Assessing Critical Thinking: A Course-Embedded Approach

Gwendolyn W. Johnson, Ph.D.
The University of Arizona
Tucson, AZ
gwj@email.arizona.edu

Anne-Marie Hall, Ph.D.
The University of Arizona
Tucson, AZ
hall@email.arizona.edu

Introduction

The purpose of the Critical Thinking Assessment Project (CTAP) is to develop a process for assessing critical thinking on the campus. The process is designed to collect baseline information about students’ critical thinking skills at the lower and upper division. The project was developed during 2008-09 with Phase I being launched in fall 2009. Five sections of first-year composition (FYC F2009) in English 101 and two large Tier Two classes were selected for participation. A critical thinking assignment that was developed by participating faculty was administered to the students in the seven classes. Students in the FYC classes were given the assignment during the first few weeks of school while the students in the Tier Two classes were given the assignment toward the end of the semester. The assignment, which included a common reading, was embedded in each course, and subsequently scored independently using a rubric (See Appendix A) that was developed for the project. As far as students were concerned, the assignment was a ‘regular’ part of the course. The grading for the assignment for course requirement purposes was left up to the instructors.

Phase II of the CTAP was implemented in spring 2010 and was developed to look at improvement in critical thinking over the course of the two semester Foundations English Composition course sequence. During the first week of May, students were given an essay and a prompt and asked to evaluate the effectiveness of an argument and to identify and analyze particular rhetorical strategies used by the writer to make their own argument.

Six sections of first-year composition (FYC S2010) in English 102 were selected for participation. A critical thinking assignment was developed by faculty in the Writing Program that assessed the skills taught in English 102. Students in the six sections representing three different instructors were given this assignment in the final two weeks of the semester and asked to read the essay, write an essay in response to the reading, and submit their essays electronically. The assignment included a reading of the type that students would be expected to analyze as part of the course goals and objectives. All of the instructors required this assignment and graded it as part of the course.

The majority (94%) of students in the FYC classes were first-time college students while 90% of students in the Tier Two classes were continuing students who were beyond the first year.

Phase III, which was implemented in fall 2010, employed the basic model as in Phases I and II. A sample of student work was collected from a large Tier Two general education class. The majority (92%) were continuing students who were beyond the first year. The assignment was not only embedded in the course but was an essay prompt that was part of the final examination. The critical thinking assessment
was performed on the written responses to that exam prompt and was scored using the revised project rubric (See Appendix B). The context for the assessment was different from Phase I and II and led to an interesting discussion about what constitutes “critical thinking” in the different writing contexts.

**Method**

**Participants**

**Instruments**

Scoring Rubric

The scoring was developed by the university’s assessment coordinating council during spring 2009 and is based on Bloom’s taxonomy of learning objectives (Bloom, 1956).

**Procedure**

**Phase I**

During spring 2010, 10 instructors participated in a norming and grading session. All 10 instructors had received an anchor set of 10 papers representing the full range of the rubric (1 to 5), a copy of the assignment and associated reading, and the rubric. The instructors were asked to read and take notes in preparation for the session. The readers discussed 6 papers for one hour, coming to agreement about the representative scores, clarifying concerns about the rubric and assignment.

All papers were read twice and composite scores were created by summing the two scores; scores ranged from 2 (lowest) to 10 (highest). Any papers with a disagreement (two or more points apart on the rubric) were read a third time and a final composite score assigned. Additionally, all papers receiving an initial composite score of 5 [3 (Satisfactory) and a 2 (Unsatisfactory)] were considered borderline and read a third time in order to move the paper to Satisfactory or Unsatisfactory. Thus, no final composite scores of 5 were assigned.

The final score distributions were grouped into three rating categories: Excellent, Satisfactory, or Unsatisfactory. The range of scores possible were 2 (lowest – meaning the paper received two scores of 1 by two different readers) to 10 (two scores of 5 by two different readers).

Scores of 9-10   Excellent  
Scores of 6-8   Satisfactory  
Scores of 2-4   Unsatisfactory
Phase II
In summer 2010, seven instructors participated in another norming and grading session. The instructors received an anchor set of 11 papers representing the full range of the rubric (1 to 5), a copy of the assignment and associated reading, and the rubric (see Appendix G - a revision from the rubric used in Phase I). The instructors were asked to read and take notes in preparation for the session. The readers discussed all 11 papers for one hour, coming to agreement about the representative scores. All papers were read twice. A total of 119 papers were collected. Any paper receiving a score more than one number apart was read a third time; additionally, all papers receiving a 2 and a 3 were read a third time since a 2 represents below competent and a 3 represents competent; thus there are no scores of 5 in the total. As in Phase I, the final scores were grouped into the three ratings categories (See Appendix G for a scatterplot of raw scores).

Phase III
In spring 2011, five instructors participated in another grading session; all five instructors received an anchor set of 11 papers representing the full range of the rubric (1 to 5), a copy of the assignment and associated reading, and the rubric (see Appendix B). The instructors were asked to read and take notes in preparation for the session. The readers discussed all 11 papers and came to agreement about the representative scores. All papers were read twice. Any paper receiving a score more than one number apart was read a third time; additionally, all papers receiving a 2 and a 3 were read a third time since a 2 represents below competent and a 3 represents competent; thus there are no scores of 5 in the total. As in Phases I and II, the final scores were grouped into the three ratings categories.

Analysis of Scores

Phase I
A total of 280 papers were collected and scored: 172 from Tier Two and 108 from First-Year Composition (FYC). Table 2 displays the number and percent of papers that fell within the three rating categories. Table 3 shows the number and percent of papers that were read a third time in order to move them into either the Satisfactory or Unsatisfactory category.

Table 1. Number of papers and percent receiving ratings of Unsatisfactory, Satisfactory, and Excellent

<table>
<thead>
<tr>
<th>Category</th>
<th>Unsatisfactory (2-4)</th>
<th>Satisfactory (6-8)</th>
<th>Excellent (9-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of papers</td>
<td>Percent</td>
<td># of papers</td>
</tr>
<tr>
<td>FYC 2009</td>
<td>47</td>
<td>43%</td>
<td>58</td>
</tr>
<tr>
<td>TIER TWO</td>
<td>33</td>
<td>19%</td>
<td>109</td>
</tr>
<tr>
<td>TOTAL</td>
<td>80</td>
<td>29%</td>
<td>167</td>
</tr>
</tbody>
</table>

Table 2. Resolution of Papers with Initial Composite Scores of 5 – number of papers and percentage

<table>
<thead>
<tr>
<th>Category</th>
<th>Papers with Scores of 5</th>
<th>Resolved to Satisfactory</th>
<th>Resolved to Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of papers</td>
<td>Percent</td>
<td># of papers</td>
</tr>
<tr>
<td>FYC 2009</td>
<td>37</td>
<td>34%</td>
<td>12</td>
</tr>
<tr>
<td>TIER TWO</td>
<td>26</td>
<td>15%</td>
<td>11</td>
</tr>
<tr>
<td>TOTAL</td>
<td>63</td>
<td>23%</td>
<td>23</td>
</tr>
</tbody>
</table>
The extent to which the readers agreed in their scoring of the papers was calculated. Table 4 displays the number of scores that were in agreement and the percentage of agreement among the readers.

Table 3. Inter-rater Agreement for Fall 2009

<table>
<thead>
<tr>
<th></th>
<th>Number of Scores in Agreement</th>
<th>Percentage of Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>FYC 2009</td>
<td>101 of 108</td>
<td>93%</td>
</tr>
<tr>
<td>TIER TWO</td>
<td>146 of 172</td>
<td>85%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>247 of 280</td>
<td>82%</td>
</tr>
</tbody>
</table>

**Phase II**

A total of 119 papers from were collected and scored. Table 5 displays a comparison of the number and percent of papers that fall in the rating categories for the FYC students that participated in fall 2009 and the FYC students that participated in spring 2010.

Table 4. Number of papers and percent receiving ratings of Unsatisfactory, Satisfactory, and Excellent

<table>
<thead>
<tr>
<th></th>
<th>Unsatisfactory (2-4)</th>
<th>Satisfactory (6-8)</th>
<th>Excellent (9-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td># of papers</td>
<td>Percent</td>
<td># of papers</td>
<td>Percent</td>
</tr>
<tr>
<td>FYCF2009</td>
<td>47</td>
<td>58</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>43%</td>
<td>54%</td>
<td>3%</td>
</tr>
<tr>
<td>FYCS2010</td>
<td>45</td>
<td>69</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>38%</td>
<td>58%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Table 5. Resolution of Papers with Initial Composite Scores of 5

<table>
<thead>
<tr>
<th></th>
<th>Papers with Scores of 5</th>
<th>Resolved to Satisfactory</th>
<th>Resolved to unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td># of papers</td>
<td>Percent</td>
<td># of papers</td>
<td>Percent</td>
</tr>
<tr>
<td>FYC F2009</td>
<td>37</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>34%</td>
<td>32%</td>
<td>68%</td>
</tr>
<tr>
<td>FYC S2010</td>
<td>29</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>24%</td>
<td>48%</td>
<td>52%</td>
</tr>
</tbody>
</table>

Table 6. Inter-rater Agreement for Spring 2010

<table>
<thead>
<tr>
<th></th>
<th>Number of Scores in Agreement</th>
<th>Percentage of Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>FYC F2009</td>
<td>101 of 108</td>
<td>93%</td>
</tr>
<tr>
<td>FYC S2010</td>
<td>115 of 119</td>
<td>97%</td>
</tr>
</tbody>
</table>

**Phase III**

A total of 108 papers were collected and scored. Table 8 displays a comparison of the number and percent of papers that fall in the rating categories for the students that participated in fall 2010.

Table 7. Number of papers and percent receiving ratings of Unsatisfactory, Satisfactory, and Excellent
<table>
<thead>
<tr>
<th>Tier Two 2010</th>
<th>35</th>
<th>32%</th>
<th>60</th>
<th>55%</th>
<th>13</th>
<th>12%</th>
</tr>
</thead>
</table>

Table 8. Resolution of Papers with Initial Composite Scores of 5

<table>
<thead>
<tr>
<th>Papers with Scores of 5</th>
<th>Resolved to Satisfactory</th>
<th>Resolved to unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td># of papers</td>
<td># of papers</td>
<td># of papers</td>
</tr>
<tr>
<td>Tier Two F2010</td>
<td>12</td>
<td>11%</td>
</tr>
</tbody>
</table>

Table 9. Inter-rater Agreement for Spring 2011

<table>
<thead>
<tr>
<th>Tier Two F2010</th>
<th>Number of Scores in Agreement</th>
<th>Percentage of Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>91 of 108</td>
<td>84%</td>
</tr>
</tbody>
</table>

Results

Phase I
The results of Phase I of the CTAP were encouraging. The majority of students received Satisfactory or Excellent scores (71%). In addition, only 19% of Tier Two students received Unsatisfactory while 43% of FYC students received Unsatisfactory scores, suggesting that students are improving in critical thinking and writing abilities between the time they enter the UA and the time they have enrolled in 45-60 units. Another promising result was the increase in numbers of Excellent scores between FYC and Tier Two students: only 3% of FYC papers received an Excellent score while 17% of Tier Two students received a score of Excellent.

Twenty-three percent of the papers received a composite score of 5 and were considered borderline, with one reader scoring the paper Unsatisfactory (2) and one scoring the paper Satisfactory (3). In FYC, 34% of the total papers received a borderline score of 5 while only 15% of the Tier Two papers received a borderline score of 5; 63% of borderline scores were resolved to Unsatisfactory. In general, the Tier Two papers received higher scores and were less likely to fall into the 3/2 scoring dilemma.

Results showed high agreement among the readers about scoring the papers. Interestingly, 26 of 172 Tier Two papers were in disagreement (15%) while only 7 of the 108 FYC papers were similarly in dispute (7%). The readers were all instructors of FYC, which could have led them to read that level of writing more accurately; also, there tended to be more disagreement in the norming session as well about the higher end papers – another possible explanation for less agreement on the Tier Two papers. Overall, the agreement was high with 82% inter-rater agreement.

General Comments about Papers
In a closing discussion with the 10 readers, many offered general statements about the FYC papers as distinguished from the Tier Two papers. In sum

- FYC papers tended to
  - Summarize more
  - Rely on simple paragraph structures (point, illustrate, explain)
  - Not grasp globalization aspect of the article
  - Focus on personal lives to make a point
Avoid the ‘big picture”
Argue in polemics
Indulge in “broad racialized stereotypes”
Not appreciate nuance – i.e., argue that all products should only be produced in the U.S.
Rely more on the authority of the author to make their own points

Tier Two papers tended to
Rely on background knowledge (content) to pull into their responses
Individualize voices and see nuanced ways to make a point
Understood the point of the article and the assignment
Had problem-solving capabilities
Saw consequences. cause-effect in reasoning
Saw grey areas, not just black and white, digital thinking about topics
Recognized flaws in the article (i.e., Chinese did not give any information so they did not infer a problem from that whereas in FYC, students made inferences from observations that were not substantiated)
Demonstrated meta-cognitive skills (able to know how and why they know what they know)

Phase II
Results of Phase II of the CTAP suggested improvement between FYC students in the first two weeks of class in Fall 2009 and in the final week of classes in Spring 2010 in their critical thinking ability. While the majority of the students received Satisfactory or Excellent scores in both semesters, there was an increase of 6% more students receiving Satisfactory or Excellent at the end of the two-semester sequence. Additionally, 5% fewer students received Unsatisfactory scores.

The borderline scores (composite of 5) decreased by 10% in spring 2010; more significantly, 16% more students in Spring 2010 had resolutions up to competent by a third reader (compared to Fall 2009). In addition, 16% fewer students had resolutions down to below competent by a third reader.

In both sets of readings, there was extremely high inter-rater agreement – 93% in Fall 2009 and 97% in Spring 2010.

General Comments about Papers
In a closing discussion at the end of the holistic scoring session, the 7 readers offered the following general statements about the rubric, the prompt, the reading, and the students’ essays:

- Rubric was significantly improved and was especially helpful in determining the difference between a 2 (below competent) and a 3 (competent) as well as between a 4 and a 5. Some instructors felt the word “problem” was difficult and preferred “thesis” or “issue” or “argument.”
- Reading “Race Over” was difficult but did lead students into immediately critical thinking. There was a certain amount of irony in the essay which almost all students did not notice. Reasons for this vary but most instructors felt texts used to teach critical analysis skills tend to be serious and that few instructors talk about “satire” or “irony.” Thus almost all the students read the essay “straight up.”
- Prompt – probably not very portable. This prompt references rhetorical effectiveness and strategies, a skill set very specific to English 102 at The University of Arizona. A better prompt might have asked students to read a researched essay and analyze the effectiveness of the argument. Also, the rhetorical analysis essay taught in English 102 is a five page essay and this prompt asked students to do a rhetorical analysis in about 2 pages.
• Readers felt the prompt encouraged students to “write an essay” rather than “answer a prompt.”
• Three instructors read the Phase I papers as well as the Phase II papers. All three felt the writing itself was improved for the FYC cohort; in addition, all the readers believed that the overall writing was improved.
• All the readers were impressed with “decently structured essays” which they felt students in August could not do.
• The subject matter of the reading (“Race Over”) suggested a very serious topic and students did address the subject matter in mostly serious ways (almost no students noted the irony in the reading).

**Phase III**

**General Comments about Papers**
In a closing discussion at the end of the holistic scoring session, the five readers offered the following general statements about the rubric, the prompt, the reading, and the students’ essays:

• Reading “Franklin D. Roosevelt and the New Deal” provided ample evidence of a significant fiscal crisis and the president’s approach to implementing programs for change.
• Readers felt the prompt asked for both a listing of details and analysis, but in the exam setting, most students focused on listing the details rather than providing rich analysis of them. Three instructors read the Phase I and II papers as well as the Phase III papers. All three suggested that the writing demonstrated a range of writing styles and approaches to the prompt. Surprisingly most students did not relate the prompt to their own lives or historic conditions, and some opted for a list rather than an essay format. The most successful writers were those that provided some context for their response and then offered integrated analysis of individual programs as they worked through their essay.
• All the readers were impressed with some students’ ability to integrate facts with analysis, and the raters were particularly impressed with the range of writing styles evidenced in the sample.
• The subject matter of the reading (“Franklin D. Roosevelt and the New Deal”) suggested that students might relate the historic moment in the reading with our current financial crisis. It was surprising that most students did not discuss such relationships.

**Rubric**
During Phase I of the Project, the readers discussed several problems with the rubric. Generally, the readers found the descriptors under the 4 rating (above average) problematic: competent, adequate, satisfactory, reasonable. These are terms better suited to a 3 rating. In general, terminology was inconsistent and sometimes too vague. All 10 raters submitted their notes on the rubrics to Dr. Hall and Dr. Johnson.

Based on input from the readers, faculty, and ACC members, the rubric (Appendix G) was revised for Phase II of the CTAP. Three of the readers who had participated in Phase I of the CTAP indicated that the revised rubric was much clearer than the earlier version and commented on its usefulness in assessing papers. The revised rubric (Appendix G) was used in Phase III.

**Conclusions**

The first three Phases of the CTAP have resulted in a model that may be used to assess critical thinking within a course. The process has yielded a sample of baseline information about students’ critical thinking skills at the lower and upper division, over a two semester course sequence, and as part of an exam. Using results from Phase I, the rubric has been revised and appears to be useful with at least two
types of assignments. Table 11 displays the number of papers and percent receiving ratings of Unsatisfactory, Satisfactory, and Excellent across the four student groups.

Table 10. Number of papers and percent receiving ratings of Unsatisfactory, Satisfactory, and Excellent across the four student groups

<table>
<thead>
<tr>
<th></th>
<th>Unsatisfactory (2-4)</th>
<th>Satisfactory (6-8)</th>
<th>Excellent (9-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of papers</td>
<td>Percent</td>
<td># of papers</td>
</tr>
<tr>
<td>FYC F2009 (n=108)</td>
<td>47</td>
<td>43%</td>
<td>58</td>
</tr>
<tr>
<td>FYC S2010 (n=119)</td>
<td>45</td>
<td>38%</td>
<td>69</td>
</tr>
<tr>
<td>Tier Two F2009 (n=172)</td>
<td>33</td>
<td>19%</td>
<td>109</td>
</tr>
<tr>
<td>Tier Two F2010 (n=108)</td>
<td>35</td>
<td>32%</td>
<td>60</td>
</tr>
<tr>
<td>TOTAL (N=507)</td>
<td>160</td>
<td>36%</td>
<td>167</td>
</tr>
</tbody>
</table>

**Next Phase**
More long-term plans for the assessment of critical thinking include developing assignments/models for assessing CT in senior capstone courses and designing a longitudinal study in which a cohort of first-year students are tracked across several years to assess their critical thinking abilities over time. The results of the CTAP will be used to inform discussions about institutional goals and standards for critical thinking on the campus.

**Appendices**
A: Critical Thinking/Problem Solving Rubric – Fall 2009
B: Critical Thinking/Problem Solving Rubric – Spring 2010
# Appendix A: Critical Thinking/Problem Solving Rubric – Fall 2009

This rubric derives from speculation on what the assignment requires: many papers/assignments incorporate only some of the requirements listed here. The score you assign should reflect your sense of the predominant effect of the paper. In some instances, a second reader will also rate the assignment.

<table>
<thead>
<tr>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The assignment demonstrates mastery of the dimensions associated with critical thinking ability.</strong></td>
<td><strong>The assignment exceeds mere competence in demonstrating critical thinking ability.</strong></td>
<td><strong>The assignment demonstrates competency to the demands of critical thinking ability.</strong></td>
<td><strong>The assignment demonstrates minimal competence in critical thinking ability.</strong></td>
<td><strong>The assignment fails to demonstrate competency to the demands of critical thinking ability.</strong></td>
</tr>
<tr>
<td>Thoroughly, defines the problem and breaks it down into components and identifies which are central or critical to the problem, organizes components in a logical manner, and places the problem in proper context.</td>
<td>Provides competent problem definition</td>
<td>Simplistic problem definition</td>
<td>Lacks adequate problem definition</td>
<td>No problem definition</td>
</tr>
<tr>
<td>Clearly, distinguishes reliable from unreliable information, recognizes bias, and identifies relevant information in relation to the problem.</td>
<td>Adequate distinction between reliable and unreliable information, adequate recognition of bias and identification of relevant information</td>
<td>Limited distinction between reliable and unreliable information, limited recognition of bias and identification of relevant information</td>
<td>Weak distinction between reliable and unreliable information, weak recognition of bias and identification of relevant information</td>
<td>No distinction between reliable and unreliable information, no recognition of bias, and no identification of relevant information in relation to the problem</td>
</tr>
<tr>
<td>Strong formulation of various hypotheses using relevant information and assumptions, selecting the more promising hypotheses for first consideration.</td>
<td>Competent formulation of hypotheses</td>
<td>Simplistic formulation of hypotheses</td>
<td>Lacks adequate formulation of hypotheses</td>
<td>No formulation of hypotheses and assumptions</td>
</tr>
<tr>
<td>Draws valid conclusions from assumptions, hypotheses and relevant information. Distinguishes a necessary from a probable inference.</td>
<td>Draws reasonable conclusions</td>
<td>Draws weak conclusions and</td>
<td>Invalid conclusions are drawn</td>
<td>No conclusions are drawn and no distinction between a necessary and a probable inference</td>
</tr>
<tr>
<td>Strong evaluation of a conclusion in terms of its applicability, recognizes conditions necessary to verify a conclusion and those which make a conclusion inapplicable.</td>
<td>Adequate evaluation of a conclusion</td>
<td>Weak evaluation of a conclusion</td>
<td>Inadequate evaluation of a conclusion</td>
<td>No evaluation of a conclusion</td>
</tr>
</tbody>
</table>
Appendix B: Critical Thinking/Problem Solving Rubric

Revised Spring 2010

This rubric derives from speculation on what the assignment requires: many papers/assignments incorporate only some of the requirements listed here. The score you assign should reflect your sense of the predominant effect of the paper. In some instances, a second reader will also rate the assignment.

<table>
<thead>
<tr>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>The assignment demonstrates mastery of the dimensions associated with critical thinking ability.</td>
<td>The assignment exceeds mere competence in demonstrating critical thinking ability.</td>
<td>The assignment demonstrates competency to the demands of critical thinking ability.</td>
<td>The assignment demonstrates minimal competence in critical thinking ability.</td>
<td>The assignment fails to demonstrate competency to the demands of critical thinking ability.</td>
</tr>
<tr>
<td>Thoroughly and effectively defines the problem and breaks it down into components and identifies which are central or critical to the problem, organizes components in a logical manner, and places the problem in proper context.</td>
<td>Provides clear problem definition with supporting argument</td>
<td>Competent problem definition even if not fully sufficient</td>
<td>Simpistic or less than adequate problem definition</td>
<td>No problem definition</td>
</tr>
<tr>
<td>Clearly and effectively distinguishes reliable from unreliable information, recognizes bias, and identifies relevant information in relation the problem.</td>
<td>Clearly distinguishes between reliable and unreliable information, adequate recognition of bias and identification of relevant information</td>
<td>Makes some distinctions between reliable and unreliable information, limited recognition of bias and identification of relevant information</td>
<td>Limited or weak distinction between reliable and unreliable information, weak recognition of bias and identification of relevant information</td>
<td>No distinction between reliable and unreliable information, no recognition of bias, and no identification of relevant information in relation to the problem</td>
</tr>
<tr>
<td>Identifies unstated, unsupported, and irrelevant assumptions with effective analysis that is sophisticated and fresh.</td>
<td>Identifies and analyzes effects of assumptions</td>
<td>Identifies assumptions</td>
<td>Limited or unsatisfactory identification of assumptions</td>
<td>No identification of assumptions</td>
</tr>
<tr>
<td>Strong formulation of various hypotheses using relevant information and assumptions, selecting the more promising hypotheses for first consideration.</td>
<td>Effective formulation of hypotheses developed with consistent detail and explanations</td>
<td>Formulation of hypotheses though may lack in sufficient analysis</td>
<td>Lacks adequate formulation of hypotheses and assumptions</td>
<td>No formulation of hypotheses and assumptions</td>
</tr>
<tr>
<td>Draws significant and valid conclusions from assumptions, hypotheses and relevant information. Distinguishes a necessary from a probable inference.</td>
<td>Draws reasonable conclusion that is generally effective</td>
<td>Draws adequate conclusions with reasoning that is generally coherent</td>
<td>Invalid conclusions are drawn and reasoning is hard to follow</td>
<td>No conclusions are drawn and no distinction between a necessary and a probable inference</td>
</tr>
<tr>
<td>Strong evaluation of a conclusion in terms of its applicability, recognizes conditions necessary to verify a conclusion and those which make a conclusion inapplicable.</td>
<td>Effective evaluation of a conclusion</td>
<td>Competent if overly general evaluation of a conclusion</td>
<td>Weak evaluation of a conclusion</td>
<td>No evaluation of a conclusion</td>
</tr>
</tbody>
</table>
References
