Research Article

A CONTRASTIVE ANALYSIS OF CONTENT-BASED INSTRUCTION ON ENGLISH FOR STUDENTS OF MATHEMATICS & COMPUTER AND COOPERATIVE ESP ASSESSMENT

*Lydia Leung

Asia University, Tokyo, Japan

Abstract

This applied research report compares and contrasts the suitability of a cooperative ESP assessment at a Japanese university to a content-based instruction course at Tel Aviv University. Five focus areas including development and design, blueprint/specifications, assessment application, wash back on curriculum/summative assessments and evaluation and fine tuning the assessment will be analyzed in further detail to make suggestions and additional improvements with a gauge on the effectiveness on each area. In conclusion, both assessment and curriculum case studies appear to embrace similar philosophies in curriculum. However, upon closer analysis of the assessment instruments used in the assessment case study it would be unsuitable for application in high-stakes test. Furthermore, critical pedagogical issues and cultural factors play a relevant role in determining the lack of fit between these two case studies. It is also much easier to satisfy national and international standards of English at other universities by indicating the results achieved in a standardized test than a cooperative written test low in reliability.

Keywords: ESP curriculum assessment instruments, Japanese university, Tel Aviv University, high-stakes test, critical pedagogy, cultural factors.

INTRODUCTION

The following applied research report will analyze the degree and fit between a curriculum and assessment case study pair. The curriculum case study is a content-based instruction (CBI) course in English for students of mathematics and computer science at Tel Aviv University and the assessment case study focuses on cooperative English for Specific Purpose (ESP) assessment on computers at a Japanese University. The philosophy of curriculum that both case studies appear to embrace is a constructivist approach to assessment in that new knowledge is constructed from a learner's experiences (Dewey, 1966). This is reflected in both case studies where the assessments applied were emphasized on the significance of experience in learning. In the curriculum case study, the course is more process-oriented in that it is "content based, learner-centered, and task-focused" (Kol, 2002). Before the exit exam in this course, learners' needs are considered in that they are given the content materials a week and a half before the end of the semester to study the topics and discuss with their peers, but without the instructor's help and may even refer to these materials and a dictionary during the exam. The rational behind this is that the main focus of content-based courses "is on the learner's interaction with the test, as the learner negotiates the text to construct meaning (Hudson as cited in Kol, 2002) and then communicates the meaning through a task" (Dale & Cuevas as cited in Kol, 2002). As for the assessment case study, an alternative approach to assessment using a cooperative written test is required of students to pass a one-credit ESP class. This involves some collaboration work as students are grouped into three to take turns writing the same test but revising and correcting the answers of their peers. The rationale behind cooperative testing, according to Dewey's principle, is to teach progressively or create a progressive vehicle for assessment (Dewey, cited in Goodman, 2003).

In view of the above statements, the curriculum and assessment case study pair is a suitable fit considering the philosophies of curriculum but in terms of the nature of the tests they lack suitability. One being a low-stakes test focused on writing and the other, a high-stakes test focused on reading.

Setting

The assessment case study focuses on an alternative assessment in a mandatory, low-stakes, one-credit university ESP class in the engineering department with special attention to the topic of computers. This course may not be their primary concern because students take other classes as well that are worth more credits. There are approximately 40 first-year students between the ages of 18-20 with females constituting 20% of the enrollment. Some exceptions that do not fit the profile are returnees from overseas and repeaters who have failed the course. In total, 26 classes of 90 minutes lessons are held in the spring and fall semesters and each semester four written tests are implemented. In the cooperative written test which is based on interdependency of groups, only two-thirds of the exam consists of discrete point items and the rest are written questions and answers. Japanese students are strong on receptive skills which enable them to pass exams that feature discrete lexical items but weak on productive skills. Two grades are assigned including a group grade and an individual grade. The assessment not only emphasizes learning over achievement but also fosters creative and interactive elements. The curriculum case study involves Content-Based Instruction (CBI) in math and computer science at Tel Aviv University in Israel. In order to meet selective entrance requirements, first-year students who fail to pass the exemption-level English course must take one or more courses to meet the requirements by the end of the third semester. Otherwise, they are not permitted to proceed into the fourth semester. The exemption-level courses are non-credit classes about 4 hours a week in a 14 week semester. The course characteristics are content-based, learner-centered and task-focused. It is also based on

***Corresponding Author:** **Lydia Leung**
Asia University, Tokyo, Japan

integrated language skills with a focus on sub skills for reading math texts. Students receive content materials in a week and a half in advance to complete the final exam in this exit-level course.

Connection to the curriculum goals & objectives

Rationale & Appendix A

The most significant goal and objective in the assessment case study is to test whether learning is actually taking place through the feedback from the cooperative tests. This alternative form of assessment can be described as a formative task that results in a summative assessment because a final grade is attached to it. As Morgan *et al.* (2004) states, "the formative assessment tasks are interwoven with the teaching and learning activities and then a summative task is set. When the interweaving of formative assessment tasks towards a summative event is formalized in a course, it may be called scaffolding (p.19). The cooperative assessment allows for several rounds of feedback among peers on the same test that enhances the process of scaffolding as they are allowed to provide, correct or revise the answers on the test to gain an overall summative group grade and individual grade. If fitted to an appropriate set of standards such as the Canadian Language Benchmarks (CLB) (Pawlikowska-Smith, 2000), writing would be the main skill being tested in the cooperative test as it has a written response section that "requires students to demonstrate their understanding of conceptual and factual information by answering questions in complete written sentences" (Goodman, 2003). Students may be at "Stage 3: Advanced Proficiency" Writing Level 7-10. Although, they may have had six year of language education in junior and senior high school, Goodman (2003) states that, "Japanese students are well versed in taking multiple-choice examinations, but they are weak at expressing themselves" (p.54). The cooperative test assessment ties in with the curriculum goal and objective in that the learners are processing and interacting with the concepts taught through the intense recycling of ideas to generate a diverse and thorough understanding of them (Goodman, 2003). Students grapple at answers, interact with their peers and struggle with their thoughts to signify that intense thinking is actively at work (Goodman, 2003).

The most significant goal and objective in the curriculum case study is to build up students' literacy skills that enable them to participate in the academic discourse of mathematics and computer science. The assessment activities used in the general assessment approach could be classified as summative because "at some point, whether students have had several opportunities to rework and improve their performance or whether they have had no opportunities at all, the performance must be graded. The allocation of a final grade on an assessment task – or an entire course – is known as summative, and sometimes terminal, assessment" (Morgan *et al.,* 2004). This includes a unit test on computers and key vocabulary tests for every mathematics and computer text studied. For the math, screen reading and class choice units, a reading comprehension, writing assignment or oral presentation is required. In addition, a final exam on all exit-level courses constitutes the conclusion of the course. Again, if the assessment activities are fitted to the CLB (Pawlikowska-Smith, 2000), they would be at a "Stage 3: Advanced Proficiency" Reading & Writing Level 7-10 being at an academic university. The purpose of the assessments is relevant to the main goal and objective which is to elevate students' literacy skills to pass the exit exam by the third semester.

Statement of Fit

The quality of fit between the curriculum goals and assessment activity choices within the assessment case study is satisfactory. However, if the cooperative written test were to be applied in the context of a high stakes test such as the exit exam in the curriculum case study, it would be a poor fit. The reason for this is due to the fact that cooperative assessment may not be suited to all cultures or high-stakes situations. Individual assessments are demanded when a large population of students must be classified according to level (Goodman, 2003).

Focus on the Assessment Case Study

Development and Design

The reason for the development of an alternative or hybrid method of testing in the assessment case study was due to the lack in usefulness of the test. Discrete-point form items of individual assessment may not be helpful to students who must communicate through writing in academic discourse. Japanese students' learning came out of passing entrance exams by rote memorization instead of viewing learning as a continual process that may lead to autonomous learning. Cooperative learning helps address the many challenges in the course such as comprehension of course content and materials, lack of motivation or interest, insufficient course time and classroom space. Furthermore, cooperative learning "promotes social and educational interaction through the use of positive interdependence" (J&J, 1994, 1999; Kluge, 1994, as cited in Goodman, p.52) whereas paper-and-pencil tests are "purely discrete-point form of individual assessment" (Goodman, 2003, p.51). In the next section, I will use Bachman & Palmer's (1996) first component of test development, the design stage, to closer analyze the details of the cooperative written test.

Purpose of the Test

The purpose of the test is to measure computer concepts and lexical items taught in the engineering course using objective questions and to evaluate students' ability to demonstrate concepts and facts through written communication in complete sentences.

Description of TLU Domain & Task Types

The characteristics of the test task are based on the characteristics of the TLU task of language instructional domains. Since this ESP course is not a primary concern for students, it may be difficult to determine what the appropriate real-life domain would be because students may not use the language outside of class. Bachman & Palmer (1996) also state that students may take the course "for no obvious reason except to fulfill an educational requirement" (p.105).

Characteristics of Test Takers

Personal Characteristics - The test takers are 40 Japanese first-year university students between the ages of 18-20 with about

20% of the students being female. There are also some returnees from overseas and repeaters who make up the learner population. Topical Knowledge – All test takers have highly specific topical knowledge in the engineering field. General Level and Profile of Language Ability - Most test takers are stronger in receptive skills than productive skills despite having received English language instruction for six years in junior and senior high school. Test takers are very familiar with discrete lexical items but are weak at expressing themselves in writing. Potential Affective Responses to Test Tasks – Students feel positive about taking the test since the topics taught in class are reflected in the questions. Furthermore, some test items are recycled from previous tests.

Defining the Construct to be Measured

The construct, broadly defined, is the ability to use language to read and write to understand academic textbooks related to computers in the engineering field. It includes the:

1. Ability to recognize written vocabulary
2. Ability to analyze, discuss and assimilate key concepts
3. Ability to communicate in writing

Plans for Evaluating the Qualities of Usefulness

Reliability – The cooperative test consists of two parts. The first section consists of objective questions such as multiple choice, true or false, word matching, cloze and item sequencing. These responses may be scored objectively with a scoring key. The second section is a written response question type of format. Subjective marking is a more complicated process than objective marking so intra-rater reliability and inter-rater reliability may be used. Intra-rater reliability refers to an examiner who gives the same marks to the written or oral task on two different occasions (Alderson, 1995). Inter-rater reliability is "the degree of similarity between different examiners" or "the agreement among markers" (Alderson, 1995, p.135). There may be variations between examiners and the standard at times but an overall high degree of consistency must be met, in order for the test to be considered reliable (Alderson, 1995). Construct Validity – The cooperative test is considered valid in that it not only measures the knowledge to recognize vocabulary but also metacognitive strategies in the written section to express the key concepts learned. Interactiveness – This alternative form of assessment is highly interactive since students must cooperate with one another to figure out the appropriate answers.

Language knowledge: a specific range of language knowledge is involved

Topical knowledge: significant involvement of topical knowledge is involved

Strategies: test takers will have to demonstrate their metacognitive strategies

Affect: affective responses to the test should be positive since test takers can cooperatively write the test

Impact – The cooperative test has a positive impact on test takers because students feel it is easier than individual tests and it builds close relationships with their classmates. For the instructors, it is considered a timesaving feature as there are fewer papers to grade.

Authenticity – The cooperative test is highly authentic because "all of the test items test class material" (Goodman, 2003, p.53) and items are also recycled from previous tests.

Practicality – This type of test is moderately practical because available resources may not exceed the required resources. The test takes about 30 minutes and is only three pages in length. Different colored pens, a timer and possibly only one classroom are needed to conduct the test. The scoring of the test may require a few more instructors to evaluate the written section.

Inventory of Available Resources and Plan for their Allocation

There is no mention of the costs associated with this test in the case study.

Blueprint/Specifications

Below is a brief summary of the assessment blueprint/specifications selected from Bachman & Palmer's (1996) model:

A. Number of parts/tasks: The purpose of the cooperative written test is organized around two tasks: one is to recognize lexical items and the other to express written communication
B. Salience of parts: the written part requires three different test takers each using different colored pens to answer the questions cooperatively
C. Sequence of parts: students A, B, and C receive pages, 1, 2, and 3 respectively. When the timer stops, A's paper is passed to B, B to C and C to A. The timer is reset for 3-4 minutes and then, the cycle starts again and continues for 30 minutes until the test is finished.
D. Relative importance of parts or tasks: two-thirds is objective questions and one-third is written response questions
E. Number of tasks per part: Five tasks for the objective questions; no mention for the written questions

There are several critiques to this alternative form of assessment in respect to reliability issues, weighting of the test sections and scoring of the test.

Reliability –since students work cooperatively to do the questions, those that are slow at taking tests but may have a good grasp of the concepts taught do not have a chance to demonstrate their learned knowledge if faster test takers have already filled in the answers.

Weighting – two-thirds of the test consists of objective questions which does not allow students to fully demonstrate their writing skills, especially, if a test is written cooperatively.

Scoring – in relation to reliability, the individual scores of the test would be unfair as some may have had a chance to fill in the correct answers but others in the group may have already done it. So, points would be awarded to the faster test taker instead of the slower test taker.

Assessment Application

The two grades are assigned to students: a group and individual grade. The group grade is numeric and dependent

upon the number of correct responses and the individual grade is symbolized by a '+' sign on a holistic contribution which is also norm referenced. A final grade is assigned as A, B, or C. There are two parts to the test including an objective questions section and written response section. A rubric or scoring guide is not provided with the case study so the following would be plausible examples.

In the first section, we could apply a score key to check the objective questions.

Table-1. Score key

Part I:	Objective Questions
Multiple choice	1 point for each correct answer
True or false	1 point for each correct answer
Word matching	1 point for each correct answer
Cloze	1 point for each correct answer
Item sequencing	1 point for each correct answer

In the second section, we could use an analytic scale to evaluate the responses (Alderman, et al, 1995).

Table 2. Analytic scale for evaluation of responses

Part II:	Written Response Questions Points
Relevance and adequacy of content	3
Compositional organization	3
Cohesion	3
Adequacy of Vocabulary for Purpose	3
Grammar	2
Mechanical Accuracy I (Punctuation)	1
Mechanical Accuracy II (Spelling)	1

Then, we could add up the total number of correct responses for Part I & Part II to arrive at a final percentage mark for the group grade. As for the individual grade, a holistic rubric could be used as such:

Table 3. Holistic grading rubric

Individual Grade	Holistic Scale	Points
+	Minimal contribution	10
+ +	Fair contribution	20
+ + +	Maximum contribution	30

The totals on the four tests could be added up differentiating between group and individual contributions. Then, the group and individual grades can also be totaled to reach a final grade for the course resulting in an A, B, or C. In respect to weighting, the objective question section makes up two-thirds of the test but it may not mean that it is given more weighting in terms of grading. It does suggest, however, that discrete lexical items are given more emphasis than the actual ability to analyze, discuss and assemble key concepts pertinent in academic discourse. Although, the case study does not explicitly mention that the written section is given more weighting, the author does emphasize that the "production of written English…is essential for academic students because it promotes thinking" (Krashen, cited in Goodman, 2003) and the "importance of including written response questions on the test cannot be stressed enough with regard to university students" (Leamnson, cited in Goodman, 2003).

Wash back to the Curriculum or Summative Assessments

The formative assessments introduced in the assessment case study is defined as having the opportunity to improve one's performance on the same task (Morgan et al, 2004) which

include elicited formal feedback and observational feedback. The ongoing assessments that take place in formal written feedback evaluates "subjects as class activities, the instructor's teaching style, course materials, tests, other group members, the school, the students' likes and habits with regard to computers, and most important, the students' learning" (Goodman, 2003) several times a year. The more observational feedback is in the form of students' participation in real time and "participation is defined as an interactive exchange of ideas as when volunteering, asking questions, or responding in English during class" (p.57). When students participate, marks are awarded beside their names. These two types of feedback created positive washback in teaching and learning to confirm whether learning is actually taking place. Summative assessment tasks, on the other hand, are not repeated or cannot be improved upon and "forms the basis for a final determination of performance" (Morgan *et al,* 2004). It is also seen negatively by educationalists that devalues personal worth and future prospects (Taras, 2008). Although, the cooperative written test has a final grade attached to it, it still formed positive washback in that it impacted attitudinal and motivational changes. Goodman (2003) states that feedback from students on the cooperative tests indicate that this alternative method is not only easier than individual tests but it also facilitates close friendships among classmates. Furthermore, the cooperative test is designed to alter students' perceptions of tests by providing positive interactive elements to it.

Evaluation of Usefulness and Fine Tuning the Assessment

In the following section, we will use the six qualities of tool usefulness in Bachman & Palmer's (1996) model to reveal several qualities that are dominant in the assessment case study.

Reliability – the test is low in reliability as collaboration is used. Since students work together, they are able to provide hints to certain questions that maybe other students in the group may not have thought of. This test is unable to gauge individual progress clearly even though color-coding is used to differentiate the answers.

Validity –it is high in validity because it tests vocabulary concepts, written academic discourse and cognitive skills as mentioned in the overall curriculum goals.

Interactiveness –the test is low in interactiveness because two-thirds of the test consists of objective questions which highly restricts involvement of language knowledge and metacognitive strategies (p.28)

Impact –the test not only created positive affective response but also solved several problems such as time constraints, physical space and difficult course content that traditional approach to teaching failed to do.

Authenticity – it is moderate in authenticity because only one-third of the test is written response and the rest relies on discrete point items. It is noted that very few language use domains in universities involve this kind of task in academic writing (p.28)

Practicality –it is highly practical because the test solved many problems related to time, human resources and material resources

The qualities that appear to be dominant are validity, impact and practicality. The reason for emphasizing these qualities is in response to the many challenges that traditional teaching approaches have posed such as negative student feedback, insufficient instructor time and difficult content. The whole point of cooperative assessment is for students to engage in learning through process-oriented approaches instead of out of meaningful contexts. The case study had been "fine tuned" according to the feedback of the students. The cooperative written tests led to assigning an individual grade instead of one group grade. A color-coding system to differentiate between the students' answers has also been incorporated because some students expressed dissatisfaction when they did most of the work on the tests.

Conclusion

Linking the Curriculum & Assessment Case-studies

In conclusion, the assessment and curriculum case studies both reflect similar philosophies of curriculum where learning is seen as process-oriented experienced through making meaning of social contexts (Dewey, 1966). However, if cooperative testing were to be applied in the curriculum case study it would be unsuitable for a high stakes test in certain cultures. As Goodman (2003) states:

Although cooperative assessment is a wonderful approach, it may not be suited to all cultures. High-stakes situations, where large numbers of students must be classified according to level, usually demand individual assessment. In such cases, computer scoring of individual assessment instruments may be more appropriate and economical (p.58). In terms of viewing whether these two case studies are a good fit, we also need to look at the cultural factors. In the assessment case study, exams are heavily used in Japan to assess the masses beginning in elementary school to university. But as Goodman (2003) states, "after one has passed the university entrance examinations, the educational environment changes from high-stakes to a low-stakes endeavor. There are no more entrance examinations to cram for because college is the top rung of the educational ladder" (p.49). In comparison to the curriculum case study, students at Tel Aviv University have "a strong instrumental disposition, meaning that they are goal-oriented and are not willing to engage in activities that do not directly promote goal achievement" (Kol, 2002, p.64). Their main goal is to reach exemption-level English by the end of the third semester. Otherwise, they are not permitted to continue on to the fourth semester as this course determines the future paths of students. The cooperative written test would not fit in the context of the Hebrew university because the degree of motivation and attitude towards testing is different as the affective responses to test taking changes in a high and low stakes environment.

In view of critical pedagogy, there is also a lack of fit between these two case studies. The final exam in the curriculum case study is deemed more important than the cooperative written test in the assessment case study as it is used to determine students' futures. Although, the curriculum case study did not explicitly state whether the final exam is a standardized test we can deduce that in high-stakes testing that this is most likely the case. It is much easier to satisfy national and international standards of English at other universities by indicating to them the results achieved in a standardized test than a cooperative written test low in reliability. To reiterate, both case studies appear to embrace similar philosophies of curriculum but if a low-stakes cooperative written test were to be applied to a high-stakes environment where students' futures are under pressure it would be unfit.

REFERENCES

Alderson, J. Charles, Clapham, Caroline, Wall, Dianne. (1995). Language Test Construction and Evaluation. Cambridge: Cambridge University Press.

Bachman, Lyle, F. & Palmer, A. S. (1996). *Language Testing in Practice.* Oxford, Great Britain: Oxford University Press.

Dewey, John. (1966). Democracy and Education. New York: Free Press.

Fulcher, Glen & Davidson, Fred. (2007). *Language testing and assessment: An advanced resource book.* Abington, Oxon: Routledge

Goodman, Mitchell. (2003). Cooperative ESP Assessment at a Japanese University. In Christine A. Coombe & Nancy J. Hubley (Eds.), *Assessment Practices.* (pp. 49-59). TESOL, Inc.

Kol, Sara. (2002). English for Students of Mathematics and Computer Science: A Content-Based Instruction Course. In JoAnn Crandall & Dorit Kaufman (Eds.). *Content-Based Instruction in Higher Education Settings.* (pp.63-77). TESOL, Inc.

Morgan, Chris; Dunn, Lee; Parry, Sharon; & O'Reilly, Meg. (2004). *Issues and themes in assessment.* London: Routledge Farmer

Pawlikowska-Smith, Grazyna. (2000). *Canadian language benchmarks 2000: English as a second language-for adults.* Available November 24, 2009 from the Centre for Canadian Language Benchmarks, http://www.language.ca/ display_page.asp?page_id=254

Taras, Maddalena. (2008) Summative & formative assessment: Perceptions and realities. *Active Learning in Higher Education*, 9, 72-192.

****

Appendix A

Comparison/Contrast Overview Chart of the Curriculum Goals/Objectives	
Cooperative ESP Assessment	**Math & Computer Science CBI**
Goodman, 2003	Kol, 2002
• "overall course aim is to broaden the students' technical English vocabulary and develop language skills by analyzing, discussing, and assimilating concepts related to the computer through the use of the course materials" (p.51)	• For students to achieve exemption level English by the end of the third semester of study (p.63)
• "The goal of all of these cooperative activities is to recycle and process ideas more intensely to generate a diverse, well-rounded understanding of them" and "to give the students an exciting educational experience to remember beyond the straightforward course content" (p.52)	• To develop literacy skills associated with academic discourse of mathematics and computer science (p.63)
• "One of the goals of the cooperative test is to create an atmosphere in which learners can be naturally engaged and less conscious of the formal nature of the event" (p.53)	• Content-based, learner-centered, task-based course (p.69)
• "One purpose of the cooperative tests is to get feedback to see if learning is taking place. In this sense, all forms of assessment are a kind of feedback for teachers" (p.56)	• Reading skills in math and computer units, screen reading of electronic texts in other units; the type of reading used is compared
• provide academic interaction and peer correction (p.54)	• Vocabulary is key component and is reviewed and tested; useful for effective reading (p.72)
	• "Computer unit closes with a test involving the retrieval of information in the articles read in that unit and integration of the known information with information from a new text" (p.73)
	• "In the math unit, students read a new mathematics text and answer reading comprehension questions about it" (p.73)
	• "The screen reading unit and the fourth unit may conclude with either a writing assignment or an oral presentation" (p.73)
	• "Students need to implement the higher level reading skills of text comparison, integration of information from a number of sources, and synthesis of new knowledge" (p.73)

****