Evidence-Centered Concept Map as a Thinking Tool in Critical Thinking Computer-based Assessment

Research Report

Yigal Rosen, Ph.D.
Maryam Tager

April 2013
About Pearson
Pearson, the global leader in education and education technology, provides innovative print and digital education materials for pre-K through college; student information systems and learning management systems; teacher licensure testing; teacher professional development; career certification programs; and testing and assessment products that set the standard for the industry. Pearson’s other primary businesses include the Financial Times Group and the Penguin Group. For more information about the Assessment & Information group of Pearson, visit http://www.pearsonassessments.com/.

About Pearson’s Research Reports
Pearson’s research report series provides preliminary dissemination of reports and articles prepared by Test, Measurement, and Research Services (TMRS) staff, usually prior to formal publication. Pearson’s publications in pdf format may be obtained at http://www.pearsonassessments.com/research.
Abstract

Major educational initiatives in the world place great emphasis on fostering rich computer-based environments of assessment that make student thinking and reasoning visible. Using thinking tools engages students in a variety of critical and complex thinking, such as evaluating, analyzing, and decision making. The aim of this study was to explore patterns in student critical thinking performance and motivation in Evidence-Centered Concept Map (ECCM) settings, compared to basic notepad settings. One hundred ninety 14-year-old students from the United States, United Kingdom, Singapore, and South Africa participated in the study. Students in both modes were able to analyze a multifaceted dilemma by using similar information resources. In the ECCM mode, students used ECCM to organize their thinking; in another mode, students were provided with a basic online notepad to make records as needed. Overall, the findings showed that students assessed in ECCM mode outperformed their peers in notepad mode in critical thinking skills. Student who worked with ECCM provided more informed recommendations by using supporting evidence from the available resources and discussing alternative points of view on the topic. The findings showed a significantly positive relationship between student recommendation score and the ability to develop ECCM. In addition, the results demonstrated that it did not matter for students’ motivation whether they analyzed the dilemma with or without ECCM. Directions for future research are discussed in terms of their implications for large-scale assessment programs, teaching, and learning.

Keywords: critical thinking, concept map, computer-based assessment
Evidence-Centered Concept Map in Critical Thinking Assessment

Concept Map as a Thinking Tool in Critical Thinking Computer-based Assessment:

Making Student Thinking Visible

Introduction

Computer-based environments are becoming more central in the classroom and have been used as intellectual partners for active participation in construction of knowledge (Dede, 2009; Jonassen, 2008; Jonassen & Reeves, 1996; Lajoie, 2000; Salomon & Perkins, 2005). However, in many cases, the technology is implemented for traditional practices, while paradigmatic change in computer-based educational assessment is rare. Some assessment designers and educators, in their enthusiasm for implementing cutting-edge advanced technology, take a technology-centered approach without sensitivity to how people learn. In contrast, other assessment designers and educators take a learner-centered approach, in which they begin with an understanding of learning processes and attempt to infuse technology as an aid to student learning and assessment (Mayer, 2001; Rosen & Beck-Hill, 2012). Qualitatively different learning environments offer different kinds of assessment experiences and thus serve different educational goals. Research shows that computer-based constructivist learning environments can more effectively promote higher-order thinking skills, learning motivation, and teamwork, than can traditional settings (Rosen & Salomon, 2007). Just as technology and learning sciences play an essential role in helping to develop more effective learning practices, they also can provide key improvements in assessment.
Evidence-Centered Concept Map in Critical Thinking Assessment

(Bennett, 1999; Bennett et al., 2007; Pellegrino, Chudowsky, & Glaser, 2001; Tucker, 2009). Measuring complex skills such as critical thinking, creativity, and collaborative problem solving requires designing and developing assessments that address the multiple facets implied by these skills. One of the possible ways to achieve these changes in educational assessment is by providing visible sequences of actions that students have taken by using various tools within the contexts of relevant societal issues and problems that people care about in everyday life. Thinking tools are computer applications that enable students to represent what they learned and know using different representational formalisms. Studying the role of thinking tools in computer-based assessment of higher-order thinking skills is crucial to determining whether these types of scaffolding tools can bring a real added-value into large-scale computer-based assessment programs. The purpose of this study was to provide empirical evidence of what can be achieved in terms of possible differences in student achievement and motivation by intertwining a thinking tool in a performance assessment of student critical thinking. This paper addresses these challenges by introducing a new methodology for scalable use of thinking tools in computer-based assessment of higher-order thinking skills, providing findings from an empirical pilot study conducted in four countries, as well as discussing implications of the findings for further research and development.
Defining Critical Thinking

Ennis (1993) defines critical thinking as “reasonable reflective thinking focused on deciding what to believe or do.” (p. 180). Critical thinking requires the component interdependent competencies of evaluating the credibility of sources, analyzing the quality of arguments, making inferences using reasoning, and making decisions or solving problems (see Lai & Viering, 2012, for a literature review). Critical thinking often appears in Program for International Student Assessment (PISA) and US National Assessment of Educational Progress (NAEP) in assessment of science, math and reading. Critical thinking was part of problem solving assessment in PISA 2012, with major emphasis on evaluation of the available information, assumptions, and possible solutions, as well as looking for additional information or clarification (OECD, 2010). These PISA-related critical thinking competencies are captured well by Kahneman’s (2011) acronym WYSIATI — What You See Is All There Is—. People make decisions on very limited evidence and often assert their proposition with great confidence. They tend to believe their current information set is the complete information set. Another consequence of WYSIATI is that people evaluate complex topics without considering a wider range of alternatives or trying to get more information that could potentially contradict the decision. The individual may understand in principle that worthless information should not be treated differently from a complete lack of information, but WYSIATI limits the ability to apply this principle. Thus, a simpler spectrum of evidence makes it easier to create a coherent interpretation of a topic, while the amount of evidence and its quality do not count for much.
Critical thinking and mindfulness are used interchangeably in the education research literature to mean similar constructs. A three-fold model of mindfulness that was proposed by Langer (1989, 1997) entails the continuous creation of new categories; openness to new information; and an implicit awareness of more than one perspective. She contrasted mindfulness with mindlessness, which is characterized as entrapment in old categories; automatic behavior that precludes attending to new signals; and action that operates from a single perspective.

According to Partnership for 21st Century Skills (2009) and Assessment and Teaching of 21st Century Skills (Binkley et al., 2012), a comprehensive set of critical thinking competencies include: use various types of reasoning as appropriate to the situation; analyze how parts of a whole interact with each other to produce overall outcomes in complex systems; examine ideas; identify and analyze arguments; synthesize and make connections between information and arguments; interpret information and draw conclusions based on the analysis; categorize, decode, and clarify information; effectively analyze and evaluate major alternative points of view; and ask meaningful questions that clarify various points of view.

There is a notable lack of consensus on whether critical thinking is generalizable or entirely domain-specific. We adopt the intermediate approach (Ennis, 1985; Smith, 2002), according to which some critical thinking skills apply to multiple domains (e.g., analyzing the logics of a claim or evaluating the reliability of evidence source), whereas others are unique to specific subject areas (e.g., experimental inquiry in science or reasoning in math). In our research, an operational definition of critical thinking refers to the capacity of an individual to effectively engage in a process of making decisions or
solving problems by analyzing and evaluating evidence, arguments, claims, beliefs, and alternative points of view; synthesizing and making connections between information and arguments; interpreting information; and making inferences using reasoning appropriate to the situation. In identifying critical thinking skills, this research attempts to incorporate skills identified in other assessment frameworks, such as the Partnership for 21st Century Skills (2009) and Assessment and Teaching of 21st Century Skills (Binkley et al., 2012).

Assessing Critical Thinking Skills

Critical thinking can be very difficult to measure in a valid and reliable manner. First, this is because of the various conceptualizations of critical thinking as domain-general as opposed to domain-specific, as well as because of the differences in the definitions of the construct (Halpern, 1998; Kuncel & Hezlett, 2010; Moss & Koziol, 1991; Norris, 1989). A narrower definition in which critical thinking is considered a finite set of specific competencies could provide a better platform for measuring critical thinking. These competencies could be useful for effective decision making for many (but not all) contexts, while their efficacy is further curtailed by students’ specific knowledge demands in the specific context. Second, it is difficult to assess critical thinking because it is an ongoing process rather than a recognizable outcome. The conventional assessment formats limit students’ ability to optimally apply their critical thinking, and restricts educators’ ability to follow students’ thinking process (Bonk & Smith, 1998; Fischer, Spiker, & Riedel 2009).

Critical thinking assessment tasks should provide adequate collateral materials to support multiple perspectives and include process as well as product indicators. Problems underlie such tasks should use ill-defined structure that often involve multiple goals that are in conflict, have more than one defensible solution and require students to go beyond recalling or restating learned information (Mayer, & Wittrock, 2006; Moss, & Koziol, 1991). Critical thinking assessment tasks should make student reasoning visible by requiring students to provide evidence or logical arguments in support of judgments, choices, claims, or assertions (Fischer, Spiker, & Riedel 2009; Norris, 1989). Embedding computer-based thinking tools in critical thinking performance assessment, which makes student thinking visible, is one of the promising approaches that should be further explored.

**Concept Map as a Thinking Tool in Critical Thinking Assessment**

According to the National Education Technology Plan (U.S. Department of Education, 2010), the educational assessments “are not using the full flexibility and power of technology to design, develop, and validate new assessment materials and processes for both formative and summative uses” (p. 25). Computer technologies such as interactive thinking tools that aid cognitive processing can support intellectual performance and enrich individuals’ assessment experience. Thinking tools (or mindtools) are computer applications that enable students to represent what they learned and know using different representational formalisms. There are several classes of thinking tools, including semantic organization tools, dynamic modeling tools, information interpretation tools, knowledge construction tools, microwords, and
conversation and collaboration tools (Jonassen, 2006; Jonassen, & Reeves, 1996). Assessment thinking tools represent thinking processes in which the student is engaged, such as evaluating, analyzing, connecting, elaborating, synthesizing, designing, problem solving, and decision making. Using Perkins’s (1993) terminology, the unit of analysis in these assessments is not the student without the technology in his or her environment — — the person-solo — but the person-plus the technology, in this case the student plus the thinking tool.

Concept maps have been widely used as thinking tools for teaching, learning, and assessment as a way to help the student think and represent his or her thinking processes (Jonassen, 1996; Kinchin et al., 2000; Novak, & Cañas, 2008; Ruiz-Primo, 2004). A concept map is a semi-formal knowledge representation tool visualized by a graph consisting of finite set of nodes, which depict concepts, and finite set of arcs, which express relationships between pairs of concepts (Novak, 1998; Novak, & Cañas, 2008). A linking phrase can specify the kind of relationship between concepts. As a rule, natural language is used to represent concepts and linking phrases. Moreover, the arcs of the graph can have the same weight or different weights. The concept maps comprise concepts and their relationships, often arranged hierarchically according to the importance of the concepts described, with the most general concepts at the top of the map and the more specific concepts below them, but cross-links can be used to indicate relationships between the concepts. Several studies have shown that concept maps are a valid and reliable medium to represent students’ understanding (Hoeft et al., 2003; McClure, Sonak, & Suen, 1999), making them a valuable pedagogical tool. Concept maps are used in educational settings in various ways. In the first, students build a
concept map on a given topic. The student must decide on the structure and the content of a concept map. Another use of a concept map is “fill-in-the-map” tasks, where the structure of a concept map is given, and a student must fill it using the provided set of concepts and/or linking phrases. In the last, students analyze a concept map previously built by an expert.

Concept mapping is a cognitively challenging task that requires various higher-order thinking processes, such as assessing and classifying information, recognizing patterns, identifying and prioritizing main ideas, comparing and contrasting, identifying relationships, and logical thinking (Jonassen, 1996; Kinchin et al., 2000). These processes require the student to elaborate and organize information in meaningful ways, which cannot be realized through simply memorizing facts without understanding their meaning and underlying associations. The thinking processes involved in concept mapping are highly related to critical thinking competency as defined by various assessment frameworks (Binkley et al., 2012; OECD, 2010; Partnership for 21st Century Skills, 2009).

While concept mapping can be used to enhance students’ thinking, it also has some constraints. Students in several studies reported that building concept maps is difficult and time-consuming especially when they first experience it (All et al., 2003; Hsu, 2004). Thus, it is necessary to understand the nature of the task and the thinking skills required while designing a concept map. In our research we use a three-phase concept map to empower the student to analyze various claims and evidence on a topic and to draw a conclusion, or Evidence-Centered Concept Map (ECCM) in short. The stages of student work with ECCM on a critical assessment task include: (a) gathering
various claims and evidence from the resources provided (some claims and evidence contradict one another); (b) organizing the claims with supporting evidence gathered in the previous phase on ECCM without hierarchical relationships; and (c) linking claims and specifying the kind of a relationship between claims. It should be noted that no hierarchical order is required in ECCM. The three-phase working structure of ECCM was designed to increase the cognitive and measurement interdependency between the three distinctive competencies in critical thinking as they are identified in our research: (a) analyzing and evaluating evidence, arguments, claims, beliefs, and alternative points of view; (b) synthesizing evidence, arguments, claims, beliefs, and alternative points of view; and (c) making connections between information and arguments. By using ECCM in a critical thinking assessment, we provide scaffolding for the student thinking process by enabling the construction of a well-integrated structural representation of the topic, as opposed to the memorization of fragmentary information, and we externalize the student’s conceptual understanding of the topic.

In summary, technology offers opportunities for assessment in domains and contexts where assessment would otherwise not be possible or would not be scalable. One of the important possibilities that technology brings to the educational assessment of critical thinking is the capacity to embed thinking tools, empowering the student’s thinking and making his or her thinking process visible. These tools could be designed in such a way that students will be fully engaged in a task without introducing difficulties irrelevant to the construct or forcing the allocation of additional time. Thus, it is
necessary to understand the nature of the task and the thinking skills required while
designing an appropriate thinking tool.

This paper addresses these challenges by studying student performance in ECCM
and notepad modes of critical thinking computer-based assessment.

**Research Questions**

The study addressed empirically the following questions regarding student performance
and motivation in critical thinking assessment in ECCM and notepad settings:

1. What are the differences in student critical thinking performance between ECCM
   and notepad modes of assessment as reflected in the student recommendation?

2. How are a student’s abilities to develop ECCM, and create a linkage within ECCM, 
   related to student performance in critical thinking assessment, as reflected in the 
   student recommendation?

3. What are the differences in critical thinking performance within ECCM and notepad 
   modes of assessment between female and male students as reflected in the student 
   recommendation?

4. How are a student’s GPA, ELA, and Math achievement, as reported by the students, 
   related to the student recommendation in ECCM and notepad modes of assessment?

5. What are the differences in student motivation while working on a critical thinking 
   assessment task with and without ECCM?

6. What are the differences in the student recommendation between ECCM and 
   notepad modes of assessment as reflected in time-on-task?
Method

The study participants included 190 students, all 14 years old, from the United States, United Kingdom, Singapore, and South Africa. The results presented in the current article came from a larger study in which students from six countries were recruited to participate in a 21st Century Skills Assessment project study investigating innovative ways of developing computer-based assessment in critical thinking, creativity, and collaborative problem solving (see Rosen, & Tager, 2013, for study of collaborative problem solving). The researchers collected data from November 2012 to January 2013. Recruitment of participating schools was achieved through collaboration with local educational organizations based on the following criteria: (a) the school is actively involved in various 21st Century Skills projects, (b) population of 14-year-old students proficient in English, and (c) sufficient technology infrastructure (e.g., computers per student, high-speed Internet).

Table 1

Research population by mode of assessment and country

<table>
<thead>
<tr>
<th>Group</th>
<th>ECCM Female</th>
<th>ECCM Male</th>
<th>Notepad Female</th>
<th>Notepad Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>12</td>
<td>16</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Singapore</td>
<td>17</td>
<td>21</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>South Africa</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Overall</td>
<td>38</td>
<td>64</td>
<td>40</td>
<td>48</td>
</tr>
</tbody>
</table>
In all, 102 students participated in ECCM mode, and 88 participated in notepad mode. Of the total students who participated, 112 were boys (58.9%) and 78 were girls (41.1%). Table 1 summarizes the country and gender distribution of participating students between the ECCM and notepad groups.

No significant differences were found in GPA, ELA, and Math average scores between participants in ECCM and notepad modes within the countries. This similarity in student background allowed comparability of student results in critical thinking assessment tasks between the two modes.

Critical Thinking Assessment

In this critical thinking computer-based assessment task, the student was asked to analyze various pros and cons of whether or not to buy organic milk for the school cafeteria and write a recommendation to a school principal. Students who participated in ECCM mode were required to use a concept map during the analysis of web-based pre-determined resources, while students who participated in notepad mode were able to take notes by using an embedded free text notepad, but were not provided any kind of thinking tool. Among the websites that were accessible to the students were: organic milk company website along with an interview script/video with the CEO of the organic milk company, independent organic milk association, dairy farmers of North America, anti-organic milk along with an interview script/video with the blogger (a past worker of an organic milk company), Disease Control Center, and a news website. The resources included various content orientations (pros and cons related to the organic milk issue), relevancy, and level of reliability. Due to the exploratory nature of the study, the students
were not limited in time-on-task. The task was checked by teachers from the four participating countries to ensure that students would be able to work on the task, and that the task could differentiate between high and low levels of critical thinking ability. Interviews were conducted with students representing the target population to validate the ECCM approach.

The following information was presented interactively to the student during the task in ECCM mode:

Episode #1: The task starts by asking the student to provide name and background information and to select the preferred avatar. Then the tool panel and the story panel were introduced to the student as presented in Figures 1-2. A similar introduction was presented to the students in the second mode of the task, while the notepad replaced the concept map.
Evidence-Centered Concept Map in Critical Thinking Assessment

Figure 1. Introducing the tool panel to the student

![Tool Panel Image]

Figure 2. Introducing the story panel to the student

Episode #2: In both modes, the task was initiated by receiving a request from the principal to conduct research on whether or not to buy organic milk for the school cafeteria (Figure 3).
Episode #3: Figures 4-7 show examples of the task screens during the resources exploration episode. The major area of the screen allows the student to view the available web-based resources. On the right side of the screen, the students were able to take notes by using drag-and-drop functionality or typing the text. In ECCM mode the student was able to classify the notes into claims and evidence in preparation for constructing the concept map. Similar resources were accessible to the students in both modes.
**Evidence-Centered Concept Map in Critical Thinking Assessment**

**Figure 4.** The list of resources available to the student

- **World’s Best Organic from Greenfield Organic Farms Milk**
  - [www.greenfieldorganicfarms.com](http://www.greenfieldorganicfarms.com)
  - Produced by family farmers in harmony with nature without antibiotics, synthetic hormones or pesticides. All Greenfield Farms dairy farms are pasture-based, meaning a major portion of the cows’ diet comes from certified organic pasture.

- **Organic Milk Association**
  - [www.organicmilkassociation.com](http://www.organicmilkassociation.com)
  - What’s the difference between organic milk and regular milk? About $3 per gallon in most stores -- but smart consumers are realizing that for health reasons alone, it’s worth paying extra for organic milk.

- **Organic Milk FAQs**
  - In terms of quality, safety and nutrition, there’s no difference between organic and regular milk. Both contain the same combination of nutrients that make dairy foods an important part of a healthy diet.

- **Anti Organic: Why Organic Isn’t Better**
  - [antiorganicblog.tumblr.com](http://antiorganicblog.tumblr.com)
  - Lost my job today, theoretically because the company is restructuring and no longer needed me. No one else was asked to leave, though. I think they may have seen my blog, and they must not have liked it.

- **The Real Truth About Organic Food.com**
  - There’s a lot of hype these days about going “organic”, especially around milk. In the grocery store, organic milk costs almost twice as much as non organic milk, so is it really worth the extra money in these tight economic times?

- **Disease Control Center**
  - [www.DiseaseControlCenter.com](http://www.DiseaseControlCenter.com)
  - Raw milk can carry harmful germs that can make you very sick or kill you. If you’re thinking about drinking raw milk because you believe it has health benefits, consider other options.

- **News From Bloomberg**
  - Cartons of Dean Foods Co. (DF)’s Horizon organic milk fortified with DHA feature a picture of a young girl to
Figure 5. Classifying notes into claims and evidence in preparation for ECCM construction
Figure 6. The student is exposed to resources that provide different points of view on the topic and vary in the reliability level and relevancy.
Figure 7. The student is exposed to resources that provide different points of view on the topic and vary in the reliability level and relevancy (cont.)

Once the student decided that the resources review was complete, he or she clicked on Next and then proceeded to the concept map construction episode (in ECCM mode) or proceeded directly to the recommendation-writing episode (in notepad mode).

Episode #4 (ECCM mode only): In this stage the student is required to construct a concept map by positioning previously gathered claims and supporting evidence
Evidence-Centered Concept Map in Critical Thinking Assessment

according to whether the claim/evidence is *for* or *against* organic milk. No relationship creation between claims is required at this stage. Figures 8-9 show examples of the task screens during this episode.

*Figure 8.* The student constructs the initial concept map without the relationships between claims
Figure 9. The student constructs the initial concept map without the relationships between claims (cont.)

Episode #5 (ECCM mode only): The student is asked to create relationships between claims. Those are created by dragging from the link icon on a claim to a second related claim and typing a short description of how they are related. Figure 10 shows an example of a screen for relationships created in a task.
Figure 10. The student-created relationships within the concept map

Episode #6: The student is asked to write a recommendation to a school principal based on the research conducted. Students in ECCM mode were able to view their concept map while typing the recommendation, while students in the notepad mode were able to see the notes previously taken.

Critical Thinking Scoring Rubrics

Following an operational definition of critical thinking, the critical thinking score was given based on the recommendation the student provided on whether to buy organic milk for the cafeteria. The written recommendation represented the capacity of an
individual to effectively engage in a process of making decisions by analyzing and evaluating evidence, arguments, claims, beliefs, and alternative points of view; synthesizing and making connections between information and arguments; interpreting information; and making inferences by using reasoning appropriate to the situation. Table 2 shows the rubric that was used to score students’ written recommendation. For the purposes of more meaningful interpretation of student scores for the teachers, the 0-3 scale was later converted into 0-100% scale.

Table 2

Scoring rubric for student written recommendation

<table>
<thead>
<tr>
<th>Point Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Student provides a recommendation and explains the decision, using supporting text from source material. The recommendation refers to at least 3 of the following dimensions of the topic: health value, animal care, cost, environmental impacts. The student discusses alternative points of view on the topic.</td>
</tr>
<tr>
<td>2</td>
<td>Student provides a recommendation and explains the decision, but may use limited supporting text. The recommendation refers to at least 3 of the following dimensions of the topic: health value, animal care, cost, environmental impacts, but doesn’t discuss alternative points of view on the topic. OR the student discusses alternative points of view, but refers to less than 3 of the dimensions.</td>
</tr>
<tr>
<td>1</td>
<td>Student provides a recommendation, but does not explain the decision, OR student explains solution but does not provide a recommendation. The recommendation refers to one of the following dimensions of the topic: health value, animal care, cost, environmental impacts.</td>
</tr>
<tr>
<td>0</td>
<td>Student does not respond, or fails to address the task</td>
</tr>
</tbody>
</table>

In addition to the scoring of student recommendations, the student-constructed concept map and the relationships within the concept map were scored in the ECCM
mode. Data on these dimensions were collected in order to enable empirical examination of a research question regarding possible correlation between the student’s abilities to develop ECCM and create a linkage within ECCM, and the student’s ability to write a recommendation. Tables 3-4 show the rubrics that were used to score student concept maps and the relationships created within the concept map.

Table 3

*Scoring rubric for the concept map constructed by a student*

<table>
<thead>
<tr>
<th>Point Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Student concept map captures key points of both sides of the argument, and pulls out supporting text. At least 5 valid claims supported by valid evidence were constructed. The claims are positioned correctly according to the ‘strongly for’ and ‘strongly against’ scale.</td>
</tr>
<tr>
<td>2</td>
<td>Student response captures a few key points of both sides, but supporting text may be limited or somewhat weak. Three or four valid claims supported by valid evidence were constructed. The claims are positioned correctly according to the ‘strongly for’ and ‘strongly against’ scale.</td>
</tr>
<tr>
<td>1</td>
<td>Student response only captures one side of issue, and neglects details and key points. One or two valid claims supported by valid evidence were constructed and positioned correctly on the scale OR one valid claim is supported by more than one valid evidence and positioned correctly on the scale OR more than two valid claims supported by valid evidence were constructed but incorrectly positioned on the scale.</td>
</tr>
<tr>
<td>0</td>
<td>Student does not respond, or fails to address the task</td>
</tr>
</tbody>
</table>
Table 4

**Scoring rubric for the relationships within a concept created by a student**

<table>
<thead>
<tr>
<th>Point Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Student effectively analyzes how parts of a whole interact with each other to produce overall complex dilemma representation system, as it reflected in the concept map relationships. At least 4 claims are connected and a valid reason for the linkage is provided.</td>
</tr>
<tr>
<td>2</td>
<td>Student analyzes how parts of a whole interact with each other to produce overall complex dilemma representation system, as it reflected in the concept map relationships. At least 3 claims are connected and a valid reason for the linkage is provided.</td>
</tr>
<tr>
<td>1</td>
<td>Student attempts to analyze how parts of a whole interact with each other to produce overall complex dilemma representation system, as it reflected in the concept map relationships. Two claims are connected and a valid reason for the linkage is provided OR more than two claims are connected but the linkage reasons are limited.</td>
</tr>
<tr>
<td>0</td>
<td>Student does not respond, or fails to address the task</td>
</tr>
</tbody>
</table>

Scoring of the student responses was provided independently by two teachers from participating schools in the United States. Inter-coded agreement of recommendation scoring was 94% and 100% for the concept map and the relationships. It should be noted that student responses were scored based on the rubrics presented in Tables 2-4, while spelling and grammar issues did not affect the student score.

**Motivation Questionnaire**

The questionnaire included 4 items to assess the extent to which students were motivated to work on the task. Participants reported the degree of their agreement with each item on a 4-point Likert scale (1 = strongly disagree, 4 = strongly agree). The items, adopted from motivation questionnaires used in previous studies, included (Rosen, 2009; Rosen, Beck-Hill, 2012): I felt interested in the task; The task was fun; The task was
attractive; I continued to work on this task out of curiosity. The reliability (internal consistency) of the questionnaire was 0.81.

Students were also asked to indicate background information, including gender, Grade Point Average (GPA), and Math and English Language Arts (ELA) average scores. It should be noted that this informant was provided directly by the students and was not validated with the teachers. The background information was collected because of potential interaction with study variables.

**Results**

All results are presented on an aggregative level beyond the countries because no interaction with country was found. First, the results of student performance in a critical thinking assessment are presented to determine whether there is a difference in the student critical thinking score as a function of working with an evidence-based concept tool. Next, the results regarding the relationship between student performance in critical thinking assessment and the ability to develop ECCM, and create a linkage within ECCM, are shown. Then, gender-related results are presented to indicate possible differences in student performance in ECCM and notepad modes, as well as the relationship with the student’s school achievement. Last, student motivation, and time-on-task, in both modes are demonstrated.

**Student Critical Thinking Performance**

The results of the critical thinking scores indicated that students who worked with ECCM on an assessment task significantly outperformed the students who were assessed
Evidence-Centered Concept Map in Critical Thinking Assessment

in notepad mode (M=69.9, SD=27.2 in ECCM mode, compared to M=54.5, SD=19.0 in notepad mode; ES=.7, t(df=188)=4.7, p<.01). Students who worked with ECCM provided more informed recommendations by using supporting evidence from the available resources and discussing alternative points of view on the topic.

**ECCM-related Performance and Student Critical Thinking**

To better understand the relationship between student critical thinking and the ability to develop ECCM, and create a linkage within ECCM, analysis of correlations between the variables was conducted. The findings showed a significantly positive relationship between student critical thinking score and both the ability to develop ECCM, and the ability to create a linkage within ECCM (r=.62, p < .01 and r=.59, p < .01, respectively). Although the student’s ability to develop ECCM and his or her ability to create a linkage are related to the same bigger construct of working with ECCM, the results indicated that these two sub-constructs are relatively distinctive (r=.40, p<.01).

**Gender and Critical Thinking Performance**

The findings indicated that female students outperformed males in the critical thinking score while working with ECCM (M=78.1, SD=24.8, in female students, compared to M=65.1, SD=3.4 in males; ES=.5, t(df=100)=2.4, p<.05). No significant difference between female and male students was found in critical thinking score in notepad mode (M=54.2, SD=19.5, and M=54.9, SD=18.8; ES=-.03, t(df=86)=-.2, p=.87, respectively).
Student School Achievement and CPS Performance

Correlations between the variables were conducted in order to determine potential relationships between student-reported GPA, ELA achievement, and Math achievement as measured by traditional school assessments and student performance in critical thinking in ECCM and notepad modes of assessment. The findings showed low positive correlation between student critical thinking score in ECCM mode and student school achievement as reflected by GPA and ELA (\( r = .20, p < .05 \) and \( r = .22, p < .05 \), respectively). No significant correlations were found between student critical thinking score and school achievement in notepad mode.

Student Motivation and Time-on-Task

Data were analyzed to determine possible differences in student motivation of being engaged in working with ECCM versus a notepad mode. The results demonstrated that it did not matter for the student’s motivation whether he or she analyzed the dilemma with or without ECCM (\( M = 2.7, SD = .6 \) in ECCM mode, compared to \( M = 2.6, SD = .6 \) in notepad mode; \( ES = .1, t(df=188)=.9, p=.37 \)). No significant difference was found in time-on-task (\( ES = .2, t(df=188)= 1.4, p=.16 \)). On average, time-on-task in ECCM mode was 33.2 minutes (SD=15.1), while students in the notepad mode each spent 2.9 minutes less on the task (\( M = 30.3, SD = 13.7 \)).
Discussion

Policymakers, researchers, and educators are engaged in vigorous debate over leveraging the power of technology to measure what matters for student college and career readiness in valid, reliable, and scalable ways. Technology can support measuring performance that cannot be assessed with conventional testing formats, providing educational systems with opportunities to enable more effective and engaging assessments of important competencies and aspects of thinking (U.S. Department of Education, 2010; National Research Council, 2011). In order to understand how students perform on critical thinking computer-based assessment with thinking tools that can provide scaffolding for the student’s thinking process, it is necessary to examine empirically student performance with these tools. The goal of this study was to explore patterns in student critical thinking performance and motivation in ECCM mode, compared to notepad mode of assessment. Students in both modes were able to analyze a multifaceted dilemma regarding whether or not to buy organic milk for the school cafeteria by using similar information resources. However, while in the ECCM mode, students used ECCM to organize their thinking; in the notepad mode, students were provided with a basic online notepad to make records as needed. The findings showed that students assessed in ECCM mode outperformed their peers in notepad mode in their critical thinking. Overall, decision making with a concept map involved significantly higher levels of analysis and evaluation of evidence, claims, and alternative points of view, as well as synthesis, making connections between information and arguments, interpreting information, and making inferences by using reasoning appropriate to the
situation. Moreover, it was found that student ability to construct ECCM and the ability to create relationships within ECCM are positively linked to student performance in critical thinking. Concept mapping as a thinking tool supports, guides, and extends the thinking process of the student. The thinking tool does not necessarily reduce information processing, but its goal is to make effective use of mental efforts of the student to create a person-plus the technology in computer-based assessment (Jonassen, 2006; Perkins, 1993). To successfully make a decision or solve a multifaceted problem, the student must mentally construct a problem space by analyzing various pieces of information, and mapping specific relationships of the problem. ECCM facilitates the analysis that students conduct and requires them to think more deeply about the multifaceted topic being analyzed than they would have without the thinking tool. The results demonstrated that it did not matter for a student’s motivation whether he or she analyzed the dilemma with or without ECCM, which suggests that the ECCM introduced no motivational obstacles for students in terms of being required to work with a thinking tool. To the degree that students do not give full effort to an assessment test, the resulting test scores will tend to underestimate their levels of proficiency (Eklöf, 2006; Wise & DeMars, 2005). One may claim that adding the ECCM-based thinking process to the assessment could be perceived negatively by the student as an additional assessment requirement and not as a scaffolding tool. Thus, the evidence of equivalent motivational level during both modes of critical thinking assessment is a positive indicator for the use of thinking tools in general and ECCM in particular in computer-based assessments.

One major possible implication of the score difference in critical thinking between the ECCM and the notepad modes is that assessments delivered in multiple
modes may differ in score meaning and impact. Each mode of CPS assessment can be uniquely effective for different educational purposes. For example, an assessment program that has adopted a vision of a conceptual change in assessment may consider the person-plus the thinking tools approach for higher-order thinking assessment as a more powerful avenue for next generation computer-based assessment, while the person-solo approach may be implemented as a more conventional computer-based assessment. While technology tools can promote fundamental improvements in assessment of higher-order thinking skills (Bennett, 1999; Bennett et al., 2007; Pellegrino, Chudowsky, & Glaser, 2001; Tucker, 2009), assessment of foundational knowledge, skills, and abilities can rely on more traditional person-solo oriented assessment approaches. Thinking tools can enable scaffolding and visibility in the student thinking process while working on complex problem solving or decision-making situations that require mindfulness and thinking beyond WYSIATI (Langer, 1989; Kahneman, 2011).

With respect to possible differences in critical thinking performance within both modes of assessment between female and male students, we found that female students outperformed males in critical thinking score while working in ECCM mode. However, no significant difference between female and male students was found in critical thinking score in notepad mode. These results suggest that female students were able to leverage the ECCM for more meaningful cognitive scaffolding by using it as a thinking tool to optimize their critical thinking process. These findings are consistent with previous studies that have shown that female students were more successful in constructing more nets per concept map (Gerstner & Bogner, 2009). Similarly to more conventional person-solo oriented assessment, students may benefit differently from qualitatively different
types of assessment item types or environments. In this assessment the thinking tool was introduced before the actual measurement of student performance started. However, no examples of a constructed ECCM or teacher-led instructions were provided as part of this pilot study. One may consider adding these introduction components to such an assessment to promote student familiarity with the tool, as well as support student metacognitive awareness of the potential benefits of using this tool in an assessment context.

Further analysis examined the extent to which student GPA, ELA achievement, and Math achievement, as reported by the students, are related to student critical thinking performance in both modes of assessment. We found evidence for a low positive relationship between GPA, ELA, and critical thinking student scores in ECCM mode, but no significant correlations were found in notepad mode of assessment. These results suggest that the current critical thinking approaches for assessment in both modes are distinctive from measurement of the conventional domains in schools. The ECCM that was new to all students allowed each student to better analyze the dilemma relatively, regardless of his or her reading, writing, or math skills. Building a concept map is a cognitively challenging task that requires assessing and classifying information, recognizing patterns, identifying and prioritizing main ideas, comparing and contrasting, identifying relationships, and thinking logically (Jonassen, 1996; Kinchin et al., 2000). These findings suggest that a semi-formal visualized information representation with a finite set of concepts and relationships between pairs of concepts reduces the cognitive complexity of analyzing a complex situation for all students (Novak, 1998; Novak & Cañas, 2008). Although a natural language is used to represent concepts and linking
phrases, it is evident that no advanced ELA proficiency is required to be able to show proficiency in critical thinking assessment with an ECCM embedded thinking tool.

The current study had several limitations. First, it is based on a relatively small and non-representative sample of 14-year-old students in four countries and self-reported school achievement. However, due to a lack of empirical research in the field of computer-based assessment of critical thinking skills with embedded thinking tools, it is necessary to conduct small-scale pilot studies in order to inform more comprehensive approaches of critical thinking person-plus assessment. Further studies could consider including a representative sample of students with a wider range of ages and backgrounds and self-reported school achievement. Second, the study operationalized the thinking tool in critical thinking assessment through ECCM, while other approaches could be considered, including semantic organization tools, dynamic modeling tools, information interpretation tools, knowledge construction tools, microwords, and conversation and collaboration tools (Jonassen, 2006; Jonassen & Reeves, 1996). Finally, it is possible that the comparability findings between ECCM and notepad performances in other critical thinking contexts will be different. Future studies could consider exploring differences in student performance in a wide range of problems and decision-making situations.
References


