

Evaluation in College English Language Education: A Consideration of Validity

Oka, Hideo

English Language Education, College of General Education, Kyushu University : Assistant Professor

<https://doi.org/10.15017/6796039>

出版情報 : 言語科学. 13, pp.1-12, 1977-03. The Group of Linguistic Studies College of General Education, Kyushu University

バージョン :

権利関係 :



Evaluation in College English Language Education

—A Consideration of Validity—

Hideo Oka

I. Introduction¹⁾

The term evaluation in its broad sense extends to the evaluation of a whole curriculum, or the evaluation of teaching methods and materials. However, in this article the term will be defined as the evaluation of learning achievement and of language proficiency—that is, the testing of foreign language competence and skill. When evaluation thus defined is applied to a particular stage of learning, such as the liberal arts course of the university, it must be viewed in the whole process of teaching and learning.

From the standpoint of a scientific approach to English language education, evaluation exists in the systematic sequence of the theoretical study of the subject: aims are carefully analysed to bring forth objectives; learner's factors are closely examined; and finally, teaching materials and methodology are established according to the foregoing categories.²⁾ Therefore, to avoid some likely pitfalls in the treatment of evaluation in an isolated manner, we must instead deal with evaluation in due relation to the objectives of the course, the materials and methods used, as well as to the fundamental issues of evaluation itself: otherwise, it would not prove to be sufficiently systematic.

English language education in the liberal arts course does not have any guidelines equivalent to the syllabuses at junior and senior high schools, so that the contents and objectives remain ambiguous. At the same time, it differs not only in its contents from the teachers' training course or literature course, but also in the sense that it is given to all the students irrespective of their majors. This characteristic of general education, together with the lack of specification, has brought about a situation where everything is totally dependent on the experience and subjectivity of the individual

¹⁾ This article is partly based on the oral presentation of a paper under the title of 「大学一般教育課程における英語評価の諸問題」 (“Various Problems of English Testing in the Liberal Arts Course of the University”) at the annual conference of JACET (The Japan Association of College English Teachers) held at Kansai University on October 10, 1976.

²⁾ See B. H. Banathy, “Current Trends in College Curriculum: A Systems Approach,” in E. M. Birkmaier (Ed.), *Britannica Review of Foreign Language Education*. Vol. 1, Encyclopaedia Britannica, Inc., 1968, p. 106. He illustrates in a diagram the systematic procedures of college curriculum development based on the concept of “systems approach.”

teacher, be it the selection of textbooks or the adoption of methods. This results in a certain ambiguity, that is, whatever he does may always be partially right, but at the same time may be unsatisfactory in other respects. This paper will attempt to clarify the complexities involved in evaluation where no systematization of English language education as a whole is established. For this purpose the procedure taken here is, first, to grasp objectively the current testing methods adopted by individual teachers, according to a survey conducted by the author; secondly, to examine closely the fundamental issue of evaluation on the basis of the results obtained therein, i.e. "validity" of evaluation in relation to overall language proficiency; and finally, to investigate how to guarantee validity in reference to course objectives, through which a possible organization of systematic evaluation in a future curriculum will be suggested.

II. Current Practice

A survey was conducted by the author to obtain information which might reflect the thinking and practice of college English teachers with respect to testing and evaluation.³⁾ The following diagram shows the outline of the results obtained from the survey.

1. Which do you usually test, achievement as regards the covered materials or general proficiency of English?

Achievement	12 (63 %)
Proficiency	0 (0 %)
Combined	7 (37 %)
2. Which do you usually adopt, subjective or objective tests?

Subjective tests	4 (21 %)
Objective tests	1 (5 %)
Combined	14 (74 %)
3. Which types of questions do you give in examinations? (3 points for those always given, 2 for those mostly given, 1 for those sometimes given, and 0 for those never given)

Translation into Japanese	39 (17 %)
Translation into English	25 (11 %)
Explanation or discussion in Japanese	19 (8 %)
Free composition or discussion in English	17 (7 %)
Multiple choice	14 (6 %)
True or false	10 (4 %)
Insertion of suitable words or phrases	15 (7 %)
Vocabulary	9 (4 %)
Pronunciation	19 (8 %)
Hearing	24 (10 %)

³⁾ The survey was conducted on evaluation and the selection of textbooks in July, 1976 at the College of General Education, Kyushu University. The questionnaires were sent to 47 full-time and part-time teachers, including 4 foreign teachers who were in charge of conversation classes, and 19 answers were returned.

Dictation	27 (12 %)
Speaking	12 (5 %)
4. Which do you usually adopt, absolute or relative evaluation?	
Absolute evaluation	7 (37 %)
Relative evaluation	1 (5 %)
Combined	11 (58 %)

Major results summarized from the survey are as follows:

- (1) As regards the content of evaluation, a majority of the teachers base their examinations on achievement alone, and some on a combination of achievement and proficiency.
- (2) As for the form of evaluation, the combined form of subjective and objective testing is dominant.
- (3) As for the actual types of questions, subjective questions, especially translation from English into Japanese, are favored. The fourth item in this category mainly applies to the so-called "English C," which stresses composition, and the last four items are chiefly related to the LL and conversation classes. It must be noted, however, that the points and percentages do not necessarily correspond to the actual number of questions given in examinations, but only indicate the relative emphasis of each item. In the case of translation into Japanese, for example, the total points of 39 (17 %) do not mean 17 points out of 100 as a full mark—it can easily be assumed that the rate is much higher—, but only illustrate that most teachers almost always give such questions.
- (4) As regards the interpretation of the testing results, a majority of the teachers adopt a combined method of absolute and relative evaluation, with some relying on a single norm of absolute evaluation.

The above-analyzed results will be regarded as indicative of present practice in evaluation. There appears to be a vague consensus of testing methods among the teachers. It must be admitted, however, that the results do not reveal anything more than the general framework of evaluation. What is vitally significant and can be considered the central issue of evaluation is not the formal aspect of subjective vs. objective testing, nor the absolute vs. relative treatment of testing results, but primarily the fundamental question of "validity" in evaluation⁴⁾: On what criteria does one determine that the learning objective has been achieved? On what ground does one determine that the learner possesses the required competences and skills in the foreign language? To exemplify the consideration of validity, translation into Japanese, which proved to be most widely adopted, is by itself far from evaluating overall language proficiency, and even if the course objective is the training of reading

⁴⁾ See B. Spolsky, "Language Testing: The Problem of Validation," in K. Croft (Ed.), *Readings on English as a Second Language*. Winthrop, 1972, p. 348. He concludes that validity, the central problem of foreign-language testing, remains a serious problem.

skill, the question still arises whether the skill of translating can be equated with reading skill itself.⁵⁾

In comparison with the importance of validity, the controversy between subjective and objective testing or absolute and relative evaluation becomes secondary. It is generally believed that, since in subjective tests it is easier to prepare questions directly suited for evaluating the objectives of the course, they are superior in validity to objective tests. However, if the uniform entrance examination to national universities (totally objective), which is now in the experimental stage, is compared with an examination comprising merely translation, such generalization cannot be applied. On the other hand, if there is any doubt about validity as to the content of testing itself, however ingeniously the testing results may be interpreted afterwards, the contrast between absolute and relative evaluation will become almost meaningless.

III. Validity from the Viewpoint of Communicative Competence

Validity, which may be defined as the relevance of an examination to the analysis of language skills and syllabuses, is not a single entity, but is subdivided into several components—face, content, construct, predicative, concurrent, etc.⁶⁾ Among them content and construct validity are most closely related to the evaluation of language proficiency. Content validity is concerned with how well the content of an examination samples the subject matter which it is supposed to measure. For example, in a test of reading skill, if the examinee can answer without reading the passage, or if he cannot answer just because he does not know a certain word, the test cannot be considered to possess content validity. On the other hand, construct validity is related to the quality an examination is actually measuring. In a speaking test, for instance, if the examiner checks only the ability in phonology without taking the overall fluency of speaking into account, the test loses its construct validity. In order to achieve these aspects of validity, the structure of communicative competence must be analyzed carefully, and an examination must be so constructed as to satisfy the structural description of communicative competence.

For this purpose, it is worthwhile to review the concept of evaluation in its historical perspective, and thereby inquire into the problem of validity. The history of evaluation in foreign language teaching will be divided into three major periods in

⁵⁾ See W. M. Rivers, *Teaching Foreign-Language Skills*. The Univ. of Chicago Press, 1968, pp. 290-296. She points out inadequacies of several traditional types of tests in her discussion of validity.

⁶⁾ Cf. R. Lado, *Language Testing: The Construction and Use of Foreign Language Tests*. Longman, 1961. A. Davies, *Language Testing Symposium: A Psycholinguistic Approach*. OUP, 1968. D. P. Harris, *Testing English as a Second Language*. McGraw-Hill, 1969. Analyses are made as to the aspects of validity, though some minor differences can be seen in classification.

accordance with the changing viewpoints of language.⁷⁾ The first period can be labeled as "traditional evaluation," corresponding to the grammar-translation method. During that period, the teaching method and consequently the testing procedures consisted chiefly of translation and exercises which dealt with the knowledge of grammar and vocabulary. Since questions which require the mechanical knowledge and manipulation of grammatical rules only compel, on the part of the learner, mental training through memorization, they are far detached from language use or the performance skills of the language, although they might be necessary as a learning activity at a certain stage.

The second stage of evaluation was the so-called "discrete-point approach," founded on structural linguistics. Even though this approach was based on a scientific view of language, it cannot be recognized as fully satisfactory for the current purpose of an inquiry into validity. For the analytical approach taken by structural linguists does not assure that the various linguistic elements constitute overall communicative competence, when they are in turn put together. On the contrary, it leaves out some significant, though maybe small part. Illustrating this deficiency, the tests which make use of minimal pairs are commonly used as an aural discrimination test in the LL, but the distinctive features which cause such contrasts are in fact limited to part of a language, and consequently the remaining part will be missing. In other words, the tests based on the discrete points of a language can encompass only partially the particular problems of a language. Although they may be useful for the solution of those particular problems, it must be admitted that they are not sufficient to evaluate overall language proficiency. Therefore, the criticism against the discrete-point approach indicates the discrepancy between linguistic competence and communicative competence, based on the recognition that linguistic knowledge does not necessarily guarantee the existence of an equivalent communicative skill.⁸⁾

The recent trend in evaluation shows attempts in various ways to overcome the deficiencies involved in the former testing methods discussed above, and evaluate overall communicative competence in a foreign language more exactly and comprehensively. J. B. Carroll illustrates a detailed schematization from the psycholinguistic point of view.⁹⁾ In order to obtain a complete assessment of the individual's knowledge and skill in a language, he distinguishes between the competence and the performance

⁷⁾ Cf. A. Davies, *op. cit.* D. P. Harris, *op. cit.* E. J. Brière, "Are we really measuring proficiency with our foreign language tests?," *FLA*, 4, 4, 1971, pp. 385-391. The same classification is adopted, though the names of each period may differ.

⁸⁾ Various authors have expressed the opinion. See especially L. A. Jakobovits, *Foreign Language Learning: A Psycholinguistic Analysis of the Issues*. Newbury House, 1970, pp. 149-160.

⁹⁾ J. B. Carroll, "Fundamental Considerations in Testing for English Language Proficiency of Foreign Students," in H. B. Allen (Ed.), *Teaching English as a Second Language*. McGraw-Hill, 1965, pp. 364-372.

of language behavior, and schematizes them separately with special emphasis on performance abilities. Criticizing Carroll's model, R. L. Cooper expands the notion from the sociolinguistic standpoint, and adds the factors of language variety and social context to his sociolinguistic framework.¹⁰⁾

Interesting strategies of evaluating overall language proficiency can be found in J. W. Oller, Jr.'s "cloze tests" and B. Spolsky's "noise tests."¹¹⁾ These two types of tests have one noticeable feature in common: they attempt to measure the global language proficiency by a single means of obstruction—brackets with missing words on the one hand, and "white noise" on the other. The reports that these tests have indicated a fairly high correlation with external criteria and other theoretical considerations¹²⁾ imply that some secret is hidden in the operation required for solving these test problems. This secret seems to reflect the communicative competences, which S. J. Savignon tried to measure in a practical and concrete manner.¹³⁾ That is, something similar to the operation contained in the encoding and decoding activities of the communication process is likely to be included in the procedure for solving these particular problems.

From the overview of the historical changes in evaluation thus far, a more valid analysis could be suggested with regard to the structure of foreign language proficiency. Foreign language proficiency will be given a complete description by approaching it from both linguistic competence at the base and communicative competence at a higher level of language behavior. Linguistic competence is related to the knowledge of linguistic elements—lexicon, syntax, and phonology/orthography—as subordinate components of communicative competence. Communicative competence is what lies beneath and enables the actual performance of integrated skills, in which linguistic elements and paralinguistic factors are organically integrated. Though linguistic elements are necessary, they cannot prove sufficient for foreign language proficiency: in addition, paralinguistic factors must be combined to make them function appropriately and effectively in a psycho-socio-linguistic milieu. Through qualitative investigation, paralinguistic factors are further classified into two distinct aspects: sub-linguistic and non-linguistic factors. Sub-linguistic factors here mean those entities contained in a language itself which function to give meaning to linguistic elements

¹⁰⁾ R. L. Cooper, "An Elaborated Language Testing Model," *LL*, Special Issue, 3, 1968, pp. 57-65.

¹¹⁾ J. W. Oller, Jr., "Cloze Tests of Second Language Proficiency and What They Measure," *LL*, 23, 1, 1973, pp. 105-118. B. Spolsky, *et al.* "Preliminary Studies in the Development of Techniques for Testing Overall Second Language Proficiency," *LL*, Special Issue, 3, 1968, pp. 79-98.

¹²⁾ See especially P. Atai and J. W. Oller, Jr., "Cloze, Dictation, and the Test of English as a Foreign Language," *LL*, 24, 2, 1975, pp. 245-252. The results show that cloze scores correlated with the combined dictation at .75 and with the TOEFL total score at .79/.78.

¹³⁾ S. J. Savignon, *Communicative Competence: An Experiment in Foreign-Language Teaching*. The Center for Curriculum Development, Inc., 1972.

or integrate them organically into a meaningful unit. In this subcategory, socio-cultural meanings and linguistic contexts should be included. On the other hand, the latter non-linguistic factors refer to the psychological and social factors which exist at the base of or behind language behavior, exercising a significant influence. Among the psychological factors the intention of a speaker and the connotation of an utterance lend great significance to the communicative role of language behavior. As for the social factors, the situational context and the role relationship are the major factors that restrict the human interaction of language behavior. It follows that the so-called "four skills"—listening, speaking, reading and writing—can duly be regarded as the performance of such underlying communicative competence, manifested explicitly by different means (spoken/written) in different forms (active/passive), in an actual communicative setting.¹⁴⁾

If the structure of foreign language proficiency is grasped as described above, in order to attain full validity in a general proficiency test, it becomes necessary in theory to measure not only linguistic competence, but also the wider scope of elements which constitute communicative competence. In this connection, "criterion-referenced tests" proposed by R. M. Valette may be considered superior in validity as a testing method of overall proficiency. By this method the mastery over the code or "langue" of a foreign language is first measured, and according to the level determined by this preliminary test, the next step is taken to measure the level of proficiency in each of the four skills. In addition to these tests of indirect communication, a test of direct communication in the form of an interview and a sub-test on vocabulary are provided.¹⁵⁾ In practice, however, unless an examination is to make a complete detailed assessment of overall proficiency, the concept of communicative competence must be adapted to coincide with the specific objectives of a course and of evaluation. On the basis of the analytical description of communicative competence, due distribution of emphasis should be given to the different elements and factors, depending on what the objectives are.

IV. Validity from the Viewpoint of Course Objectives

Together with the consideration of validity from the standpoint of communicative competence discussed in the previous chapter, there is another aspect indispensable for the systematization of future curriculum development: that is, an assignment to

¹⁴⁾ 岡秀夫「外国語能力とその評価」『広島大学大学院教育学研究科博士課程論文集』第1巻, 1975, pp. 77-86. (H. Oka, "Foreign Language Ability and Its Evaluation") The author has shown a schematization of communicative competence, and has given a detailed explanation of each element and factor.

¹⁵⁾ R. M. Valette, "Evaluating Oral and Written Communication: Suggestions for an Integrated Testing Program," *LL*, Special Issue, 3, 1968, pp. 111-120.

guarantee validity from the viewpoint of course objectives.¹⁶⁾ In other words, just as teaching materials and methods must be closely connected with and be relevant to the course objective, so evaluation must be correlated to the course objectives; otherwise, it would not prove systematic in the organization of a whole curriculum. The recent realization with respect to teaching methodology is that teaching methods are subject to change according to variables such as objectives, learner's factors, materials, and so on. As a consequence there is no single "panacea" method which works best for any learner in any situation. In the same way, evaluation must also be so devised as to meet the objective of the course, or to be goal-oriented. For example, to ask for appreciation of a complicated literary text may play an important part in the specialized course of the literature department, but in the general English course of the liberal arts department there arises some doubt as to its validity. Total reliance on translation exercises of isolated sentences from Japanese into English in a so-called "composition" class does not go beyond the manipulative level of writing skill, and consequently validity of evaluation along the line of course objective cannot be fully realized.

The traditional notion as regards English teaching in the liberal arts department is that it aims at "culture," and that the methodological procedure takes the form of "reading." However, the fundamental question arises as to what the vague concept of "culture" really represents. Concurrently, the unbalanced emphasis and deemphasis on language skills might be criticized from the global view of communicative competence. Furthermore, the traditional framework does not seem to give us a satisfactory answer to the educational context of the liberal arts course from the perspective of the whole educational system—that is, it is placed halfway between high school and a specialized course of the university.

In such an ambiguous and perplexing situation, the specification of aims and the corresponding formation of subdivided courses with specific objectives are primarily and vitally necessary as a starting point of the systematization of a future curriculum of English language education. A set of courses with different objectives will be so organized as to be integrated into the curriculum of English to form a systematic unity. Then, evaluation should also comply with the setup of specified courses. It may be meaningful to measure the student's overall proficiency in a foreign language for certain purposes—for example, for the sake of comparative research between audio-lingual habit theory and cognitive code-learning theory—but when specific course objectives are established, each objective must be carefully analyzed, and

¹⁶⁾ See F. G. Healey, *Foreign Language Teaching in the Universities*. Manchester U. P., 1967, p. 174. He presents a suggested series of categories into which university language teaching can be divided in order to devise a coherent research program: (1) Course objectives, (2) Course content, (3) Teaching methods, (4) Testing methods.

according to the analysis, due emphasis should be laid on respective aspects of communicative competence in order to assure validity of evaluation.

From a careful examination of intrinsic values and aims of English language education, the author has suggested a future curriculum organization.¹⁷⁾ In the recognition of inadequacies of the prevailing rigid dichotomy between academic, literary objectives and pragmatic, social objectives, a new schema has been proposed in pursuit of pluralism and synthesis. The design of three courses and their respective course objectives, drawn from the analysis of aims, and the criteria of forming a class, based on an inquiry into learner's factors, is shown in the diagram below:¹⁸⁾

	Linguistic Value	Cultural Value	Pedagogical Value
Courses	Linguistic course	Cultural course	Technical course
Course Objectives	Development of communicative competence	Understanding of culture in its broad sense	Preparation for the specialized subject
Class Organization	According to abilities	Selective	According to the major

This development of a curriculum seeks the specification of course objectives and the subdivision of courses with specifically defined objectives. The basic concept of these courses in relation to course objectives is as follows: The linguistic course is devised for the systematic development of language competence and skills, aiming at communicative competence, through which linguistic awareness, including the contrast between the mother tongue and the target language, is fostered. The cultural course—culture here being defined in a broad anthropological sense—promotes the understanding of culture, including not only the phenomenal aspects, but also the pattern of thinking (a psychological aspect of culture). The course would also enable the learner to come into contact with “Culture” as a historical heritage, through which an insight into culture could be acquired. The technical course is based on the recognition of the pedagogical value, which derives from the educational consideration of the place of the liberal arts department in the whole educational system. In addition to general educational purposes, it is designed to bridge the student's educational path from high school to faculty, and to prepare the student for his future career in a particular specialized subject by the medium of English. In order

¹⁷⁾ 岡秀夫「大学教養課程における英語教育——目的論と学習者論」『英語英文学論叢』第26集, 九州大学英文学研究会, 1976, pp. 167-192. (H. Oka, “English Language Education in the Liberal Arts Course of the University—Aims and Learner's Factors”) The author has given a summary in English, pp. 201-206.

¹⁸⁾ *Idid*, p. 206.

to make these specified courses work for the fulfilment of their respective objectives, the organization of the classes must also be goal-oriented: (1) according to the student's ability in the first category, due to the varied levels of attainment as a result of his previous learning experience; (2) the adoption of a selective system in the second category, taking the student's motivation and interest into consideration; (3) according to the major in the last category. Such division of courses and organization of classes promote not only the clarification of objectives, but also the learner's motivation and the relevance of teaching and learning to both the learner and the objective.

If the objectives and courses are organized as proposed above, evaluation will also be affected by each course objective, and accordingly validity of evaluation will become varied in this systematic approach to the process of English language education. In the linguistic course as a whole, overall proficiency in English ought to be evaluated, as has been asserted in the previous chapter. However, from the nature of the course, it may be difficult to give a combined training in all the four skills. This leads to a classification of classes according to the skills: conversation and LL classes to train aural/oral skills; a reading class; and a composition class. The teaching procedure of a reading class here does not mean the traditional pattern of translation and explanation in the name of reading, but a systematic training of the reading skill by means of rapid reading and extensive reading. A composition class does not deal exclusively with traditional translation exercises, but with developmental exercises of the writing skill by means of the controlled or paragraph writing technique. In the same way as in the teaching method, evaluation should lay emphasis on language activities rather than learning activities, and on the content and effectiveness of communication beyond the manipulative knowledge of linguistic elements.

Once the classes are formed and specified according to the attainment levels of skills, the consideration of validity in evaluation will become comparatively easy to cope with. As was concluded at the end of the last chapter evaluation should then encompass both the assessment of linguistic competence—knowledge of linguistic elements—, and of functional knowledge and skills based on the concept of communicative competence. In a practical program deriving from this theory an objective test may be prepared, based on linguistic elements for the purpose of the classification of students into classes of different levels according to the results. Their functional knowledge and skills may be measured in an achievement test at the end of the course. An assignment for future research in this category is the establishment of proficiency levels and the development of proficiency tests. If proficiency tests are so constructed as to reflect the basic idea of levels—levels being prescribed by behavioral objectives, as well as vocabulary, speed, subject, etc.¹⁹⁾—, achievement goals

¹⁹⁾ Cf. 安藤昭一「大学英語教育の目標と内容」『新英語教育論』講座・新しい英語教育1, 大修館, 1976, pp. 248-266. (S. Ando, "Objectives and Contents of College English Language Education")

will be clarified, and consequently the systematization of evaluation as well as teaching and learning will be facilitated.

The cultural course, from the nature of the course objective, should differ from the pragmatic nature of the linguistic course. Overemphasis or entire reliance on the pragmatic theory of instrumental skill-training tends to ignore the pedagogical basis of the liberal arts department, paying little attention to the significance of its educational role. Under the present system, the unordered mingling of linguistic and cultural aspects makes it difficult to give a clear definition of objectives or validity of evaluation. In spite of the commonly acknowledged aim of "culture," the actual situation is more often than not restricted to evaluating the level of a passive learning activity of linguistic elements. Thus, the separation of the cultural and the linguistic aspects will facilitate the formation of a teaching/learning coordination. To illustrate this, a cross-cultural text may be chosen. On the basis of comprehension of the text, the target language serves as a medium of communication, and the strategical stress has to be on promoting learner thinking and discussion in connection with subject matter, through which cultural insight will be acquired. Evaluation should also accede to the objective and the pattern of teaching and learning: that is, in evaluation the mere assessment of passive understanding of the text will prove inadequate. In addition, the awareness and understanding of culture ought to be evaluated. This implies a transfer from the narrow concept of instrumental bilingualism to the broader educational context of "biculturalism." What seems crucial to the validity of evaluation is that the attainment goal is to be explicitly defined not in terms of linguistic features, but in terms of the cultural content. To assure validity, then, a well-defined objective should be reflected in the examination to assess the achievement level of the objective. Such types of questions that require the explanation or discussion of some cultural aspect will be included in a written examination.

The third category of the technical course is entirely missing in the present curriculum. However, when the role of the liberal arts department is viewed in a wide perspective of the whole system of formal education, it becomes clear that the special role of the department in general must be taken into account, in which English language education should also take part. That is, in addition to general educational purposes, some introductory preparation for the student's future career in a specialized subject ought to be undertaken by the medium of English; otherwise, the curriculum would lack in the sequential consistency and goal-orientedness of the whole university education. This could also serve to bridge the gap between the general nature of English teaching in high school and the requirement for specialization at the faculty level. The establishment of this course depends largely on teaching materials—not of general or literary nature, but of a specific field. For this purpose the first preliminary work in the development of an English curriculum is the selection of textbooks and

the drawing of a list of recommended textbooks in cooperation with the faculties involved. According to each faculty, several textbooks of an introductory level should be included in the list, from which the teaching material will be selected by the individual teacher. If the technical course is arranged in this manner, it will become necessary, for a valid evaluation, to measure the student's basic understanding of the subject by the medium of English. This course is designed to familiarize the student with the language of his major field—in particular, emphasizing style and logic, as well as terminology. Evaluation should reflect the design of the course in order not to fail in the consistency of course organization. Practical procedures of evaluation may chiefly measure the logical and exact comprehension of technical phenomena and processes, in addition to the understanding of technical terms in the field. It must be remembered, however, that the primary concern of validity depends on the selection of relevant materials.

V. Conclusion

This study of the validity of evaluation consistent with course objectives has been an attempt to overcome the lack of unity and systematization of the present curriculum of English language education. It presupposes that the present framework of the liberal arts department and the requirement of English as a foreign language will remain constant. This will, therefore, be considered a suggestion for the direction of a reform based on new theoretical considerations. It is hoped that through it, a systematic organization of the English curriculum will be developed and validity of evaluation in particular will be clarified. Admittedly, how these ideas can be put into practice is an assignment for future research. The proportion of three distinct courses, for example, will vary, though not drastically, according to the situation. The proficiency level of the student should be considered to be of primary importance, and other relevant factors should also be paid attention to. The basic ideas set forward throughout this article must be adapted to the particular learning situations and conditions for their most effective functioning.