ASIAN JOURNAL OF THE SCHOLARSHIP OF TEACHING AND LEARNING VOLUME 4 ■ NUMBER 4 ■ DEC 2014

ORIGINAL ARTICLE

Measuring Critical Thinking as Learning Outcome: A Pilot Study from the Ideas and Exposition Module

W. Patrick WADE[1], Ya Hui Michelle SEE[2]

[1] Centre for English Language Communication
National University of Singapore
[2] Department of Psychology
National University of Singapore

Address for Correspondence: Dr Walter Patrick Wade, Writing Unit, #02-16E Education Resource Centre, University Town, 8 College Avenue West, Singapore 138608

Email: wpatrickwade@nus.edu.sg

Recommended citation:
Wade, W. P., and See, Y. H. M. (2014). Measuring Critical Thinking as Learning Outcome: A Pilot Study from the Ideas and Exposition Module. *Asian Journal of the Scholarship of Teaching and Learning, 4*(4), 220-236.
https://doi.org/10.24112/ajsotl.43318

Measuring Critical Thinking as Learning Outcome: A Pilot Study from the Ideas and Exposition Module

Abstract

This paper presents results from a pilot study of the effectiveness of an *Ideas and Exposition 1* module in fostering critical thinking (CT) skills amongst students with reference to a particular CT instrument, the Halpern Critical Thinking Assessment (HCTA). The study had two major goals: (1) to determine whether or not the HCTA is an appropriate tool for measuring learning outcomes in *I&E 1*, as determined by whether or not the 5 CT skills measured by the HCTA (argument analysis, verbal reasoning, hypothesis testing, using likelihood and uncertainty, and decision making and problem solving) are taught during the course; and (2) to determine which student CT skills improved (if any) after taking a particular *I&E 1* module. The hypothesis of the study was that there would be substantial, although not complete alignment of the HCTA and the course learning outcomes in *I&E 1*, specifically with respect to the course's ability to promote argument analysis, verbal reasoning, and decision making and problem solving. The pilot study found a statistically significant improvement in argument analysis, no improvement in hypothesis testing, decision making and problem solving, or managing likelihood and uncertainty, and a decrease in verbal reasoning skills. Possible explanations are provided, as well as further suggestions regarding the limits of the pilot study and the lessons it might offer to instructors who hope to improve the CT skills of their students.

INTRODUCTION

The University Town College Programme (UTCP) is a multidisciplinary academic programme offered at the National University of Singapore by the new residential colleges: the College of Alice and Peter Tan, Tembusu College, and Residential College 4. The curriculum of UTCP comprises five modules that students take over the course of four semesters, which concurrently fulfill their University Level Requirements. The UTCP aims to provide students with an academically rigorous environment, small class sizes to encourage closer interaction between students and faculty, and increased access to informal learning opportunities outside of the classroom.

An important part of the UTCP curriculum is administered by the Centre for English Language Communication: two *Ideas and Exposition (I&E)* modules, *I&E 1* and *I&E*

2, which aim to develop critical thinking and writing skills at the outset of students' undergraduate careers. These modules scaffold together to form the complete *Ideas and Exposition* course: *I&E 1* aims to teach students expository writing skills and *I&E 2* aims to teach research-based writing skills. These modules are designed to be rhetorically intensive, as they ask students to consider their motives for writing, their academic and non-academic audience(s), and the rigor of their argumentation as they produce polished essays over multiple drafts. Based on this consistent focus on improving the quality of student writing across both levels of the module, they fit the major criteria for "writing intensive" courses offered at other academic institutions as they have a low student-teacher ratio, require students to complete a draft-revision process, spread writing assignments across the semester, and have the majority of the student's grade determined by the success of writing assignments (Farris and Smith, 1992).

In *I&E 1*, students are introduced to academic expository writing through engagement with a specific course topic, i.e., "Photography and Society," "Heroes," "The Politics of Food."[1] Course topics are intended to define interdisciplinary domains of study rather than assume knowledge and expertise in a particular discipline. This is done through the course design process: readings from different academic fields that each treat the topic of the course are included on the reading list and made subject for discussion. In "Photography and Society," for example, readings are drawn from Media Studies, Visual Communication, Cultural Studies, History, and Political Science. By focusing on an interdisciplinary topic, these courses contribute to the breadth of the UTCP curriculum.

Students learn skills associated with the genre of expository writing by producing three papers for *I&E 1*. The first is a reflective summary paper, in which students demonstrate understanding of an academic source and engage it critically. The reflective summary asks students to demonstrate their ability to effectively summarize and evaluate sources, which are important for their success in subsequent written assignments. The second is a critical comparison paper, in which students evaluate how two different authors treat a common topic or show how one author's perspective might give the reader an improved understanding of the work of a second. This paper asks students to summarize, synthesize, and evaluate the arguments of secondary source material. The third assignment is an expository essay, in which students explain something significant about a topic related to course themes by critically engaging with the larger scholarly and/or public conversation surrounding that topic.

Course instruction has both content and writing objectives, with students expected to delve into course topics so that they have an appropriate basis for nuanced, thoughtful writing. Students are expected to learn how to understand and engage with sources on the course topic, as well as identify how authors in the field of study critically engage one another. Students are expected to employ cogent logical connections to support their claims and to learn appropriate organization and phrasing to successfully translate

[1] Since the start of the *I&E* programme, seventeen different course topics have been offered at the level of *I&E 1*.

explanation into writing. They are expected to properly document sources as well as to learn revision skills that will help them to reevaluate their own work and productively comment on the work of their peers (Appendix 1). These revision skills are taught through frequent peer reviews and one-on-one conferences with the instructor on preliminary paper drafts.

I&E 2 builds on the expository writing skills taught in *I&E 1* with a focus on developing research-based argumentative papers. As in *I&E 1*, students develop their writing skills through engagement with a specific, interdisciplinary course topic, i.e., "Interpreting the Vampire," "Ethics and Outer Space," "Visualizing Southeast Asian Cities."[2] In these courses, students should learn three significant kinds of academic writing pertinent to academic research: the annotated bibliography, the research proposal, and the research paper. Students choose a research topic relatively early in the course and orient each assignment toward the successful development and execution of a final paper on their research topic. Along the way, they are taught generic expectations for annotated bibliographies, research proposals, and research papers within the content area. The course thus takes a rhetorical approach to genre that that recognizes that genres cannot be taught outside of a particular context within which writing has specifiable social purposes, audiences, and argument forms (Wardle, 2009).

I&E 2's shift from expository writing to research-based argumentative writing involves a corresponding shift in expected course learning outcomes. Students are expected to perform both primary and secondary research on individual research topics. While conducting secondary research, they assess the relevance and reliability of multiple sources. Furthermore, they synthesize these sources to identify research problems in the area of study and to support hypotheses about their individual research topic. While conducting primary research, they are expected to employ basic research methods (gathering primary texts, conducting simple surveys or interviews, observing field sites) and to subject gathered data to analysis by using concepts from the relevant literature as an interpretive lens. Students are expected to continue to develop revision skills, expanding them to consideration of issues in research design and execution.

The diversity and range of these learning outcomes, coupled with the breadth of course topics, as well as the substantial differences between the two levels of the module, pose a significant challenge to uniform assessment of the course's impact on students. In order to arrange for an appropriate measure of course effectiveness, I hoped to find a single measure capable of tracking the major improvements course instructors aim to instill in student writing across the two modules.

A common element across *I&E 1* and *I&E 2* that this paper proposes to measure is critical thinking (CT). CT refers to a suite of higher order thinking skills pertaining to inference, evaluation, analysis, argumentation, creativity, and reflectiveness. An

[2] Since the start of the *I&E* programme, eight different course topics have been offered at the level of *I&E 2*.

interdisciplinary group of 46 scholars convened in 1990 and issued the Delphi report, defining CT as "purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which judgment is based" (Facione et al., 1990). Other skills associated with CT include the ability to ask appropriate questions, define terms, and identify assumptions (Ennis, 1985) the ability to take into account probability and uncertainty in evaluating a persuasive message, or the ability to formulate and test hypotheses in the face of problems (Halpern, 2003).

Instead of engaging the larger literature about the nature of CT, we hope to find out if existing instruments can be used to measure the development of such discrete CT skills in *I&E* modules. After considering the specific course learning outcomes for *I&E* modules (Appendix 1 and 2) and other possible assessment tools, we found the most potential for alignment with Diane Halpern's critical thinking assessment tool, the Halpern Critical Thinking Assessment (HCTA). This tool stands out for being one of the few CT measures that has had cross-cultural validation with Asian populations (Hau et al., 2006; Ku et al., 2010). It is comparatively simple to administer and is unique amongst CT assessments as it also claims to measure transfer of CT skills by considering participants' disposition to use them in everyday scenarios (Butler et al., 2012).

Halpern's test measures five discrete sets of CT skills: verbal reasoning, argument analysis, thinking as hypothesis testing, using likelihood and uncertainty, and decision making and problem solving. Each of these CT-related skills and their relationship to learning outcomes in *I&E* modules is discussed below.

1) Verbal reasoning skills.

This dimension of CT reflects the ability to identify and assess persuasion techniques as they are deployed in ordinary language. It assesses the ability to identify and evaluate the assumptions, definitions, labels, and connotations at work in persuasive uses of ordinary language. Approximately 12% of the questions in the HCTA measure verbal reasoning skills. This could be aligned with outcomes in *I&E 1* and *I&E 2*, namely, "analyzing how authors problematize what other authors say" (Appendix 1) and "assessing the relevance and reliability of multiple sources" (Appendix 2).

2) Argument analysis skills.

This dimension of CT reflects the ability to identify the logical connections between premises and conclusions in elaborated arguments, as well as evaluate the strength of reasons and evidence given as premises within arguments. Approximately 21% of the questions in the HCTA measure argument analysis skills. These thinking skills seem closely aligned with program learning outcomes in *I&E 1* and *I&E 2*, namely,

"entering conversations between authors by problematizing their arguments and arguing why one's problem and response are reasonable with available evidence (Appendix 1)," "organizing and wording ideas to help readers understand a line of reasoning" (Appendix 1)," and "assessing the relevance and reliability of multiple sources" (Appendix 2).

3) Thinking as hypothesis testing.

This dimension of CT reflects skills that might otherwise be associated with the scientific method: making observations, asking questions, making predictions, and gathering evidence to address questions. It also reflects the ability to evaluate data sets to see if they support given conclusions. Approximately 24% of the questions in the HCTA measure thinking as hypothesis testing. This thinking skill could be aligned with learning outcomes from *I&E 2*, namely, student's ability to "identify and articulate exigent research questions" and "to understand how theory and method inform and produce knowledge" (Appendix 2).

4) Using Likelihood and uncertainty.

This dimension of CT reflects the ability to determine the degree of certainty with which certain premises and conclusions are given. It involves understanding the appropriate or inappropriate use of hedging in the support of claims, as well as judging the probability of certain statements. Approximately 12% of the questions in the HCTA measure the ability to use likelihood and uncertainty. This thinking skill is not aligned with the learning outcomes of the *I&E* program.

5) Decision making and problem solving skills.

This dimension of CT reflects the ability to define problems, generate alternatives, and compare solutions to come up with the best decision to undertake with respect to a given problem. Approximately 31% of the questions in the HCTA measure decision making and problem solving. This thinking skill could be aligned with learning outcomes in *I&E 1* and *2*, namely, "Entering the conversations between authors by problematizing their arguments and arguing why one's problem and response are reasonable with available evidence (Appendix 1)," "identifying exigent research questions (Appendix 2)," and "synthesizing multiple sources to construct and support hypotheses" (Appendix 2).

There is some risk, however, that the alignment proposed between the HCTA and *I&E* courses may not be complete. There are two reasons for potential misalignment. First, standardized CT assessments like Halpern's are based on particular understandings of CT that may or may not correspond with the way teachers understand and teach it. Second, because the development of CT skills as a learning outcome in *I&E* is integrated into student's engagement with a particular content area, CT instruction is not the sole focus of the course. There is an on-going debate over the effectiveness of general vs. integrated critical thinking courses (Hatcher, 2006), and it remains to be seen whether or not CT skills have been effectively integrated into *I&E* course instruction.

With these issues in mind, this paper presents results from a pilot study of the effectiveness of *I&E* modules in fostering CT skills amongst students with reference to the HCTA. The study asked two questions:

1. To what extent is the HCTA an appropriate tool for measuring learning outcomes in *I&E 1* with respect to the 5 CT skills measured by the HCTA?

2. Which student CT skills improved (if any) after taking a particular *I&E 1* module?

The hypothesis of the study was that there would be substantial, although not complete alignment of the HCTA and the course learning outcomes in *I&E 1*, specifically with respect to the course's ability to promote argument analysis, verbal reasoning, and decision making and problem solving. This is because these are the skills most likely to be associated with the major course writing assignments. Thinking as hypothesis testing is a CT skill more appropriate to the second level of the module, when students are asked to formulate research questions and practice research design, and thus we did not expect *I&E 1* to improve students' performance in this skill. We predicted that the course would not improve students' ability to use likelihood and uncertainty, because this skill is least pertinent to the design of *I&E* modules, which do not ask students to engage with mathematical uses of reason (i.e., considering sample size, probability, or statistical claims).

In the general discussion, further suggestions will be offered regarding the lessons the pilot study might give to instructors who hope to improve the critical thinking skills of their students through their teaching.

METHOD

Participants and Overview of Procedure

Twenty-four students in their first year at the University Town College Programme (17

females and 7 males) completed two versions (version S1 and version S2) of the HCTA during two sessions, which were 9 weeks apart. The sessions were conducted during class. One student was more than two standard deviations (SDs) away from the sample mean in the time taken to complete the pre-test (i.e., 49.48 mins). Another student was more than two SDs from the sample mean in the time taken to complete the post-test (i.e., less than five minutes).[3] The data from these two students were excluded in the analyses because the students would not have had enough time to read through and address the test prompts—the times and scores suggested instead that students were clicking through the test without completing it.

Students were given the HCTA twice in a pre-test/post-test experimental design (Ennis, 1993). In the pre-test administration, students were asked to complete version S1. This assessment took place at the National University of Singapore at the start of the 3rd week of Semester 1, AY2013-14. In the post-test administration, the same students were asked to complete version S2. This assessment took place at the National University of Singapore in the 12th week of Semester 1, AY2013-14. Test version S1 includes two kinds of questions: open-ended questions to measure free recall of the CT variables and multiple choice questions to measure recognition of the CT variables. Test version S2's only difference with version S1 is that it removes the open ended free recall questions—otherwise the tests are identical. Because of the differences between the two tests, only the recognition variables associated with the 5 CT-related skills and the aggregate CT score were collected and analyzed.

Materials

Participants took the Halpern Critical Thinking Assessment, which measures five discrete sets of CT skills: Verbal Reasoning, Argument Analysis, Hypothesis Testing, Using Likelihood and Uncertainty, and Decision Making and Problem Solving. The means and standard deviations for participants' scores for Verbal Reasoning were 5.18 and 1.01 respectively in the pre-test, and 4.64 and 1.002 respectively in the post-test. The means and standard deviations for participants' scores for Argument Analysis were 14.64 and 2.01 respectively in the pre-test, and 15.86 and 1.73 respectively in the post-test. The means and standard deviations for participants' scores for Hypothesis Testing were 17.32 and 2.46 respectively in the pre-test, and 17.50 and 3.47 respectively in the post-test. The means and standard deviations for participants' scores for Using Likelihood and Uncertainty were 5.00 and .93 respectively in the pre-test, and 4.73 and 1.12 respectively in the post-test. The means and standard deviations for participants' scores for Decision Making and Problem Solving were 25.77 and 2.78 respectively in the pre-test, and 26.18 and 3.11 respectively in the post-test. Standard errors of the measurement can be calculated from these figures.

[3] The mean amount of time students took to complete the pre-test was 85.26 minutes (SD=15.11). The mean amount of time for the post-test was 20.58 minutes (SD = 6.27).

Participants completed the HCTA on the Vienna Test System Online. Because the system provides only aggregated instead of raw scores, we were unable to compute the internal consistencies for each subscale or the overall scale. The HCTA presents participants with 25 real-world scenarios and poses questions aimed at probing the kind of thinking involved in addressing them. The HCTA is an academically rigorous critical thinking (CT) test. It has been assessed positively for its construct validity (Butler, 2012), inter-rater reliability (Halpern, 2013), and cross-cultural validity (Hau et al., 2006; Ku et al., 2010).

RESULTS AND DISCUSSION

Correlations among Critical Thinking Skills

Scores on the five CT skills were mostly not correlated with one another. This was true for pre-test scores, ps > .14 and post-test scores, ps > .24. Table 1 shows the correlations for pre-test scores, where the only significant correlation was a negative one between Negotiating Likelihood and Uncertainty and Decision making and Problem Solving, r = -.57, p = .005. In addition, Table 2 shows the correlations for post-test scores, where Verbal Reasoning was positively related to Hypothesis Testing, r = -.57, p = .005, and Decision making and Problem Solving was positively related to Argument Analysis. There was also a marginally significant positive correlation between Argument Analysis and Hypothesis Testing, r = .37, p = .09. Overall, the results suggest that the five skills are not redundant with one another. This is because scores on the components are largely uncorrelated, and even when they are correlated, the correlation is modest and not stable across the pre-test and the post-test.

Table 1. Correlations Among Scores on the Five CT Skills in Pre-Test Results

	1	2	3	4	5
1. Hypothesis Testing	---	0.072	0.053	0.167	-0.170
2. Verbal Reasoning	---	---	-0.107	-0.256	0.322
3. Argument Analysis	---	---	---	0.051	0.180
4. Likelihood and Uncertainty	---	---	---	---	-0.574*
5. Decision making and Problem Solving	---	---	---	---	---

*p <.05

Table 2. Correlations Among Scores on the Five CT Skills in Post-Test Results

	1	2	3	4	5
1. Hypothesis Testing	---	0.479*	0.369	0.037	0.225
2. Verbal Reasoning	---	---	0.355	0.204	-0.054
3. Argument Analysis	---	---	---	0.029	0.448*
4. Likelihood and Uncertainty	---	---	---	---	0.261
5. Decision making and Problem solving	---	---	---	---	---

*p <.05

Measuring Student Performance in the HCTA

Paired-samples T-tests were conducted to compare participants' scores on the post-test and pre-test for each CT skill. Participants' scores on Argument Analysis in the post-test (M= 15.86, SD = 1.73) were significantly higher than their scores in the pre-test (M= 14.64, SD = 2.01), t (21) = -3.354, p = .003. Note that the maximum score possible for Argument Analysis is 19. Unexpectedly, there was a marginally significant tendency for participants' scores on Verbal Reasoning in the post-test (M= 4.64, SD = 1.00) to be lower than their scores in the pre-test (M= 5.18, SD = 1.00), t (21), p = .05. Note that the maximum score possible for Verbal Reasoning was 7.

Furthermore, scores across the two test administrations for Thinking as Hypothesis Testing and Argument Analysis were relatively stable, with performance on the pre-test significantly correlated with performance on the post-test (hypothesis testing r=0.611, p=.003; argument analysis r=0.588, p=.004). No significant correlations between pre- and post-test performance were found for the other three CT skills.

GENERAL DISCUSSION

The lack of correlation between the various CT skills measured by the HCTA provides preliminary evidence to support the idea that CT should be treated as a suite of distinct skills that need to be taught independently. Just because a course endeavors to teach CT does not mean it will achieve this learning objective in a broad fashion: instead, learning outcomes for CT instruction should be tailored to specific CT skills that students are expected to develop during a course or an academic career. This is because it is possible to become better at argument analysis or hypothesis testing without necessarily becoming better at decision making or verbal reasoning.

A more direct implication of our findings is that the assessment of learning outcomes for CT instruction would benefit from the consideration of multiple possible outcomes rather than a single outcome. That is, these preliminary results caution against reliance on a single assessment, perhaps out of convenience, and instead, suggests that extra efforts to use a variety of assessments would be worthwhile.

As mentioned earlier, HCTA has been evaluated positively for its construct validity (Butler, 2012), inter-rater reliability (Halpern, 2013), and cross-cultural validity (Butler et al, 2012; Hau et al., 2006; Ku et al., 2010). That said, we would like to clarify that the purpose of this pilot test is not to suggest that the HCTA be used indiscriminately and solely for assessing critical thinking. In fact, before further research efforts are put forth to rely on this instrument heavily in the context of *I&E*, it would be worthwhile to examine whether the underlying latent factors that were identified in previous research are generalizable to the context of *I&E*. In other words, a larger sample could be recruited to provide data for the purpose of exploratory or confirmatory factor analyses (MacCaullum, Widaman, Zhang, & Hong, 1999; see also Fabrigar & Wegener, 2011).

The significant difference between the pre- and post-test results with respect to participants' improvement in argument analysis skills provides preliminary support for our hypothesis that the course boosts CT skills in this area. This is consistent with the efforts to increase CT in *I&E 1* via the course assessments and the course format. That is, in *I&E 1*, the reflective summary assignment, critical comparison assignment, and expository paper assignment all require students to understand, evaluate, and compare argumentation from multiple academic sources. Such skills are also reinforced during seminar style discussion of academic readings. It is worth noting that the improvement in argument analysis over time occurred despite a high correlation between students' pre-test and post-test scores in argument analysis. Thus, taken together, the pre-test and post-test correlation and the pre-test and post-test difference provide preliminary evidence that although students were stable in their argument analysis when compared to their peers, they were still able to show improvement

in their own argument analysis skills over time. In other words, as a group, all students benefitted from the course regardless of whether they were initially lower or higher than their peers in argument analysis. Nonetheless, it would be worthwhile to see if these results could be replicated with a different assessment of argument analysis skills, and with a different I&E 1 course.

The lack of significant results regarding improvement with respect to hypothesis testing, likelihood and uncertainty, and decision making and problem solving could suggest at least two conclusions: either that the *I&E 1* course does not teach these skills (as measured by the HCTA) or that *I&E 1* has failed to teach these skills. If the former were true, then the HCTA would not completely align with course learning objectives and, if the latter were true, then course instruction would need to be revised to emphasize these skills.

Some support for the first conclusion is suggested by the discussion of the course in the introduction. As mentioned earlier, although *I&E 1* and *I&E2* are designed to enhance critical thinking skills, each level emphasizes a specific domain. As a reminder, *I&E 1* teaches expository writing skills while *I&E 2* teaches research based writing skills. Hypothesis testing is one domain that is emphasized in *I&E 2* instead of *I&E 1*. Therefore, it would be useful for future research to examine whether pre-test and post-test scores in hypothesis testing would emerge for students in an *I&E 2* course, even though hypothesis testing did not improve in the current pilot study. Indeed, further studies could be conducted to compare and contrast *I&E 1* and *I&E 2* in the extent to which they enhance the learning of similar or distinct domains of CT.

Managing likelihood and uncertainty is another domain where pre-test and post-test scores did not differ in the current pilot. This is likely because *I&E* modules do not require student engagement with mathematical probabilities or statistical samples, as they are based on a humanistic model of research and argument. Again, further research could examine the extent to which *I&E 1* modules overlap with or are distinct from other courses that teach quantitative reasoning, in order to compare their impact on various domains in CT.

Decision making and problem solving is also a domain where pre-test and post-test scores did not differ in the current pilot. It is noteworthy that the questions assessing this domain were designed to be correlated with all the other domains, as suggested in the test manual: "In some sense, all the subtypes of critical thinking skills are involved in decision making." However, in the current pilot, decision making and problem solving was only positively related to argument analysis in the post-test and surprisingly, negatively related to managing likelihood and uncertainty in the pre-test. This suggests that, at least in the current sample, the questions assessing decision making and problem solving might not capture the construct as intended. Therefore, it would be useful to see if other ways to assess decision making and problem solving might allow

improvements in this CT domain to emerge for *I&E 1* students.

The most troubling results from the pilot study had to do with the statistically significant decrease in verbal reasoning scores from a mean score of 5.18 on the pre-test to a mean score of 4.64 on the post-test. A number of factors could be at play here. It is possible that student performance on the post-test was impacted by the mounting demands and pressures facing students at the end of the semester relative to the start of the semester. Assuming that demands and pressures undermine learning of all CT, however, this possibility is unlikely in the current pilot because students did show an improvement in argument analysis. A second explanation has to do with the nature of the course itself: as a course examining issues of truth and manipulation in photography, students would be expected to engage with photographs and other images as everyday, ordinary texts to unpack and critique as opposed to analyzing ordinary language in speeches, editorials, or other public statements. As a result, the course would be more likely to increase visual reasoning or visual literacy (New London Group, 1996) than verbal reasoning. This explanation, although interesting, would only account for the decrease in verbal reasoning in the current pilot study if visual reasoning and verbal reasoning were somehow negatively correlated such that improvements in visual reasoning lead to decreases in verbal reasoning. All things considered, it is hard to imagine how participation in a college level writing course could decrease verbal reasoning skills, so these results should be taken with a healthy dose of skepticism, and it would be worthwhile to see if they are replicated in future work.

There are some limitations to the study as well that should be taken into account. First, a control condition was not used to compare the results of I&E instruction to instruction in other courses. Future studies should include a control condition to ensure that any gains in CT skills are not attributable to maturation of the sample population or other extraneous variables. Adding a control group would also allow researchers to examine the potential interaction effects, that is, the extent to which the I&E course intervention is particularly effective for students with higher than those with lower pre-existing critical thinking skills, or vice-versa. Such potential interaction effects have implications for identifying subsets of students for which the I&E course intervention is less effective, and generate new questions regarding how to tailor the intervention for these subsets of students. Second, the results gathered only examine one I&E module, and thus the sample for the study is small. Future studies should examine multiple modules offered as part of *I&E* and examine a larger number of students.

In spite of these limitations, the pilot study of the possible effectiveness of *I&E* courses on student CT skills provides some guidance for future research efforts. First, the HCTA can be seen to partially align with course learning outcomes with respect to argument analysis, but it may or may not align in other domains. Prior to any course revision, it would be prudent to see if other CT

instruments—for example, the Ennis-Weir Critical Thinking Test (Ennis and Weir, 1985) or the Critical Thinking Analytic Rubric (Saxton, Belanger, and Becker, 2012)—replicated the results of the pilot study with respect to similar CT skills, or if CT gains were produced with respect to CT skills not measured by the HCTA. Second, the pilot offers preliminary support for the conclusion that teachers of other courses that aim to improve student CT skills should think carefully about which specific skills they seek to improve and how they should tailor classroom assignments to those specific improvements. This is because CT is not a single skill, and it may prove difficult to learn when taught in broad strokes.

REFERENCES

Butler, H. A. (2012). Halpern Critical Thinking Assessment predicts real-world outcomes of critical thinking. *Applied Cognitive Psychology* 26, 721-9.

Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., and Almedia, L. S. (2012). The Halpern Critical Thinking Assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity*, 7, 112-21.

Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership*, 43(2), 44-8.

Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practice* 32, 179-86.

Ennis, R. H. and Weir, E. (1985). *The Ennis-Weir critical thinking essay test*. Pacific Grove, CA: Midwest Publications.

Fabrigar, L. R. and Wegener, D. T. *Exploratory factor analysis*. New York, NY: Oxford University Press, 2011.

Facione, P.A., Sanchez, C.A., Facione, N.C., & Gainen, J. (1990). The disposition toward critical thinking. *The Journal of General Education*, 44 (1), 1-25.

Farris, C. and Smith, R. (1992). "Writing intensive courses: Tools for curricular change" in S. H McLeod and M. Soven (eds.), *Writing across the curriculum: A guide to developing programs* (71-86). Newbury Park, CA: Sage.

Halpern, D. F. (2003). *Thought & Knowledge: An Introduction to Critical Thinking*. Fourth edition. New Jersey: Lawrence Erlbaum Associates.

Halpern, D. F. (2013). *The Halpern Critical Thinking Assessment manual*. Version 22. Moedling, Austria: Schuhfied (Vienna Test System).

Hatcher, D. L. (2006). Stand-alone versus integrated critical thinking courses. *The Journal of General Education*, 55 (3/4), 247-72.

Hau, K., Halpern, D. F., Marin-Burkhart, L., Ho, I. T., Ku, K. Y. L., Chan, N., & Lun, V. M. (2006). *Chinese and United States students' critical thinking: Cross-cultural construct validation of a critical thinking assessment*. Paper presented at the American Educational Research Association Annual Meeting, San Francisco, CA.

Ku, K. Y. L. and Ho, I. T. (2010). Dispositional factors predicting Chinese students' critical thinking performance. *Personality and Individual Differences* 48, 54-8.

MacCaullum, R. C., Widaman, K. F., Zhang, S., and Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 84-99.

New London Group (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review* 66 (1), 60-92.

Saxton, E., Belanger, S., and Becker, W. (2012). The Critical Thinking Analytic Rubric (CTAR): Investigating intra-rater and inter-rater reliability of a scoring mechanism for critical thinking performance assessments. *Assessing Writing* 17, 251-70.

Wardle, E. (2009). "Mutt genres" and the goal of FYC: Can we help students write the genres of the university? *College Composition and Communication*, 60 (4), 765-89.

Appendix 1. Common Learning Outcomes for *I&E 1*.

UTWP modules help students to produce expository writing that readers will recognize as increasing their understanding of a given topic. These modules develop five sets of core strategies that underlie successful scholarly writing in the arts, humanities, social sciences, life sciences, physical sciences, and mathematics:

- Analyzing how authors problematize what other authors say and how they argue their responses to these problems,

- Entering the conversations between these authors by problematizing their arguments and arguing why one's problem and response are reasonable with available evidence,

- Organizing and wording ideas to help readers understand a line of reasoning,

- Documenting sources so readers can check one's use of other scholars' ideas, and

- Revising the content, wording, and organization of a paper, as well as surface features such as spelling, punctuation, etc.

Appendix 2. Common Learning Outcomes for I&E 2.

This module will provide students with the opportunity to learn and apply five core strategies that underlie successful scholarly research and writing:

- assessing the relevance and reliability of multiple sources,
- understanding how theory and method inform and produce knowledge,
- identifying and articulating exigent research questions,
- synthesizing multiple sources to construct and support hypotheses,
- revising one's own thesis, methodological orientation, evidence, and argumentation.

These strategies will be applied as students produce research-based writing.