the feedback value. However, a few students felt that the feedback was a bit lengthy or sometimes difficult to understand.

Student Engagement

Roughly 71% of the students perceived the training modules engaging or very engaging. Students pointed out various features they liked about the modules, including interesting texts and scenarios, good examples, animated videos, various task formats, length of the modules, and questions that require thoughtful responses. A junior in civil engineering commented: "I believe the beginning two modules were very engaging, alternating between videos teaching you the material and the example problems." A graduate student in computer science and engineering wrote: "The video explanation was short and precise hence it conveyed what need to be without losing the interest of the audience mostly due to the long video length. Also, the passages and aspects provided for critical thinking training were of very good quality."

Perceived Helpfulness and Suggested Improvements

Most of the students (83.4%) perceived the training modules as helpful or very helpful in supporting the development of their critical thinking skills. Their comments suggested that students viewed the modules as a promising tool to support the development of their critical thinking skills, for example,

- "The tasks and questions helped students think critically and the feedback expanded the students' thoughts." – a junior in business marketing
- "Yeah, the training module helped sharpen the critical skills. The questions were thought-provoking and also, the feedback helped in deeper understanding." – a graduate student in computer science and engineering
- "I feel as though I have new tools to use when thinking critically about the structure of an argument and the assumptions made, as well as how to properly assess its value." – a junior in aerospace engineering

Only two students rated the modules as "of little help" (2 on a 5-point scale), and neither provided comments. The remaining students took a neutral position regarding the helpfulness of the training modules. Suggested improvements included offering additional explanations for challenging questions, reducing the reading load, incorporating more pictures and videos to present material more effectively, and adding opportunities for collaborative learning.

Discussion

Our research aimed to address a need to support college students in developing critical thinking skills essential for successfully completing their coursework and thriving in their future careers. Specifically, we designed multiple critical thinking training modules grounded in established critical thinking frameworks and refined through content review by 12 faculty experts across multiple disciplines. Then we conducted a pilot study with 84 college students to evaluate the usability and feasibility of the training modules. Our analyses examined item functioning and reliability, preliminary validity evidence, changes in performance across two parallel task sets, factors contributing to student performance, and students' perceptions of the training modules. The pilot study results provided preliminary evidence supporting the potential use of the critical thinking training modules as tools for supporting and assessing college students' critical thinking skills. Our analyses revealed that the tasks were reliable, with a relatively strong correlation between the Argument and Research modules, adequate item-total correlations, reasonable variations of difficulty, and an adequate correlation with students' GPA. Various studies reported that critical thinking scores by various measures had moderate correlations with course grades and GPA (e.g., Hawkins, 2012; Liu & Roohr, 2013). For example, Martín-Raugh et al. (2023) showed that noncognitive skills and critical thinking skills

predict college GPA, in addition to high school GPA and SAT (Scholastic Aptitude Test)/ACT (American College Test) scores. Similarly, our results indicated that students with strong critical thinking skills are more likely to excel academically. Success in a course often requires students to engage deeply with the subject matter, think critically about concepts, and apply their knowledge in various contexts.

Further, students performed better on the second Research task set than on the first set after receiving immediate explanatory feedback. This finding, together with students' positive perceptions of the feedback, indicates the potential usefulness of explanatory feedback embedded within the modules. However, because the study did not include a condition without feedback, the performance improvement cannot be attributed specifically to feedback Overall, the findings provide preliminary evidence that the modules may serve as meaningful learning and assessment tools across diverse educational contexts and student backgrounds. Yet, we recognize that the short duration of the training likely limited the extent of these improvements, as critical thinking development requires continuous, targeted practice over an extended period (Hyytinen et al., 2024; van Peppen et al., 2021). Critical thinking is inherently difficult because it requires a diverse range of cognitive strategies, each of which may be context dependent (Butler, 2024). Students arrive at college with varying levels of prior exposure to critical thinking, making it essential to bridge these disparities in foundational knowledge (Liu et al., 2022). Also, the diversity of academic disciplines and the complexity of subject matter require educators to adapt their teaching strategies to suit different contexts (Dwyer et al., 2011). While we adopted a general approach designed to be applicable to students across various majors and levels of academic performance, future work may benefit from tailoring interventions to account for students' diverse academic backgrounds and experiences. Lastly, engaging students actively in critical thinking often demands significant time and effort (Behar-Horenstein & Niu, 2011). This challenge can lead to prioritize content delivery over skill development (Huber & Kuncel, 2016). Through this line of research, we aim to address this issue by offering a variety of opportunities for students to practice critical thinking skills consistently throughout their academic journey.

Due to the use of a small convenience sample across multiple institutions, these results are best viewed as exploratory. Additionally, the student profiles from the three institutions were so different, making it difficult to disentangle the influence of students' backgrounds from institutional effects. Future research can address this limitation by intentionally selecting

comparable samples to explore the relationship between student backgrounds and their critical thinking performance. Another limitation is lack of a control group for examining learning gains; in future research, we aim to conduct experimental examinations to further explore the impact of the training modules activities on students' critical thinking skill development by including a control condition. Also note that the course GPA only provided an indirect measure of students' critical thinking skills. Future work could include more direct measures of critical thinking skills to strengthen the validity argument about the training modules. Future extensions could also examine how students apply these critical thinking skills when evaluating AI-generated explanations, summaries, or arguments. AI-enhanced tools may support critical thinking development, while over-reliance on AI-generated content may hinder deep engagement in critical thinking (Rothinam et al., 2025).

Despite these limitations, this research provides empirical evidence of the promise of the critical thinking tool in terms of its reliability, usefulness, and engagement. Informed by the study results, we subsequently modified the modules. Based on the improved second-set performance and students' positive perceptions of the feedback, we revised the Argument module to include two sets of parallel tasks that will support a more systematic examination of performance changes in future research. We also revised multiple tasks in the Research module because some students expressed confusion about those tasks. Further, we developed a new module focusing on critically evaluating information in social media or people's conversation on the topic of climate change. In this module, students will listen to conversations and read texts and then evaluate the trustworthy of the information about climate change provided in the conversations and texts. Moving forward, we plan to collaborate with instructors from multiple institutions to integrate the critical thinking training modules with selected courses and identify effective supports that can promote students' development of critical thinking. Through these collaborative efforts, we will gain insights into how the training modules can be seamlessly integrated into college-level instruction, creating a supportive learning environment.

References

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226028576.001.0001

Arum, R., & Roksa, J. (2014). *Aspiring adults adrift: Tentative transitions of college graduates*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226197142.001.0001

Aston, K. J. (2024). 'Why is this hard, to have critical thinking?' Exploring the factors affecting critical thinking with international higher education students. *Active Learning in Higher Education, 25*(3), 537–550. https://doi.org/10.1177/14697874231168341

Bailin, S. (2002). Critical thinking and science and education. *Science & Education, 11*(4), 361–375. https://doi.org/10.1023/A:1016042608621

Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning (TLC)*, *8*(2). https://doi.org/10.19030/tlc.v8i2.3554

Bensley, D. A. (1998). *Critical thinking in psychology: A unified skills approach.* Thomson Brooks/Cole Publishing Co.

Blaich, C., & Wise, K. (2011). *From gathering to using assessment results: Lessons from the Wabash national study* [Occasional Paper No. 8]. National Institute for Learning Outcomes Assessment. https://www.bu.edu/provost/files/2015/09/From-Gathering-to-Using-Assessment-Results_Lessons-from-the-Wabash-Study-C.-Blaich-K.-Wise1.pdf

Butler, H. A. (2024). Predicting everyday critical thinking: A review of critical thinking assessment. *Journal of Intelligence, 12*(2), Article 16. https://doi.org/10.3390/jintelligence12020016

Davies, M. (2013). Critical thinking and the disciplines reconsidered. *Higher Education Research & Development, 32*(4), 529–544. https://doi.org/10.1080/07294360.2012.697878

DeAngelo, L., Hurtado, S., Pryor, J. H., Kelly, K. R., Santos, J. L., & Korn, W. S. (2009). *The American college teacher: National norms for the 2007–2008* [Research brief]. Higher Education Research Institute at UCLA. https://www.heri.ucla.edu/PDFs/pubs/briefs/brief-pr030508-08faculty.pdf

Dondi, M., Klier, J., Panier, F., & Schubert, J. (2021, June 25). *Defining the skills citizens will need in the future world of work.* McKinsey & Company.

https://www.mckinsey.com/~/media/mckinsey/industries/public%20and%20social%20sector/our%20insights/defining%20the%20skills%20citizens%20will%20need%20in%20the%20future%20world%20of%20work/defining-the-skills-citizens-will-need-in-the-future-of-work-final.pdf

Duro, E., Elander, J., Maratos, F. A., Stupple, E. J. N., & Aubeeluck, A. (2013). In search of critical thinking in psychology: An exploration of student and lecturer understandings in higher education. *Psychology Learning & Teaching*, *12*(3), 275–281. https://doi.org/10.2304/plat.2013.12.3.275

Dwyer, C., Hogan, M. J., & Stewart, I. (2011). The promotion of critical thinking skills through argument mapping. In C. P. Horvath & J. M. Forte (Eds.), *Critical thinking* (1st ed., pp. 97–122). Nova Science Publishers.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity, 12,* 43–52. https://doi.org/10.1016/j.tsc.2013.12.004

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, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction* (ED315423). ERIC. https://files.eric.ed.gov/fulltext/ED315423.pdf.

Finley, A. (2021). *How college contributes to workforce success: Employer views on what matters most*. Association of American Colleges and Universities. https://dgmg81phhvh63.cloudfront.net/content/user-photos/Research/PDFs/AACUEmployerReport2021.pdf

Garside, C. (1996). Look who's talking: A comparison of lecture and group discussion teaching strategies in developing critical thinking skills. *Communication Education*, *45*(3), 212–227. https://doi.org/10.1080/03634529609379050

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist*, *53*(4), 449–455. http://dx.doi.org/10.1037/0003-066X.53.4.449

Halpern, D. F. (2003a). Thinking critically about creative thinking. In M. A. Runco (Ed.), *Critical creative processes* (pp. 189–207). Hampton Press.

Halpern, D. F. (2003b). *Thought and knowledge: An introduction to critical thinking* (4th ed.). Lawrence Erlbaum Associates Publishers.

Halx, M. D., & Reybold, L. E. (2006). A pedagogy of force: Faculty perspectives of critical thinking capacity in undergraduate students. *The Journal of General Education*, *54*(4), 293–315. https://doi.org/10.1353/JGE.2006.0009

Harris, A. (2021). *The State must provide: Why America's colleges have always been unequal—and how to set them right*. Harper Collins Publishers.

Hart Research Associates. (2018). *Fulfilling the American dream: Liberal education and the future of work*. https://www.aacu.org/sites/default/files/files/LEAP/2018EmployerResearchReport.pdf

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Haviland, S., Robbins, S., Kirova, D., Bochenek, J., & Fishtein, D. (2022). *Noncredit career and technical community college programs as a bridge to employers: Report on NYC region study* (Research Report No. RR–22-09). ETS. https://doi.org/10.1002/ets2.12351

Hawkins, K. T. (2012). *Thinking and reading among college undergraduates: An examination of the relationship between critical thinking skills and voluntary reading* [Doctoral dissertation, University of Tennessee in Knoxville]. TRACE Institutional Repository. https://trace.tennessee.edu/utk_graddiss/1302/

Huber, C. R., & Kuncel, N. R. (2016). Does college teach critical thinking? A meta-analysis. *Review of Educational Research*, *86*(2), 431–468. https://doi.org/10.3102/0034654315605917

Hyytinen, H., Jämsä, M., Tuononen, T., & Kleemola, K. (2025). A systematic-narrative review of performance-based assessments of critical thinking in higher education. *Assessment & Evaluation in Higher Education, 50*(8), 1293–1310. https://doi.org/10.1080/02602938.2025.2553341

Hyytinen, H., Nissinen, K., Kleemola, K., Ursin, J., & Toom, A. (2024). How do self-regulation and effort in test-taking contribute to undergraduate students' critical thinking performance? *Studies in Higher Education*, *49*(1), 192–205. https://doi.org/10.1080/03075079.2023.2227207

Liu, O. L., Frankel, L., & Roohr, K. C. (2014). *Assessing critical thinking in higher education: Current state and directions for next-generation assessment* (Research Report No. RR–14-10). ETS. https://doi.org/10.1002/ets2.12009

Liu, O. L., Mao, L., Frankel, L., & Xu, J. (2016). Assessing critical thinking in higher education: The HEIghten™ approach and preliminary validity evidence. *Assessment & Evaluation in Higher Education*, *41*(5), 677–694. https://doi.org/10.1080/02602938.2016.1168358

Liu, O. L., & Roohr, K. C. (2013). *Investigating 10-year trends of learning outcomes at community colleges* (Research Report No. RR–13-34). ETS. https://doi.org/10.1002/j.2333-8504.2013.tb02341.x

Liu, O. L., Roohr, K. C., Seybert, J. M., & Fishtein, D. (2022). Are fourth-year college students better critical thinkers than their first-year peers? Not so much, and college major and ethnicity matter. *Educational Measurement: Issues and Practice, 41*(1), 64–79. https://doi.org/10.1111/emip.12430

Martín-Raugh, M., Kell, H., Ling, G., Fishtein, D., & Yang, Z. (2023). Noncognitive skills and critical thinking predict undergraduate academic performance. *Assessment & Evaluation in Higher Education, 48*(3), 350–361, https://doi.org/10.1080/02602938.2022.2073964

McGrew, S., Breakstone, J., Ortega, T., Smith M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education, 46*(2), 165–193. https://doi.org/10.1080/00933104.2017.1416320

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

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Fuad, Z. A. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Educational Process: International Journal, 14*, Article e2025031. https://doi.org/10.22521/edupij.2025.14.31

National Association of Colleges and Employers. (2017). *Job outlook 2018*.

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

O'Dwyer, E. P., Sparks, J. R., & Oláh, L. N. (2023). Enacting a process for developing culturally relevant classroom assessments. *Applied Measurement in Education, 36*(3), 286–308. https://doi.org/10.1080/08957347.2023.2214652

Paul, R. W., Elder, L., & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendations* (ED437379). ERIC. https://files.eric.ed.gov/fulltext/ED437379.pdf

Rothinam, N., Vengrasalam, R., Naidu, S., Nachiappan, S., & Jabamoney, S. (2025). Systematic literature review on critical thinking in higher education. *Edelweiss Applied Science and Technology, 9*(5), 2046–2063. https://doi.org/10.55214/25768484.v9i5.7377

Spector, J. M., & Ma, S. (2019). Inquiry and critical thinking skills for the next generation: From artificial intelligence back to human intelligence. *Smart Learning Environments, 6*, Article 8. https://doi.org/ 10.1186/s40561-019-0088-z

Stupple, E. J. N., Maratos, F. A., Elander, J., Hunt, T. E., Cheung, K. Y. F., & Aubeeluck, A. V. (2017). Development of the critical thinking toolkit (CriTT): A measure of student attitudes and beliefs about critical thinking. *Thinking Skills and Creativity, 23*, 91–100. https://doi.org/10.1016/j.tsc.2016.11.007

Van der Kleij, F. M., Eggen, T. J. H. M., Timmers, C. F., & Veldkamp, B. P. (2012). Effects of feedback in a computer-based assessment for learning. *Computers & Education*, *58*(1), 263–272. https://doi.org/10.1016/j.compedu.2011.07.020

van Gelder, T. (2005). Teaching critical thinking: Some lessons from cognitive science. *College Teaching, 53*(1), 41–48. https://doi.org/10.3200/CTCH.53.1.41-48

van Peppen, L. M., Verkoeijen, P. P. J. L., Heijltjes, A. E. G., Janssen, E. M., & van Gog, T. (2021). Enhancing students' critical thinking skills: Is comparing correct and erroneous examples beneficial? *Instructional Science, 49*(6), 747–777. https://doi.org/10.1007/s11251-021-09559-0

Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes.* Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Notes

[1] https://interstatepassport.wiche.edu/overview-interstate-passport/learning-outcomes/

Suggested Citation

Song, Y., Bridgeman, B., Yang, Z., & Li, C. (in press). *Cultivating critical thinking: Design and evaluation of training modules for college students*. ETS Research Report Series. https://doi.org/10.64634/emhbze14.

Copyright © 2026 by Educational Testing Service. All rights reserved.

ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS). All other trademarks are the property of their respective owners.