Higher Education Quality Council of Ontario

Building Capacity to Measure Essential Employability Skills:
A Focus on Critical Thinking
Appendix

Gary Kapelus, Nancy Miyagi, Valerie Scovill
George Brown College

Cite this publication in the following format:

Kapelus, G., Miyagi, N., Scovill, V. (2017) *Building Capacity to Measure Essential Employability Skills: A Focus on Critical Thinking Appendix.* Toronto: Higher Education Quality Council of Ontario.

Appendices

Appendix A: Essential Employability Skills

All graduates with the following Ontario college credentials, Ontario college certificate, Ontario college diploma and Ontario college advanced diploma, must be able to reliably demonstrate the essential employability skills required in each of the following six categories:

Skill category	Defining skills: Skill areas to be demonstrated by graduates	Learning Outcomes: The levels of achievement required by graduates The graduate has reliably demonstrated the ability to:
Communication	• Reading • Writing • Speaking • Listening • Presenting • Visual literacy	*Communicate clearly, concisely and correctly in the written, spoken and visual form that fulfills the purpose and meets the needs of the audience.* *Respond to written, spoken or visual messages in a manner that ensures effective communication.*
Numeracy	• Understanding and applying mathematical concepts and reasoning • Analyzing and using numerical data • Conceptualizing	*Execute mathematical operations accurately.*
Critical thinking & problem solving	• Analyzing • Synthesizing • Evaluating • Decision making • Creative and innovative thinking	*Apply a systematic approach to solve problems.* *Use a variety of thinking skills to anticipate and solve problems.*
Information management	• Gathering and managing information • Selecting and using appropriate tools and technology for a task or a project • Computer literacy • Internet skills	*Locate, select, organize and document information using appropriate technology and information systems.* *Analyze, evaluate and apply relevant information from a variety of sources.*
Interpersonal	• Teamwork • Relationship management • Conflict resolution • Leadership • Networking	*Show respect for the diverse opinions, values, belief systems and contributions of others.* *Interact with others in groups or teams in ways that contribute to effective working relationships and the achievement of goals.*
Personal	• Managing self • Managing change and being flexible and adaptable • Engaging in reflective practices • Demonstrating personal responsibility	*Manage the use of time and other resources to complete projects.* *Take responsibility for one's own actions, decisions, and consequences.*

Source: Ontario Ministry of Advanced Education and Skills Development.
http://www.tcu.gov.on.ca/pepg/audiences/colleges/progstan/essential.html

Appendix B: Literature Review

Executive Summary

The literature review focused on addressing three relevant research questions:

- What are the common measurable elements or constructs of critical thinking?
- How is critical thinking generally measured? What should be considered when developing rubrics specific for critical thinking?
- Can critical thinking skills be self-assessed?

It was found that there is no single widely accepted definition of critical thinking (CT). Definitions range from in-depth descriptions of declarative knowledge, procedural knowledge and metacognition (Kurfiss, 1988) to the habits undertaken when thinking critically (Marzano et al, 1993) to a simple broad description that critical thinking is "the king of thinking that professionals in the discipline use when doing the work of the discipline" (Taylor, 2004). Though definitions diverge, most share some commonalities, especially about CT being a skill or process that involves attitudes, habits, values, and behaviour (Rickles, Schneider, Slusser, WIlliams, & Zipp, 2013). As a result, there are different constructs of critical thinking.

In terms of measuring critical thinking, the literature specified a number of types of tests (multiple choice, short answer and multi-format) for a total of 15 tools across three studies and meta-analyses. Two of the studies did not recommend one specific type of tool; the third (Ku, 2009) noted the importance of multi-format response tools, specifically the Halpern Critical Thinking Assessment Using Everyday Situations.

Literature on developing and validating rubrics shows that certain factors are important, specifically clear definitions and operationalization, the use of benchmarks and anchors, and using tools that are validated and reliable. In terms of measuring reliability, most studies looked at inter-rater reliability, measuring percentage agreement or correlation coefficients between raters.

Other studies used ANOVAs to determine whether there are significant differences between raters. The example CT rubrics focus on most of the same values: clarity, accuracy, logic and understanding, and vocabulary.

Finally, studies looking at student self-assessment found generally positive results. Positive results were also reported for peer assessment. When using student self-assessment, one study recommended that students should be provided with feedback on their performance periodically and students should practice self-assessing and receive feedback on the accuracy of their self-assessments as it may be a skill that students can obtain through practice and feedback. The major drawbacks include the findings that certain subgroups of students tended to inflate their own grade, resulting in reliability issues.

I. **Critical Thinking**

Definition of Critical Thinking

There is no single widely accepted definition of critical thinking. Definitions range from in-depth descriptions of declarative knowledge, procedural knowledge and metacognition (Kurfiss, 1988) to the habits undertaken when thinking critically (Marzano et al, 1993) to a simple broad description that critical thinking is "the king of thinking that professionals in the discipline use when doing the work of the discipline" (Taylor, 2004). Though definitions diverge, most share some commonalities. Many agree that critical thinking is a skill or process and involves attitudes, habits, values, and behaviour (Rickles, Schneider, Slusser, Williams, & Zipp, 2013).

During the initial consultations with a group of GBC faculty who teach college English (COMM1007), it was revealed that professors do not follow a standard definition of critical thinking. For the purpose of this project, COMM1007 faculty were given definitions of critical thinking and asked to reflect on these definitions and their own understanding of critical thinking to provide recommendations on the elements of this skill that we should measure. Faculty favoured two definitions. The first states that critical thinking is "purposeful, self-regulatory judgment, which results in interpretation, analysis, evaluation, and inference as well as explanation of the evidential, conceptual and methodological considerations on which a judgment is based" (Facione, 1990). The second, simpler definition came from the American Association of College and Universities. The AACU CT Rubric states: "Critical thinking is a habit of mind characterized by the comprehensive exploration of issues, ideas, artifacts, and events before accepting or formulating an opinion or conclusion." (American Association of College and Universities, n.d.).

Faculty also provided input into a list of indicators for critical thinking for students. For students to demonstrate critical thinking, they must:

- Be able to state what the issue is;
- Defend/support their arguments using objective sources (learning about how to differentiate the value of different sources was considered very important, especially in the internet world) and using proper documentation;
- Interpret findings;
- Show self-awareness of their own thoughts and biases and the bias of the writer (or presenter, etc.);
- Be aware of and use the language of critical thinking — persuasive words, words that show gender bias, benevolent racism, etc.;
- Analyze, question, understand that a piece of writing is made up of multiple perspectives. Understanding that they bring a perspective to the reading. As writers, understand their audience and how to make their message clear; and
- Formulate different solutions to problems — give different possibilities, rank the problems.

Current Practices in Teaching Critical Thinking in Higher Education.

Brookfield (2011) suggests that the best way to teach critical thinking is to model it. Hemming (2000) also stated that critical thinking must be modeled using a "walk-through," showing the process critical thinking takes. Broadbear (2003) posited that teaching critical thinking should be structured on four elements: ill-structured problems (such as a case study or scenario), criteria for assessing thinking ("why" questions), student assessment of thinking (providing feedback and criteria to assess thinking), and improvement of thinking (practising and revising thinking techniques). Halpern (1998) touched on some of the same elements as Broadbear (2003) and Bensley (2010, see table below) did, noting the importance of dispositions and attitudes and metacognitive monitoring (assessing one's own thinking).

Table 6: Guidelines for teaching and assessing critical thinking from Bensley (2010)

1. Motivate your students to think critically	This can be completed by discussing issues relevant to real life that are impacted by critical thinking. This is also affected by attitudes and dispositions.
2. Clearly state the CT goals and objectives for your class	Be transparent about what will be accomplished in the class.
3. Find opportunities to infuse CT that fit content and skill requirements of your course	Engage students and allow them to practice their CT skills.
4. Use guided practice, explicitly modeling and scaffolding CT	Model critical thinking and allow students to practise critical thinking while giving feedback.
5. Align assessment with practice of specific CT skills	Integrate critical thinking into other content.
6. Provide feedback and encourage students to reflect on it	Give feedback when necessary in order to increase awareness.
7. Reflect on feedback and assessment results to improve CT instruction	Instructors can use feedback from students to improve instruction.

It appears that there are many common themes in teaching critical thinking. The major commonalities noted by researchers include modeling critical thinking, finding problems and allowing students to practise critical thinking, and giving feedback to students to allow them to improve.

Assessment of CT

In terms of specific CT assessment tools already in existence, there are several measures that are currently available (Paul & Elder, 2007; Ku, 2009; Bers, 2005). Ku notes the importance of measuring critical thinking using multiple response formats, such as The Halpern Critical Thinking Assessment Using Everyday Situations (HCTAUES), as they tend to capture critical thinking to a greater degree (see Table 7). The HCTAUES maps onto most of the criteria set out by Paul and Nosich's 21 criteria for the assessment of higher order thinking (n.d.). Authors agreed that critical thinking is highly contextual, but needs to be highly transferable in a variety of subjects and situations, and the assessment should allow learners to document, demonstrate and justify their own engagement in critical thinking (Paul & Nosich, n.d.; Brookfield, 1997).

Table 7: Available CT measures found in Paul and Elder (2007), Ku (2009), and Bers (2005)

Type	Name	Description	Validity
Multiple Choice Measures	Watson–Glaser Critical Thinking Appraisal	80 items that measure skills in five aspects of critical thinking: inference, recognition of assumptions, deductions, interpretation and evaluation of arguments.	.69-.85
	Cornell Critical Thinking Test	Story-based test in only multiple choice questions. Level X contains 71 items designed for Grade 4 to college students, and Level Z contains 62 items to be used for gifted high school and college students. Altogether, the two forms of the test measure seven aspects of critical thinking including induction, deduction, credibility, assumptions, semantics, definition and prediction.	.67-.90
	California Critical Thinking Skills Test	34 items measuring five categories of skills including interpretation, analysis, evaluation, inference and explanation.	.61-.72
	Collegiate Assessment of Academic Proficiency	Measures academic skills in five general education skill areas, among them critical thinking. Measures the ability to clarify, analyze, evaluate, and extend arguments.	n/a
	Critical Thinking Subtest: Analytic Reasoning	Provides evidence of whether, and to what extent, students are able to reason analytically. Machine-scorable (currently being developed).	n/a
	Critical Thinking: Concepts and Understanding	Provides evidence of whether, and to what extent, students understand the fundamental concepts embedded in critical thinking (and hence tests student readiness to think critically). Machine-scorable.	n/a
	Test of Everyday Reasoning	This 35-item multiple-choice test is produced by Insight Assessment and is designed to assess an individual's or group's basic reasoning skills. The test takes 50 minutes and yields a total score for overall reasoning skills, as well as three subscale scores in the categories of analysis, inference and evaluation.	n/a
Open- Ended	Ennis–Weir Critical Thinking Essay Test	Popular essay test of the general CT ability of high school or college students. It is a highly structured test that examines students' ability to identify built-in reasoning flaws in an argumentative passage, as well as their ability to defend their own arguments.	.82-.86
	Critical Thinking Assessment Test	Short answer test measures those components of critical thinking and problem solving that faculty across disciplines think are most important.	>.80
	Collegiate Learning Assessment	The Collegiate Learning Assessment is an innovative performance-based assessment that measures institutions' contributions to their students' CT and written-communication skills.	n/a
	Critical Thinking Reading	Provides evidence of whether, and to what extent,	n/a

	and Writing Test	students can read closely and write substantively (and hence tests student ability to read and write critically). Short Answer.	
	International Critical Thinking Test	Provides evidence of whether, and to what extent, students are able to analyze and assess excerpts from textbooks or professional writing. Short Answer.	n/a
	Commission Study Protocol for Interviewing Faculty Regarding Critical Thinking	Provides evidence of whether, and to what extent, critical thinking is being taught at a college or university (Can be adapted for high school). Based on the California Commission Study. Short Answer.	n/a
	Foundation for Critical Thinking: Consequential Validity using Assessment to Drive Instruction	It provides guidelines for gathering evidence of whether, and to what extent, critical thinking is being taught at a college or university (can be adapted for high school). Short Answer.	n/a
Multi-Response Format	The Halpern Critical Thinking Assessment Using Everyday Situations	Measures CT ability using questions set in authentic and believable contexts. The test consists of 25 scenario-based questions; each asks for open-ended responses as well as multiple-choice responses, totaling 50 questions. The test measures five categories of skills: verbal reasoning (e.g., recognizing the use of pervasive or misleading language), argument analysis (e.g., recognizing reasons and conclusions in arguments), hypothesis testing (e.g., understanding sample size, generalizations), using likelihood and uncertainty (e.g., applying relevant principles of probability, base rates), as well as decision-making and problem solving (e.g., identifying the problem goal, generating and selecting solutions among alternatives).	.81-.82

II.   Methods for Developing and Validating Rubrics

The use of scoring rubrics can have many benefits, such as increased consistency of scoring, the possibility to facilitate valid judgment of complex competencies, and promotion of learning. Using a meta-analysis, Jonsson and Svingby (2007) found that rubrics enhance reliability scoring in performance analyses, especially if the rubrics are analytical, topic-specific, and used with exemplars and rater training. They also found that rubrics may potentially promote learning and improve instruction.

Methods of rubric development seem to follow similar patterns, focusing on factors such as clear definitions and operationalization, the importance of using benchmarks or anchors, methods of ensuring reliability and validity. Allen and Knight (2009) proposed an eight-step process for rubric development. They used this method to develop a rubric for news release writing at Columbia College in Chicago.

Table 1: Eight steps in a process to collaboratively develop and validate a rubric (Allen & Knight, 2009)

	Description	Generalized Model	Specific Example
Step 1	Develop learning objectives for course	Discussion among faculty who teach course; input from professionals and the sources related to subject matter; research professional, certification requirements and exams	Discussions among faculty who teach public relations writing; input from professionals and academicians from other institutions; literature review, including information from professional organizations like Public Relations Society of America (PRSA)
Step 2	Identify sample of work to link learning outcomes to professional competencies	Review professional sources, including special reports, conference proceedings, etc.; academic literature search	Examined PRSA Port of Entry Report and PRSA requirements for professional accreditation; professional-source review; frequent meetings with and discussions among faculty teaching PR writing
Step 3	Develop evaluation rubric	Primary and secondary research; professional input; repeated discussions with faculty about work sample	Regular faculty meetings to discuss PRSA model and certification requirements, primary and secondary research, one on one meetings with academics, interviews with professionals
Step 4	Longitudinally test student learning as measured by rubric	Establish professional baseline; refine with academic and professional input	Two-test process: Test 1: PR professionals used rubric to evaluate student work sample; Test 2: professionals evaluated second sample of student work; faculty analyzed results
Step 5	Identify problems with sub-optimal performance	Use longitudinal data to assess rubric and pedagogy; evaluate uniformity in approach in multiple sections taught by many faculty	Faculty evaluated rubric against test results; then altered rubric's structure, expanded content, added detailed description to each of 5 categories; identified process to add weights to categories
Step 6	Improve construct validity of rubric	Determine weights for dimensions; collect and distill knowledge from professionals and academicians	Used the Delphi method first with professionals and then with faculty to reach consensus on weights in each category of the rubric
Step 7	Determine ability of rubric	Assess rubric's ability to differentiate individual	Faculty and professionals separately used rubric to assess and grade a weak and a

	Description	Generalized Model	Specific Example
	to differentiate between submissions	student work of varying performance; assess stability of rubric's standards; analyze variance to develop statistical evidence of rubric's accuracy and consistency	strong news release; used two-way ANOVA to determine if significant differences exist; used F test for equal variances
Step 8	Analyze data to improve reliability and further validate rubric	Establish inter-rater reliability; determine a baseline; test for rater repeatability and reproducibility	Faculty graded one news release; used Delphi method to reach consensus on grade; used F test for equal variances; looking ahead; faculty will grade same paper; we will then test for repeatability and reproducibility

These authors used the Delphi method for reaching consensus about the weights for each area of the rubric. The same method is used in Allen and Knight (2009) and Knight and Allen (2012). The Delphi method is a process where participants express an opinion on the issue at hand. For the first round, a facilitator collects feedback and then reports the feedback to the group without identifying the participant behind the opinion. Participants are encouraged to use the feedback to modify their opinion for the next round. Rounds continue until a unanimous consensus is reached. Allen and Knight also evaluated reliability using 10 strong and 10 weak news releases. Faculty and professionals were asked to mark the news releases using the rubric they had developed. An ANOVA showed that there were significant differences between the strong and weak news releases, showing the rubric is accurate in differentiating between strong and weak work. A more concise approach was taken by Wolf and Stevens (2007), specifying just three steps to rubric development (see Table 2).

Table 2: Wolf and Stevens's three steps to rubric development

Step 1: Identifying Performance Criteria	Define the criteria that defines the performance. It should be three to six criteria, all of which are observable and measurable. Some criteria may be weighted. Criteria may be determined by the professor alone or in conjunction with students.
Step 2: Setting Performance Levels	Typically, rubrics have between three and six rating levels. The fewer the levels of performance for the rater to consider, the greater the reliability and efficiency in scoring the performance. The more levels, the lower the reliability in scoring and the more time it will take for raters to make the decision.
Step 3: Creating Performance Descriptors	Provide enough information to guide the creation and scoring of the project, but not so much that it overwhelms the reader or the performer. Parallel structure across descriptions for each criterion (e.g., delivery) is important. The more parallel the descriptions are in form and content, the more dependable and efficient the scoring will be.

Validating the Rubric

A meta-analysis from Jonsson and Svingby (2007) show that most articles on rubric development look at reliability using percentage agreement among raters or correlation among raters. In the studies they reviewed, the range of correlations is .27–.98, with the majority between .55 and .75. In consistency estimates, values above .70 are deemed acceptable. The authors noted these factors in rubric reliability based on their review (p. 135):

- Benchmarks are most likely to increase agreement, but they should be chosen with care since the scoring depends heavily on the benchmarks chosen to define the rubric.
- Analytical scoring is often preferable, but perhaps not so if the separate dimension scores are summarized in the end.
- Agreement is improved by training, but training will probably never totally eliminate differences.
- Topic-specific rubrics are likely to produce more generalizable and dependable scores than generic rubrics.
- Augmentation of the rating scale (for example, the raters can expand the number of levels using + or – signs) seems to improve certain aspects of inter-rater reliability, although not consensus agreements. For high levels of consensus agreement, a two-level scale (for example, competent – not competent performance) can be reliably scored with minimal training, whereas a four-level scale is more difficult to use.
- Two raters are, under restrained conditions, enough to produce acceptable levels of inter-rater agreement.

For the purpose of this project, it was felt that assessing both inter-rater agreement (IRA) and inter-rater reliability (IRR) of blind markings would provide useful information to inform future changes to the assessment tool. While agreement and reliability are used interchangeably in research studies, each represents a different concept. "Agreement is defined as the degree to which scores/ratings are identical, whereas reliability relates to the extent of variability and error inherent in a measurement" (Gisev, Pharm, Bell, & Chen, 2013, p. 331). IRA, on the other hand, represents the level of exact agreement among markers. As a rule-of-thumb, we used 75% as the minimum acceptable score for absolute agreement (IRA) and .8 for IRR (Graham, Milanowski and Miller, 2012).

III. Student Self-assessment

Sadler and Good (2006) found that self-grading appeared to result in increased student learning in a study on Grade 7 students, though lower performing students tended to inflate their grades. Kruger and Dunning (1999) found that when people self-assess, those who score in the bottom quartile tend to overestimate themselves by 50% on average. The authors posited that this effect was due to the participants' low cognition; not only were participants unskilled in the domains in which they were tested, but their "incompetence robs them of the metacognitive ability to realize it" (Kruger & Dunning, 1999). Raupach, Münscher, Beibbarth, Burckhardt and Pukrop (2011) and Darrow, Johnson, Miller, and Williamson (2002) found that student self-assessment was an appropriate tool. Sarin and Headley (2002) found the same generally positive results, also reporting that while student's self-assessment grades were correlated with test grades, the correlation was only moderate .18-.26. Muñoz and Alvarez (2007) found from moderate to

high correlations between student self-assessment and teachers' assessment and recommended that self-assessment be ongoing, a culture acceptance of self-assessment be raised, students be provided with assistance in the process of self-assessment, and teachers be trained for student autonomy.

A meta-analysis on self-assessment in higher education found that certain factors affected the degree of agreement between student self-assessment and teacher assessment. These factors included study design (well-designed studies, as determined by the authors, had higher student-teacher agreement) course level (study participants in more advanced courses tended to be a better judge of their performance) and study area (students in science courses tended to have higher agreement with teachers) (Falichov & Boud, 1989). Overall, the major benefits to self-assessments include time and cost-savings, as well as the aforementioned results that in general, self-assessments are similar to faculty assessments. The major drawbacks include the findings that certain subgroups of students tended to inflate their own grade, resulting in reliability issues.

Recommendations to strengthen the relationships between student self-assessment and objective assessments include periodic feedback on student performance throughout the course. As well, students should practise self-assessing and receive feedback on the accuracy of their self-assessments as it may be a skill that students can obtain through practise and feedback (Sitzmann, et. al., 2010). For the purposes of this project, self-assessment should be used if students have the chance to practise self-assessing while receiving feedback on their self-assessment. Special attention should be paid to students who have lower scores on objective measures to ensure their self-assessments are not inflated during practice self-assessments.

Appendix C: George Brown College CT Rubric — Final Version

Criteria	Descriptions of levels			
	Exemplary/exceeds expectations	Meets expectations	Below expectations	Inadequate
(1) *Clarifies the issue to be discussed and/or the position to be argued in this paper*	✓ Clearly states the specific issue/position to be discussed or argued in the paper ✓ Provides additional information about why and how this issue was selected	✓ Clearly states the specific issue/position to be discussed or argued in the paper	✓ Identifies but does not clearly state the specific issue/position to be discussed or argued in the paper	✓ Does not acknowledge the issue/position to be discussed or argued in the paper
(2) *Identifies the sources of ideas or evidence used in developing the argument or conclusions*	✓ Identified sources of ideas/evidence are relevant to the assignment and credible ✓ Identifies a broader or more comprehensive range of sources than is required	✓ Identified sources of ideas/evidence are relevant to the assignment and credible	✓ Some identified sources are either irrelevant to the assignment, questionable or not credible	✓ Does not identify the sources of ideas/evidence
(3) *Analyzes the ideas or evidence to develop the argument or conclusions*	✓ Analysis both integrates and synthesizes all of the identified sources of ideas/evidence ✓ Analysis is logical. ✓ Analysis is on topic ✓ Analysis leads to, strengthens and focuses the argument or conclusions	✓ Analysis includes all the identified sources of ideas/evidences. ✓ Analysis is logical ✓ Analysis is on topic ✓ Analysis leads to and supports the argument or conclusions	✓ Analysis reflects some but not all of the identified sources of ideas/evidences ✓ Analysis is not logical ✓ Analysis is not on topic ✓ Analysis does not lead to or support the argument or conclusions	✓ Does not include an analysis
(4) *Critiques contradictory evidence, information, experts' opinions and/or methodologies*	✓ Identifies and questions/challenges contradictory evidence, information, experts' opinions and/or methodologies as presented ✓ Analyzes the strengths and limitations of the evidence being challenged	✓ Identifies and questions/ challenges contradictory evidence, information, experts' opinions and/or methodologies as presented	✓ Accepts contradictory evidence, information, experts' opinions and/or methodologies without question or criticism	✓ Does not acknowledge contradictory evidence, information, experts' opinions and/or methodologies

(5) *Acknowledges personal biases or assumptions*	✓ Articulates/explains personal biases or assumptions ✓ Acknowledges the relevance or impact of personal biases or assumptions on their analysis or conclusions	✓ Articulates/ explains personal biases or assumptions	✓ Acknowledges but does not articulate personal biases or assumptions	✓ Does not mention or recognize personal biases or assumptions
(6) *Describes conclusions*	✓ Conclusion is logical ✓ Conclusion arises from effectively synthesized and thoroughly evaluated evidence in the argument ✓ Solutions are prioritized where there is more than one	✓ Conclusion is logical ✓ Conclusion is rooted in/linked to the evidence in the argument ✓ Solutions are prioritized where there is more than one	✓ Conclusion lacks logic ✓ Conclusion is weakly linked, if at all, to evidence in the argument ✓ Where an attempt has been made to prioritize solutions, there are gaps in logic and/or understanding	✓ Conclusion is absent, incorrect or irrelevant to the evidence in the argument ✓ Where there is more than one solution presented, these are incorrectly or not at all prioritized

Appendix D: Critical Thinking — Blind Marking Recording Form (Version 3)

General instructions:

You are being asked to use our generic CT assessment rubric to mark 6 anonymous samples from 3 different assignments, each of which required students to use CT skills. For context, you have also been given the actual description of each assignment (what the students read), with the original marking scheme and logistics removed. In setting each of these assignments, the instructors were guided by the six CT constructs in our rubric. Your task, in reviewing each sample, is to identify whether there is evidence of each of these CT constructs in the sample and to determine the level at which each of these constructs is demonstrated in the sample by placing an "**X**" where appropriate.

Once this form is complete, email it to *CONTACT* by *DATE*

Section One

ASSIGNMENT 1: [NAME]
STUDENT #1

Criteria	Description of levels			
	Exemplary/ exceeds expectations	Meets expectations	Below expectations	Inadequate
Clarifies the issue to be discussed and/or the position to be argued in this paper				
Identifies the sources of ideas or evidence used in developing the argument or conclusions				
Analyzes the ideas or evidence to develop the argument or conclusions				
Critiques contradictory evidence, information, experts' opinions and/or methodologies				
Acknowledges personal biases or assumptions				
Describes conclusions				

Section Two

OPTIONAL: Please comment on your experience with the rubric when assessing different aspects of critical thinking

The Criteria:

Are the criteria clear? Are the criteria different from each other?

```
```

The Descriptions (refer to the rubric form):

Do the descriptions match the criteria? Are the descriptions clear and different from each other?

```
```

The Scale:

Do the descriptors under each level ('Inadequate', 'Below expectations' etc.) truly represent that level of performance?

```
```

The OVERALL rubric:

Does the rubric clearly connect to the outcomes that it is designed to measure? Can the rubric be understood by external audiences? Does it reflect teachable skills?

```
```

Appendix E: Qualitative Analysis — Faculty Feedback About Their Experience with the CT Rubric in Assessing CT Skills

Qualitative analysis CT Rubric #2

Faculty participating in blind marking shared their experience with the rubric tool by answering specific questions in the space provided in the blind marking recording form. A summary of their answers are provided below. Input from this round of marking was used to adjust the rubric for testing in the third and final blind marking scheduled.

Criteria - Are the criteria clear? Are the criteria different from each other?

Markers were almost equally split between those who found the criteria sufficiently clear and those who felt that some criteria needed adjustments. While this question was in reference to the six criteria or constructs being distinguishable from one another, comments from the latter group of faculty were mainly about other aspects of the rubric, especially about the mismatches they encountered between the assignment (what students were asked to demonstrate) and the rubric (what markers were asked to assess), which made marking very difficult. Faculty also noted a lack of clarity in the rubric about what exactly students were being asked to demonstrate and a few questioned whether the rubric was missing the assessment of technical skills, such as the ability to cite sources appropriately, for example. Based on the markers' feedback, included below are comments from select markers:

"…while the criteria in the rubric are clear, they are also very specific and the assignments do not all seem to require achievement of all elements of the rubric." — Teacher volunteer #8

"…choice of support sources, I was not sure if simply a list of references was sufficient or if references should be cited in the text…" — Teacher participant #2

"…personal biases…should students be stating their own biases clearly or can this be implied in the content of the paper? ... [conclusions] Are you asking students to write a strong conclusion at the end, or to draw logical, well-reasoned conclusions throughout the paper?" — Teacher volunteer #3

"…shouldn't there be some consideration given to sources referenced or cited appropriately?"
— Teacher volunteer #6

Description - Do the descriptions match the criteria? Are the descriptions clear and different from each other?

There was a general consensus among respondents that the descriptions matched the criteria and were distinct enough; however, additional work was still needed. Faculty noted that greater variability across levels was required. In some cases the assessment fell in between levels or the level descriptors lack indicators to guide the marker. This, in turn, forced some markers to check off a level they were unsure about and/or made them question their choice. This may help to explain the low IRA results (exact agreement among markers) when there is little discernable difference between close levels. Respondents noted that wording in the descriptors under each level were sometimes too vague to help assess students'

level of performance. One respondent questioned the importance given to the number of sources chosen as a CT skill, rather than to use and analyze sources to support a position taken.

"I find myself trying to decide which component of the criteria is more important, if student's performance falls between two levels." — Teacher participant #4

"Overall I find [descriptions] too vague and lacking specific indicators upon which to evaluate learners' mastery of material." — Teacher volunteer #1

"I actually felt that the quality of the paper was quite different across the two of the same assignment, but I was forced to mark them similarly due to the low variability in the scale." — Teacher participant #5

"[support sources] Level 3 indicates 'sufficient' in number and Level 4, 'appropriate' in number. For a student to clearly achieve a Level 4, he/she would need to know the distinction between sufficient and appropriate." — Teacher volunteer #8

"It's not the number of references that is important but whether or not the knowledge the student has obtained from the reference can be used to support the position they are taking in their paper." — Teacher volunteer #5

Scale - Do the descriptors under each level truly represent that level of performance?

Overall, faculty comments were very similar to the ones provided about the descriptions, especially around the lack of specificity that made levels close to each other not being distinguishable enough. In other cases, faculty noted that the level descriptors, especially "level 1-missing" did not seem appropriate for the CT skill being assessed (for example, personal bias missing or the absence of personal bias could be interpreted as a good outcome, but the scale in the rubric suggests otherwise). The conclusion descriptors for missing information did not follow the same pattern as descriptors from other levels to indicate progression. In summary, some markers felt that the number in the scale needed to increase and/or that more specificity was needed in the level descriptors; one faculty specially made references to the numbers in the scale (4) not aligning with the 5 letter grading system (A, B, C, D, F) that GBC professors follow when assigning marks.

Overall rubric - Does the rubric clearly connect to the outcomes that it is designed to measure? Can the rubric be understood by external audiences? Does it reflect teachable skills?

Comments were in line with the feedback provided, earlier. Except for a few cases, faculty generally found the rubric and its content easy to understand but some had concerns and/or reservations about using the same standard rubric for grading purposes. The mismatch faculty observed between the instructions and the rubric requirements led some to believe that having a standard CT rubric that could be applied across the curriculum/assignments is not possible. Some respondents also noted that the rubric does not allow the marker to evaluate the whole paper. Moreover, the attempt to separate CT from skills, such as language, is problematic — especially in cases when the writing skill is so poor that it impedes the assessment of CT.

"It was often unclear whether the problem was poor critical thinking skills when the language and writing were almost incomprehensible. If you can't write clearly, it's hard to demonstrate that you can think clearly, and that may be part of what's at the heart of this matter." — Teacher volunteer #8

"I found the rubric challenging to use, again because of the one-size-fits-all approach. It is easier to use a rubric that accompanies an assignment with well crafted, specific and discrete questions, with criteria and descriptors that reflect assignment expectations … For the most part, I found the student writing examples to be of poor quality … but there was no rubric criteria to reflect this skill area. I was therefore forced to push that evaluation into the Missing Information description level." — Teacher volunteer #1

Qualitative analysis CT Rubric #3

Nineteen faculty markers provided feedback about their experience with the rubric.

Criteria - Are the criteria clear? Are the criteria different from each other?

The majority of respondents felt the criteria were clear overall and different from each other; a notable improvement from the second blind marking. Of note is that all six criteria were mentioned once or twice by faculty who suggested areas for change, such as:

- Add, revise criteria descriptors/use action verbs
- Expand criterion name: C2 *Identifies <u>and</u> properly documents;* C4 *Acknowledges and critiques*

Few respondents reported a mismatch between the assignment and the rubric requirements, the inapplicability of the rubric to certain criteria which made marking difficult (for example, difficult to grade bias) or the lack of clarity in the descriptors to guide marking.

> **Description -** Do the descriptions match the criteria? Are the descriptions clear and different from each other?

While faculty generally found the descriptions matched the criteria, some of their concerns centred around the lack of discernable differences between levels that are next to each other, especially the levels either side of "below": below vs. meets; below vs. adequate. In addition, faculty felt that certain definitions and descriptors were either missing, or inapplicable and/or hard to interpret.

"… Analysis is a particular skill that involves identifying salient issues, grouping/categorizing them, drawing similarities and differences, etc. This needs to be reflected in the descriptions ... [same with] conclusion."
— Teacher volunteer #14

"… below expectations description for acknowledges personal biases and assumptions. I'm not … sure how you can acknowledge the assumptions without describing what they are." — Teacher participant #4

"I'm finding the phrase 'identified' sources confusing. How would the meaning be different if you took out identified?… [for] *'acknowledges but does not articulate personal biases'; how would I know if she acknowledges her biases if she doesn't present them?" — Teacher volunteer #8*

Scale - Do the descriptors under each level truly represent that level of performance?

Only seven out of 18 faculty who responded to this question fully agreed that the descriptors represent the level of performance. The remaining faculty either reiterated their previous comments or did not answer this question likely because they had already addressed it throughout the document.

Overall rubric - Does the rubric clearly connect to the outcomes that it is designed to measure? Can the rubric be understood by external audiences? Does it reflect teachable skills?

The majority of respondents (14 out of 19) agreed that the rubric connected to the outcomes being measured; although fewer specifically stated that the rubric reflects teachable skills. Comments about the scale and descriptors, in particular, were somewhat similar to those from the second blind marking. The most salient points include:

- Instructions (what students were asked to do) and the rubric (what faculty were asked to assess) did not always match, especially criterion C5 (acknowledges personal biases or assumptions). The demonstration of this criterion as described in and assessed by the rubric was not made overt in the assignment. Markers were left wondering whether students did not meet the requirement or were not asked to fulfill it.
- Interestingly, despite that, C1 (Clarifies the issue…) generated the least agreement among markers; none mentioned it as being problematic.
- Faculty responses indicate that additional editing to the levels and descriptors is still needed to make them clearly distinguishable from each other and applicable to all CT constructs. In addition, a definition of skills and/or additional descriptors (i.e., bias, analysis) was recommended to provide greater clarity to those using the rubric. One respondent suggested collapsing the last two levels as they appear to be synonymous: Inadequate and Below Expectations.
- Lastly, few faculty also commented on the difficulties in assessing CT when student writing is so poor that they are unsure if the problem relates to language or reasoning and added that the rubric needed a criterion for assessing the quality of writing.

Blind marking — Unintended positive outcomes

While the intent of the blind marking activity was to develop a valid tool to assess the acquisition of CT among GBC students, the process engaged faculty participants and volunteers in different ways; resulting in unintended positive outcomes, especially around changes in attitudes (acceptance/intention), behaviour (application of knowledge/new concepts), and self-reflective thought about one's teaching practice.

 "I have made Critical Thinking assignments explicit in several of my courses to highlight them and make students aware of what they are learning." — Teacher participant #2

"As a result of this blind marking assignment, I am focusing so much more on the critical thinking skills you have raised in this rubric, and, therefore, I will have a better idea of how/whether the rubric reflects the teachable skills at the end of this term!" — *Teacher volunteer #8*

"[The rubric] has left me wondering how I might change my teaching to address some of these skills (specifically, critiquing contradictory evidence)." — *Teacher volunteer #9*

"My first reaction as I read this latest iteration of the rubric was to say – aloud – "Wow, I'm going to use this!" ….Though I am still half convinced this is actually an assessment of whether faculty assign grades based on skills in writing and presentation rather than as an objective response to the stated assignment criteria." — *Teacher volunteer #12*

Appendix F: Faculty Participants' Reflection/Experience with CT and Project Overall

Thinking about the <u>six CT constructs</u> that were specifically chosen for the CT rubric…

At what point and how did you explicitly talk about any of these six CT constructs, or about CT generally, to your students?

Faculty indicated that CT constructs were discussed with students at multiple time points during the semester. CT constructs and the rubric were introduced at the beginning of the course as well as throughout the semester during specific classroom activities and assignments. Faculty noted numerous opportunities to highlight the importance of CT for students and the benefits/purpose of developing such skills. This helps increase the number of opportunities for students to grasp the concept when it is presented in different situations throughout the course. Faculty also noted that use of the rubric was integral to the process of teaching CT constructs. As instructors, they benefited from having the rubric as a teaching tool.

"I projected the rubric on the screen and went through the concept of critical thinking, its importance in liberal studies and beyond the classroom…" — General Education/Liberal Arts faculty participant

What changes did you make in what you actually taught/demonstrated as a result of these CT constructs?

Faculty described changing teaching practice so as to make CT more explicit in the classroom. While a number indicated that they had always incorporated CT concepts into their courses, they were now intentionally placing more overt emphasis on CT for their students. The language provided in the rubric became a vital teaching tool as it facilitated student understanding of the concepts and improved professor feedback to students on their assignments. Faculty also noted integrating CT constructs in new and different ways into the classroom. For example, new CT assignments and classroom exercises were developed. This signals a behaviour change on the part of faculty to incorporate more CT into the classroom.

"Included several exercises … to help students identify the issue at hand and practised having students taking and defending their position …" — General Education/Liberal Arts faculty participant

"I have created three more assignments based on the CT rubric for another course." — General Education/Liberal Arts faculty participant

"The language provided in the rubric was also extremely helpful in communicating these skills to the students." — General Education/Liberal Arts faculty participant

"I found that I started to think about how I was teaching differing points of view and the type of information I was including as examples. I think I really became more aware of my own biases and perspective in my teaching." — General Education/Liberal Arts faculty participant

How did you use these CT constructs in designing/revising your assignments?

The rubric was reported as helping to refine or alter assignments to better reflect CT skills and constructs. Faculty also noted altering the rubric to better fit the parameters of certain assignments. The rubric helped

faculty to improve the assignment instructions given to students, which in turn provided students with increased clarity of their expectations on how they would be graded.

"I have taken some of the wording from the six constructs and matched it up with my assignment instructions." — General Education/Liberal Arts faculty participant

"I used it to actually help clarify assignment instructions. Before this project I did not have a concrete way of deciding when students were being critical. I had an idea of what I was looking for but found it difficult to communicate this to the students." — General Education/Liberal Arts faculty participant

Were there any of the six constructs that you chose not to use or found not relevant to your needs? Please explain.

Faculty provided a range of responses. Some found all constructs relevant while others dropped certain constructs according to the needs of particular assignments. The Personal Bias construct was perceived as not applicable to all courses (for example, science courses).

Were there other CT constructs that you used instead? If so, how did you use them in your teaching or assignments or marking?

No additional CT constructs were added by respondents.

Please describe the benefits and challenges of using these CT constructs in your teaching.

Numerous benefits to incorporating the CT constructs into teaching were noted such as: Improved quality and consistency of grading (i.e., less subjectivity, more clarity for both professor and student); enhanced student depth of understanding of CT constructs and increased student confidence in their own CT skills; improved faculty understanding of CT constructs and how to apply them to the classroom and better convey concepts to students. Challenges reported by faculty included time constraints, such as needing extra time during class to teach concepts and for students to adequately grasp concepts. Some challenges incorporating the rubric were noted, where it was found to not be applicable to every assignment.

"The biggest benefit was that my students became highly aware that they are actually good at CT, so they could comfortably say; "I am good at critical thinking." — General Education/Liberal Arts faculty participant

"The greatest challenge is not the constructs or the rubric in its entirety, but my own level of complacency when attempting to judge if critical thinking was happening. I guess what I'm saying is that I was hesitant to implement the rubric and I felt like we were trying to put a square peg in a round hole. But once accepting that the tool had to evolve with my teaching and assessments, I began to see its worth and value." — General Education/Liberal Arts faculty participant

Thinking about the <u>CT rubric tool</u> itself …

At what point and how did you introduce the CT rubric tool to your students during the course? How did you actually use the CT rubric tool in marking your students' work?
Introduction of the CT rubric tool ranged considerably from not at all, to early in the course, to not until specific assignments were discussed. With respect to use of the rubric for marking, faculty noted that they frequently used a modified version for marking. For example, some faculty reported setting benchmarks for "meets expectations" then adjusting based on student performance, others noted adjusting the score range on the rubric, or omitting certain constructs from the rubric that did not fit the assignment. Faculty also reported providing a copy of the rubric to students and discussing it during class.

Please describe the benefits and challenges in using the CT rubric tool

A number of benefits to using the CT rubric tool were noted, such as helping to make concepts more tangible for students, clarifying expectations, and helping to provide more accessible language with which to discuss CT concepts in the classroom. It was noted as a helpful teaching tool for professors.
Some challenges to using the rubric included needing to adjust the rubric to meeting the specific needs of the assignments. Some faculty also noted difficultly in applying the rubric in a consistent and fair manner to student work.

"Simplified my work by helping me to structure my assignments in the most useful way. In other words, I could get the most out of my assignments. This allowed me to reduce the number of assignments I gave and at the same time maximize student benefits in terms of critical thinking. In all, I think that this is a win-win situation."— General Education/Liberal Arts faculty participant

Looking ahead …

What concrete changes would you make to your course outline as a result of using the CT constructs and rubric?

Faculty reported changing course outcomes or assignment objectives to better align with CT constructs. They also reported making CT more explicit in the classroom for students and placing more emphasis on these concepts at the beginning of their courses.

After this CT experience, how would you say your own teaching practice or approach has been affected?

Faculty reported positive impacts on their teaching practice. For example, it allowed faculty to achieve greater balance between the teaching of content and the process of developing CT skills. It was also noted that use of CT has brought the course more in-line with the department's philosophy (specifically Liberal Arts). Faculty also reported placing more explicit focus on CT and use of the rubric with students so that they could benefit from clearer expectations.

"As a result of this project and in consultations with [name]*, I have slightly modified my course outcomes to better include a CT focus. I am certainly more aware of critical* thinking *skills myself, and I am still fine-tuning my delivery and inclusion of CT skills within my online courses."* — General Education/Liberal Arts faculty participant

How do you think your students have benefited from this different way of focusing on CT?

Students were perceived as benefiting in multiple ways such as improved clarity of assignment expectations and more concrete understanding of CT concepts. The use of CT skills was noted as supporting students in their future studies and increasing student confidence in own abilities. It was also noted that use of the rubric made marking less subjective and more transparent for students.

What concrete suggestions would you have for another teacher who is interested in integrating CT into their teaching and courses?

A number of practical suggestions for other teachers were noted such as: taking the time to understand the rubric and how to adjust course assignments and marking procedures, providing additional and explicit instructions about CT to students and how the rubric will be used, and being open to modifying the rubric to meet the needs of the course.

Blind marking volunteers' experiences

In addition to gathering feedback from General Education/Liberal Arts faculty participants, teachers who volunteered to mark student samples also provided feedback about their experience using the CT rubric. While the intent of the blind marking activity was to develop a valid tool to assess the acquisition of CT among GBC students, the process engaged faculty participants and volunteers in different ways, resulting in unintended positive outcomes, especially around changes in attitudes (acceptance/intention), behaviour (application of knowledge/new concepts), and self-reflective thought about one's teaching practice.

"I have made Critical Thinking assignments explicit in several of my courses to highlight them and make students aware of what they are learning." — Teacher participant #2

"As a result of this blind marking assignment, I am focusing so much more on the critical thinking skills you have raised in this rubric, and, therefore, I will have a better idea of how/whether the rubric reflects the teachable skills at the end of this term!" — Teacher volunteer #8

"[The rubric] has left me wondering how I might change my teaching to address some of these skills (specifically, critiquing contradictory evidence)." — Teacher volunteer #9

"My first reaction as I read this latest iteration of the rubric was to say – aloud – "Wow, I'm going to use this!" …. Though I am still half convinced this is actually an assessment of whether faculty assign grades based on skills in writing and presentation rather than as an objective response to the stated assignment criteria." — Teacher volunteer #12

Appendix G: Student Self-Assessment of Critical Thinking – Entry Survey

First three letters of your mother's first name.			First three letters of **your** month of birth			**Your** day of Birth		

1. In which program are you currently enrolled?

2. In which **year** of your program are you enrolled?

3. What is the HIGHEST level of education you completed **before entering this program**?

    a. Less than High School Diploma
    b. High School Diploma/GED
    c. Some college/university (no certificate, diploma or degree earned)
    d. College diploma, trade/vocational certificate, or registered apprenticeship
    e. University bachelor's degree (eg. BA, BSc, B.Ed.)
    f. Graduate degree (Master's, PhD)
    g. Other:_____

4. What is your age? _____

SCENARIO: Your instructor asks you to write a paper that requires you to learn about a controversial topic in enough detail to form and describe your own opinion. You need to provide supporting evidence on how you arrived at your conclusion.

Please rate your current knowledge level in regard to completing such an assignment, ranging from '**0**' (Not at all knowledgeable) to '**6**' (Extremely knowledgeable).

1. Given a situation or topic to consider, I know how to identify the main issue(s).
    0   1   2   3   4   5   6

2. When given a paper to read, I know how to identify the author's views and objectives in the issue presented.
    0   1   2   3   4   5   6

3. I know how to tell if the evidence presented by the author is relevant to the argument.
    0   1   2   3   4   5   6

4. I know how to tell when additional information is needed in order for me to confirm or reject a given claim.
    0   1   2   3   4   5   6

5. I know how to determine the similarities and differences among the various opinions presented for a given problem.

        0   1   2   3   4   5   6

6. I know how to consider opposing views in support of information when controversial issues are examined.

        0   1   2   3   4   5   6

7. I know how to systematically analyze the problem using multiple sources of information to draw conclusions.

        0   1   2   3   4   5   6

8. I know how to tell if the information supporting the arguments is reliable.

        0   1   2   3   4   5   6

9. I know how to explain how my personal preference/bias may affect my analysis.

        0   1   2   3   4   5   6

10. I know how to develop logical conclusions.

        0   1   2   3   4   5   6

11. I know how to arrive at conclusions that are supported with strong evidence.

        0   1   2   3   4   5   6

12. I know how to clearly explain my reasons for accepting or rejecting an author's claim.

        0   1   2   3   4   5   6

13. I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions.

        0   1   2   3   4   5   6

To answer the following, think in terms of the above skills **in general.**

14. How much instruction or training have you received previously relative to the above skills?

    a. I haven't received instruction/training
    b. I have received some instructions/training
    c. I have received substantial instruction/training

Thank you!

This project has been approved by the George Brown College Research Ethics Board, Approval No. 6003842

Appendix H: Student Self-Assessment of Critical Thinking – Exit Survey

First three letters of your mother's first name.			First three letters of **your** month of birth			**Your** day of Birth		

1. In which program are you currently enrolled?

2. In which year of your program are you enrolled? _____

3. In your own words, please define the term '*Critical Thinking*'

SCENARIO: Your instructor asks you to write a paper that requires you to learn about a controversial topic in enough detail to form and describe your own opinion. You need to provide supporting evidence on how you arrived at your conclusion.

1. Given a situation or topic to consider, I know how to identify the main issue(s).
    0   1   2   3   4   5   6

2. When given a paper to read, I know how to identify the author's views and objectives in the issue presented.
    0   1   2   3   4   5   6

3. I know how to tell if the evidence presented by the author is relevant to the argument.
    0   1   2   3   4   5   6

4. I know how to tell when additional information is needed in order for me to confirm or reject a given claim.
    0   1   2   3   4   5   6

5. I know how to determine the similarities and differences among the various opinions presented for a given problem.
    0   1   2   3   4   5   6

6. I know how to consider opposing views in support of information when controversial issues are examined.
    0   1   2   3   4   5   6

7. I know how to systematically analyze the problem using multiple sources of information to draw conclusions.
    0   1   2   3   4   5   6

8. I know how to tell if the information supporting the arguments is reliable.
    0   1   2   3   4   5   6

9. I know how to explain how my personal preference/bias may affect my analysis.

    0   1   2   3   4   5   6

10. I know how to develop logical conclusions.

    0   1   2   3   4   5   6

11. I know how to arrive at conclusions that are supported with strong evidence.

    0   1   2   3   4   5   6

12. I know how to clearly explain my reasons for accepting or rejecting an author's claim.

    0   1   2   3   4   5   6

13. I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions.

    0   1   2   3   4   5   6

14. Have your critical thinking skills, as described above, improved over the semester as a result of taking this course?

    a. substantially improved
    b. somewhat improved
    c. no change
    d. don't know

Thank you!

This project has been approved by the George Brown College Research Ethics Board, Approval No. 6003842

Appendix I: Student Self-Assessment of Critical Thinking – Results

STUDENTS WHO COMPLETED BOTH ENTRY AND EXIT SURVEYS

Table 1: Characteristics of the students who completed BOTH the entry and exit surveys

Factors	Phase-II (Control) (n = 77) n (%)	Phase-III (Intervention) (n = 94) n (%)	p-value
Area			
Arts and Humanities	46 (59.7%)	2 (2.1%)	
Science and Technology	24 (31.2%)	72 (76.6%)	< 0.001
Social Sciences	7 (9.1%)	20 (21.3%)	
Centre			
Arts & Design	7 (9.1%)	8 (8.5%)	
Business	20 (26.0%)	25 (26.6%)	
Community Serv. & ECE	4 (5.2%)	3 (3.2%)	
Health Sciences	10 (13.0%)	12 (12.8%)	0.002
Construction & Eng. Tech.	15 (19.5%)	11 (11.7%)	
Hospitality & Culinary Arts	16 (21.0%)	7 (7.4%)	
Preparatory & Liberal Studies	4 (5.2%)	28 (29.8%)	
Year			
One	18 (23.4%)	42 (44.7%)	
Two	42 (54.5%)	26 (27.7%)	< 0.001
Three	9 (11.7%)	9 (9.6%)	
Four	0 (0.0%)	17 (18.1%)	
Education			
High School or Less	45 (58.4%)	55 (58.5%)	
Some PSE	12 (15.6%)	17 (18.1%)	0.915
PSE	19 (24.7%)	22 (23.4%)	
Age, mean (SD)	24.3 ± 7.5	23.1 ± 5.8	0.281
Age groups			
< 21 years	33 (42.9%)	40 (43.0%)	
21–25 years	20 (26.0%)	32 (34.0%)	0.318
26–30 years	8 (10.4%)	12 (12.8%)	
> 30 years	13 (16.9%)	8 (8.5%)	
How much instruction or training have you received previously relative to the above skills? [entry survey]			
I haven't received instructions/training	5%	6%	
I have received some instructions/training	67%	70%	

I have received substantial instructions/training	28%	23%	0.800

Have your critical thinking skills, as described above, improved over the semester as a result of taking this course? [exit survey]

Substantially improved	9%	15%	0.236
Somewhat improved	59%	66%	
No change	23%	13%	
Don't know	9%	6%	

Differences in proportion are assessed by using Pearson Chi-square test and differences in mean by independent sample t-test.

Table 2: Self-assessment of CT skills among Phase II and III participants in <u>ENTRY</u> survey.

Factors	Phase-II (Control) (n = 77) Mean (SD)	Phase-III (Intervention) (n = 94) Mean (SD)	p-value
Given a situation or topic to consider, I know how to identify the main issue(s)	4.6 (1.0)	4.5 (1.0)	0.847
When given a paper to read, I know how to identify the author's views and objectives in the issue presented	4.6 (0.9)	4.5 (1.1)	0.585
I know how to tell if the evidence presented by the author is relevant to the argument	4.5 (1.0)	4.5 (1.0)	0.948
I know how to tell when additional information is needed in order for me to confirm or reject a given claim	4.6 (1.1)	4.4 (1.1)	0.261
I know how to determine the similarities and differences among the various opinions presented for a given problem	4.9 (0.9)	4.7 (1.0)	0.156
I know how to consider opposing views in support of information when controversial issues are examined	4.5 (1.0)	4.6 (1.0)	0.324
I know how to systematically analyze the problem using multiple sources of information to draw conclusions	4.4 (1.1)	4.4 (1.3)	0.960
I know how to tell if the information supporting the argument is reliable	4.4 (1.1)	4.4 (1.2)	0.994
I know how to explain how my personal preferences/bias may affect my analysis	4.7 (1.1)	4.4 (1.2)	0.167
I know how to develop logical conclusions	4.6 (1.1)	4.6 (1.2)	0.924
I know how to arrive at conclusions that are supported with strong evidence	4.6 (1.1)	4.6 (1.2)	0.899
I know how to clearly explain my reasons for accepting or rejecting an author's claim	4.4 (1.1)	4.4 (1.3)	0.958
I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions	4.6 (1.1)	4.4 (1.2)	0.213

Differences in mean by independent sample t-test.

Table 3: Self-assessment of CT skills among Phase II and III participants in EXIT survey.

Factors	Phase-II (Control) (n = 77) Mean (SD)	Phase-III (Intervention) (n = 94) Mean (SD)	p-value
Given a situation or topic to consider, I know how to identify the main issue(s)	4.9 (0.8)	4.7 (0.8)	0.300
When given a paper to read, I know how to identify the author's views and objectives in the issue presented	4.8 (1.0)	4.6 (1.0)	0.188
I know how to tell if the evidence presented by the author is relevant to the argument	4.8 (1.0)	4.8 (1.0)	0.877
I know how to tell when additional information is needed in order for me to confirm or reject a given claim	4.7 (1.0)	4.7 (0.9)	0.996
I know how to determine the similarities and differences among the various opinions presented for a given problem	4.8 (0.8)	4.9 (0.9)	0.836
I know how to consider opposing views in support of information when controversial issues are examined	4.7 (0.9)	4.6 (1.1)	0.814
I know how to systematically analyze the problem using multiple sources of information to draw conclusions	4.7 (0.9)	4.6 (1.1)	0.504
I know how to tell if the information supporting the argument is reliable	4.8 (1.0)	4.7 (1.1)	0.767
I know how to explain how my personal preferences/bias may affect my analysis	4.8 (0.9)	4.8 (1.0)	0.788
I know how to develop logical conclusions	4.7 (1.1)	4.9 (1.0)	0.239
I know how to arrive at conclusions that are supported with strong evidence	4.9 (0.9)	4.8 (1.0)	0.541
I know how to clearly explain my reasons for accepting or rejecting an author's claim	4.7 (1.0)	4.6 (1.0)	0.490
I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions	4.7 (1.0)	4.5 (1.1)	0.170

Differences in mean by independent sample t-test.

Table 3a: Self-assessment of CT skills among Phase II and III participants in the <u>entry</u> survey by age and education.

Factors	Prior education level				Age groups					
	High School or less		Some PSE / PSE		< 21 years		21 – 25 years		> 25 years	
	Phase-II (n = 45)	Phase-III (n = 55)	Phase-II (n = 31)	Phase-III (n = 39)	Phase-II (n = 33)	Phase-III (n = 40)	Phase-II (n = 20)	Phase-III (n = 32)	Phase-II (n = 21)	Phase-III (n = 20)
Given a situation or topic to consider, I know how to identify the main issue(s)	4.5 (0.9)	4.6 (1.0)	4.7 (1.0)	4.5 (0.9)	4.5 (0.9)	4.7 (1.0)	4.6 (1.0)	4.6 (0.8)	4.5 (1.1)	4.2 (1.0)
When given a paper to read, I know how to identify the author's views and objectives in the issue presented	4.4 (0.9)	4.5 (1.1)	4.9 (0.9)	4.5 (1.0)	4.4 (0.9)	4.6 (1.0)	4.8 (1.1)	4.5 (1.1)	4.7 (0.9)	4.3 (1.1)
I know how to tell if the evidence presented by the author is relevant to the argument	4.5 (0.9)	4.5 (1.1)	4.9 (0.9)	4.5 (1.0)	4.5 (0.9)	4.6 (1.0)	4.8 (1.0)	4.6 (1.0)	4.4 (1.0)	4.2 (1.1)
I know how to tell when additional information is needed in order for me to confirm or reject a given claim	4.5 (1.1)	4.3 (1.1)	4.8 (0.9)	4.6 (1)	4.5 (0.9)	4.4 (1.2)	4.7 (1.3)	4.6 (1)	4.7 (1.1)	4.1 (1.1)
I know how to determine the similarities and differences among the various opinions presented for a given problem	4.7 (0.9)	4.6 (1)	5.1 (0.9)	4.7 (0.9)	4.7 (0.8)	4.7 (1)	5 (1.1)	4.7 (0.9)	5 (0.9)	4.7 (1.1)
I know how to consider opposing views in support of information when controversial issues are examined	4.3 (1)	4.6 (1)	4.8 (1)	4.7 (1.1)	**4.1 (0.9)**	**4.7 (1.0)**	4.6 (1.3)	4.8 (0.9)	**4.9 (0.7)**	**4.3 (1.2)**
I know how to systematically analyze the problem using multiple sources of information to draw conclusions	4.2 (1.0)	4.1 (1.3)	4.6 (1.1)	4.7 (1.3)	4.2 (1.1)	4.4 (1.1)	4.6 (1.2)	4.4 (1.5)	4.5 (0.9)	4.2 (1.4)

I know how to tell if the information supporting the argument is reliable	4.2 (1.1)	4.3 (1.2)	4.6 (1.1)	4.5 (1.1)	4.3 (1.0)	4.5 (1.1)	4.6 (1.1)	4.5 (1.3)	4.3 (1.2)	4 (1.3)
I know how to explain how my personal preferences/bias may affect my analysis	4.5 (1.2)	4.2 (1.2)	4.9 (0.9)	4.7 (1.1)	4.5 (1.1)	4.3 (1.2)	4.7 (1.1)	4.8 (1.0)	**4.9 (1.0)**	**4.1 (1.3)**
I know how to develop logical conclusions	4.5 (1.1)	4.5 (1.3)	4.9 (1.0)	4.8 (1.1)	4.2 (1.1)	4.7 (1.1)	5.1 (1.0)	4.7 (1.4)	4.9 (0.8)	4.6 (1.3)
I know how to arrive at conclusions that are supported with strong evidence	4.4 (1.1)	4.4 (1.2)	4.8 (1.0)	4.8 (1.2)	4.3 (1.1)	4.5 (1.1)	4.8 (1.1)	4.9 (1.1)	4.8 (1.0)	4.2 (1.4)
I know how to clearly explain my reasons for accepting or rejecting an author's claim	4.3 (1.1)	4.2 (1.3)	4.6 (1.0)	4.7 (1.1)	3.9 (1.0)	4.4 (1.2)	4.8 (1.0)	4.5 (1.4)	4.7 (1.0)	4.3 (1.1)
I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions	4.4 (1.2)	4.2 (1.3)	4.8 (0.9)	4.6 (1.0)	4.5 (1.1)	4.3 (1.2)	4.9 (0.9)	4.6 (1.2)	4.5 (1.2)	4.2 (1.1)

Differences in mean by independent sample t-test. Text in bold indicate that the differences observed are statistically significant.

Table 3b: Self-assessment of CT skills among Phase-II and III in <u>exit</u> survey by age, and education level.

Factors	Prior education level				Age groups					
	High School or less		Some PSE / PSE		< 21 years		21 – 25 years		> 25 years	
	Phase-II (n = 45)	Phase-III (n = 55)	Phase-II (n = 31)	Phase-III (n = 39)	Phase-II (n = 33)	Phase-III (n = 40)	Phase-II (n = 20)	Phase-III (n = 32)	Phase-II (n = 21)	Phase-III (n = 20)
Given a situation or topic to consider, I know how to identify the main issue(s)	4.7 (0.8)	4.7 (0.7)	5 (0.8)	4.8 (1.0)	4.7 (0.8)	4.7 (0.7)	5 (0.9)	4.7 (0.9)	5 (0.7)	4.8 (1.1)
When given a paper to read, I know how to identify the author's views and objectives in the issue presented	4.6 (1.0)	4.5 (1.1)	**5 (0.8)**	**4.6 (0.9)**	4.5 (0.9)	4.6 (1.1)	**5.1 (1.0)**	**4.4 (0.9)**	5 (0.7)	4.6 (0.9)
I know how to tell if the evidence presented by the author is relevant to the argument	4.7 (1.0)	4.7 (1.1)	5 (0.9)	4.8 (0.9)	4.7 (0.9)	4.7 (1.1)	4.9 (1.1)	4.8 (0.9)	4.8 (1.0)	4.8 (1.1)
I know how to tell when additional information is needed in order for me to confirm or reject a given claim	4.7 (1.0)	4.7 (1.0)	4.7 (1.0)	4.7 (0.9)	4.6 (1.0)	4.7 (1.0)	4.9 (0.9)	4.7 (0.8)	4.7 (1.0)	4.6 (1.0)
I know how to determine the similarities and differences among the various opinions presented for a given problem	4.7 (0.8)	4.8 (1.0)	5 (0.8)	4.9 (0.8)	4.7 (0.8)	4.9 (1.0)	5.1 (0.8)	4.9 (0.8)	4.8 (0.7)	4.8 (1.1)
I know how to consider opposing views in support of information when controversial issues are examined	4.6 (0.8)	4.7 (1.0)	4.8 (0.9)	4.6 (1.1)	4.5 (0.9)	4.7 (1.1)	4.9 (0.7)	4.7 (0.8)	4.9 (0.8)	4.4 (1.4)

I know how to systematically analyze the problem using multiple sources of information to draw conclusions	4.7 (0.9)	4.5 (1.2)	4.7 (1.0)	4.7 (1.0)	4.6 (1.0)	4.5 (1.3)	4.9 (1.0)	4.6 (0.9)	4.8 (0.7)	4.7 (1.0)
I know how to tell if the information supporting the argument is reliable	4.6 (1.0)	4.6 (1.2)	4.9 (0.9)	4.8 (0.9)	4.5 (1.1)	4.7 (1.2)	5.1 (1.0)	4.6 (1.2)	4.8 (0.5)	4.9 (0.9)
I know how to explain how my personal preferences/bias may affect my analysis	4.7 (0.8)	4.7 (1.1)	4.9 (1.0)	4.9 (0.7)	4.7 (0.9)	4.7 (1.0)	5 (1.0)	4.8 (1.1)	4.8 (0.8)	4.8 (0.6)
I know how to develop logical conclusions	4.6 (1.1)	4.8 (1.1)	4.9 (1.0)	5.1 (0.8)	4.5 (1.1)	4.9 (1.1)	5.1 (0.9)	4.8 (1.1)	4.9 (1.0)	5 (0.9)
I know how to arrive at conclusions that are supported with strong evidence	4.8 (0.9)	4.7 (1.1)	5.1 (0.9)	4.9 (0.8)	4.6 (1.0)	4.8 (1.1)	5.1 (0.9)	4.6 (0.9)	5.1 (0.8)	4.9 (1.0)
I know how to clearly explain my reasons for accepting or rejecting an author's claim	4.6 (1.0)	4.5 (1.0)	4.9 (0.9)	4.7 (0.9)	4.5 (1.1)	4.6 (1.1)	5 (0.8)	4.5 (0.9)	4.9 (0.8)	4.6 (0.8)
I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions	4.6 (1.0)	4.4 (1.2)	4.9 (1.0)	4.6 (0.9)	4.4 (1.0)	4.5 (1.2)	**5.1 (1.0)**	**4.5 (0.9)**	4.9 (0.8)	4.5 (1.0)

Differences in mean by independent sample t-test. Text in bold indicate that the differences observed are statistically significant.

Table 4: Self-assessment of CT skills among entry and exit survey in Phase II (control)

Factors	Entry survey (n = 77) Mean (SD)	Exit survey (n = 77) Mean (SD)	p-value
Given a situation or topic to consider, I know how to identify the main issue(s)	**4.6 (1.0)**	**4.9 (0.8)**	**0.015**
When given a paper to read, I know how to identify the author's views and objectives in the issue presented	4.6 (0.9)	4.8 (1.0)	0.113
I know how to tell if the evidence presented by the author is relevant to the argument	**4.5 (1.0)**	**4.8 (1.0)**	**0.009**
I know how to tell when additional information is needed in order for me to confirm or reject a given claim	4.6 (1.1)	4.7 (1.0)	0.477
I know how to determine the similarities and differences among the various opinions presented for a given problem	4.9 (0.9)	4.8 (0.8)	0.813
I know how to consider opposing views in support of information when controversial issues are examined	4.5 (1.0)	4.7 (0.8)	0.054
I know how to systematically analyze the problem using multiple sources of information to draw conclusions	**4.4 (1.1)**	**4.7 (0.9)**	**0.006**
I know how to tell if the information supporting the argument is reliable	**4.4 (1.1)**	**4.7 (1.0)**	**0.010**
I know how to explain how my personal preferences/bias may affect my analysis	4.7 (1.1)	4.8 (0.9)	0.257
I know how to develop logical conclusions	4.7 (1.1)	4.7 (1.1)	0.626
I know how to arrive at conclusions that are supported with strong evidence	**4.6 (1.1)**	**4.9 (0.9)**	**0.013**
I know how to clearly explain my reasons for accepting or rejecting an author's claim	**4.4 (1.1)**	**4.7 (0.9)**	**0.027**
I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions	4.6 (1.1)	4.7 (1.0)	0.359

Differences in mean by independent sample t-test. Text in bold indicate that the differences observed are statistically significant

Table 5: Self-assessment of CT skills among entry and exit survey in Phase-III (intervention)

Factors	Pre-course survey (n = 94) Mean (SD)	Post-course survey (n = 94) Mean (SD)	p-value
Given a situation or topic to consider, I know how to identify the main issue(s)	4.5 (1.0)	4.7 (0.8)	0.065
When given a paper to read, I know how to identify the author's views and objectives in the issue presented	4.5 (1.1)	4.6 (1.0)	0.603
I know how to tell if the evidence presented by the author is relevant to the argument	4.5 (1.0)	4.8 (1.0)	0.052
I know how to tell when additional information is needed in order for me to confirm or reject a given claim	**4.4 (1.1)**	**4.7 (0.9)**	**0.016**
I know how to determine the similarities and differences among the various opinions presented for a given problem	**4.7 (1.0)**	**4.9 (0.9)**	**0.034**
I know how to consider opposing views in support of information when controversial issues are examined	4.6 (1.0)	4.6 (1.1)	0.818
I know how to systematically analyze the problem using multiple sources of information to draw conclusions	4.4 (1.3)	4.6 (1.1)	0.066
I know how to tell if the information supporting the argument is reliable	**4.4 (1.2)**	**4.7 (1.1)**	**0.007**
I know how to explain how my personal preferences/bias may affect my analysis	**4.4 (1.2)**	**4.8 (1.0)**	**0.002**
I know how to develop logical conclusions	**4.6 (1.2)**	**4.9 (1.0)**	**0.015**
I know how to arrive at conclusions that are supported with strong evidence	4.6 (1.2)	4.8 (1.0)	0.087
I know how to clearly explain my reasons for accepting or rejecting an author's claim	4.4 (1.3)	4.6 (1.0)	0.092
I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions	4.4 (1.2)	4.5 (1.1)	0.262

Differences in mean by independent sample t-test. Text in bold indicate that the differences observed are statistically significant.

STUDENTS WHO COMPLETED <u>ONE OR BOTH</u> SURVEYS

Table 6: Overall Self-assessment of CT skills and demographic characteristics among entry and exit survey in Phase II (control)

Factors	Entry survey (n = 267) Mean (SD)	Exit survey (n = 165) Mean (SD)
Age in years	24.8 (8.2)	
Age groups		
< 21 years	103 (39.5%)	
21–25 years	81 (31.0%)	
26–30 years	34 (13.0%)	
> 30 years	43 (16.5%)	
Education Level		
High School or Less	139 (39.0%)	
Some PSE	48 (13.5%)	
PSE	76 (21.3%)	
Prior training / instruction in CT		
Substantial / Somewhat training	233 (89.6%)	
No training	27 (10.4%)	
Improvement in CT		
Substantial / Somewhat improvement	111 (69.4%)	
No Change	25 (15.6%)	
Don't know	24 (15.0%)	
Given a situation or topic to consider, I know how to identify the main issue(s)	4.6 (1.0)	4.7 (0.9)
When given a paper to read, I know how to identify the author's views and objectives in the issue presented	4.5 (1.1)	4.7 (1.0)
I know how to tell if the evidence presented by the author is relevant to the argument	4.4 (1.1)	4.7 (1.0)
I know how to tell when additional information is needed in order for me to confirm or reject a given claim	4.5 (1.1)	4.6 (1.0)
I know how to determine the similarities and differences among the various opinions presented for a given problem	4.7 (1.0)	4.8 (0.9)
I know how to consider opposing views in support of information when controversial issues are examined	4.4 (1.0)	4.6 (1.1)
I know how to systematically analyze the problem using multiple sources of information to draw conclusions	4.4 (1.1)	4.6 (1.1)
I know how to tell if the information supporting the argument is reliable	4.4 (1.2)	4.6 (1.0)
I know how to explain how my personal preferences/bias may affect my analysis	4.6 (1.1)	4.7 (1.0)

I know how to develop logical conclusions	4.6 (1.1)	4.6 (1.1)
I know how to arrive at conclusions that are supported with strong evidence	4.5 (1.1)	4.8 (1.0)
I know how to clearly explain my reasons for accepting or rejecting an author's claim	4.4 (1.1)	4.5 (1.0)
I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions	4.3 (1.2)	4.5 (1.1)

Table 7: Overall Self-assessment of CT skills and demographic characteristics among entry and exit survey in Phase III (intervention)

Factors	Pre-course survey (n = 320) Mean (SD)	Post-course survey (n = 164) Mean (SD)
Age in years	24.8 (8.2)	
Age groups		
< 21 years	136 (42.9)	
21–25 years	118 (37.2%)	
26–30 years	42 (13.2%)	
> 30 years	21 (6.6%)	
Education Level		
High School or Less	186 (58.3%)	
Some PSE	71 (22.3%)	
PSE	62 (19.4%)	
Prior training / instruction in CT		
Substantial / Somewhat training	283 (89.6%)	
No training	33 (10.4%)	
Improvement in CT		
Substantial / Somewhat improvement	130 (79.3%)	
No Change	20 (12.2%)	
Don't know	14 (8.5%)	
Given a situation or topic to consider, I know how to identify the main issue(s)	4.5 (1.0)	4.8 (0.9)
When given a paper to read, I know how to identify the author's views and objectives in the issue presented	4.4 (1.1)	4.6 (1.0)
I know how to tell if the evidence presented by the author is relevant to the argument	4.4 (1.1)	4.7 (1.1)
I know how to tell when additional information is needed in order for me to confirm or reject a given claim	4.3 (1.2)	4.7 (1.0)
I know how to determine the similarities and differences among the various opinions presented for a given problem	4.6 (1.0)	4.9 (1.0)

I know how to consider opposing views in support of information when controversial issues are examined	4.4 (1.1)	4.6 (1.1)
I know how to systematically analyze the problem using multiple sources of information to draw conclusions	4.3 (1.2)	4.7 (1.1)
I know how to tell if the information supporting the argument is reliable	4.2 (1.2)	4.7 (1.1)
I know how to explain how my personal preferences/bias may affect my analysis	4.5 (1.1)	4.7 (1.0)
I know how to develop logical conclusions	4.6 (1.1)	4.9 (1.1)
I know how to arrive at conclusions that are supported with strong evidence	4.5 (1.1)	4.8 (1.1)
I know how to clearly explain my reasons for accepting or rejecting an author's claim	4.3 (1.1)	4.6 (1.0)
I know how to organize my paper and to select certain vocabulary that will enhance my message and strengthen my position and conclusions	4.3 (1.2)	4.6 (1.1)