Journal of the Scholarship of Teaching and Learning, Vol. 25, No. 4, December 2025, Article 6
doi: 10.14434/josotl.v25i4.37606

Measuring Impacts of Andragogical Treatments: An Experimental Study of Self-perceived Critical Thinking Behavioral Change in Undergraduate Students

Barry Boyd
Texas A&M University
b-boyd@tamu.edu

Jennifer R. Strong
Texas A&M University

Abstract: Critical thinking is often noted as lacking in today's workforce, but research has demonstrated that critical thinking can be taught while improving students' course content information retention. However, new reports reveal inadequate inclusion of critical thinking andragogy in the curriculum. Given this dichotomy, this study aimed to use a behaviorally-anchored competency-based model to asses students' self-perceived critical thinking behavioral change in a one-semester treatment. Instructors used the QUEEN Critical Thinking Model, a behaviorally-anchored competency-based approach critical thinking model and assessment, to teach and evaluate critical thinking behaviors to students in an upper-level leadership ethics course. QUEEN Student Self-Assessment indicates students' critical thinking behaviors increased significantly over one semester when instructors integrated critical thinking with course material. To prepare students to close the essential skills gap, we recommend instructors purposefully design learning strategies that teach critical thinking behaviors. Based on the results of this study, it can be concluded that using high-impact andragogical methodology to apply the QUEEN Model of Critical Thinking is effective for integrating critical thinking behaviors into established college classes to increase students' self-perceived critical thinking behaviors.

Keywords: Critical thinking assessment, teaching strategies, critical thinking behaviors, critical thinking models

A recent phenomenon, coined the skills gap, sheds light on the increasing difficulty for organizations to reconcile the disparity between rapid innovation and employee skills needed to fulfill ever-changing demands. Critical thinking is noted as one of the most deficient skills within this gap (Burner et al., 2019), but is essential for organizational success. In the 2018 World Economic Forum, critical thinking was ranked as a top indicator to assess global economic productivity (Schwab, 2018). In their biannual employers' survey, the Association of American Colleges and Universities (2021) found that 60% of employers describe critical thinking as an essential skill for success in the workplace. However, only 39% believe students graduate with these skills. Establishing partnerships with higher education institutions is recognized as a primary method to reform the knowledge base of job-seeking graduates; however, over 50% of organizations need more real change occurring at the educational level (Burner et al., 2019). In a survey of business leaders, the American Management Association (2012) reported that over half the population believed it easier to teach critical thinking to recent graduates than experienced workers.

Critical thinking is a higher-order skill that does not come naturally for many people. Kuhn (1991) compares critical thinking to ballet. Just as humans do not naturally dance on their toes, humans do not naturally think critically, but they can be trained to be proficient in that skill. Van Gelder (2005) agreed that critical thinking is not a natural process for humans, especially undergraduate students. When given a problem, students tend to go with the first solution that seems correct, rarely pursuing the matter further (Shermer, 2002).

Research has demonstrated that critical thinking can be taught while improving students' information retention (Grussendorf & Rogol, 2018; Lilly et al., 2022). However, other research reveals inadequate inclusion of critical thinking in the curriculum, thus potentially lowering the capacity of new employees to practically apply skills in the workplace (Ennis, 2018). To counter, studies suggest the importance of considering in-class experiences and out-of-class activities, such as reading non-assigned books and cumulative study hours, to help explain statistically significant improvements in critical thinking behaviors after just one year of college (Terenzini et al., 1995). Moreover, Abrami et al. (2008) and Huber and Kuncel (2015) substantiate these findings through two separate meta-analyses, which defend a general increase in critical thinking disposition after instruction, adding that explicit direction to hone these skills is imperative to produce positive results. A globally recognized Master of Science in Leadership program demonstrated that critical thinking could be developed at the university level and is optimized when introduced in the student's first year and reinforced throughout all courses (Watkins & Earnhardt, 2015).

Therefore, a noticeable gap exists between the potential of higher education to improve students' critical thinking skills and an evident change in graduates' exhibition of these skills through their behaviors in the workplace. This research sought to explore the effectiveness of intense critical thinking instruction and practice in a one-semester treatment. The QUEEN model of critical thinking was used to teach behaviorally-anchored competencies. The QUEEN Student Self-Assessment is based on the QUEEN Model and was used to evaluate critical thinking behaviors inside and outside the classroom (Boyd & Strong, 2023; Boyd & Strong, 2021).

Literature Review

Perhaps the most agreed-upon conceptualization of critical thinking came from the 1990 APA Delphi study led by Facione (1990), who developed a definition still referenced today. From this study, a consensus of experts defined critical thinking as the "purposeful, self-regulatory judgement which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgement is based" (Facione, 1990, p. 3).

Characteristics of a proficient critical thinker include but are not limited to curiosity, open-mindedness, awareness of intrinsic bias, a thorough examination of information, and reliable judgment (Facione, 1990). Penkauskienė et al. (2019) claim managers perceive critical thinking as an employee's ability to avoid mistakes, make correct decisions, practice self-control, self-directed action, and social responsibility. Additionally, promulgating the idea of proper disposition toward critical thinking has been deemed equally necessary as reinforcing the skill since a relative affinity regarding higher-order thinking is essential for voluntary practice outside of an educational context (Facione et al., 1995).

A disconnect between the definition of critical thinking and its application in the classroom is widely recognized in literature; as Lee et al. (2014) propose, there is a need to re-evaluate the alignment of instructional objectives with educational practices, as some are more inclined to cultivate passive learners as opposed to innovative thinkers. However, to improve the dissemination of critical thinking curriculum, institutions should primarily ensure instructors lead by example, adequately demonstrating critical thinking behaviors and disposition (Stedman & Adams, 2012). Additionally, research reveals a disparity between teachers' comprehension of higher-order thinking and their awareness of when students exhibit these qualities. This suggests a need for explicit teaching objectives with listed critical thinking behaviors (Choy & Cheah, 2009). Pressure from the overarching educational system to meet standardized learning outcomes within a limited time discourages a self-directed, interactive environment, thus curtailing the development of critical thinking skills (Fadhlullah & Ahmad, 2017).

As with any other skill, van Gelder (2005) ascertains that critical thinking behaviors must be practiced. They will never improve unless students are actively "doing" critical thinking. Furthermore, student engagement, positive feedback, and peer interaction influence critical thinking learning outcomes (Smith, 1977). Research reports significant gains in critical thinking after one semester when students are provided questions that provoke higher-order thinking on exams or assignments more frequently (Renaud & Murray, 2007). The former results within a limited time frame are promising, considering Pascarella and Terenzini (2005) found a 19% increase in students' critical thinking after four years of university education.

Given these parameters, some instructional strategies have proven more effective in the classroom than others. Ericson and Charness (1994) describe using deliberate practice where the goal is to improve thinking. Lilly et al. (2022) found that critical thinking disposition could be increased in a one-semester treatment that included specific critical thinking learning objects (CTLO) and practice. Exercises must be designed to improve specific behaviors, graduating from less to more complex problems. Such practice requires guidance and timely feedback on students' performance.

Another example involves using productive query, defined by questions stimulating higher-order thinking, instead of reproductive query or referring to questions prompting recollection and regurgitation (Tienke et al., 2012). Student achievement has also improved when critical thinking assessment criteria are observed throughout the course and shared with the student for personal evaluation (Broadbear, 2003). As opposed to intelligence tests, critical thinking evaluations should include criterion sampling of behavioral traits and competency measurements, which include more generalizable skills, such as communication and leadership, which reach beyond specific subject content (McClelland, 1973). To counter, Tedesco-Schneck (2013) suggests organizations should avoid promulgating a narrow positivist approach in which evidence-based, observable outcomes are the sole means of criterion formulation and course objectives but instead favor a dynamic process of active engagement with students willing to venture outside of the typical passive lecture experience. Research shows that compared to intelligence tests, the ability to think critically is a more effective predictor of how an individual will behave or react in a given situation (Butler et al., 2017).

Conceptual Framework

The QUEEN Critical Thinking Model takes the complicated and multifaceted concept of critical thinking and breaks it down into five tenets, or behavioral competencies. The QUEEN Model was developed from a Delphi study of critical thinking experts. The expert panel was asked what behaviors students exhibit when thinking critically. After completing three rounds of inquiry, 28 critical thinking behaviors were identified. Through deductive analysis, these behaviors fell into five distinct categories: Questioning, Understanding, Evaluating, Explaining, and Neoteric Thinking (Boyd & Strong, 2023). An exploratory factor analysis (EFA) further refined the categories, reducing the behaviors to 24. This study used The QUEEN model as the framework for teaching and assessing students' critical thinking behaviors.

Questioning

For students to exhibit critical thinking behaviors, they must not be afraid to ask the Five Ws many of us were trained in as children. However, *No Child Left Behind* and similar government programs in the United States slowly suspended the drive to engage in such questioning (Gentry, 2006). For the QUEEN model, *Questioning* is not just about asking the "if" of their actions but the depth and understanding of the questioning. The *Questioning* tenet includes the behaviors students exhibit by (1) questioning the what and why of the issue, (2) seeking accurate information on the problem, (3) seeking relevant information, and (4) questioning all positions on the issue (Boyd & Strong, 2023).

Understanding

Looking for behaviorally anchored examples of *Understanding* may be difficult, but instructors should delve deeper as they engage their students in written and oral discussions and projects. Understanding the holistic context of situations is paramount. *Understanding* involves the understanding of self and others as an integral step of critical thinking. Students exhibit *Understanding* by (1) understanding the purposes of others, (2) understanding the background of the issue, (3) recognizing alternative views, and (4) knowing their assumptions (Boyd & Strong, 2023).

Evaluating

As Choy and Cheah (2009) note, engaging students in higher-order thinking is imperative. Bloom's Taxonomy, which many universities rely on to incorporate action verbs into course objectives for evaluation purposes, uses *Evaluation* as an indicator of higher-order thinking (Lewis & Smith, 1993). For the QUEEN model, the *Evaluation* behaviors students should be able to present to proficiency in are (1) judging the credibility of sources, (2) judging if the information is relevant or "white noise," (3) drawing appropriate conclusions using the data/information chosen, (4) identifying potential consequences of decisions, (5) identifying and examining one's assumptions as well as others', and (5) recognizing the differences between fact and value claims (Boyd & Strong, 2023).

Explaining

As leaders and fully functional members of society, it is not enough for us to merely understand and critically think about topics; we must be able to communicate our thoughts and thought processes to those around us (Cunningham, 1985). *Explaining* allows students to show their ability to communicate their thinking to others effectively. Students exhibit *Explaining* by (1) explaining their reasoning process, (2) justifying or defending their positions based on fact, (3) creating alternative answers/solutions, (4) using examples to connect points, (5) discussing their reasoning process, and (5) using models or theoretical frameworks to back their conclusions (Boyd & Strong, 2023).

Neoteric Thinking

Neoteric Thinking, often attributed to poets moving past traditional methods of episodic writing into a different mode of storytelling, is defined as creating and implementing new ideas (Lyne, 1978). Critical thinkers who engage in neoteric critical thinking can not only connect the information to stable application examples, as those high in explanation can, but they also have the added layer of advocacy. These new and critical thinking ideas go nowhere if they do not have a champion to advocate for these new ways of thinking and doing. Students who exhibit *Neoteric Thinking* show proficiency by (1) synthesizing new solutions/products/ideas, (2) promoting new concepts, (3) staying open-minded to new ideas or crucial viewpoints, (4) creating alternative solutions, and (5) engaging in self-examination and reflection to become a more profound critical thinker and more effective learner (Boyd & Strong, 2023).

Behavioral Anchors

Behavioral anchors are defined as characteristics of core competencies associated with the mastery of content. Behavioral anchors are performance capabilities needed to demonstrate the acquisition of knowledge, skill, and ability (competency) (Buford & Lindner, 2002). Knowledge is reliable information learners use to perform successfully (Buford & Lindner, 2002). Skills are recognizable aptitudes needed to accomplish tasks, and abilities are expressed as an individual's capacity to display observable behaviors or behaviorally-driven outcomes. Behavioral anchors delineate characteristics that encourage knowledge attainment and understanding (Dooley & Lindner, 2002). Competency-based behavioral anchors benchmark performance progress using observable, evidence-based markers (Buford & Lindner, 2002).

Classroom motivation and participatory engagement are enhanced when students are made aware of personal competency attributes and methods to improve other areas of critical thinking (Drawbaugh, 1972). Institutions implementing competency-based behavioral anchored assessments are better equipped to produce workforce-ready students capable of reflective decision-making and innovative problem-solving (Muller-Frommeyer et al., 2017). To remain competitive in an evolving business environment, employers more often seek competency-based professional development to refine skills needed for job-specific responsibilities, and employees are more competent in evaluating behavior than disposition (Dooley et al., 2004). Moreover, behavioral traits within competency models assist in maximizing output by increasing the intersectionality of individual performance and organizational strategy (Rothwell & Lindholm, 1999). The 24 behaviors of the QUEEN Critical Thinking Model serve as the behavioral anchors for this study.

Purpose and Objectives

Given the importance of increasing students' critical thinking knowledge, skills, and abilities, as well as faculties' struggle to find critical thinking assessments that are based on behaviors, not inclination, the purpose of this study was to use a behaviorally-anchored competency-based model to teach critical thinking behaviors within the confines of a semester-long course. An assessment tool based on this same critical thinking model examined students' self-perceived critical thinking behavioral change.

The research objective of this study was to measure the change in students' self-perceived critical thinking behaviors in upper-level undergraduate leadership courses. Specific objectives were:

1. Determine the methodological effectiveness of the QUEEN Student Self-Assessment as a post-then-pre-test.
2. Determine if a one-shot critical thinking intervention changes students' self-perceived critical thinking behaviors.
3. Determine if intentional critical thinking teaching strategies change students' self-perceived critical thinking behaviors.

This study was conducted with the support of an NIFA Higher Education Challenge Grant. The grant project aimed to create and utilize a behaviorally-anchored, competency-based critical thinking assessment to measure undergraduate students' critical thinking behaviors.

Methods

The QUEEN Critical Thinking Model consists of 24 behaviors grouped into five categories. Samples of behavioral statements for each category are provided in Table 1.

Table 1. Sample QUEEN Critical Thinking Behaviors.

Questions	I ask questions to clarify the what and why of the situation I look for accurate and relevant information about the situation
Understanding	I understand the background of the issue I seek to understand alternative views I understand others' assumptions
Evaluation	I judge the credibility of sources I draw appropriate conclusions using data/information available I distinguish between fact and value claims
Explanation	I state results or conclusions clearly I use examples to connect points I can justify/defend my positions
Neoteric	I synthesize new solutions/products/ideas I am open to new ideas or viewpoints I promote new concepts

The population for this study included all students majoring in leadership at an AAC&U university. The sample consisted of a census of two upper-level leadership courses. The sample was then sub-divided into three groups: (A) a control group who received no critical thinking instruction; (B) a group who were introduced to the critical thinking model but were not given opportunities to practice the behaviors (Treatment One); and (C) a group who received instruction in the critical thinking model and were given opportunities to practice critical thinking throughout the semester (Treatment Two).

Group A

Students randomly assigned to Group A (n=58) were considered the control group. These students were taught course content only. There was little to no mention of critical thinking as a practice and no QUEEN instruction. Group A was administered the QUEEN Student Self-Assessment (Boyd & Strong, 2021) as a pre-test at the beginning and again as a post-then-retrospective pre-test at the end of the semester. Group A's lack of critical thinking instruction was to mitigate inaccuracies due to acquiring new skills and knowledge in critical thinking over the semester.

Group B

Students randomly assigned to Group B (n =85) received Treatment One. These students were taught the QUEEN model via a 30-minute lecture. This lecture broke down the QUEEN competencies by describing the behaviors associated with each construct and giving concrete examples of students exhibiting those critical thinking behaviors. Group B was also given the QUEEN Student Self-Assessment (Boyd & Strong, 2021) pre-test at the beginning of the semester, prior to the QUEEN instruction. They also completed the QUEEN Student Self-Assessment post-then-retrospective pre-test at the conclusion of the semester. This post-then-retrospective pre-test was given 10 weeks after the instruction on the QUEEN model.

Group C

Students randomly assigned to Group C (n=302) received Treatment Two. Treatment Two consisted of various critical thinking instructional methods and evaluation assignments to assess the progress of students' application of critical thinking behaviors throughout the semester. These methods were chosen based on van Gelder's (2005) recommendation that students be taught critical thinking theory to enhance their understanding and Strong et al.'s (2021) findings that students benefitted greatly in their understanding when critical thinking was defined and models used to frame the concept. The two instructors began teaching the QUEEN Critical Thinking Model to Group C in the first week of class.

Instructor A Strategies for Group C

Instructor A used both overt and covert means to engage students in critical thinking behaviors during ethical topic and case study discussions. Students were prompted to utilize behaviors from the QUEEN critical thinking model on each course unit reflection to analyze the course material. Each assigned case study required students to question sources and information, determine if relevant information was used in the case decision, and discuss alternative solutions. Another assignment prompted students to reflect on assigned videos regarding religion and ethics and answer various free-response questions requiring them to define terms, state personal beliefs, compare and contrast, draw conclusions, and search for alternative sources of information. The purpose of the final assignment, developing a personal code of ethics, was to encourage students to use critical thinking behaviors as they considered their personal beliefs and motives, situational nuances, and moral theories that guide their beliefs and code.

Instructor B Strategies for Group C

While lecturing on the QUEEN, Instructor B engaged the students in public pedagogy (Williams & McClure, 2010) by using current events (campus politics, national politics) and pop culture references (movies, TV shows, songs). The class discussion focused on the examples given by the instructor and was facilitated both face-to-face and via Zoom for those joining the class. To aid in students' understanding of QUEEN, the reflective writing assignment focused on a current "hot topic" event happening on campus, being an epicenter of renewed discussion on the impact of statues as historical reminders of racial strife. Each student was asked to research, analyze, synthesize, and develop their position using the tenets of the QUEEN model. Specifically, they were to: (Q) determine what questions needed to be asked and describe their perceptions of the issue, (U) discover relevant information and deduce its purpose (looking at multiple points of view), (E) evaluate the relevant information, credibility of the source, and consequences, (E) explain their conclusions based on the information, and (N) propose a solution to the problem detailing how they reflected and synthesized their information. This reflective writing piece was overtly and directly tied to QUEEN, but other assignments and discussions asked students to apply QUEEN as they completed the work. Instructor B also used a personal code of ethics as a culminating assignment where students developed their code of ethics and engaged in critical thinking behaviors to complete the assignment.

Student Self-Assessment

As noted, the control group (A) and the one-shot Treatment One group (B) were administered the QUEEN Student Self-Assessment (Boyd & Strong, 2021) as a pre-test at the beginning of the semester and again as a post-then-retrospective pre-test at the end of the semester. The QUEEN Student Self-Assessment was administered to Group C, who only received Treatment Two as a post-then-retrospective pre-test. The assessment was delivered and administered to all groups electronically via Qualtrics.

For the post-then-retrospective pre-assessment, students were asked to evaluate their critical thinking behaviors at the end of the semester. At the same time, they were asked to reflect on the beginning of the semester and rate their critical thinking behaviors before being exposed to the treatment. Measuring the change is most often completed with a paired samples t-test. The paired samples t-test presents a null hypothesis, which posits that mean scores for the post-then-retrospective pre-test will be equal, with an alternative hypothesis suggesting significant differences between the two types of retrospective questions (Zimmerman, 1997). Nicoll-Senft and Seider (2010) successfully used the post-then-retrospective pre-test when evaluating the impact of 4MAT teaching methods in increasing self-perceived learning. Taking the posttest and the pretest simultaneously makes it more likely that both ratings will be made from the same frame of reference, thus eliminating the effects of response shift bias (Fernandez et al., 2014; Marshall et al., 2007; Rohs & Langone, 1997). Response shift bias occurs when students are asked to self-assess their behaviors before understanding how they are defined. As a result, students tend to overestimate those behaviors on the pre-test, then on the post-test, rate themselves lower than on the pre-test (Rohs & Langone, 1997). This is often explained as students needing to understand the intricacy of the behaviors; thus, not knowing what they do not know, they rate themselves higher on a behavior until they truly understand what it means. They realize they do not know as much as they thought or exhibit that behavior as much as they thought (Nicoll-Senft & Seider, 2010).

For each QUEEN competency, students were asked to rank the frequency to which they applied critical thinking behaviors on an ordinal scale ranging from one (never) to four (often). The SPSS was used to analyze the students' responses.

Cronbach's Alpha was performed on the instrument and each competency subscale. Cronbach's Alpha calculates the average of all possible split-half coefficients and ensures questions within the instrument are reliable and related by design (Bryman, 2012). The higher the alpha result, the stronger the correlation among participants' responses to items within the scale. Given the alpha is a correlation statistic, results can range from 0 (no correlation among items) to 1 (perfect correlation). An instrument with an alpha greater than .70 is considered to have good internal reliability (Bryman, 2012). The QUEEN Student Self-Assessment has high internal reliability with a Cronbach's Alpha of .937 for the complete instrument. The Cronbach's Alpha for each subscale was also very strong, as seen in Table 2.

Table 2. Cronbach's Alpha for Student Self-Assessment, N = 302.

Scale	Chronbach's Alpha	Number of Items
Total Instrument	.937	24
Questioning	.787	4
Understanding	.874	4
Evaluating	.831	6
Explaining	.856	5
Neoteric Thinking	.826	5

Findings

Objective One

Objective one was to determine the methodological effectiveness of the QUEEN Student Self-Assessment as a post-then-retrospective pre-test. This was accomplished by testing the scores of Group A and B students' pre-test self-perceived QUEEN behaviors compared to the retrospective pre-test of the same behaviors. The paired-sample t-test revealed that students who took the assessment as a pre-test at the beginning of the semester and then as a retrospective pre-test at the end significantly overestimated their critical thinking behaviors at the beginning of the semester. Their scores on the pre-test were significantly higher than on the retrospective pre-test taken at the end of the semester.

When analyzing the significance further, researchers found the effect size of each QUEEN competency was also different. Table 3 shows Cohen's d scores, which ranged from 0.3140 to 0.984 on the five scales. An effect size of 0.20 indicates a small effect, d=.0.50 indicates a medium effect, and d=.0.80 or greater indicates a large effect size (Lakens, 2013). In this analysis, the categories of Questioning and Evaluating have small effect sizes. Neoteric thinking has a medium effect size, while Understanding and Explaining have a large effect size. This means students overestimated their self-perception of their critical behaviors of Understanding and Explaining more than Questioning and Evaluating.

Table 3. Pre-test Scores Compared to Retrospective Pre-test Scores given Post, N=143.

Scale	Mean Difference	Std. Dev.	Standard Error Mean	T-value	Sig.	Cohen's d
Questioning - Pre-test Retrospective Pre-test	.27035	.55996	.06038	4.477*	.000	0.482
Understanding - Pre-test Retrospective Pre-test	.48983	.49787	.05369	9.124*	.000	0.984
Evaluating - Pre-test Retrospective Pre-test	.50803	.61810	.06785	7.488*	.000	.3140
Explaining - Pre-test Retrospective Pre-test	.52941	.61355	.06655	7.955*	.000	.8629
Neoteric Thinking - Pre-test Retrospective Pre-test	.48235	.64410	.06986	6.904*	.000	.7489

*Statistically significant at .05 level.

Objective Two

Objective Two was to determine if a one-shot critical thinking intervention changes students' self-perceived critical thinking behaviors. Group B's one-shot intervention was a 30-minute lecture on the QUEEN model. The post-test portion of the QUEEN Student Self-Assessment post-then-retrospective pre-test was used to answer this research objective. The independent-sample t-test (N=143) indicates students receiving the introduction to critical thinking behaviors (one-shot treatment) did not rate their critical thinking behaviors significantly higher than the students in the control group, except for the Evaluating behaviors (Table 4). The difference on this scale was statistically significant, with Group B scoring higher than students in Group A.

Table 4. Independent Samples T-test between Control and One-Shot Treatments, N=143.

Scale	Mean Difference	Std. Dev.	Standard Error Mean	T-value	Sig.
Questioning Control	.03489	.4642	.0504	.456	.649
Treatment		.4263	.0560		
Understanding Control	-.1210	.4600	.0499	-1.583	.116
Treatment		.4321	.0567		
Evaluating - Control	-0.175	.4928	.0548	-2.014*	.046
Treatment		.5220	.0685		
Explaining Control	-0.1035	.5010	.0553	-1.167	.245
Treatment		.5325	.0705		
Neoteric Thinking - Control	-.1495	.5067	.0553	-1.710	.089
Treatment		.5200	.0683		

*Statistically significant at .05 level.

Objective Three

Objective Three was to determine if intentional critical thinking strategies change students' self-perceived critical thinking behaviors. A paired-sample t-test (N=302) was used to analyze differences in post-then-retrospective pre-test scores for the entire QUEEN student self-assessment instrument. In this case, the t-value is reported as 16.822 with a significance value of 0.000 ($p<0.05$), indicating that a statistically significant difference exists between students' self-perceived critical thinking behaviors post-then-retrospective pre-test scores (Table 5). In addition, the t-value for each of the five subscales shows statistically significant differences between the post-then-retrospective pre-test scores. All differences were significant at the .05 level.

Cohen's d was calculated to determine the effect size or practical significance of differences between means. The complete instrument and all scales had Cohen's d scores above 1.0, indicating a large effect size (see Table 4). In other words, post-test means were more than one standard deviation greater than pre-test means, and students saw significant behavioral changes before learning and applying the QUEEN Model of Critical Thinking and the treatment.

Table 5. Paired-Samples t-test and Cohen's d between Post-then-Pre-test scores, N = 302.

Scale	Mean Difference	Std. Dev.	Standard Error Mean	T-value	Sig.	Cohen's d
Full instrument - Post Full instrument – Then Pre	.58116	.41034	.03488	16.822*	.000	1.42
Questioning - Post Questioning – Then Pre	.52461	.46331	.04124	13.001*	.000	1.13
Understanding - Post Understanding - Then Pre	.62347	.46344	.03922	15.342*	.000	1.36
Evaluating - Post Evaluating - Then Pre	.65734	.53254	.04712	13.989*	.000	1.23
Explaining - Post Explaining - Then Pre	.57119	.48377	.04355	13.619*	.000	1.18
Neoteric Thinking - Post Neoteric Thinking - Then Pre	.62983	.57557	.04861	13.151*	.000	1.09

*Statistically significant at .05 level.

Additionally, Group C (treatment group) scores were compared to Group A's (control group) scores on the QUEEN Student Self-Assessment. Group C's scores were significantly higher than Group A's at the .05 level on all five scales. These findings can be seen in Table 6.

Table 6. Independent Samples T-test between Control (A) and Full Treatment (C), N= 360.

Scale	Mean Difference	Std. Dev.	Standard Error Mean	T-value	Sig.	Cohen's d
Questioning Control n=302 Treatment n=58	.2747	.4579 .3326	.0359 .0437	4.189*	.001	.4288
Understanding Control Treatment	.2087	.4460 .2977	.0349 .0391	3.308*	.001	.4126
Evaluating - Control Treatment	.3050	.4107 .3434	.0322 .0451	5.508*	.001	.3943
Explaining Control Treatment	.1869	.4749 .3151	.0372 .0414	2.785*	.006	.4389
Neoteric Thinking - Control Treatment	.1972	.4908 .3506	.0385 .0460	2.813*	.005	.4584

*Statistically significant at .05 level.

Discussion

Regarding Objective 1, analysis indicates that participants tend to overestimate their use of critical thinking behaviors before being exposed to the definitions of those behaviors. The difference between students' pre-test and retrospective-pre test scores was statistically significant. This finding supports the findings of Fernandez et al. (2014), Marshall et al. (2007), and Rohs and Langone (1997) regarding response-shift bias in pre and post-test instruments and supports the use of the post-then-retrospective pre-test format.

The results of Objective Two found that exposing students to a one-time lecture over a critical thinking model does not significantly improve their critical thinking behaviors. Educators who adhere to principles of andragogy in their classes, as well as active and experiential learning, might not be surprised by this finding. Multiple studies have shown that for students to internalize material and change their behavior, there must be more to learning than traditional short-term lectures (Townsend, 2002; Williams & McClure, 2010). This finding also supports the work of various studies that concluded that one-shot training is ineffective in modifying behavior or increasing learning application (Cummins, 1995; Montgomery, 2002; Gross et al., 2012).

The final objective of this study was to determine if intentional critical thinking teaching strategies increased students' self-perceived critical thinking behaviors. Based on the statistically significant differences between Post-Then-Pre scores for the combined QUEEN questionnaire as well as each tenet, coupled with the large effect size, it is evident that students' self-perceived critical thinking behaviors increased over one semester when allowed to practice those behaviors repeatedly. The tenet of Evaluating saw the most change from the students' reflection before and after the course, followed by Neoteric Thinking, Explaining, Understanding, and Questioning. Structured, purposeful critical thinking interventions effectively increased students' critical thinking behaviors across all five tenets of the QUEEN Critical Thinking Model. This supports several studies (Lilly et al., 2022; van Gelder, 2005) that ascertained students must be given adequate practice in critical thinking for improvement to occur. Student growth in critical thinking behaviors may also be attributed to instruction and application of critical thinking theory (van Gelder, 2005). This finding is further supported when comparing the control group that received no instruction in critical thinking behaviors to the treatment group. Treatment group scores were significantly higher at the .05 level, showing a change in self-perceived critical thinking behavior. Lilly et al. (2022) and van Gelder (2005) had similar findings but focused on changing critical thinking disposition rather than behaviors.

The Post-then-retrospective pre-test format of the student self-assessment provided students with a more objective perspective to evaluate their critical thinking behaviors through retrospective analysis and current observation. Additionally, researchers found that purposefully designed and integrated strategies and various application assignments successfully increased students' self-perceived critical thinking behaviors. Both instructors consistently applied critical thinking behaviors and allowed students to practice such behaviors.

Conclusion and Recommendations

As researchers, practitioners, and employers have found, the essential skills gap in undergraduate students is extensive. Universities are increasing their efforts to develop and assess these skills, including leadership and critical thinking (Abrami et al., 2008; Ennis, 2018; Huber & Knuckle, 2015). However, more purposeful teaching methodology and evaluation strategies must be implemented. To meet this need, instructors must be ready, willing, and able to integrate essential skills into their established courses and evaluate the competency gained by their students. Based on the findings of this study, it can be concluded that purposeful andragogy is an effective way to increase self-perceived critical thinking behaviors in undergraduate students. This supports Strong et al.'s (2021) findings that students' perceptions of their critical thinking behaviors can grow when they are taught and able to apply specific critical thinking models, yet takes that finding a step further by noting that course material and critical thinking must be integrated throughout the semester and not just as a small portion of course content.

Many higher education institutions stress the importance of increasing critical thinking capacity as an essential outcome in assessment. In fact, for many institutions, critical thinking assessment and evaluation are part of accreditation (Boyd et al., 2017). Behavioral skills training is an empirically established method for helping students acquire skills. It involves instruction in the behavior, modeling, role-play or practice, and feedback (Parsons, et al., 2012). Utilizing a behaviorally anchored measure of critical thinking allows instructors to quantify the change in critical thinking skills and abilities. It moves past the limitations of evaluations that measure disposition only. In addition, instructors can focus instruction on specific critical thinking behaviors, helping students fill gaps to improve their thinking. Students can also use the QUEEN self-assessment to evaluate their critical thinking behavior development. Based on the findings of this study, it can be concluded that the QUEEN Post-then-retrospective pre-test student self-assessment is a reliable method for assessing the self-perceived impact of critical thinking behavior shift within a university setting by eliminating the limitations of self-reported data, including response shift bias. Post-then-retrospective pre-tests should be considered an effective way to evaluate students' self-perceived growth in essential behaviors.

To prepare students to close the essential skills gap, we recommend instructors purposefully design learning strategies that teach critical thinking behaviors. Based on the results of this study, the tenets of the QUEEN Model of Critical Thinking are an effective way to pair existing activities, such as cases, simulations, current events, and reflections, with critical thinking behaviors in established college classes. Including activities, both graded and non-graded, that are structured based on the QUEEN behaviors and applied to course content, students can reflect and integrate behavioral development.

For future research, we recommend replicating these studies using experimental design to test the impact of teaching the QUEEN model as a one-shot training, as a course-long application, using control groups to determine the most effective instructional methods. Further investigation is also recommended into the teaching and application methods that impact the most. Additional qualitative methods can move this study from self-perceived changes from the student to faculty-observed changes. We also suggest using the faculty QUEEEN check sheet assessment to gauge the instructor-rated change in critical thinking behaviors (Busa et al., 2021). And finally, we recommend longitudinal studies to assess the lasting effect of teaching critical thinking behaviors.

Acknowledgments

Funding: This research was supported by a 2017 Higher Education Challenge Grant from the National Institute for Food and Agriculture (NIFA). The university's Institutional Review Board approved this study: IRB2017-0026M.

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research, 7*(4), 1102–1134. https://doi.org/10.3102/0034654308326084

American Management Association. (2012). Critical skills survey. https://www.amanet.org/assets/1/6/2012-critical-skills-survey.pdf

Associations of American Colleges and Universities. (2021). *How college contributes to workforce success.* https://www.aacu.org/research/how-college-contributes-to-workforce-success

Boyd, B. L., Dooley, K. E., & Kurten, J. S. (2017). Creating a faculty cadre of critical thinking fellows. *NACTA Journal, 61*(3), 202-207.

Boyd, B. L. & Strong, J. (2021). Establishing the royal standard for critical thinking: The QUEEN model. Paper presented at the annual meeting for the Association for International Agricultural Extension Education 36th Annual Conference, Virtual Conference. p. 371. https://drive.google.com/drive/folders/ 1v_NrvQ6jJPrRdZ6CzQw9y1SRZBks_sT9

Boyd, B. & Strong, J. (2023). *Ethical leadership: Theory to practice* (2nd ed.). Kendall Hunt Publishing.

Broadbear, J. T. (2003). Essential elements of lessons designed to promote critical thinking. *Journal of Scholarship of Teaching and Learning, 3*(3), 1-8.

Bryman, A. 2012. *Social research methods*. 4th ed. Oxford University Press.
Buford, J. A., Jr. & Lindner, J. R. (2002). *Human resource management in local government: Concepts and applications for students and practitioners*. Southwestern.
Burner, T., Supinski, L., Zhu, S., Robinson, S., & Supinski, C. (2019). The global skills gap: Bridging the talent gap with education, training, and sourcing. *Society for Human Resource Management.* https://www.shrm.org/about-shrm/Documents/SHRM%20State%20of%20Workplace_Bridging%20the%20Talent%20Gap.pdf
Busa, B. A., Boyd, B. L., and Strong, J. (2021). Check yes or no: Assessing students' critical thinking behaviors using the QUEEN checklist assessment. Poster presented at the annual meeting for the Association of Leadership Educators (Virtual).
Butler, H. A., Pentoney, C., & Bong, M. P. (2017). Predicting real-world outcomes: Critical thinking ability 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
Choy, S. C. & Cheah, P. K. (2009). Teacher perceptions of critical thinking among students and its influence on higher education. *International Journal of Teaching and Learning in Higher Education, 20*(2), 198-206.
Cummins, R. A. (1995). An assessment of attitudes toward leadership by participants of selected leadership labs at Texas A&M University. Unpublished dissertation, Texas A&M University.
Cunningham, L. (1985). Leaders and leadership: 1985 and beyond. *Phi Delta Kappan, 67*(1), 17-20.
Dee, T. S. & Jacob, B. A. (2010). The impact of no child left behind on students, teachers, and schools. *Brookings Papers on Economic Activity*, 149-207. Brookings Institution Press.
Dooley, K. E. & Lindner, J. L. (2002). Competency-based behavioral anchors as authentic tools to document distance education competencies. *Journal of Agricultural Education, 63*(1), 24-35. https://doi.org/10.5032/jae.2002.01024
Dooley, K. E., Lindner, J. R., Dooley, L. M., & Alagaraja, M. (2004). Behaviorally anchored competencies: Evaluation tool for training via distance. *Human Resources International, 7*(3), 315-332. https://doi.org/0.1080/1367886042000210958
Drawbaugh, C. C. (1972). A framework for career education. *Journal of the American Association of Teacher Educators in Agriculture, 13*(2). 16-23.
Ennis, R. H. (2018). Critical thinking across the curriculum: A vision." *Topoi, 37*, 165-184. https://doi.org/10.1007/s11245-016-9401-4
Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. *American Philosophical Association*. ED 315 423. https://files.eric.ed.gov/fulltext/ED315423.pdf
Facione, P. A., Sanchez, C. A., Facione, N. C., & Gainen, J. (1995). The Disposition Toward Critical Thinking. *The Journal of General Education*, 44(1), 1-25. https://www.jstor.org/stable/27797240
Fadhlullah, A. & Ahmad, N. (2017). Thinking outside of the box: Determining students' level of critical thinking skills in teaching and learning. *Asian Journal of University Education, 13*(2), 51-70. http://ir.uitm.edu.my/id/eprint/ 21917/1/AJ_ AFIFAH%20FADHLULLAH%20AJUE%2017.pdf
Fernandez, C. S. P., Noble, C. C., Jensen, E., & Steffen, D. (2014). Moving the needle: A retrospective pre- and post-analysis of improving perceived abilities across 20 leadership skills. *Maternal and Child Health Journal, 19*(2), 343–352. https://doi.org/ 10.1007/s10995-014-1573-1.
Gentry, M. (2006). No child left behind: Neglecting excellence. *Roeper Review, 29*(1), 24-27.

https://doi.org/10.1080/02783190609554380.
Goldberg, C. & Gallimore, R. (1991). Changing teaching takes more than a one-shot workshop. *Educational Leadership*. 68-72.
Gross, M., Latham, D., & Armstrong, B. (2012). Improving below-proficient information literacy skills: Designing an evidence-based educational intervention. *College Teaching 60*(3), 104-111. https://doi.org/10.1080/87567555.2011.645257
Gruden, N. & Stare, A. (2018). The influence of behavioral competencies on project performance. *Project Management Journal, 49*(3), 98-109. https://doi.org/ 10.1177/8756972818770841
Grussendorf, J. & Rogol, N. C. (2018). Reflections on critical thinking: Lessons from a quasi-experimental study. *Journal of Political Science Education, 14*(2), 151-166. https://doi.org/10.1080/15512169.2017.1381613
Huber, C. R. & Kuncel, N. R. (2015). Does college teach critical thinking? A meta-analysis. *Review of Educational Research, 86*(2), 431-468. https://doi.org/10.3102/ 0034654315605917
Kuhn, D. (1991). *The Skills of Argument.* Cambridge University Press.
Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and AVOVAs. *Frontiers in Psychology, 4*, 1-12. https://doi.org/10.3389/fpsyg.2013.00863
Lee, H., Lee, J., Makara, M., Fishman, B. J., & Hong, Y. (2014). Does higher education foster critical and creative learners? An exploration of two universities in South Korea and the USA. *Higher Education Research and Development, 34*(1), 131-146. https://doi.org/10.1080/07294360.2014.892477
Lewis, A. & Smith, D. (1993). Defining higher order thinking. *Theory Into Practice, 32*(3), 131-137. https://doi.org/10.1080/00405849309543588
Lilly, T., Darr, P., Schmolesky, M., Lindley, T., Lundolph, P., Schilpzand, M., Shim, Y., Higgins, R., Weinstein, A., Burko, L. M., & Deutsch, D. V. (2022). Intending to teach critical thinking: A study of the learning impacts over one semester of embedded critical thinking learning objects. *Journal of the Scholarship of Teaching and Learning, 22*(3), 34-54. https://doi.org/10.14434/josotl.v22i3.31801
Lyne, R. O. A. M. (1978). The neoteric poets. *The Classical Quarterly, 28*(1), 167-187. https://doi.org/10.1017/S0009838800037848
Marshall, J. P., Higginbotham, B. J., Harris, V. W., & Lee, T. R. (2007). Assessing program outcomes: Rationale and benefits of posttest-then-retrospective-pretest designs. *Journal of Youth Development, 2*(1). https://doi.org/10.5195/jyd.2007.366
McClelland, D. C. (1973). Testing for competence rather than for intelligence. *American Psychologist, 28*(1), 1-14. https://doi.org/10.1037/h0034092
Montgomery, B. C. (2002). The relationship between training and leadership self-perception of selected Hispanic university students. [Unpublished Master Thesis]. Texas A&M University.
Muller-Frommeyer, L. C., Aymans, S. C., Bargmann, C., Kauffeld, S., & Herrmann, C. (2017). Introducing competency models as a tool for holistic competency development in learning factories: Challenges, examples and future application. *Procedia Manufacturing, 9,* 307-314. https://doi.org/10.1016/j.promfg.2017.04.015
Nicoll-Senft, J. M. & Seider, S. N. (2010). Assessing the impact of the 4MAT teaching model across multiple disciplines in higher education. *College Teaching, 58*(1), 19-27. https://doi.org/10.1080/87567550903245623
Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: a guide for practitioners. *Behavior Analysis in Practice, 5*, 2–11.
Pascarella, E. T. & Terenzini, P T. (2005). *How college affects students: A third decade of research (Vol 2).* Jossey-Bass.

Penkauskienė, D., Railienė, A., & Cruz, G. (2019). How is critical thinking valued by the labor market? Employer perspectives from different European countries. *Studies in Higher Education, 44*(5), 804-815. https://doi.org/10.1080/03075079.2019.1586323

Renaud, R. D. & Murray, H. G. (2007). The validity of higher-order questions as a process indicator of educational quality. *Research in Higher Education, 48,* 319-351. https://doi.org/10.1007/s11162-006-9028-1

Rohs, F. R. & Langone, C. A. (1997). Increased accuracy in measuring leadership impacts. *Journal of Leadership Studies, 4*(1), 150-158. https://doi.org/10.1177/ 107179199700400113

Rothwell, W. J. & Lindholm, J. E. (1999). Competency identification, modelling and assessment in the USA. *International Journal of Training and Development, 3*(2), 90-105.

Schwab, K. (2018). The global competitiveness report. *World Economic Forum.* http://www3.weforum.org/docs/GCR2018/05FullReport/TheGlobalCompetitivenessReport2018.pdf

Shermer, M. (2002). *Why people believe weird things: Pseudoscience, superstition, and other confusions of our time.* W. H. Freeman.

Smith, D. G. (1977). College classroom interactions and critical thinking." *Journal of Educational Psychology, 69*(2), 180-190. https://doi.org/10.1037/0022-0663.69.2.180

Stedman, N. L. P. & Adams, B. L. (2012). Identifying faculty's knowledge of critical thinking concepts and perceptions of critical thinking instruction in higher education." *NACTA Journal, 56*(2), 9-14.

Strong, J. R., Boyd, B. L., & Sanderson, H. (2021). The impact of participation in a college led critical thinking academy." *NACTA Journal, 65,* 280-284.

Tedesco-Schneck, M. (2013). Active learning as a path to critical thinking: Are competencies a roadblock?" *Nurse Education in Practice, 13,* 58-60. https://doi.org/10.1016/ j.nepr.2012.07.007

Terenzini, P. T., Springer, L., Pascarella, E. T., & Nora, A. (1995). Influences affecting the development of students' critical thinking skills. *Research in Higher Education, 36*(1), 23-39. https://doi.org/10.1007/BF02207765

Tienken, C. H., Goldberg, S., & Dirocco, D. (2012). Questioning the questions. *Kappa Delta Pi Record, 46*(1), 39-43. https://doi.org/10.1080/00228958.2009.10516690

Townsend, C. D. (2002). Leadership education: Fantasy or reality. *Journal of Leadership Education, 1*(1), 35-40.

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

Vial, A. C., Napier, J. L, & Brescoll, V. L. (2016). A bed of thorns: Female leaders and the self-reinforcing cycle of illegitimacy. *The Leadership Quarterly, 27*(3), 400-414. https://doi.org/10.1016/j.leaqua.2015.12.004

Watkins, D. V. & Earnhardt, M. P. (2015). Developing critical thinking within a master of science in leadership program. *The Academy of Educational Leadership Journal, 19*(1), 184-194.

Williams, J. R. & McClure, M. (2010). The effects of teaching methods in leadership knowledge retention: An experimental design of lecture, experiential, and public pedagogy. *Journal of Leadership Education, 9*(2), 86-99. https://doi.org/10.12806/V9/I2/RF7

Zimmerman, D.W. (1997). Teacher's corner: A Note on interpretation of the paired-samples t-test. *Journal of Educational and Behavioral Statistics, 22*(3), 349-360. https://doi.org/10.3102/10769986022003349