A Study of Task Type for L2 Speaking Assessment

Teng, Huei-Chun
National Yunlin University of Science & Technology

The purpose of the present study is to investigate the effect of task type on the performance of EFL speaking tests for Taiwanese college students. The major research questions explored in the study include: (1) Will test takers perform differently on various task types of EFL speaking tests? (2) Are there any differences in the accuracy, complexity, and fluency of test takers' discourse in terms of different task types? (3) What are test takers' perceptions toward the three speaking tasks? Subjects in the study were 30 students of English major at a university in Taiwan. The three task types adopted in the study consisted of answering questions, picture description, and presentation. The subjects were tested in a language-lab setting and responded on an audiotape. After completing the speaking test, subjects answered a questionnaire designed to elicit their affective reactions toward the three tasks. The tapes were scored independently by two English teachers of native speaker. The taped protocols were also transcribed for the analysis of accuracy, complexity, and fluency. Results of the study can provide empirical evidences for the effects of L2 speaking assessment tasks. Results are also expected to offer some implications for designing EFL speaking tests.

Introduction

With the prevalence of Communicative Language Teaching (CLT), a considerable amount of the teaching and learning of a second language (L2) today is done orally. Consequently, developing speaking proficiency rates high among the objectives of most L2 programs. As pointed out by Shohamy, Reves, and Bejerano (1986), the earlier tests of oral proficiency can be termed 'precommunicative' since the speaking tasks the test-takers were required to perform were mostly mechanical repetition of words and sentences, the supplying of pattern answers to pattern questions, and substitution drills. However, these tests were viewed as unauthentic by language teachers and testers with the growing emphasis on CLT. As a result, direct tests of speaking proficiency have been developed by involving a test setting in which the examinee and one or more human interlocutors engage in communicative oral interactions. (Clark, 1975). Yet, according to Shohamy (1994) a number of variables in direct speaking tests tend to affect test-takers' scores, including the role relationship, personality and grades of testers and respondents, the purpose of the interaction, the topic, and the setting. Therefore, there is a need to control those variables by conducting oral tests in a more uniform way.

Semi-direct oral tests were developed to ensure reliability and validity without compromising the communicative features of oral tests. In these tests, test-takers respond to authentic recorded and visual tasks which require the production of discursive reactions. The oral tests are uniform tests because all test-takers perform similar language tasks. On the

other hand, they involve a variety of communicative characteristics as they elicit a wide range of oral interactions and discourse strategies.

For the past decades, a great deal of attention has been devoted to the development of tests of oral language proficiency for use with foreign language learners. However, compared with paper and pencil testing, the field has been largely neglected. Due to the practicability of oral testing (Cohen, 1980). As a result, many problems remain to be examined, such as the subjectivity of the rating process, the noninterval nature of the scales adopted for rating, the absence of high demonstrated validity across a variety of instruments and language abilities, and a paucity of testing methods beyond the oral interview (Henning, 1987).

The purpose of the present study is to investigate the effect of task type on the performance of EFL speaking test for Taiwanese college students. The major research questions explored in the study will be: (1) Will test takers perform differently on various task types of EFL speaking test? (2) Are there any differences in the accuracy, complexity, and fluency of test takers' discourse in terms of different task types? (3) What are test takers' perceptions toward the three speaking tasks?

A major goal of foreign language learning is to acquire oral facility in the target language. Although a great deal of attention has been devoted to the assessment of L2 oral proficiency, scant efforts has been paid to developing valid and reliable oral testing methods (Robinson, 1992). According to Skehan & Foster (1999), one area for language testing research seems very promising is to see whether task characteristics have interesting effects on the nature of speaking performance. It is important to conduct research on task types, and to explore the predictability of the language characteristics associated with such tasks. In the last few years, only some studies have looked into the impact of task type on L2 speaking assessment. Among them, very few have dealt with the implementation of EFL speaking tests to Taiwanese students. Thus, by providing empirical evidences and descriptions of speaking assessment tasks, the present study will seek to contribute to our understanding of L2 speech performance, and further to offer implications for designing EFL speaking tests.

Literature Review

As indicated by Bachman & Palmer (1981), one of the areas of most persistent difficulty in language testing continues to be the measurement of oral proficiency. From the review of research literature, a number of studies have been conducted on the validation of oral tests. For example, Dandonoli & Henning (1990) examined the construct validity of the ACTFL Proficiency Guidelines and oral interview procedures. The results provided strong support for the use of the Guidelines as a foundation for the reliability and validity of the Oral Proficiency Interview (OPI). Stansfield & Kenyon (1992) conducted a study to develop and validate a simulated oral proficiency interview (SOPI) as an alternative method to the face-to-face procedure employed by OPI. Moreover, Shohamy (1994) examined the validity of direct versus semi-direct oral tests. Results showed that concurrent validity of the two types of tests was high, yet the two tests still differed in a number of aspects, such as the elicitation tasks and the language samples obtained. A study by O'Sullivan, Weir, & Saville (2002) addressed the relatively neglected area of validating the match between intended and actual test-taker language with respect to the language functions representing the construct of

spoken language ability.

One of the main problems associated with oral tests is that they are subjective in nature and that there are no clear criteria for correctness. Some researchers on second language testing have looked into the issue of oral test rating. Shohamy (1983) examined inter-and intra-rater reliability of the oral interview test. She suggested that speaking tests such as the Oral Interview can be used reliably by decision-makers in spite of their subjective nature. Besides, Chalhoub-Deville's study (1995) contended that researchers might need to reconsider employing generic component scales. She recommended a research approach that derives scales empirically according to the given tests and audiences, and the purpose of assessment. Halleck (1995) also investigated the relationship between holistic and objective measures in the OPIs of 107 EFL students in China. Results indicated significant main effects for proficiency level and interview task, and provided some support for the holistic rating system put forth in the ACTFL proficiency guidelines. Furthermore, Kenyon & Tschirner (2000) compared test reliabilities for the German Speaking Test, a semi-direct tape-mediated oral proficiency test, and the ACTFL OPI. Results revealed a high score equivalency between ACTFL proficiency ratings obtained on both tests. In O'Loughlin's study (2002), eight female and eight male test-takers undertook a practice IELTS interview on two different occasions, once with a female interviewer and once with a male interviewer. Results showed that gender did not have a significant impact on the IELTS interview.

In addition, several studies were found to be related to the purpose of the present research, i.e., to examine the effects of task type for oral assessment. First, in Henning's study (1983) the three oral testing methodologies of imitation, completion, and interview were compared for reliability and validity by employing an initial sample of 143 adult Egyptian EFL learners. He found that the pronunciation component of the imitation method exhibited highest overall validity across all indexes. Comparison of the three oral testing methods showed the ranking order in terms of available validity indexes, i.e., (1) imitation, (2) interview, and (3) completion. Carpenter, Fujii, & Kataoka (1995) designed a new oral interview procedure for eliciting a representative sample of spontaneous Japanese language abilities from children aged 5-10. The test included six subtests and made use of realia, role playing, information gap activities and naturalistic conversation, all designed to comprise an oral interview. Results showed that the procedure elicits a language sample that is superior in quality and quantity to other existing Japanese oral test instruments for children. Moreover, Foster & Skehan (1996) investigated the effects of planning time and three different tasks (personal information exchange, narrative, and decision-making) on the variables of fluency, complexity, and accuracy. Interactions were found between task types and planning conditions, such that planning had more influence on narrative and decision-making tasks than on personal information exchange task. Skehan & Foster (1999) also explored the effects of inherent task structure and processing load on the performance on a narrative retelling task. They suggested that more structured tasks generated more fluent language, and complexity of language was influenced by processing load. A study by Jeng *et al* (2000) used experimental design methods to compare three tasks of oral assessment. Results show that individual interviews took more time and effort, but were perceived to have higher value largely due to its interactive features between examinees and examiners. There were more problems with

paired discourse and taped recording methods. Besides, Wu, R. (2002) investigated the effects on task difficulty of performance conditions associated with the code complexity of written input in the read-aloud tasks of a semi-direct speaking test.

Finally, there are a number of studies which can provide useful information for the current research. Some studies have looked into the affective reactions to speaking tests (e.g., Scott, 1986; Orr, 2002). Several researchers have analyzed the discourse in speaking test performance, such as Gelderen (1994), Douglas (1994), and O'Loughlin (1995). A few studies have been conducted to examine the influence of planning time (e.g., Mehnert, 1998; Ortega, 1999). Teng (2002) and Wu, H. (2002) have also studied the implementation of EFL speaking tests to Taiwanese students.

Method

Subjects

Subjects in the current study were 30 students at a university in Taiwan. They studied at the Department of Applied Foreign Languages. They had approximately a high-intermediate level of EFL proficiency.

Instrumentation

The instruments used in the present study consisted of an EFL speaking test and an affective questionnaire. The test was a semi-direct speaking test with Chinese instructions printed in the test booklet and recorded on the audiotape. There were three task types adopted in the speaking test, including answering questions, picture description, and presentation. In the first task, the test taker was required to respond to three questions recorded on the test tape, each question being heard once. The test taker was given 30 seconds to answer each of the questions. In the second task, the test taker studied a picture accompanied by three guided questions written in Chinese. The test taker was given 30 seconds to look over the picture and questions, and the given 90 seconds to complete a description of the picture. In the third task, the test taker read the statement printed on the test paper. The test taker was given 90 seconds to think about what he/she planned to say about the statement, and then given 90 seconds to make a presentation on the statement.

The second instrument adopted in the study was an affective questionnaire, which was mainly based on Scott's (1986) work. The questionnaire was designed to elicit test takers' affective reactions toward the speaking test and the three assessment tasks, ie., answering questions, picture description, and presentation. The questionnaire included four parts and 35 questions in total.

Procedures

Before the experiment begins, subjects were be told in detail what they were required to do in the study. In order to counterbalance the practice effect of task type, the 30 subjects were randomly assigned to three groups with different presentation order of the three speaking tasks. Each of the three subject groups were tested in a language-lab setting and responded in an audiotape. It took about 10 minutes for the subjects to complete the speaking test. Then subjects answered the affect questionnaire.

Data Analysis

Two English teachers of native speakers, who are both trained raters, independently assessed each subject's answer tape and assigned a score based on Shohamy's (1985) holistic rating scale for speaking test (see Table 1). The computed interrater reliability was 0.76. Besides, the present study adopted the analytic approach to analyze subjects' performance data. The recorded speech samples were transcribed and coded to measure the accuracy, complexity, and fluency of subjects' performance. Accuracy was measured by calculating the number of error-free clauses as a percentage of the total number of clauses (Skehan & Foster, 1999). Complexity was indexed by dividing the number of clauses by the number of c-units (communication units). Accordning to Foster & Tonkyn (1997), c-unit is defined as a simple clause, or an independent subclausal unit, together with the subordinated clauses associated with them. Fluency was measured by dividing the number of syllables in a given speech sample by the time taken to produced them (measured in seconds) and multiplying the result by 60 (Mehnert, 1998). The statistical procedure, ANOVA, was conducted to test the hypotheses concerning the research questions.

Table 1. Holistic Rating Scale for Speaking Test (Shohamy, 1985)

Rating	Interpretation
1	Unintelligible No lg. Produced No interaction possible
2	Hardly intelligible Very poor lg. produced Only simplest, fragmentary interaction possible
3	Clearly intelligible Simple lg. produced Interaction possible Not articulate
4	Responsive in interaction Slightly more sophisticated language produced Consistent errors: but do not interfere with fluency Strong MT interference (translated patterns, etc.)
5	Almost effortless in expression Adequate in interaction Errors: NOT consistent
6	Facility of expression Comfortable, initiating in interaction Sporadic mistakes
7	No limitation whatsoever Near-native

Results

Subjects' Performances on Speaking Tasks

The main intent of the present study is to empirically investigate the effect of task type on the performances of EFL speaking tests for Taiwanese college students. Based on the research purpose, subjects' performances were analyzed in terms of rating, accuracy, complexity, and fluency. Table 2 demonstrates the descriptive statistics of subjects' speaking test performances. In terms of rating assessed by two raters on a 7-point scale, subjects got the highest average score (M = 4.19) for the task of presentation, followed by answering questions (M = 3.94), and then picture description (M = 3.81). Besides, three analytic scoring methods were adopted to analyze subjects' performance. With regard to accuracy measured by calculating the number of error-free clauses as a percentage of the total number of clauses, subjects got the highest score (M = 0.78) for the task of answering questions, followed by presentation (M = 0.74) and then picture description (M = 0.66). As for complexity indexed by dividing the number of clauses by the number of c-units, subjects got the highest score (M = 1.73) for the task of answering questions, followed by presentation (M = 1.66) and then picture description (M = 1.32). In regard to fluency measured by dividing the number of syllables by the seconds to produce them, subjects got the highest score (M = 2.15) for answering questions, followed by presentation (M = 1.75) and then picture description (M = 1.49).

Table 2. Descriptive Statistics of Subjects' Performance

Performance / Task	N	Rating		Accuracy		Complexity		Fluency	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Answering Questions	30	3.94	1.24	0.78	0.20	1.73	0.54	2.15	0.41
Picture Description	30	3.81	0.98	0.66	0.24	1.32	0.30	1.49	0.24
Presentation	30	4.19	1.33	0.74	0.16	1.66	0.42	1.75	0.39

To determine if there were any significant differences in subjects' speaking test performance due to the effect of task type, a one-way ANOVA on the four dependent variables was conducted respectively. Results in Table 3 show that there are significant main effects for the two variables, i.e., complexity (F = 3.286, p = 0.023) and fluency (F = 14.140, p = 0.000).

Table 3. ANOVA of Subjects' Performance

SV	Variable	SS	df	MS	Error	F	p-value
Task	Rating	1.167	2	0.583	63.812	0.411	0.665
	Accuracy	0.118	2	0.014	1.795	1.481	0.238
	Complexity	1.317	2	0.66	8.411	3.286*	0.023
	Fluency	3.551	2	1.776	5.650	14.140**	0.000

* $p<0.05$ ** $p<0.01$

With significant main effects for complexity and fluency, to further investigate the difference among the three task types of speaking test, post-hoc tests with Tukey's procedure were conducted to make pairwise comparisons of group means. As shown in Table 4, subjects got significantly higher complexity scores for answering questions than for picture description ($p = 0.028$). As for the performance on fluency, subjects scored significantly higher for answering questions than for the other two task types ($p = 0.000$, $p = 0.008$).

Table 4. Post Hoc Test of Subjects' Performance

Performance	Task Comparison	Mean Difference	SE	p-value
Complexity	Answering Questions vs. Picture Description	0.407*	0.153	0.028
	Answering Questions vs. Presentation	0.066	0.153	0.904
	Presentation vs. Picture Description	0.341	0.153	0.090
Fluency	Answering Questions vs. Picture Description	0.662**	0.125	0.000
	Answering Questions vs. Presentation	0.397**	0.125	0.008
	Presentation vs. Picture Description	0.265	0.125	0.098

* $p<0.05$ ** $p<0.01$

Subjects' Perceptions of the Speaking Test

Table 5. Subjects' Perceptions of Tasks

Statement \ Task	Answering Questions	Picture Description	Presentation
I think this task can assess my speaking ability accurately.	3.82	3.65	3.81
I feel nervous before the task.	4.00	3.71	3.88
I feel nervous during the task.	3.88	3.76	3.93
I think I did well on the task.	2.29	2.41	2.56
I think the task should be included in the speaking test.	3.82	4.06	4.00
I think I had an adequate opportunity to demonstrate my ability to speak English with the task.	3.12	3.53	3.62
I think the task was too short.	2.76	2.94	2.75
I prefer the task to others.	3.06	3.53	2.88
I understand what I was supposed to do during the task.	3.29	3.65	3.38
I think the task corresponds to what I learn in class.	3.53	3.18	3.38
I think the task is too difficult.	2.71	3.06	3.5

In the present study, a questionnaire was designed to elicit subjects' affective reactions toward the speaking test and the three assessment tasks. Subjects were required to indicate their agreement on a 5-point scale. As shown in Table 5, subjects had higher agreement for the task of answering questions on the three statements, i.e., I feel nervous before the task ($M = 4.00$), I think this task can assess my speaking ability accurately ($M = 3.82$), I think the task corresponds to what I learn in class ($M = 3.53$). In terms of picture description, subjects

had higher agreement on the three statements, i.e., I think the task should be included in the speaking test (M = 4.06), I understand what I was supposed to do during the task (M = 3.65), I prefer the task to others (M = 3.53). With regard to the task of presentation, subjects had higher agreement on the three statements, i.e., I feel nervous during the task (M = 3.93), I think I had an adequate opportunity to demonstrate my ability to speak English with the task (M = 3.62), I think the task is too difficult (M = 3.5). As for subjects' perceptions of the whole test (see Table 6), they had higher agreement on the statement, i.e., I think the oral test should include more tasks (M = 3.75).

Table 6. Subjects' Perceptions of the Test

Statement	N	Mean	SD
I would rather take a written test than an oral test.	30	2.94	1.06
If I take the same test on another day, the result will be the same.	30	2.19	1.05
I would like my English teacher to be present during the test.	30	2.56	0.96
I feel more comfortable when I take an oral test by talking to a real person.	30	2.94	1.29
I think the oral test should include more tasks.	30	3.75	0.77

Discussion

In the current research, results indicated that there was no significant difference in the subjects' holistic rating scores for the three task types, including answering questions, picture description, and presentation. That is, test takers did not perform differently on various task types of EFL speaking test. However, significant main effects were found for task type on the two analytic measures, i.e., complexity and fluency. The findings may be explained by the difference in scoring methods. Although the holistic rating in the present study was conducted by two raters based on the holistic rating scale for speaking test (Shohamy, 1985), the rating itself is still mostly subjective due to raters' intuition and general impression. As a result, holistic rating did not seem to be so sensitive to different task types as the three analytic measures which are more objective by formula calculation.

Furthermore, results of the present study showed that there were significant differences in the complexity and fluency of test takers' discourse in terms of different task types. Post hoc analyses revealed that subjects performed better in complexity for the task of answering questions than for that of picture description. According to Skehan & Foster (1999), complexity of language was influenced by processing load. They suggested that complexity was mainly affected by the conditions under which tasks were done, especially related to the processing demands that they entailed. In the current study, there seemed to be more processing load for answering questions since test takers were expected to provide answers directly related to the questions. As for picture description, less processing load was involved because of the flexible nature of description.

Besides, the current study found that subjects got higher scores in fluency when they took the task of answering questions than the scores for the other two tasks. Several possible explanations may be proposed for subjects' better performance in answering questions in terms of fluency. One possibility is that answering questions seems to be the most common

task of EFL speaking tests and activities for college students in Taiwan among the three task types. There are more opportunities for subjects to answer questions in English than to describe pictures or make presentations in English. Such a practice effect has also been indicated by Halleck (1995) when she compared different oral tasks. Another possible explanation is that more structured tasks may generate more fluent language. Skehan & Foster (1999) proposed that tasks containing clearer inherent sequential structure would lead to more fluent performance than tasks not structured in this way. In the present research, the task of answering questions could offer more structured framework for test takers to exhibit their oral proficiency than picture description or presentation.

Finally, results of the affective questionnaire demonstrated that subjects averagely felt more nervous before the task of answering questions. This finding supports Tarone & Parrish's (1988) claim that task-related variability in interlanguage is caused by different degrees of communicative pressure upon the speaker. For test takers, answering questions appeared to be a kind of semi-interview, and they might feel like communicating with real people. That is why answering questions was perceived to be more stressful than the other two tasks. Subjects' perceptions of tasks also revealed that they preferred the task of picture description to others and that they thought the task should be included in all speaking tests. It may be implied that test takers seem to be more interested in the oral task of picture description because of the visual cues provided by the task.

Conclusion

The present research proposes that Taiwanese college students performed better in the EFL speaking task of answering questions by exhibiting more fluency and complexity. It is important to see whether task characteristics have interesting effects on the nature of speaking performance. By providing empirical evidences and descriptions of speaking assessment tasks, the study can contribute to our understanding of L2 speech performance, and further to offer implications for designing EFL speaking tests.

REFERENCES

Bachman, L., & Palmer, A. (1981). The construct validation of the FSI oral interview. *Language Learning*, 31, 67-86.

Carpenter, K., Fujii, N., & Kataoka, H. (1995). An oral interview procedure for assessing second language abilities in children. *Language Testing*, 12, 157-175.

Chalhoub-Deville, M. (1995). Deriving oral assessment scales across different tests and rater groups. *Language Testing*, 12,16-33.

Clark, J. L. D. (1975). Theoretical and technical considerations in oral proficiency testing. In R. L. Jones & B. Spolsky (Eds.), *Testing language proficiency* (pp. 10-28). Arlington, VA: Center for Applied Linguistics.

Cohen, A. (1980). *Testing language ability in the classroom*. Rowley, MA: Newbury House.

Dandonoli, P., & Henning, G. (1990). An investigation of the construct validity of the ACTFL proficiency guidelines and oral interview procedure. *Foreign Language Annals,* 23, 11-22.

Douglas, D. (1994) Quantity and quality in speaking test performance. *Language Testing*, 11, 125-143.

Foster, P. & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299-323.

Foster, P., Tonkyn, A. (1997). *Measuring spoken language: A unit for all reasons*. Paper presented at the Joint CALR/CALS Symposium, University of Reading, UK.

Gelderen, A. (1994). Prediction of global ratings of fluency and delivery in narrative discourse by linguistic and phonetic measures—oral performances of students aged 11-12 years. *Language Testing*, 11, 291-319.

Halleck, G. (1995). Assessing oral proficiency: A comparison of holistic and objective measures. *Modern Language Journal*, 79, 223-34.

Henning, G. (1983). Oral proficiency testing: Comparative validities of interview, imitation, and completion methods. *Language Learning*, 33, 315-332.

Henning, G. (1987). *A guide to language testing: Development, evaluation, research*. Boston, MA: Heinle & Heinle.

Jeng, S., Ai, L., Sung, C., Iuan, C., & He, C. (2000). A comparative study of EFL speaking assessment tasks. *Selected Papers from the Ninth International Symposium on English Teaching*, 661-667.

Kenyon, D. (2000). The rating of direct and semi-direct oral proficiency interviews: comparing performance at lower proficiency levels. *The Modern Language Journal,* 84, 85-101.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 83-108.

O'Loughlin, K. (1995). Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test. *Language Testing*, 12, 217-37.

O'Loughlin, K. (2002). The impact of gender in oral proficiency testing. *Language Testing,* 19, 169-192.

Orr, M. (2002). The FCE speaking test: using rater reports to help interpret test scores. *System*, 30, 143-154.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 20, 109-48.

O'Sullivan, B., Weir, C., & Saville, N. (2002). Using observation checklists to validate speaking-test tasks. *Language Testing,* 19, 33-56.

Robison, R. (1992). Developing practical speaking tests for the foreign language classroom: a small group approach. *Foreign Language Annals,* 25, 487-496.

Scott, M. L. (1986). Student affective reactions to oral language tests. *Language Testing*, 3, 99-118.

Shohamy, E. (1985). *A practical handbook in language testing for the second language teacher*. Israel: Tel-Aviv University.

Shohamy, E. (1983). Rater reliability of the oral interview speaking test. *Foreign Language Annals,* 16, 219-222.

Shohamy, E. (1994). The validity of direct versus semi-direct oral test. *Language Testing*, 11, 99-123.

Shohamy, E., Reves, T., & Bejerano, Y. (1986). Introducing a new comprehensive test of oral proficiency. *English Language Teaching Journal*, 40, 212-220.

Skehan, P. & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49, 93-120.

Standfield, C. & Kenyon, D. (1992). The development and validation of a simulated oral proficiency interview. *The Modern Language Journal*, 76, 129-141.

Tarone, E. & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning*, 38, 1-44.

Teng, H. (2002). A study of the procedures for conducting EFL oral tests. *Newsletter for Teaching the Humanities and Social Sciences*, 12(5), 35-41.

Wu, H. L. (2002). Assessing knowledge of the English discourse structure. *Selected Papers from the Eleven International Symposium on English Teaching/Fourth Pan-Asian Conference*, 572-81.

Wu, J. (2002). Task difficulty in semi-direct speaking tests code complexity. *Selected Papers from the Eleven International Symposium on English Teaching/Fourth Pan-Asian Conference*, 244-54.