DOCUMENT RESUME

ED 044 426 TM 000 095

AUTHOR Lofgren, H.
TITLE Measuring Proficiency in the German Language: A Study of Pupils in Grade 7.
INSTITUTION School of Education, Malmo (Sweden). Dept. of Educational and Psychological Research.
PUB DATE Nov 69
NOTE 56p.

EDRS PRICE EDRS Price MF-$0.25 HC-$2.90
DESCRIPTORS Curriculum, Factor Structure, *German, *Grade 7, Language Ability, Language Instruction, *Language Proficiency, Language Research, Language Skills, *Language Tests, Measurement Instruments, Measurement Techniques, Pronunciation, Second Language Learning, *Test Construction, Test Reliability, Vocabulary
IDENTIFIERS *Sweden

ABSTRACT

An attempt was made to further the understanding of foreign language proficiencies and to formulate methods of testing them. Language testing is needed as a criterion of the efficacy of teaching. Consideration of the curriculum and analysis of the texts used would aid in test construction. A discussion of the design and content of language tests and a detailed summary of their subject-matter is provided. This study employed a battery constructed on the basis of Lado's theory regarding the relationship between language variables and integrated skills. Factor analysis of the resulting data produced a factor structure which included knowledge of vocabulary, structure pronunciation, intelligence, and language fluency. Analyses indicated that language proficiency could be seen as a closely knit unit and that correlations between different language tests are high. Suggestions are offered for further research. (PR)

special-topic
bulletin from

DEPARTMENT OF
EDUCATIONAL AND
PSYCHOLOGICAL RESEARCH
SCHOOL OF EDUCATION
MALMÖ, SWEDEN

didakometry

Löfgren, H.:

MEASURING PROFICIENCY IN THE GERMAN LANGUAGE: A STUDY OF PUPILS IN GRADE 7

No. 26 November 1969

Communications should be addressed to
Department of Educational and Psychological Research
School of Education
Fack
200 45 Malmö 28
Sweden
Director: Åke Bjerstedt

MEASURING PROFICIENCY IN THE GERMAN LANGUAGE: A STUDY OF PUPILS IN GRADE 7

Abstract. - Starting from the objective of language teaching and a theory of the relationship between language variables and integrated skills this study tries to contribute to the understanding of the foreign language ability and how to test it. It was found that the language ability can be seen as a rather self-contained unit and the correlations between the various language tests were high. However, it was possible to separate certain parts of the language ability as the factors of knowledge of vocabulary and structure pronunciation, intelligence and language fluency.

1. The need for language tests

The rapid expansion of contacts on an international level has led to a growing need for knowledge of and proficiency in foreign languages. The interest for language-teaching is increasing and with it the efforts made to improve the quality of the teaching and render it more effective. Material and methods have often been evaluated from a subjective view of what is suitable and effective. A brief survey of the development in language-teaching during recent years shows that the question of objectives has figured prominently in the discussion.

This has lead to changes being made in the aims stated in more recent curricula. Unfortunately, the desired end-product has not been so well defined that it can be converted directly into terminal tests. Thus our objectives are defined in rather general terms. We often use methods, chosen on a subjective basis, which appear to lead to our predetermined goals. On the other hand, we do not use particularly good instruments to evaluate our teaching. Objectives and criteria ought to have been developed in conjunction, but that has not been the case. When following the present lively debate on teaching methods - there is general agreement about the objectives - one feels a strong desire to cut the Gordian knot with the sword of criteria. If we only had a number of suitable criteria to refer to, we ought to be able to free the debate on teaching methods from subjective evaluations.

Thus language testing is needed as a criterion of the efficacy of the teaching. These evaluation instruments can also be used for assessing pupil-achievement, when setting marks, for example. We have today in Sweden a few norm-tests of this kind, such as the standard tests for the

comprehensive school and the "central tests" for the "gymnasium". The content of these tests is discussed in section 2.3. We can point out here, however, that these tests do not cover what is perhaps the most important of the language skills, namely the ability to speak the foreign language. If it is not possible to measure ability in speaking by means of other proficiency tests, then we must try to construct suitable speech tests. Our knowledge of total language proficiency is still too limited for us to be able to decide the relation between ability to speak a foreign language and the other language skills. One of the more important issues in the study presented here is whether it is possible to evaluate speaking proficiency without using speech tests, which are often difficult to administrate and assess.

The language tests can be divided into two main groups, namely initial tests and terminal tests. The first group can be sub-divided into prognostic tests and placing tests. The prognostic tests are used to try to predict the pupil's progress in his language studies, while the placing tests aim at placing the pupil in a suitable learning sequence. Among the terminal tests we can differentiate between achievement tests, which measure the progress made in special teaching sequences and which thus refer to the course read, and proficiency tests, which attempt to measure general language ability, independent of the course followed. As was made clear in the introduction, we are here primarily interested in terminal language tests.

Instruments are needed therefore, to evaluate terminal behaviour, both achievement and proficiency tests. When constructing these instruments, one must take into consideration whether the results are to be used for group or individual comparisons, for a diagnostic or differentiating assessment. Higher standards of accuracy are naturally necessary when the results are to be used for individual decisions. Moreover, there should preferably be different kinds of raw score distributions, depending on whether the results are to be used for diagnosis or differentiation. Often in a diagnosis there is only one single point where absolute accuracy is required. Therefore a bimodal distribution is desirable and at best a dichotomy.

In evaluating terminal behaviour, we can differentiate between the following types of test:

/

1) group diagnosis

The results are used for deciding whether the group has achieved the target set for the teaching. Such results often function as a feed-back to the teacher and can then be regarded as a diagnosis of the teaching.

2) individual diagnosis

The results can be used in the same way as in point 1, with the addition that conclusions can be drawn about individuals and certain actions recommended. In this case we must demand a high level of accuracy.

3) group differentiation

The results can here be used, for example, for comparative method experiments. The measurements should make it possible to compare groups of differing ability.

4) individual differentiation

This type of test has the same qualities as that in point 3, with the addition that it is possible to make individual decisions. This type is often used as an instrument for the setting of marks.

Finally, there is one way in which any type of test can be used. Or perhaps we should instead say, one way in which any type of test can exert influence. We know that the type of achievement tests may influence the way in which the pupil studies and thus modify the learning process. Teachers are also influenced by this type of feed-back (Bloom, 1965).

2. The skills represented in the tests

Irrespective of the ways, discussed in Section 1, in which an evaluation instrument can be used, the aim is to construct an instrument with as high a degree of validity as possible. What then should the tests contain as far as skills and material are concerned? Which variables should be included in a test battery intended to evaluate total language proficiency? What we demand of a measuring instrument is that it should be effective, i.e. it should be easy to administrate and correct, take only a short time to carry out and have a high degree of validity and reliability.

The problem is not so simple that it can be solved by merely converting the given objectives into suitable terminal tests. Our ultimate aim, the command of total language proficiency, is so complex that at the beginner stage we must divide the target into sub-targets. What then

are these sub-targets? What are the component parts of language proficiency? No attempt will be made to answer these questions here, but instead a study will be made of how the teaching can be evaluated, once the target has been decided and a particular teaching-method chosen. Irrespective of whether we try to assess our teaching at the end of a course for beginners or after one at a more advanced level, we must try, if our measurements are to be efficient, to translate the many specific behavioral objectives into a reduced number of fundamental dimensions.

2.1. Methodologists and language psychologists

"Up to the present time there has been no systematic synthesis, based on a careful examination of what language is and how it is learned, of a testing program that will measure success in second-language learning" (Brooks, 1960, p. 157).

Despite this, Brooks tries to give certain recommendations. His idea is that one must distinguish between the four skills, listening comprehension, speaking, reading comprehension and writing, all of which are of a different nature. A test programme must therefore, still according to Brooks, test the four skills separately and in various combinations.

The first consideration when deciding the content of the test is whether the test is to be used for prognosis, placing, achievement or proficiency measurement. A prognosis test must probably, at least for those just starting to learn a foreign language, contain quite different variables than the other types of test.

The Harvard Language Aptitude Project has, led by Carroll, developed and produced prognosis tests. The purpose of these tests was to predict the success of English-speaking pupils in learning a target language. In this study factor-analysis was used and Carroll considered that he could discern six factors which influence achievement in language-learning:

Factor A - verbal knowledge (knowledge of words and structures in the mother tongue)
Factor B - linguistic interest (roughly, special motivation, interest and adroitness with linguistic material)
Factor C - associative memory
Factor D - sound-symbol association (the ability in language to connect sound and symbol plus the ability to learn new connections of this kind)
Factor E - inductive language ability (the ability to draw conclusions about grammatical rules, when suitable material is offered)
Factor F - grammatical sensibility or syntactic fluency (feeling for functions and ability to produce syntactically correct verbal material)

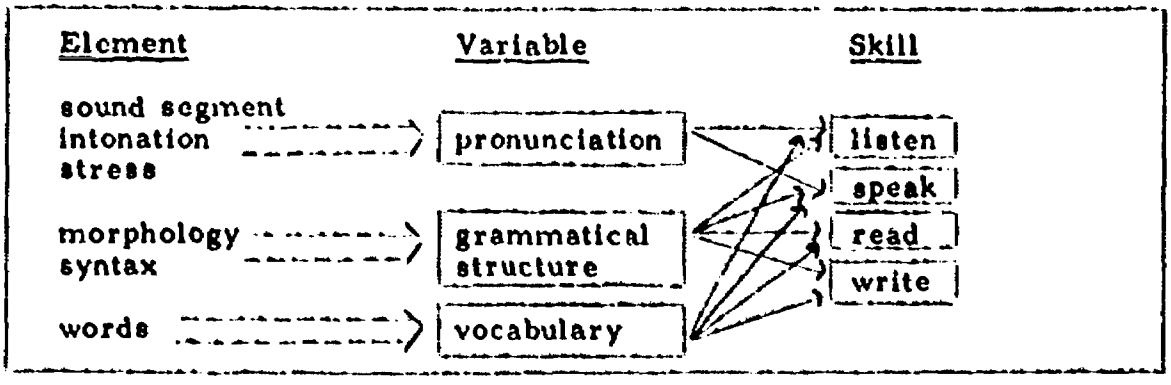
All these factors have great significance for the learning of foreign languages. Factors B, C and E are probably more important than factors A, D and F, in Carroll's opinion (1958).

Smith and Berger (1968) have also tackled the problem of prediction in the study of different teaching-methods and language laboratory systems, which was carried out in the subjects German and French. They found significant connections between intelligence, attitudes and marks in other subjects and the criterion, language proficiency test. One of the best predictors proved to be the results obtained from the Modern Language Aptitude Test developed by Carroll after his factor-analytic studies.

Lado (1962) has tried to break down the language skills into smaller components, in order to search for common elements. In discussing the relation between language tests and language learning, he claims that the tests ought to be a representative selection of what has been learnt. One can start with a random selection but with the omission of the easier items, since it is particularly the difficulties which should be tested. Lado often mentions the phrase, "Testing the problem is testing the language". He discusses in detail the strategy and variables necessary for language tests. His idea is that the language consists of elements, which constitute variables, which in their turn are not independent but are integrated with the total skills, speak, listen, read and write, plus possibly translate.

The diagram in Box 1 attempts to summarize his discussion.

Box 1. A summary of Lado's discussion of the relations between elements, variables and skills.



The synthesis in Box 1 does not include the variable 'cultural meanings', which Lado thinks should be included.

What strategy should we use when we test the pupils' language skills? According to Lado, we must test the elements and the skills separately and may choose between the one or the other, or a combination of both,

depending on the purpose of the test. Lado criticises some common types of language test, such as translation, for example, which is often used as a general language test for the setting of marks. They are easy to construct and appear to be valid, but are probably not. Translation is a special talent, even if it correlates with speak, read, listen and write. Further we have essay tests, which again measure a special ability. Being able to write a pleasant essay is not the same as being able to speak, listen to, read and write the language. Moreover, essays are difficult to judge. A third common, but bad, type of test is dictation, which really measures very little of language proficiency (Lado, 1962, Chap. 4).

Smith and Berger share Lado's opinion that the skills can be broken down into common variables. "All skills make use of the common core of structure, morphology and vocabulary" (Smith et al, 1968, p. 70).

Under the guidance of Scherer and Wertheimer (1964), a teaching experiment was carried out, in which an audio-linguistic method (special emphasis on the oral skills) was compared with a conventional method (grammar and translation). The leaders of the study complained that there were too few language tests with German as the target language. Despite intensive searching, no tests suitable for the purpose could be found, i.e. tests which measured the skills listen, speak, read and write. "Rumours about ideal batteries in existence and use in various organizations and places, at home and abroad, were relentlessly pursued, but to no avail" (op. cit., p. 108). Not until after the study was completed did, for example, MLA Foreign Language Proficiency Tests for Teachers and Advanced Students appear, together with a similar battery produced under Brooks' guidance for lower levels.

Scherer and Wertheimer also take up the question of whether it is suitable to regard language proficiency as being composed of four different skills, or possibly six if translation in both directions is included. The division into listen, speak, read and write, they say, is strongly rooted in the world of the language teacher, as is the difference between active and passive ability. For decades the language teacher has spoken of language proficiency in just these terms. It is possible, however, that empirical studies would show that proficiency in a foreign language ought not to be divided into these a priori categories.

"Factor analytic approaches would be still more powerful in helping to decide what separable aspects can be empirically demonstrated to compose proficiency in a foreign language. Had computer capacity permitted, intercorrelations among all the subtests of all the proficiencies would have

yielded direct evidence on the question of whether the traditional fourfold classification makes empirical sense. Perhaps in the future time and funds will become available for us to undertake such analyses." (Scherer & Wertheimer, 1964, p. 113)

In order to carry out the study, Scherer and Wertheimer constructed a series of tests. They tried to make them as pure as possible, i.e. constructed tests which were to measure the different skills separately. Thus an ideal listening test ought not to include reading, which is often the case, but should be limited to hearing comprehension. But although a great deal of work was put into the construction of the tests, in order to make them as pure as possible, they were not wholly successful.

Scherer and Wertheimer concluded from an intercolleration study that language proficiency can be divided into two clusters, one with the hearing and speaking skills and one with the reading and writing skills, plus translation in both directions. The conclusion seems doubtful, however, considering firstly the connections reported and secondly the not wholly perfect tests. This interpretation should perhaps also be seen in the light of the fact that the purpose of the study was not to investigate language proficiency, but to compare two methods of teaching, one with a certain overemphasis on oral training and the other with a certain over-emphasis on written training.

Swedish methodologists also commonly use the fourfold division of language proficiency.

"To know a language means to understand when it is spoken and to be able to speak it oneself, to be able to read it and to be able to write it. ... Two of the four skills are active or expressive, to speak and write, and two passive or receptive, to understand the spoken language and to read and understand the written language. ... The four skills are obviously interwoven and they neither can nor should be regarded or treated in isolation." (quotation translated from Swedish; Thorén, 1960)

By combining the six skills (including the translation skills) with the language variables suggested by Lado, the summary presented in Box 2 is obtained.

Box 2. A synthesis of the relation between integrated skills and variables.

Variables / Integrated skills	Pronunciation		Gram. structure		Vocabulary	
	recog-nition	pro-duction	recog-nition	pro-duction	recog-nition	pro-duction
Speak		x		x		x
Write				x		x
Hear	x		x		x	
Read			x		x	
Translation a) mother tongue-target language				x		x
b) target language mother tongue			x		x	

Several important questions should be asked when confronted with this survey, however. Is the variable division suitable, considering the components of the integrated skills? Is speaking proficiency made up of the expressive forms of the variables pronunciation, grammatical structures and vocabulary? Is writing proficiency made up of the expressive forms of the variables grammatical structure and vocabulary and so on? Are there high intercorrelations between tests measuring the skills speak, write and translate from the mother tongue to the target language, and high intercorrelations between the skills hear, read and translate from the target language to the mother tongue? A complicating factor, however, is that recognition and production are not separate, but have a certain bearing on each other. We can say that recognition is simpler than production, that is to say, anyone who can produce can also recognize. This argument leads to the formation of a hierarcy between the skills.

"Since the passive is easier than the active, since the static is easier to understand than the mobile, and since it is easier for a person able to write to write a sentence correctly than to produce it orally (the concept 'writing' implies that the person writing has a better opportunity for reflection than the person speaking), the ability to read is the simplest of the four skills and the ability to speak the most complicated." (Quotation translated from Swedish; Thorén, 1960, p. 20)

There are some interesting results obtained from comparisons made between listening and reading comprehension of the mother tongue. The investigations can be divided into two groups, correlation studies and

studies of different levels. Correlations of 0.50 - 0.60 were found when the persons tested had gone to elementary school, and of 0.70 when they had gone to high school. Moreover, it was found that the correlation increases if the listening and reading comprehension tests are based on equivalent forms of existing reading comprehension tests rather than on independently constructed tests (Larsen & Feder, 1940).

After correcting for reliability faults, the author of this report found almost complete agreement between equivalent forms of listening and reading comprehension tests measuring achievement in German in grade 8 (Löfgren & Hall, 1968).

The studies of different levels have investigated the comprehension of verbal material. Which gives the best comprehension, reading or hearing certain information? Up to grade 7, listening comprehension seems to be better (Beery, 1954; Hampleman, 1955). In one study at seinor high school level, reading comprehension was better of medium difficult to difficult material, otherwise no differences were found (Rulon, 1943).

The writer of this present report has found that pupils in the 8th grade achieved somewhat better results in reading comprehension tests than in listening comprehension tests in German (Löfgren & Hall, 1968).

The summary in Box 2 shows how listening and reading comprehension have common elements, which could explain the high rate of correlation. The certain difference in level which leads the older pupils to better results in reading comprehension tests, can probably be explained by the opportunity the pupils have of review when they are in a position to read the question over again. At lower levels, where the actual process of reading can cause problems, listening comprehension can sometimes be simpler than reading comprehension. One way to remove the influence of repetition on the reading comprehension test is to give the pupils the chance to listen to the task again if they so wish. Another possibility is to prevent repetition in the reading comprehension test.

Spearritt (1962) carried out a factor analysis, which included listening and reading comprehension. The analysis was based on 12 year olds and concerned their mother tongue, in this case English. Spearritt claimed to have found a separate listening comprehension factor. He also tested the hypothesis that listening comprehension should have the same loadings as reading comprehension on reasoning and verbal comprehension, but higher loadings on the attention and memory factors. The argument mentioned above also lay behind this hypothesis, namely that it is possible to

repeat more difficult passages in a reading comprehension test while this possibility does not exist in a listening comprehension test. The results showed that the hypothesis on similar loadings in the factor in inductive and deductive reasoning was acceptable. On the other hand the similarity between listening and reading comprehension was not so great when it came to verbal comprehension and it was felt that the hypothesis on higher loadings for memory factors in listening comprehension could not be accepted. This last conclusion is rather dubious, since it was also found that some tests of range memory span correlated more with listening comprehension than with reading comprehension.

2.2. Curricula, syllabuses and teaching projects

The skills aimed at in language teaching in the Swedish comprehensive school are "a basic good pronunciation, so that they (the pupils) can catch and understand clear speech, read and understand simple text, express themselves in simple, everyday situations and to some extent express themselves in writing." (quotation translated from Swedish; National curriculum for the comprehensive school, 1962, p. 190.)

No suggestions as to suitable terminal tests are given, which would define the objectives more precisely. Language proficiency is divided into speak, hear, read and write, but these skills are to be co-ordinated so that the foreign language appears as a living entity (p. 201). Special emphasis is to be laid on the oral skills. "Language teaching in the comprehensive school must start from the fact that the language is primarily a means of verbal contact between individuals. Particularly the basic teaching should aim at letting the pupils experience the language as sound, rythm and melody, since proper acoustic conceptions of the sounds of the foreign language and proficiency in pronouncing them is prerequisite for both the ability to catch and understand the foreign language and the ability to speak it" (p. 194).

The view of language study which characterizes the curriculum for the comprehensive school is also essentially valid for the 'gymnasium' and the vocational training school. The task of modern language teaching, as far as continuation languages are concerned, is to develop the ability to understand the language in speech and writing, to extend and consolidate skill in speaking, reading and writing the language ..." (quotation translated from Swedish; National Curriculum for the 'gymnasium', 1965, p. 112).

The three curricula speak of the skills speak, hear, read and write, but make no mention, on the other hand, of translation as a desirable skill.

No operational definition of the objectives in the form of suitable tests is given in any of the three curricula.

The curricula for the vocational training school and the 'gymnasium' speak however, of the desirability of the teaching material being supplied with exercises for checking and drill. Under the heading "Assessment", (quotation translated from Swedish; Curriculum for the vocational training school, 1965, p. 115; Curriculum for the 'gymnasium', p. 123), certain recommendations are given. Thus it is stated in Cvts that testing should be limited and a warning is given against letting checking exercises dominate the teaching. "Since knowledge of and proficiency in language consists of so many different components, however, the kind of control called 'tests' cannot be avoided, if assessment is to be as objective as possible. In addition, tests are always of diagnostic value to the teacher, who gains clear information on the success of his teaching in the class as a whole, and where further teaching and training are needed" (Cvts, p. 116). Rather short, well-prepared tests are recommended, of the type requiring the pupil to explain a number of words and expressions in a text, to answer certain questions or to fill in a number of blanks. The longer tests should consist of several different items, so that the pupils can do their various skills justice and so that any possible weakness in some detail involved in acquiring the language can be compensated by greater proficiency in another part. When marks are set, it is considered to be of the greatest importance that oral work is also taken into consideration. The teacher is recommended to use the same material when testing all the pupils orally, and record it, so that their proficiency in speaking may be assessed more objectively. In the 'gymnasium', written tests are included in the timetable and some of these are prepared centrally. "When setting marks at the end of term, it is again the combined knowledge and skills which are to be assessed. Consequently, the written tests should not be decisive. All the objectives should be taken into consideration" (Cgy, p. 124).

Thus the curricula make it quite clear that the total proficiency in a language is seen as being composed of different skills, each of which should be tested in order to give as fair a picture as possible of a pupil's achievement. Plenty of words are used to describe how pupils' achievements should be assessed, but no practical examples are given. The lack of well-constructed tests must be substantial considering how little time teachers have to spend on composing tests and how relatively deficient their training in the theory of measurement probably is.

2.3. Standard tests and "central tests"

"Standard tests have been introduced in order to set certain norms for marks and are intended to act as a common yardstick over the whole country" (quotation translated from Swedish; The Swedish Board of Education, 1968).

The "central tests" for the 'gymnasium' have been introduced for the same purpose. How then have these tests been constructed?

"The curriculum for the comprehensive school states what is the aim and the main content of the teaching in each subject.

The standard tests cannot, however, 'cover' the entire course of study in a subject. Such a test would be much too bulky. Moreover, some parts of the course have to be excluded, since it is difficult to construct tests for them. This applies, for example, to oral work in Swedish and modern languages. The choice of the items that are to be included in the test has been made in cooperation with experts in each subject." (SBoE, 1966.)

In the sample selection of central tests for the 'gymnasium' (1968), we are given the following answer: "The content of the tests is decided by means of an analysis of the general or, for a given subject, specific objectives for the teaching laid down in the national curriculum."

The idea is that a central test should contain all the sub-sections of the objectives, but this is impossible for reasons of time. As far as modern languages are concerned, it has been decided to omit such important items as oral production and orientation (culture and background).

Thus a central test in a modern language will be composed of several parts, including reading comprehension, listening comprehension and written production (essay). In addition a fourth section is planned, which will be a vocabulary test or a test on forms and structures. The standard tests is divided into reading comprehension, listening test and sentence test. The last-named is a kind of vocabulary test.

It is interesting to see how the mother-tongue is avoided in both the standard tests and the central tests. Could it be the direct method's fears of mixing the target-language with the mother-tongue which is responsible for this? The reading comprehension test consists of a text with questions. These questions, which are intended to measure understanding of the text, are written in the target-language, which means that a pupil can very well understand the text without being able to answer the questions which follow. To say that the questions are part of the reading comprehension is no defence. The difficulties are even greater in the listening comprehension test. The pupils have to listen to a text and are then given written multiple choice questions. What is it really that is

being tested - listening comprehension or reading comprehension? The use of multiple choice tests is also questionable. Since the alternatives are so close, the result can be that attention and memory are tested rather than listening comprehension.

2.4. Summary of the skills represented in the tests

We lack a synthesis of language proficiency based on empirical studies. Language ability has usually been divided into the five skills: read, write, speak, hear and translate. If these skills are of different natures, as Brooks thinks, for example, it ought to be possible to test them separately or in different combinations. If each skill can be divided into component parts, such as the variables suggested by Lado, pronunciation, grammatical structure and vocabulary, there ought to be tests which measure these components.

In the national curricula, total language proficiency is seen as consisting of different skills, each of which should be tested separately so as to give as fair a picture as possible of a pupil's achievements.

The standard tests and central tests are naturally a consequence of the view taken in the curricula of language proficiency. The reason why oral production is not tested in these tests is that it has been thought to be impossible for practical reasons. No attempt has been made to measure speaking proficiency with non-oral tests.

Correlation studies have produced some interesting results, but since the results are not clear-cut we still know little of how total language ability is built up. Given improved tests, further correlation studies would be of interest.

3. The subject-matter used in the tests

We come now to the question of what material the tests should contain. A distinction must again be made here between achievement and proficiency tests. If we want a test based directly on the course studied, then the objectives of this course will form the basis of the test construction. If we instead want a general language test independent of any course, then the material used can be decided by the objectives of the language teaching, as stated in, for example, curricula, text-book analyses, frequency studies and mistake analyses.

3.1. Curricula

The national curriculum for the comprehensive school has, under the main items for the three-year course in German, the following:

"Listening practice, mainly in conjunction with pictures or text, as a foundation for skill in understanding spoken German.
Reading of simple texts with a gradually increased stock of common words and phrases and with a slowly increasing degree of difficulty.
Talking practice: Simple questions and answers, also between the pupils themselves, concerning every-day situations, pictures or text, in order to train an elementary stock of words, phrases and grammatical phenomena.
Learning dialogues by heart. ...
Grammar: Practice of basic grammatical phenomena." (Cgy, p. 191-192)

This part of the curriculum obviously does not give us much help in deciding what material to use. What is meant by simple texts, common words, common phrases, simple questions and answers and by elementary grammatical phenomena?

Directions and comments are provided, which try to give a more precise picture of the general aim. Thus it is said that one of the most important tasks for language teaching must be to give the pupils the ability to make use of media such as radio, film, television, newspapers, magazines and books. As a clear and unequivocal description of objectives, however, this still leaves a good deal to be desired. As far as grammatical phenomena are concerned, the curriculum suggests items of grammar which can be taken up in the various years of the course (p. 203). This can be seen as a suggestion for material to be used in the evaluating instrument.

3.2. Text-book analyses

An analysis of the material used in a number of accepted text-books ought to be of help in deciding the content of the evaluating instruments. As part of the language project UMT, an analysis has been made of the most widely used text-books in German for grades 7 and 8, with regard to the existence, sequence and frequency of grammatical items. In addition, the content of the text-books has also been analysed from the point of view of culture, background and fields of interest (Lindell, 1966; Schwandt, 1968). As far as the existence of points of grammar is concerned, there is general agreement between the books, but when it comes to sequence and frequency there are both greater and lesser differences between the various items in the text-books. The existence analysis simply involved seeing whether or not the item was included; the sequence analysis involved studying the

order in which the various items were introduced and the frequency analysis involved counting the number of times the iten in question was dealt with.

The aim of the UMT project's analyses was admittedly the deciding of objectives, but as pointed out earlier the evaluating instrument is closely related to the objectives, so that it is quite possible to use the text-book analyses to decide the content of the evaluating instrument.

3.3. Frequency studies and analyses of mistakes

It has been mentioned earlier that a selection of the components of the language should be tested, in order to evaluate total language ability. The fact that certain parts of a language are easy to learn, because of earlier linguistic experience, for example, should be taken into consideration when choosing tasks for tests, so that particularly the difficulties are tested. It is meaningless to place among the objectives for a beginners' course command of linguistic phenomena which are seldom used. Just as we in our teaching concentrate on recurrent and difficult items of language proficiency, so we should do the same in our instruments of evaluation.

Studies of word-frequency have long been accepted as one of the foundations for word content. The main interest has now been transferred from written literature to newspaper text and the spoken language (Meier, 1964; Oehler, 1966; Pfeffer, 1964; Rosengren, 1968; Nilsson, 1969).

More recently, interest has also been focused on frequency studies of grammatical structures. Grammatical frequency studies are today being carried out with the help of data processing within the UMT project. Among the questions for which answers are being sought are: "How common are the various case constructions? Rather special forms of dative structure are said to occur in German everyday language. How common are they? The genitive is said to be uncommon now and to have been replaced by prepositional forms. Is this correct? How common are the different tenses? Does the perfect replace the imperfect as past tense in everyday German? How often does the passive occur?" (quotation translated from Swedish; Lindell, 1968.)

Thus we should test frequently-recurring difficulties. How are we then to decide what the difficulties are? One way is to study the contrasts between the mother-tongue and the target language. We then work on the assumption that the pupils run into difficulties on points where the two languages differ.

"In the light of the problems involved in a complete contrast analysis, it seems likely that a study of the pupils' mistakes could give more direct information on where they find the greatest difficulty. Such a study of mistakes should be based on the pupils' spontaneous production and not on specially arranged texts for translation, where certain difficult points can be inserted deliberately, and should furthermore include both written and oral production." (quotation translated from Swedish; Lindell, 1968.)

Such analyses of the pupils' free production have been carried out and new experiments are in progress within the UMT project (Engh, 1968).

3.4. Summary of the subject-matter used in the tests

We have a better idea of what subject-matter we want in our tests than we have of what skills should be included. We know that we are going to convert our objectives into terminal tests. This is rather more complicated to put into practice, however. Since the setting of objectives and the construction of terminal tests go hand in hand, the same methods can be used, e.g. text-book analyses, language frequency studies and mistake analyses.

4. Presentation of problems

From the introductory discussion, the following problems emerge:

1. What different factors make up language proficiency?
2. Which measurements should be included in a test-battery intended to measure total language proficiency as comprehensively as possible?
3. Are the different factors of which language proficiency consists influenced by the teaching method?
4. Are the measurements included in a test battery, which is intended to measure total language proficiency, dependent on the teaching method?
5. How should a language test directly connected with a specific course and a specific objective be constructed?

This report will attempt to contribute to the understanding of problems 1-4. Problem 5 will be dealt with in a later paper, in which the UMT project's course diagnoses will be reported. Within the UMT project, a number of teaching booklets are being prepared. The idea is that, after completting each booklet, the pupils should be given individual advice and help. For this purpose, diagnostic tests are constructed for the different booklets. UMT's study material system complete with diagnostic tests has been tested during three school-years (1966-1969) and the material for grade 7 was ready for the autumn term 1969 (German, grade 7, 1968-1969).

5. Test constructions and test data

The starting-point chosen for the test constructions is the curricula and the synthesis of skills and variables in Box 2. Box 3 gives a survey of what the different sub-tests are intended to measure.

Box 3. Survey of what the different sub-tests are intended to measure

test no.	test	intended to measure
1	Listening 1	ability to understand spoken language
2	Listening 2	ability to understand spoken language
3	Reading comprehension 1	ability to understand written text
4	Reading comprehension 2	ability to understand written text
5	Write 1	ability to write the language
6	Write 2	ability to write the language
7	Speak 1	ability to speak the language
8	Speak 2	ability to speak the language
9	Translate 1	ability to translate from the target language to the mother tongue
10	Translate 2	ability to translate from the mother tongue to the target language
11	Vocabulary 1	the passive vocabulary
12	Vocabulary 2	the active vocabulary
13	Vocabulary	the active vocabulary
14	Pronunciation 1	phoneme pronunciation
15	Pronunciation 2	intonation and stress
16	Pronunciation 3	phoneme pronunciation, intonation and stress
17	Fluency 1	fluency in target language
18	Fluency 2	fluency in mother tongue; Thurstone's w-factor
19	Intelligence 1	the verbal factor; Thurstone's v-factor
20	Intelligence 2	power of reasoning; Thurstone's r-factor
21	Mark in German	total language proficiency as assessed by the school
22	Grammar test 3	grammatical structure; verbs
23	Grammar test 4	grammatical structure; verbs
24	Grammar test 9	grammatical structure; verbs
25	Grammar test 1, 2, 6, 7	grammatical test; case

Since there is such a great lack of usable language tests, all the tests used in the investigation are newly-constructed. The author has taken the liberty of borrowing ideas and even certain items from various text-books. We have tried in the test battery to convert the objectives set by the curriculum into suitable language tests. A survey comparing the test battery's composition in relation to the curriculum is presented in Box 4.

The content of the test battery can best be seen if the sub-tests are presented. Here we simply give instructions, sometimes an example and one item from each sub-test. More detailed information is given in a separate report (Löfgren, 1969).

Box 4. A comparative presentation of the objectives laid down in the curriculum for the comprehensive school (p. 190) and the content of the test battery.

The wording of the Ccs	Content of the test battery
The task of the teaching is to make the pupils ...	
familiar with a limited, central stock of words,	vocabulary of various kinds
expressions and grammatical patterns	elementary case and tense structures
and to found a good pronunciation,	pronunciation tests
so that they can catch and understand clear speech,	listening to sentences and texts
read and understand easy text,	silent reading of sentences and texts
express themselves orally in simple, everyday situations	colloquial speech in various situations and with different demands
to some extent express themselves in writing	written production of various kinds

As far as skills and variables are concerned, the test-battery tries to cover as much as possible. The subject-matter has been limited to the items which were shown to have a high frequency rating (see p. 16).

Test no. 1. Listening 1

/In this test, you will first hear a sentence. After that, you will hear four alternative ways of continuing that sentence. You should put a cross by the best alternative on your answer sheet. You will hear each question twice. Here are two examples:/ x)

1. A B C D /Wie heisst er denn
   A Anna
   B Ich heisse Paul
   C Fritz
   D Frau Braun /

2. A B C D /Ich habe kein Heft
   A Das macht nichts. Ich habe Papier
   B Um so besser. Ich habe eine Schwester
   C Das macht nichts. Ich belege einen Platz
   D Gehst du jetzt dorthin? /

This part of the test consists of 35 items, which is probably a maximum with regard to the pupils' powers of concentration. One point which is open to discussion is whether the pupils should hear each task more than

x) Text between these signs / / is given orally.

once. Since the alternative choices are relatively short but lie very close to each other, we have decided that a repetition is essential. The preliminary investigations also showed that the pupils experienced considerable frustration when not allowed repetition.

Some kind of disturbance, such as coughing, can result in the pupils' not managing to hear all the alternatives. Our intention was also to minimise the possible significance of the pupils' immediate ability to memorise. The test was administered with a classroom tape-recorder and the timing was identical for all pupils, since the pauses were included in the tape.

Test no. 2. Listening 2

Now you are going to hear some rather short pieces of German text. For each piece of text you have some questions on your answer sheet which you should try to answer in Swedish. You will hear each piece twice. Here is an example:

/Eigentlich hören in Deutschland nicht so viele Beat wie in Schweden. Schlager aber hört man oft. Meistens sind sie furchtbar sentimental. Sie erzählen nur von der Liebe. Und oft vergisst man sie schnell. Manchmal gibt es auch "Evergreens", Schlager, die man immer wieder hört. So ein Schlager ist "Auf Wiedersehen" den du bestimmt schon gehört hast./

Tasks:

1. Which country appreciates beat music most, Germany or Sweden?

   _____

2. What are German popular songs often about?

   _____

3. What can one call a song like "Auf Wiedersehen"?

   _____

The test consists of 14 varied texts which can give a maximum of 35 points. The tasks are intended to measure understanding both in detail and in outline. In this test the pupils were again allowed to hear the text twice, with a pause for answering between. The same voices were used as in section 1.

Test no. 3. Reading comprehension 1

Instructions: In each question you are given a sentence. This is followed by four alternative ways in which this sentence could be continued. Put a cross by the best alternative.

Example: 1. Wie heisst er denn?
- A Anna
- B Ich heisse Paul
- C Fritz
- D Frau Braun

2. Ich habe kein Heft.
- A Das macht nichts. Ich habe Papier
- B Um so besser. Ich habe eine Schwester
- C Das macht nichts. Ich belege einen Platz
- D Gehst du jetzt dorthin?

This test consists of 30 items. Listening 1 and this test have similar instructions and the two tests are also constructed in the same way and have similar subject-matter.

Test no. 4. Reading Comprehension 2

Here are a number of rather short German texts which you are to read. After each text you will be given some questions which you should try to answer in Swedish.

Example: (Here the same example was given as in Listening 2, and so another text has been chosen from the test as illustration.)

Mit dem Rad durch die Schweiz.
An allen Stationen der Schweizer Bundesbahnen und der Privatbahnen können sich Touristen Fahrräder leihen. Leihgebühr: 6- DM für den ersten Tag, jeder weitere kostet 2, 70 DM. Übrigens: Wer unterwegs angeln will, erfährt aus einer Broschüre des Schweizer Verkehrsbureaus, wo er es kann und was es kostet.

1. What does it cost to hire a bicycle for two days?

_____

2. What does it say in the brochure?

_____

The test consists of 11 texts, which can give a maximum of 36 points. The tasks are intended to measure understanding both in detail and in outline.

We have tried to make these listening and reading comprehension tests as "pure" as possible, as far as the different skills are concerned. These skills are tested in Listening 2 and Reading Comprehension 2 by letting the pupils answer in Swedish, in order that their skill in writing will not interfere.

Test no. 5. Write 1

Express yourself as well as you can in German.
You are given a number of situations below and your task is to think of a suitable German sentence and write it down. Try to express yourself as well as you can, but do not try to take too difficult sentences. We will start with a couple of examples, so that you understand what you are to do.

Examples:

1. Tell someone that we drive on the right in Sweden now.

_____

_____

2. Ask your friend Karl where he is going.

_____

_____

This section consists of 20 tasks, which can give a maximum of 71 points. The tasks offer no opportunity for word-to-word translation and so the answers have been judged according to a key, assessing information and

correct language. The tasks are presented in Swedish in order to fulfil the condition that each section of the test should test only one skill. If the tasks had been given in German, the reading and writing skills would have been mixed.

Test no. 6. Write 2

Write down as much as you can about this picture: what you can see, what is happening, whatever you like. Try to make your German as correct as possible. /Don't simply write down a list of words, but try to make complete sentences./

Since this is a question of free production, certain problems arise in marking. The following norms have been used:

a) the number of words is counted and 1 point given for each word (only 1/2 point for words in lists, however)

b) wrong choice of word - 1 point
mistakes in structure minor - 1 point
major - 2 points

c) the result as a whole is calculated by adding the points.

As a result of this system of marking, a pupil can, by achieving more words, make a proportionally greater number of errors, and yet still gain the same total number of points.

Test no. 7. Speak 1

Express yourself as well as you can in German.
Now we are going to imagine that you become involved in a number of situations in which you need to be able to express yourself in German. So you will be given a situation and you are to try to find a suitable sentence to say. We are going to record your answers, so that we can mark them. Try to express yourself as well as you can without using too difficult sentences.

We will start with a couple of examples so that you understand what you are to do. When the tape-recorder in front of you starts you say: "Question 1", and then you say your answer. We will try with the first example.

You meet your teacher in the street. Greet him and ask him how he is.
Question 1. (Say your answer)
Example 2: You meet Mrs. Weise. Greet her and ask her where she is going.
Example 3: By accident you knock into a lady in a shop. What do you say?

Answer as well as you can. Take your time. When you have finished all the questions, put your hand up. We will then turn off your tape-recorder. After that sit as quietly as you can so that you do not disturb the others who are still working. As soon as your tape-recorder starts, you may begin. Don't forget to say the number of the question.

This test consists of 16 tasks, which can give a maximum of 59 points.

As in Write 1, the pupils' achievements have been judged according to a key, assessing information and correct language. Analogically with Write 1, we have chosen to present the tasks in Swedish.

Test no. 8. Speak 2

Say as much as you can about this picture: what you can see, what is happening or whatever you like. You will be given a little time to think about what you are going to say. Then when the tape-recorder starts, you can begin to talk about the picture. Speak clearly and try to use as correct German as you can. /Don't simply say a lot of words, but try to form sentences./

The same picture is used here as in the section, Write 2. The marking of the pupils' achievements was by far the most difficult in this sub-test. As we have already said earlier in the discussion, the teacher's greatest problem is the assessment of skill in speaking. This test was given simultaneously to all the pupils taking part in the study. The recordings were made in a language laboratory. The pupils' recorded answers were written down and then judged by a single person. There are many difficult problems involved in this kind of test. Admittedly the language laboratory permits all the pupils to be tested at the same time, but it also has disadvantages, such as the differing quality of the tape-recorder units, and this leads to differences between the pupils in the administration of the test and the assessment of it. Overhearing between the pupils' seats is another example of a source of disturbance. The actual marking is also difficult, even once

the answers are safely recorded. How much importance should be placed on quantity and quality respectively? It is very often impossible to catch the details of what is said. The difference between m and n, which is essential in grammatically correct German, is often drowned in other outside noises. As we see it, this kind of marking must often be inaccurate. If we could instead measure speaking proficiency with the help of written tests, we would gain a lot.

The test has been marked in the same way as the section Write 2.

Test no. 9. Translation 1 and Test no. 10. Translation 2

In this test you will be given a number of sentences to translate. First you will get 10 sentences, which are to be translated into Swedish and them you will get 10 sentences which are to be translated into German. Even if you think the sentences are difficult, try to do as much as you can of each sentence.

Translate into Swedish:
1. Eines Abends klingelt es an der Haustür.

_____

Translate into German:
1. Wilhelm goes to school every morning.

_____

These sections have been evaluated in the same way as the sections Speak 1 and Write 1, with 36 and 34 points respectively as the highest score attainable.

In the main investigation, these tests proved to have rather too short a time limit, and since both tests were given at the same time, this can have influenced the result. The pupils felt frustrated at having to go over to T2 halfway, when they did not feel they had completed T1 satisfactorily.

Test no. 11. Vocabulary 1

What do the words in the lefthand column mean? To help you, you will be given the words put in sentences.

Example:

der Bleistift	Ich habe einen Bleistift.	_____
klopfen	Wer klopft an die Tür?	_____

The test contained 53 words and each correct answer gave 1 point. The words have been taken at random from a modern frequency word-list (Oehler, 1966).

Test no. 12. Vocabulary 3

What are the words on the left in German? To help you, you will be given the words put in sentences.

Example:

write He writes a letter. _____

This test contains 41 tasks, mainly verbs and adjectives. They have been chosen in the same way as the words in Vocabulary 1. Minor spelling mistakes have been accepted.

Test no. 13. Vocabulary 2

What are the following words in German? Write the definite article as well.

Example:

pencil _____ _____

This section contains only nouns and is constructed in the same way as the two preceding Vocabulary tests. Separate points were given for the article and the word, a point for the article was given only in conjunction with a point for the word, however. Moderate spelling mistakes were accepted.

Test no. 14. Pronunciation 1

In this test we are going to assess your pronunciation of the different sounds. Read the words written below and try to pronounce them as well as you can. If you notice that you have pronounced something wrongly, you can correct yourself by saying the word again. You will always be marked on your last attempt. You have plenty of time.

1. alles
2. auch
.....
43. reise

This sub-test, which was given in the language laboratory, contains 60 words and in each word one particular phoneme is marked. The test contains a balanced selection of the most difficult German phonemes.

Test no. 15. Pronunciation 2

Now you are to read some sentences aloud, so that we can judge your stress and intonation. We are not going to judge your pronunciation of the separate words, but only the intonation. You will be given a few minutes to practise the sentences and then we will start recording.

1. Wo wohnst du Peter?

Each sentence is marked as right or wrong. All the pupils were marked by the same person. This section, which was administered in the language laboratory, contained 20 sentences.

Test no. 16. Pronunciation 3

The pronunciation of the answers in the sub-test Speak 2 were assessed on a five-point marking scale.

Test no. 17. Fluency 1 and Test no. 18. Fluency 2

You are to write down as many words as you can which begin with a certain letter of the alphabet. You can choose any words you like, so long as they are part of the German language. You need not think so much about spelling, as we are not concerned with that. The main thing is that we understand which words you mean.
You may use an inflected form of a word, but you will be given no extra points if you write several inflected forms of the same word. The article is not needed with nouns. If you are asked to give as many words as you can beginning with A (a), we mark your answers as follows:

	1. arbeiten
1 point	2. arbeitest
	3. arbeitet
1 point	4. Arbeit
1 point	5. Affe
	6. ———
	7. ———
	etc.

W1: Write as many words as you can beginning with the letter S (s).
W2: Write as many words as you can beginning with the letter L (l). /in Swedish/

The two tests were marked as explained above in the instructions.

Test no. 19. Intelligence 1

The F test, Similarities, was used as a measure of verbal intelligence. Instruction and marking was carried out according to the constructor's directions (Härnqvist et al., 1959).

Test no. 20. Intelligence 2

The DBA figure series was used as a measure of logical-inductive intelligence. Instruction and marking was carried out according to the constructor's directions (Härnqvist, 1960).

Test no. 21. German marks

The pupils' marks for the Spring Term were collected. The marks were set at the same time as the test-battery was administered and were quite independent of the test results. It is questionable whether marks should be included as a variable in the factor analysis. They can, however, be accepted as a validity criterion for the various sub-tests.

Test no. 22. Grammar test 3

Put in the missing German words. They are all equivalent to the Swedish "är" (am, are, is).
1. Ich _____ der Lehrer.
2. _____ Sie auch Lehrer?

This section is restricted to the present tense of "sein" and has 24 sentences with gaps to be filled in.

Test no. 23. Grammar test 4

Fill in the right form of the verb.
Example: Ich heiss___ Per.
1. Wir arbeit___ viel.

This section is restricted to the present tense of regular verbs and is a "fill-in" test with 17 tasks.

Test no. 24. Grammar test 9

Fill in the right forms of the verbs on the left in the sentences below.

lesen Ich _____ nicht gerne, aber mein Bruder _____ gerne Bücher und Zeitungen. Mein Vater und meine Mutter _____ jeden Morgen die Zeitung. Vati _____ meistens die Sportnachrichten. Was _____ du in der Zeitung?

This test contained the Umlaut verbs lesen, fahren, essen and helfen and since there were 33 blank spaces, the test gave a top score of 33 points.

Test no. 25. Grammar tests 1, 2, 6 and 7

G1.
Change the sentences below according to the pattern shown:
The boy is strong. He is a strong boy.
The boys are strong. They are strong boys.
1. Der Junge ist stark. Das ist _____ _____ Junge.
Die Jungen sind stark. Das sind _____ Jungen.

G2:
Fill in the words underlined. Use the correct form.
1. ich bekomme ein Buch. Das Buch ist für _____ (me).
.....
8. Der gute Freund bekommt ein Buch.
Das Buch ist für _____ _____ Freund.

The test includes the accusative forms of pronouns, articles and adjectives, with 7, 8 and 9 points respectively obtainable.

G6:
Fill in the right form of the words in brackets in the following sentences.
1. (das Auto) Sie fahren mit _____ _____.
(die Autos) Sie fahren mit _____ _____.

The test contains the dative forms of articles and nouns, with 12 points obtainable.

G7:
Who is Peter playing with? Answer with the right form of the underlined word.
1. Er will spielen. Peter spielt mit _____.

The test contains the dative forms of the personal pronouns with 7 points obtainable.

By adding the results of the grammar tests 1, 2, 6 and 7, a measurement of the achievements in case declension can be obtained.

Test data

Table 1 shows the average marks and the standard deviations for the different groups, plus reliability and the time allowed for the test. The reliability has been calculated from a smaller group, consisting of 3 classes and including about 50 pupils. By reliability is meant split-half reliability when nothing else is stated. (Raw score distributions for the various sub-tests and for the sub-groups have been reported elsewhere; cf. Löfgren, 1969.)

Table 1. Average marks, standard deviations, reliability, length of testing time and maximum points. (e = experimental group of 130 pupils, c = control group of 151 pupils, b = boys 138 pupils, g = girls 143 pupils, t = total 281 pupils)

test no.	test	group	Mean	S.D.	rel.	test time in min.	max. score
1	Listening 1	e	19,0	5,0	0,65	21	35
		c	20,6	4,7			
		b	19,3	5,0			
		g	20,4	4,8			
		t	19,9	4,9			
2	Listening 2	e	20,0	5,1	0,77	20	35
		c	18,9	5,0			
		b	18,8	5,0			
		g	20,0	5,1			
		t	19,4	5,1			
3	Reading Comprehension 1	e	15,4	4,1	0,78	15	30
		c	15,9	4,4			
		b	14,9	4,3			
		g	16,5	4,0			
		t	15,7	4,2			
4	Reading Comprehension 2	e	22,0	5,6	0,86	20	36
		c	20,8	5,5			
		b	21,1	5,6			
		g	21,6	5,6			
		t	21,3	5,6			
5	Write 1	e	32,3	13,2	0,90	20	71
		c	27,1	11,0			
		b	27,3	11,9			
		g	31,6	12,4			
		t	20,5	12,3			
6	Write 2	e	40,0	9,9	(compare	5	
		c	36,6	12,6	speak 2)		
		b	34,8	12,1			
		g	37,7	10,5			
		t	36,3	11,4			
7	Speak 1	e	29,0	8,3	0,86	5	59
		c	26,5	9,4			
		b	26,6	9,5			
		g	28,6	8,8			
		t	27,6	9,2			

test no.	test	group	Mean	S.D.	rel.	test time in min.	max. score
8	Speak 2	e	34,8	17,0	0,98[x]	2	
		c	37,4	19,4			
		b	35,8	20,3			
		g	36,5	16,3			
		t	36,2	18,3			
9	Translate 1	e	13,5	7,2	0,72	7	36
		c	14,3	6,6			
		b	12,6	6,8			
		g	15,2	6,6			
		t	13,9	6,8			
10	Translate 2	e	12,7	6,4	0,80	8	34
		c	9,8	5,7			
		b	9,9	6,0			
		g	12,4	6,2			
		t	11,2	6,2			
11	Vocabulary 1	e	23,5	7,1	0,86	13	53
		c	24,1	7,6			
		b	22,8	7,6			
		g	24,8	6,9			
		t	23,8	7,3			
12	Vocabulary 3	e	14,5	4,8	0,82	15	41
		c	11,2	5,5			
		b	11,9	5,7			
		g	13,5	5,0			
		t	12,7	5,4			
13	Vocabulary 2	e	27,6	11,8	0,82	15	50
		c	26,9	12,6			
		b	25,4	12,7			
		g	29,1	11,5			
		t	27,3	12,2			
14	Pronunciation 1	e	40,1	6,1	0,88	2	60
		c	39,1	6,9			
		b	39,1	6,6			
		g	40,0	6,4			
		t	40,0	6,5			
15	Pronunciation 2	e	12,7	3,8		1	20
		c	12,3	4,6			
		b	12,2	4,4			
		g	12,8	4,2			
		t	12,5	4,3			
16	Pronunciation 3	e	2,4	0,8			
		c	2,3	0,9			
		b	2,3	0,8			
		g	2,4	0,9			
		t	2,4	0,8			

x) between raters

test no.	test	group	Mean	S.D.	rel.	test time in min.	max. score
17	Fluency 1	e	11,2	4,1		3	
		c	10,2	4,0			
		b	9,7	4,1			
		g	11,3	3,9			
		t	10,5	4,1			
18	Fluency 2	e	14,8	4,1		2	
		c	15,0	4,7			
		b	14,7	4,0			
		g	15,1	4,8			
		t	14,9	4,4			
19	Intelligence 1	e	38,6	7,9	0,95	15	70
		c	38,5	7,8			
		b	39,3	7,6			
		g	37,9	8,1			
		t	38,6	7,9			
20	Intelligence 2	e	24,2	4,2	0,90	15	30
		c	24,0	4,8			
		b	23,9	4,8			
		g	24,4	4,1			
		t	24,1	4,5			
21	German mark	e	3,4	1,1			
		c	3,1	1,1			
		b	3,0	1,1			
		g	3,5	1,0			
		t	3,3	1,1			
22	Grammar test 3	e	16,4	6,2	0,94	3	24
		c	19,5	5,8			
		b	17,0	6,4			
		g	19,1	5,7			
		t	18,1	6,1			
23	Grammar test 4	e	12,3	3,4	0,90	4	17
		c	13,5	3,1			
		b	12,4	3,3			
		g	13,5	3,2			
		t	13,0	3,3			
24	Grammar test 9	e	22,1	7,4	0,91	14	33
		c	18,9	5,9			
		b	19,3	6,4			
		g	21,4	7,0			
		t	20,4	6,8			
25	Grammar tests 1, 2, 6, 7	e	56,3	15,0	0,80-0,85	16	81
		c	45,1	15,0			
		b	47,4	16,5			
		g	53,1	15,0			
		t	50,3	16,0			

6. Groups and procedure

The study was conducted at the turn of the month May/June 1968. Thus the testbattery was administered to the pupils taking part during the last three weeks of the Spring term. Of the 15 classes participating in the study, 7 had been taught with the UMT study material system (mimeographed version), and 8 had been taught from another text-book. The first-named group will hereafter be called the experimental group and the last-named the control group. All the pupils have five periods of German a week. The experimental and control groups are not equivalent, but we tried to make the groups as similar as possible with regard to certain background variables. Thus schools have been chosen from areas with similar socio-economic status. In some cases the experimental and control groups have been taken from the same school. The groups of pupils are probably rather similar, as the intelligence data also suggests, but they have on the other hand been taught with different material and by different teachers. Since only the ordinary German lessons were used to carry out the test, it was impossible to obtain complete data for all the pupils, even though we got in supplementary data later. Relatively speaking, the oral tests were responsible for the greatest gaps, as a result of disturbances in the technical equipment and the difficulty in gathering this data at a later date. The oral tests were carried out in the same language laboratory, so the pupils had to be transported to and from the studio. Complete data was received from 130 pupils in the experimental group and 151 pupils in the control group. The wastage in the first-named group was 19 % and in the last-named 22 %.

The tests were led by the ordinary teachers, plus staff from the Department of Educational and Psychological Research at the School of Education in Malmö. The timetable of the test is presented in Box 5.

Box 5. Test timetable

Lesson 1	Listening 2; Write 2
Lesson 2	Listening 1; Vocabulary 1
Lesson 3	Intelligence 1; Intelligence 2
Lesson 4	Write 1; Translation 1 and 2
Lesson 5	Vocabulary 2; Vocabulary 3
Lesson 6	Reading comprehension 1; Reading comprehension 2
Lesson 7	Grammar tests 1, 2, 3, 4, 6, 7 and 9
Lesson 8	Pronunciation 1 and 2; Speak 1 and 2; Fluency 1 and 2

The test battery was corrected during the summer and autumn of 1968 and the data recorded on punched cards for further analysis.

7. Method of treatment

Since most sections of the test could be corrected with the help of a key, only sections 8, and 14-16 could be affected by the judgement of the person marking. Test number 8 ist the free speech test and tests 14-16 concern pronunciation. Each pupil's result was put on a punched card, after which the material could be data processed.

The factor analysis was carried out with the help of the standard program BMDO3M (Dixson, 1967). The program uses the principal axis method with varimax rotation in the factor analysis. A more detailed description of the method can be obtained from the handbooks on factor analysis, e.g. Fruchter (1954) and Kaiser (1958). The following data was written out: average marks and standard deviations, correlation matrix, eigenvalues and the cumulative proportion of the total variance, eigenvectors, factor matrices which are controlled by summation of the series' cross products (=correlation matrix), rotated factor matrix and finally a check of the estimated communality values.

One of the problems is to decide how many factors should be extracted. There are a number of rules of thumb for when extraction should be stopped. Here we haven chosen to extract the maximum number of factors, in order to get a preliminary idea of the data structure, i.e. the data instruction has been to extract all factors with positive eigenvalues.

The most interesting, but also the most difficult part of a factor analysis is the psychological interpretation of the factors obtained. There are descriptive, materialistic and functional interpretations. A descriptive interpretation can sometimes be used when a preliminary study is being made in a certain field. This means simply stating which tests belong together and makes no attempt to explain the factors which have emerged. In a materialistic interpretation, one looks for the characteristic features of the tests, which show the highest loadings for the factor in question. Some common denominator is sought for the tests which have loadings on a common factor. A functional interpretation attempts to make some statement about human ability. One tries to describe how the people who have done well have solved the tasks in the test and information is often taken from introspective accounts. The two last-named methods of interpretation can supplement each other and can be used to give as comprehensive information as possible about the factor.

8. Results and discussion

8.1. Problem 1: What are the component factors of language proficiency?

The 300 correlation coefficients between the 25 tests are shown in Table 2. All the correlations are positive. The squared multiple correlations between each one of the tests and the remaining tests have been calculated and used in estimating the communality values. Thus a low diagonal value indicates that the test is specific.

The correlation matrix shows how there is often a close connection between sub-tests measuring the integrated skills, vocabulary and certain grammar tests. Pronunciation, intelligence and fluency tests, on the other hand, show little, or only a moderate, connection with the other sub-tests. The tests which correlate most highly with the others are:

test no.	test	mult. corr. squared
5	Write 1	0,85
13	Vocabulary 2	0,82
11	Vocabulary 1	0,78
12	Vocabulary 3	0,76
21	German mark	0,76
7	Speak 1	0,73
10	Translate 1	0,72

The tests with the lowest correlation with the others are:

test no.	test	mult. corr. squared
20	Intelligence 2	0,24
18	Fluency 2	0,28
19	Intelligence 1	0,41
15	Pronunciation 2	0,43
17	Fluency 1	0,45
14	Pronunciation 1	0,47

A high multiple correlation need not mean that the sub-test correlates highly with all the other tests. It is enough for two sub-tests to correlate highly. The correlation matrix, shows however that the structured written test and the vocabulary tests have a close connection with most of the tests included in the test battery.

A high rate of agreement can be seen between the marks in German and the language tests. This can be taken as a validity criterion that the tests measure language proficiency such as it is interpreted by the teache

Translation is a skill which is not included in the school's objectives, and its use as a means of testing has been strongly criticised. The high rates of agreement here show, however, that translation from German to Swedish can measure rather accurately at least some parts of language proficiency.

ERIC

ble 2. Correlation matrix (the total group)

Section of test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1. Listening 1	586																								
2. Listening 2	585	593																							
3. Reading comprehension 1	650	595	665																						
4. Reading comprehension 2	603	643	602	672																					
5. Write 1	632	620	675	733	851																				
6. Write 2	497	461	534	525	630	579																			
7. Speak 1	576	636	621	691	778	569	727																		
8. Speak 2	403	412	504	448	470	586	567	580																	
9. Translate 1	624	642	701	712	738	592	668	543	716																
10. Translate 2	506	517	502	576	749	552	650	409	545	621															
11. Vocabulary 1	672	684	718	694	762	609	736	510	743	621	780														
12. Vocabulary 3	577	566	650	633	734	584	705	492	639	645	725	759													
13. Vocabulary 2	639	605	723	683	787	645	748	561	727	682	776	817	817												
14. Pronunciation 1	403	409	477	466	568	433	487	278	481	438	570	517	531	466											
15. Pronunciation 2	357	390	408	451	514	480	504	399	469	377	510	466	449	520	432										
16. Pronunciation 3	498	466	505	487	593	562	553	616	561	462	555	573	613	449	449	561									
17. Fluency 1	354	395	414	444	554	445	524	379	444	475	467	510	546	400	358	455	447								
18. Fluency 2	260	253	251	275	316	316	273	257	295	251	376	246	276	321	262	308	424	235							
19. Intelligence 1	438	431	370	516	455	307	397	287	454	260	505	378	359	305	357	331	259	278	411						
20. Intelligence 2	293	295	276	310	263	133	293	042	287	177	290	239	244	164	152	143	192	126	314	241					
21. German mark	550	534	634	670	752	590	670	462	691	609	707	720	735	527	431	507	490	269	409	332	762				
22. Grammar test 3	443	384	530	543	579	481	492	337	535	415	524	438	540	372	343	320	286	175	313	237	625	551			
23. Grammar test 4	476	371	527	494	554	463	468	298	551	416	520	441	540	368	367	391	264	173	280	287	633	644	561		
24. Grammar test 9	482	552	510	600	754	529	620	367	636	585	611	682	698	430	396	443	413	162	353	264	684	513	543	674	
25. Grammar tests 1,2,6,7	401	484	493	541	691	517	569	320	540	545	551	651	647	415	313	399	421	206	347	314	730	465	490	693	658

Among the more specific tests in the test battery are the intelligence, fluency and pronunciation tests. It is of particular interest to note that in many cases the intelligence tests show low rates of agreement with the language proficiency tests also considering the somewhat selected group.

In the discussion on the various factors, an account will be given of the German mark's and the verbal intelligence test's loadings on the respective factors.

Table 3. Rotated factor matrix (the whole group)

Section	Factor												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Listening 1	-25	-30	-40	-27	15	-44	-11	-10	-02	04	-09	-04	-00
2. Listening 2	-36	-29	-50	-11	13	-30	-15	-12	-03	-12	05	02	-02
3. Reading comprehension 1	-35	-35	-29	-36	13	-45	-16	04	06	-05	03	04	02
4. Reading comprehension 2	-40	-29	-47	-30	16	-19	-19	-19	-07	-01	02	13	01
5. Write 1	-64	-27	-28	-32	24	-21	-27	-19	-07	-04	-09	05	05
6. Write 2	-36	-50	-08	-33	26	-11	-19	-05	-10	-02	01	-08	-02
7. Speak 1	-50	-41	-32	-22	18	-18	-22	-25	13	-00	08	05	03
8. Speak 2	-19	-73	-08	-15	17	-11	-10	-04	-01	-00	05	03	01
9. Translate 1	-36	-41	-38	-36	17	-29	-18	-03	-05	-11	02	12	-02
10. Translate 2	-59	-24	-10	-20	23	-21	-14	-28	-03	00	-02	-03	-02
11. Vocabulary 1	-40	-33	-39	-28	26	-39	-28	-08	-06	03	12	-01	-03
12. Vocabulary 3	-66	-34	-22	-16	15	-26	-24	01	03	12	-00	02	01
13. Vocabulary 2	-60	-41	-18	-30	20	-36	-16	-03	07	06	03	04	-00
14. Pronunciation 1	-30	-13	-15	-21	31	-17	-48	01	01	02	-00	01	-02
15. Pronunciation 2	-18	-31	-18	-19	21	-05	-49	-05	-01	-02	00	00	02
16. Pronunciation 3	-28	-57	-15	-14	26	-15	-23	01	03	02	-13	-02	-01
17. Fluency 1	-38	-25	-13	-08	50	-07	-12	-06	10	-01	-01	05	02
18. Fluency 2	-06	-16	-16	-06	52	-06	-13	-01	-03	00	01	-01	-01
19. Intelligence 1	-12	-18	-57	-12	17	-07	-18	-01	-12	00	-00	02	01
20. Intelligence 2	-15	06	-45	-19	09	-04	00	03	13	02	-01	-03	-00
21. German mark	-57	-25	-30	-50	21	-09	-15	03	02	08	06	03	-05
22. Grammar test 3	-26	-17	-20	-65	08	-11	-13	-08	-03	00	04	04	02
23. Grammar test 4	-27	-16	-21	-66	06	-12	-14	-00	04	-01	-05	-04	-01
24. Grammar test 9	-65	-20	-25	-35	03	-14	-16	02	-01	-14	-03	02	03
25. Grammar test 1, 2, 6, 7	-69	-13	-28	-32	14	02	-07	09	-01	-01	02	-04	-03

The factor matrix showed that a very great portion of the total variance was already reflected in the very first factor. Thus the ends of the different test vectors can be described as a close dot pattern in the factor space. As mentioned earlier, we have here chosen to carry the factor analysis right to the bottom, i.e. down to the eigen value 0, so as to try to extract as much information as possible about the internal relationship of the language skills. Seven of the factors extracted have been considered interpretable.

In the account given of the factors, the only sub-tests included are those showing the highest loadings on the respective factors. This means that other sub-tests can also have significant loadings. The borderline for significant loadings can be set at 0, 30, i.e. 4-5 times the correlation's standard error.

Factor 1

Many sub-tests show high loadings on this factor. The following tests show the highest loadings:

test no.	test	loading
25	Grammar tests 1, 2, 6, 7	0, 69
12	Vocabulary 3	0, 66
24	Grammar test 9	0, 65
5	Write 1	0, 64
13	Vocabulary 2	0, 60
10	Translate 2	0, 59
21	German mark	0, 57
7	Speak 1	0, 50
19	Intelligence 1	0, 12

The point which all the above sections of the test have on common is that they all test active ability. One is tempted to compare the factor with the language prognosis factor interpreted by Carroll, "verbal knowledge; knowledge of words and structures in the mother tongue." (cf. p. 4 f.). The mother tongue tests, Fluency 2 and Intelligence 1, do not, however, show any high loadings on this factor in German language proficiency. Thus this factor may be named: an active knowledge of words and structures in the foreign language.

Factor 2

The sub-tests which have the highest loadings on this factor are:

test no.	test	loading
8	Speak 2	0, 73
16	Pronunciation 3	0, 57
6	Write 2	0, 50
21	German mark	0, 25
19	Intelligence 1	0, 18

The two tests Speak 2 and Write 2 are both intended to measure free production. In the assessment of the pupils' efforts, more weight has been laid on quantity than on quality (cf. above). The fact that one of the pronunciation tests has got such a high loading on this factor is probably largely due to the halo effect. Pronunciation 3 is, as explained above, an assessment of the pupils' pronunciation in the test Speak 2. The assessor has probably subconsciously judged the pupils who were better in other respects more positively. It would have been to the advantage of future test constructions if no separate speech factor had been found, but here we are forced, for the time being, to interpret this factor as a kind of language fluency, which mainly emerges in an oral situation. It should be pointed out, however, that at the moment this factor must be considered doubtful, since the same picture has been used in Speak 2 and Write 2, and this can have resulted in what is known an instrument factor.

Factor 3

The sub-tests which show the highest loadings on this factor are:

test no.	test	loading
19	Intelligence 1	0, 57
2	Listening 2	0, 50
4	Reading comprehension 2	0, 47
20	Intelligence 2	0, 45
21	German mark	0, 30

Here we have found an intelligence factor. It is incidentally the only factor in which the intelligence tests have significant loadings. The language tests found here are those measuring the passive skills. It is a characteristic fact that high loadings are shown in Listening 2 and Reading comprehension 2, both of which consist of rather long texts followed by questions on the content. In these tests the pupils can use verbal comparison and reasoning and by help of their intelligence solve the items. The intelligence element is not so great in Listening 1 and Reading comprehension 1. In these tests the pupils are not given so many clues by the content, but instead often have to understand details in the short stimulus phrase in order to be able to answer the questions.

Factor 4

A few of the sub-tests show high loadings on this factor.

test no.	test	loading
23	Grammar test 4	0, 66
22	Grammar test 3	0, 65
21	German mark	0, 50
19	Intelligence 1	0, 12

The two grammar tests have here obviously broken away from the others. It is difficult to explain why this has happened. Possibly the type of test represented in these sub-tests are of a more passive nature. In both the pupils were required to fill in the correct verb form, respectively verb, in blank spaces in the text. There were only a few forms to choose from and it was therefore necessary for them to use the context as a guide. The test battery has no entire sub-tests measuring passive grammatical knowledge and, considering the other sections which show loadings on this factor, this is possibly a hint of a passive grammar factor.

As in factor 1, the German mark has a high loading here. Both factors measure grammatical knowledge. This possibly signifies that teachers often base their marks on a pupil's command of the grammatical part of the course

Factor 5

This factor is obviously only represented in two tests, namely

test no.	test	loading
18	Fluency 2	0,52
17	Fluency 1	0,50
21	German mark	0,21
19	Intelligence 1	0,17

This is obviously a fluency factor. It is comparable to Thurstone's w-factor. Fluency in German shows rather high correlations with especially the vocabulary test, while the fluency test in Swedish shows low correlations with the other tests. This factor is of little interest in the field of language proficiency.

Factor 6

The tests which show the highest loadings on this factor are:

test no.	test	loading
3	Reading comprehension 1	0,45
1	Listening 1	0,44
11	Vocabulary 1	0,39
21	German mark	0,09
19	Intelligence 1	0,07

These three tests measure passive proficiency, primarily passive knowledge of words and structures. The factor matrix shows that these three sub-tests have loadings on the intelligence factor, but here the intelligence element has been excluded. The intelligence tests' loadings on this factor are nil. Since we called factor 1 an active knowledge of words and structures factor, we can here speak of a passive knowledge of words and structure factor.

Factor 7

This factor is represented by two tests, namely

test no.	test	loading
15	Pronunciation 2	0,49
14	Pronunciation 1	0,48
21	German mark	0,15
19	Intelligence 1	0,18

This factor is easy to interpret as a pronunciation factor. Since these pronunciation tests have no loadings in any other factor, pronunciation occupies a special place in language proficiency. Obviously pronunciation has not been assessed in setting the German marks, although it should have been according to the objectives set out in the curriculum.

Box 6 presents a summary of the factor structure for the whole group.

Box 6. A presentation of the factor structure for the whole group.

Factor	test no.
1. active knowledge of words and structures	25, 12, 24, 5, 13, 10, 7
2. language fluency	8, 6
3. intelligence	19, 2, 4, 20
4. knowledge of structures	23, 22
5. word fluency	18, 17
6. passive vocabulary	3, 1, 11
7. pronunciation	15, 14

8.2. Problem 2: Which measurements should be included in a test battery intended to measure total language proficiency?

A factor can be described as an ideal test. Thus, when measuring language proficiency, each factor should be measured, so that the different results can then be brought together to form a total measurement. It is, however, very difficult to construct a pure factor test, so we have to be content with setting an upper limit for the measurements of the different factors. Which test best represents the various factors can be decided by calculating the correlation between the tests and the factor. Then, with the help of regression coefficients, the individual factor scores can be estimated, and finally these can be added up to make individual factor scores (Werdelin, 1958). This type of calculation is so complicated, however, that it is hardly compatible with the objective of finding a simple means of assessing total language proficiency.

Instead, we try to construct tests, which measure the various language factors, in order to be able to make direct use of the raw score data. As a result of this investigation, we are using the data obtained by choosing, for our projected test battery, the tests, or improved versions of these tests, which have high loadings on the factor and which seem to measure different skills. For example, if we find that two listening tests and a reading comprehension test have a high correlation with a certain factor and we only wish to include two of these tests in our planned battery, we omit one of the listening tests. The reason for this is that we also desire that the pupils should feel that the test battery has a high degree of validity. We know, namely, that the test battery can influence pupils in various ways (cf. p. 3).

From Factor 1 a grammar test can be chosen, which could, for example, consist of questions on verb forms and case declension. This sub-test could be rather short, since Table 1 shows that reliable measurements can be obtained even from brief tests. In addition, a vocabulary test would be chosen.

It can also be rather short. Finally, we would include in our test battery a written test of the same type as Write 1, which measures accurately and is easy to correct. From Factor 2 the most obvious choice is the free production speaking test. A special study of this skill might possibly help us to arrive at a more easily administered and easily corrected test than Speak 2. At a meeting held in Malmö, April 1969, with representatives present from other language research projects and the Swedish Board of Education, the validity of various language tests for the measurement of total language proficiency was discussed. The group considered that a pupil was hardly likely ever to be placed in a situation like the one described in Speak 2. The argument was that in a speaking situation there is also another person involved, who stimulates the answers. Thus, a dialogue would have been a more natural situation. The same criticism was not made of a test of the type Write 2, and the group considered, therefore, that completely free production should be limited to the field of writing skill. It is impossible to say definitely where a dialogue test would be located in the factor structure, but probably such a test would prove similar to Speak 1. From Factor 3 a listening test is chosen of the type called in the test battery, Listening 2. Factor 4 has already been covered by a grammar test and Factor 5 has no relevance for language proficiency. From Factor 6 is chosen a reading comprehension test of the multiple choice type similar to Reading comprehension 1 and finally a pronunciation test is given so that the seventh factor in language proficiency is also covered. It is possible that we shall succeed in developing written tests which measure pronunciation, which would mean that the task of the teacher assessing the test would be simplified. Such tests are under construction and have proved to be usable (Löfgren, 1968; Kjellmer, 1968).

We can now link up with the theoretical division made in Box 2 (cf.above) and with the objectives drawn up in the national curriculum as shown in Box 4. It then becomes clear that, not only have all the factors been represented by tests with high loadings, but the test battery also appears to have a good validity. Moreover, the test battery permits coverage of all the combinations between the language variables and the four integrated skills. The translation tests have not been included in the recommended test battery. The reason for this is the demand for external validity and not that skill in translation has proved to be a specific skill. The position of the two translation tests in the factor space shows that the tests, in the form given here, can very well be used to measure parts of language proficiency.

In connection with Box 2, the internal relationship of the various skills was also discussed. As can be seen from the factor matrix, the points put forward there are in quite close agreement with the result obtained.

8.3. Problem 3: Does the teaching method influence the factoral composition of language proficiency?

The pupils taking part in the investigation can be divided into two subgroups, an experiment group and a control group. As mentioned earlier, the experiment group has been taught with the study material system, "German 7", produced by the UMT project, while the control group has been taught with other material. It is possible, with the help of the factor analysis, to compare the different kinds of teaching material. One question to be answered is whether the teaching method influences the factoral composition of language proficiency. If that is the case, both the factor structure and the achievement level in different tests can be studied and in that way the different teaching material can be evaluated. To take an example, if two groups differ in that, in a grammar test, one shows higher loadings on a logical argument factor and at the same time achieves a better result, then this can show that a particular method of learning grammar is efficient. If different teaching methods result in largely similar factor structures, this widens our opportunity for applying the results from this sample to other groups and the same test battery could then be used for different groups.

The first step is to study the factor structure of the experiment group, then that of the control group and finally to compare the two factor structures. Table 4 can be referred to for the correlation matrix for the experimental group and Table 5 presents the rotated factor matrix.

Table 4. Correlation matrix (experimental group)

ction of test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1. Listening 1	668																								
2. Listening 2	611	610																							
3. Reading comprehension 1	661	594	705																						
4. Reading comprehension 2	722	662	682	767																					
5. Write 1	708	663	754	778	911																				
6. Write 2	537	483	568	555	649	535																			
7. Speak 1	619	646	685	715	816	589	798																		
8. Speak 2	432	393	492	493	487	470	619	627																	
9. Translate 1	673	649	757	723	812	605	697	502	767																
10. Translate 2	535	521	522	613	746	500	611	334	553	670															
11. Vocabulary 1	688	696	711	712	798	533	751	468	734	633	790														
12. Vocabulary 3	683	586	664	653	772	583	700	525	723	538	738	776													
13. Vocabulary 2	659	626	691	732	832	537	765	489	758	625	734	806	822												
14. Pronunciation 1	467	355	475	486	546	367	463	232	475	423	539	425	456	445											
15. Pronunciation 2	417	389	456	437	491	472	470	351	470	397	501	467	414	515	445										
16. Pronunciation 3	518	390	479	500	583	527	544	613	571	418	529	574	558	409	460	599									
17. Fluency 1	389	358	409	412	499	347	473	333	433	411	429	477	505	349	357	392	402								
18. Fluency 2	245	195	195	267	274	200	272	247	248	295	319	193	139	257	229	280	379	369							
19. Intelligence 1	488	506	396	560	456	371	383	359	481	287	533	432	414	266	364	338	249	323	494						
20. Intelligence 2	235	242	288	285	229	127	257	002	326	172	342	212	241	155	097	155	146	024	260	402					
21. German mark	662	576	709	717	791	572	674	507	739	620	734	774	740	448	448	486	451	215	472	365	823				
22. Grammar test 3	429	446	608	643	672	460	590	395	574	597	495	560	623	320	315	297	275	017	279	186	682	730			
23. Grammar test 4	489	438	562	547	673	490	542	347	596	497	542	541	565	362	392	432	274	129	234	258	651	644	601		
24. Grammar test 9	566	573	664	642	825	602	698	416	739	593	666	689	749	420	430	492	375	123	362	292	746	712	707	793	
25. Grammar tests 1,2,6,7	566	498	620	657	790	543	663	408	700	533	653	685	753	406	367	463	428	135	392	402	807	600	654	775	792

Table 5. Rotated factor matrix (experimental group)

Section	Factor														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Listening 1	-35	-29	-14	25	-49	-25	15	01	-13	-20	15	02	-04	-02	-01
2. Listening 2	-36	-18	-14	17	-56	-18	12	-21	-11	04	09	-05	-07	06	-03
3. Reading comprehension 1	-51	-31	-19	28	-33	-18	10	-11	-01	-02	32	01	-01	-03	-00
4. Reading comprehension 2	-52	-23	-16	24	-53	-14	19	-07	-10	-07	06	19	07	00	01
5. Write 1	-69	-30	-10	23	-32	-28	25	-15	-09	-14	06	-03	-08	10	07
6. Write 2	-44	-29	-02	36	-26	-09	13	-02	-04	-02	04	-01	-25	01	00
7. Speak 1	-51	-24	-14	41	-28	-23	21	-38	-09	-00	06	02	-00	-00	01
8. Speak 2	-27	-07	07	70	-23	-09	18	-08	03	03	05	02	03	-07	00
9. Translate 1	-52	-27	-25	31	-39	-23	16	-04	01	-02	20	-03	-08	12	02
10. Translate 2	-59	-23	-02	08	-20	-14	33	-11	-34	-02	00	01	-02	00	00
11. Vocabulary 1	-41	-35	-27	23	-46	-24	24	-18	-13	-04	07	-17	04	-02	02
12. Vocabulary 3	-48	-25	-11	36	-33	-48	11	00	-03	-02	05	-10	-04	-08	-02
13. Vocabulary 2	-57	-21	-15	29	-31	-51	09	-12	-08	-02	03	04	04	08	02
14. Pronunciation 1	-25	-57	-10	09	-15	-13	22	-05	-05	-07	06	03	04	02	02
15. Pronunciation 2	-24	-55	-02	23	-20	-06	19	-01	01	07	-02	-03	-07	-03	-02
16. Pronunciation 3	-23	-32	-09	60	-15	-16	22	05	-08	-08	-01	-03	-06	12	00
17. Fluency 1	-23	-21	-09	19	-08	-32	44	-06	-01	04	04	07	-04	-01	-02
18. Fluency 2	01	-14	-01	13	-17	00	60	-01	-02	-02	00	-02	-00	00	01
19. Intelligence 1	-15	-14	-20	19	-60	-02	24	05	07	03	-06	-00	-01	-03	01
20. Intelligence 2	-16	-05	-62	-02	-15	-04	02	-02	-01	00	02	00	00	00	-00
21. German mark	-65	-20	-30	26	-31	-26	18	06	-01	-02	05	-03	-00	-20	01
22. Grammar test 3	-82	-10	-04	14	-18	-05	-02	-05	-05	10	07	10	10	-08	-03
23. Grammar test 4	-68	-24	-19	20	-13	-01	05	04	01	-06	01	-07	-04	04	-03
24. Grammar test 9	-75	-23	-20	22	-20	-19	05	-07	02	-00	-00	-08	-12	11	01
25. Grammar test 1, 2, 6, 7	-65	-15	-40	22	-17	-30	12	-03	08	-11	-07	02	-08	-02	07

Factor 1

Most sub-tests show high loadings on this factor.

test no.	test	loading
22	Grammar test 3	0, 82
24	Grammar test 9	0, 75
5	Write 1	0, 69
23	Grammar test 4	0, 68
25	Grammar test 1, 2, 6, 7	0, 65
21	German mark	0, 65
10	Translate 2	0, 59
13	Vocabulary 2	0, 57
4	Reading comprehension 2	0, 52
9	Translate 1	0, 52
7	Speak 1	0, 51
3	Reading comprehension 1	0, 51
12	Vocabulary 3	0, 48
19	Intelligence 1	0, 15

This important factor can be said to be a knowledge of word and structures factor, with the main emphasis on the latter. The fact that the teaching method had made this group more homogeneous has obviously resulted in this factor becoming even more marked than for the group as a whole. The correlation matrix also showed that in this group there was a higher correlation between the variable and the proficiency tests.

Factor 2

In this factor both the pronunciation tests appear.

test no.	test	loading
14	Pronunciation 1	0, 57
15	Pronunciation 2	0, 55
21	German mark	0, 20
19	Intelligence 1	0, 14

The interpretation of this factor is simple. It is quite obviously a pronunciation factor.

Factor 3

In this factor we really only find the tests which measured the ability for logical deduction.

test no.	test	loading
20	Intelligence 2	0, 62
25	Grammar test 1, 2, 6, 7	0, 40
21	German mark	0, 30
19	Intelligence 1	0, 20

This factor can be interpreted as a reasoning factor. It is in evidence only in the grammar tests intended to measure case declension. Even if care must be taken in interpreting these results, it seems reasonable here to mention the UMT material's emphasis on understanding when learning grammar. Table 1 showed that in these particular tests the experiment group had much better results than the control group. This possibly provides some indication that learning with understanding is of value when learning structures.

Factor 4

Again there are only a few tests with loadings on this factor.

test no.	test	loading
8	Speak 2	0, 70
16	Pronunciation 3	0, 60
7	Speak 1	0, 41
6	Write 2	0, 36
21	German mark	0, 26
19	Intelligence 1	0, 19

If the halo effect's influence on the pronunciation test is again accepted here, we can, with the same reservations as earlier, interpret this factor as a language fluency factor, which is mainly noticeable in free oral production.

Factor 5

The sub-tests which have the highest loadings on this factor are:

test no.	test	loading
19	Intelligence 1	0, 60
2	Listening 2	0, 56
4	Reading comprehension 2	0, 53
1	Listening 1	0, 49
21	German mark	0, 31

This factor can be interpreted as an intelligence factor. A verbal intelligence test shows the highest loading, followed by text comprehension. This factor may be compared to Thurstone's v-factor.

Factor 6

test no.	test	loading
13	Vocabulary 2	0, 51
12	Vocabulary 3	0, 48
21	German mark	0, 26
19	Intelligence 1	0, 02

The two vocabulary tests also have similar loadings on Factor 1. These two tests measure the active vocabulary and taking into consideration the other tests' loadings on this factor, it can be interpreted as an active vocabulary factor.

Factor 7

This factor is to be found in both the fluency tests and is thus called a fluency factor.

test no.	test	loading
18	Fluency 2	0, 60
17	Fluency 1	0, 44
21	German mark	0, 18
19	Intelligence 1	0, 24

Table 6 shows the correlation matrix for the control group and Table 7 the rotated factor matrix.

Table 6. Correlation matrix (control group)

Section of test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1. Listening 1	625																								
2. Listening 2	619	655																							
3. Reading comprehension 1	643	617	691																						
4. Reading comprehension 2	550	618	559	658																					
5. Write 1	673	567	670	685	334																				
6. Write 2	481	462	515	522	673	697																			
7. Speak 1	607	619	599	663	743	574	718																		
8. Speak 2	374	443	510	437	513	656	559	636																	
9. Translate 1	575	656	656	723	720	596	672	580	714																
10. Translate 2	599	496	545	528	725	648	676	533	594	699															
11. Vocabulary 1	667	693	722	699	794	659	748	537	754	668	822														
12. Vocabulary 3	656	547	728	621	793	643	707	548	666	699	799	824													
13. Vocabulary 2	650	590	753	647	781	720	742	619	710	763	811	877	888												
14. Pronunciation 1	390	445	491	446	597	479	498	320	502	448	602	581	584	525											
15. Pronunciation 2	340	390	385	462	554	488	526	437	484	368	522	477	473	521	475										
16. Pronunciation 3	503	528	531	477	623	590	562	627	561	515	573	602	654	477	442	593									
17. Fluency 1	388	409	447	455	585	536	549	446	483	504	522	506	585	433	358	509	500								
18. Fluency 2	273	307	289	291	386	385	284	262	335	241	415	311	373	367	284	330	478	345							
19. Intelligence 1	404	464	353	482	477	268	414	236	431	247	485	368	316	336	355	325	272	246	426						
20. Intelligence 2	356	336	271	330	301	138	319	071	260	184	256	264	246	168	185	133	228	193	356	298					
21. German mark	501	485	596	618	710	623	660	452	663	538	705	689	736	585	419	524	509	321	356	307	763				
22. Grammar test 3	414	412	470	548	669	522	515	283	506	406	564	559	520	478	413	379	402	303	367	306	678	653			
23. Grammar test 4	431	362	496	504	550	459	473	248	505	457	506	522	548	415	381	382	325	209	288	332	686	614	576		
24. Grammar test 9	518	520	636	547	634	535	542	389	587	526	626	655	689	444	383	421	416	223	366	249	615	494	502	578	
25. Grammar tests 1,2,6,7	431	461	495	437	561	582	481	342	503	489	568	560	622	426	283	373	369	299	349	273	684	623	550	568	626

Table 7. Rotated factor matrix (control group)

Section	Factor												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Listening 1	-58	19	-43	13	22	07	-14	-04	00	08	-04	-07	-07
2. Listening 2	-42	15	-50	17	31	13	-00	04	-15	-02	-02	-28	01
3. Reading comprehension 1	-65	30	-26	15	28	13	12	-07	-08	02	-02	-07	04
4. Reading comprehension 2	-34	32	-44	13	28	22	-10	-00	-36	-03	01	01	01
5. Write 1	-49	40	-30	28	31	32	-22	04	-11	21	-05	08	00
6. Write 2	-30	42	-03	31	55	17	-15	08	-03	03	10	-02	03
7. Speak 1	-43	28	-36	16	39	26	-27	-02	-14	-01	01	-01	-03
8. Speak 2	-29	11	-05	16	74	13	-03	02	-08	-01	03	-01	02
9. Translate 1	-42	32	-32	17	41	21	-03	-00	-35	-00	-02	-09	03
10. Translate 2	-54	30	-10	16	38	10	-41	-00	-06	01	00	-01	00
11. Vocabulary 1	-57	32	-30	27	30	30	-09	14	-21	01	04	-12	-06
12. Vocabulary 3	-68	37	-17	18	31	25	-08	06	-03	02	-00	11	-04
13. Vocabulary 2	-68	40	-12	27	41	16	-12	00	-06	-11	00	03	03
14. Pronunciation 1	-31	32	-11	32	14	47	-02	04	-03	-03	-09	-08	-00
15. Pronunciation 2	-15	20	-21	17	32	53	-04	-03	-06	02	04	02	01
16. Pronunciation 3	-34	18	-15	26	56	22	-01	-04	01	04	-18	-05	-05
17. Fluency 1	-25	21	-16	53	32	11	-15	-05	-06	02	-03	05	05
18. Fluency 2	-10	13	-16	56	14	12	01	02	-02	00	01	-03	-02
19. Intelligence 1	-14	16	-56	13	11	23	03	15	-08	05	-03	-03	03
20. Intelligence 2	-11	21	-52	11	-04	00	-04	-09	03	-03	02	04	-01
21. German mark	-34	69	-19	21	25	18	-09	-01	-14	-07	-09	-01	-01
22. Grammar test 3	-16	67	-27	20	11	23	-03	03	-07	16	02	02	-04
23. Grammar test 4	-24	65	-24	06	11	17	-05	-17	-06	-03	-04	03	-03
24. Grammar test 9	-50	43	-23	10	21	14	01	05	-08	-01	04	-02	20
25. Grammar test 1, 2, 6, 7	-28	66	-20	19	19	01	-04	17	04	-04	05	-11	06

Factor 1

The following sub-tests have the highest loadings on this factor:

test no.	test	loading
13	Vocabulary 2	0, 68
12	Vocabulary 3	0, 68
3	Reading comprehension 1	0, 65
1	Listening 1	0, 58
11	Vocabulary 1	0, 57
10	Translate 2	0, 54
24	Grammar test 9	0, 50
5	Write 1	0, 49
21	German mark	0, 34
19	Intelligence 1	0, 14

In the experiment group we found a factor, which we called knowledge of words and structures. This appears to be a similar factor, though here the emphasis is on the knowledge of words. Many of the tests included in the battery show high loadings on this obviously essential factor.

Factor 2

The following sub-tests have the highest loadings on this factor:

test no.	test	loading
21	German mark	0, 69
22	Grammar test 3	0, 67
25	Grammar test 1, 2, 6, 7	0, 66
23	Grammar test 4	0, 65
19	Intelligence 1	0, 16

Many of the tests under Factor 1 also have loadings on this Factor 2. Clearly it is some kind of grammar factor. Here we must again refer to the earlier discussion on the connection between marks and knowledge of grammar. Teachers appear to set far too much store by grammar proficiency in relation to the other parts of language proficiency.

Factor 3

The following sub-tests show high loadings on this factor:

test no.	test	loading
19	Intelligence 1	0, 56
20	Intelligence 2	0, 52
2	Listening 2	0, 50
4	Reading comprehension 2	0, 44
21	German mark	0, 19

We interpret this factor as an intelligence factor. As might be expected, text comprehension has loadings on this ability to make linguistic deductions.

Factor 4

Both the fluency tests are to be found here and the factor is therefore called a fluency factor.

test no.	test	loading
18	Fluency 2	0, 56
17	Fluency 1	0, 53
21	German mark	0, 21
19	Intelligence 1	0, 13

Factor 5

In this factor are the following sub-tests:

test no.	test	loading
8	Speak 2	0, 74
16	Pronunciation 3	0, 56
6	Write 2	0, 55
21	German mark	0, 25
19	Intelligence 1	0, 11

Here, as in the earlier groups, we find a language fluency factor, which has loadings in tests measuring free oral and written production.

Factor 6

test no.	test	loading
15	Pronunciation 2	0, 53
14	Pronunciation 1	0, 47
21	German mark	0, 18
19	Intelligence 1	0, 23

A separate pronunciation factor has also been found in this control group.

Several factors are almost impossible to interpret. Factor 6 could possibly be some kind of active language factor, since the active sections of the test, Translate 2, Speak 1 and Write 1, show moderate loadings in this factor. These three tests have higher loadings on the first factor, however.

The language proficiency tests are well-collected in the factor space for both the control group and the experiment group. The traditional parts of language proficiency, speak, read, write and hear, have not appeared as separate factors. There are methods by which the correlation matrises of two groups may be compared and Ahmavaara (1954) has in addition developed a method of directly comparing factor structures. For the time being, we shall restrict ourselves to a comparative discussion of the factor structures which have appeared in the experiment and control groups.

Box 7. Summary of the outcome in the experiment and control groups.

Experiment group		Control group	
factor	test no.	factor	test no.
1. structure proficiency	22, 24, 5, 23, 25	1. vocabulary	13, 12, 3, 1, 11
2. pronunciation	14, 15	2. structure proficiency	22, 25, 24
3. reasoning	20, 25	3. intelligence v+r	19, 20, 2, 4
4. language fluency	8, 7, 6	4. word fluency	18, 17
5. verbal intelligence	19, 2, 4, 1	5. language fluency	8, 6
6. active vocabulary	13, 12	6. pronunciation	15, 14
7. word fluency	18, 17		

The above shows how comparable factors have been obtained in the two groups. In both the factors for knowledge of words and structures prove to be very important. The factors pronunciation, language fluency and word fluency are the same in both groups. For the experiment group, the intelligence factor divides into one reasoning part and one verbal part, while for the control group it appears only as a single intelligence factor.

In answer to the question put at the beginning of this section, it can be said that, in the two groups studied, there are only minor differences in the factor structure. One interpretation of the resul is that the teaching methods in the two groups were largely similar.

8.4. Problem 4: Are the measurements included in a test battery, which is designed to measure total language proficiency, dependent on the teaching method?

The two test batteries can here be set together on the basis of the analyses made of the experiment group and the control group and them compared. In composing the test batteries, the face validity should, as before, be taken into consideration.

Box 8. Suggested new test battery for the experiment and control groups.

Experiment group

factor	test no.	suggestion for new test battery
1	5, 22-25, 3	written test of type Write 1 grammar test on verb forms and case declension reading comprehension test of type Reading Comprehension 1
2	14, 15	pronunciation test
3	25	----------
4	8, 7	oral production test
5	2, 4, 1	listening comprehension test of type Listen 2
6	13, 12	vocabulary test
7	18, 17	----------

Control group

factor	test no.	suggestion for new test battery
1	13, 12, 3	vocabulary test reading comprehension test of type Reading Comprehension 1
2	22, 25, 24	grammar test on verb forms and case declension
3	2, 4	listening comprehension test of type Listen 2
4	18, 17	----------
5	8, 6	oral production test
6	15, 14	pronunciation test
1+2	5	written test of type Write 1

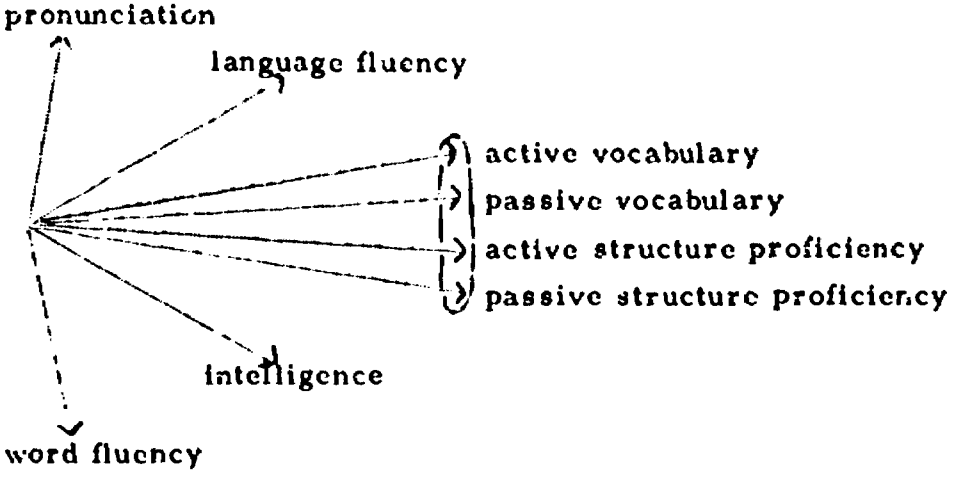
Box 8 shows how the same tests can be chosen to measure total language proficiency in both groups. The result for Problem 4 can be compared with that for Problem 3, i.e. the same test battery can be used if the teaching methods are not too dissimilar.

9. Summary

The aim of this study has been to contribute to the understanding of how language proficiency is built up and to find out which measurements should be included in a test battery, which is to measure total language proficiency as comprehensively as possible. Moreover, we have tried to shed some light on how far differences in teaching method can influence the factors incorporated in language proficiency and thus influence the composition of the test battery. The study was conducted with two groups of pupils from Grade 7, of which one group had used the UMT project's study material "German 7". During the last few weeks of the Spring Term, all the pupils participating had to go through a test battery, which was constructed on the basis of a theory on the relationship between language variables and integrated skills. The data received has been worked up with factor analytical techniques, so that a factor structure has been obtained for both the groups and the combined material.

The analyses have shown that language proficiency can be seen as a rather closely-knit unit and the correlations between the different language tests are high. It is possible, however, to sort out certain aspects of language proficiency and Figure 1 gives an outline of the results of the three factor analyses.

Fig. 1. Fundamental outline of the factoral composition of language proficiency.



The figure shows that language proficiency is divided into four factors, excluding the word fluency factor. Thus the traditional division into speak, read, write and hear do not reflect the factoral composition of language proficiency. Proficiency in vocabulary and structures has proved to be a

central factor. Hence, simple vocabulary and grammar tests can give a very accurate measurement of language proficiency. When it comes to the vocabulary and structure proficiency factor, the factor structures were not quite the same in the two groups. This factor can possibly be divided into four smaller parts, as shown in the figure, namely into active and passive vocabulary and active and passive structure proficiency. Pronunciation, intelligence and language fluency factors are more similar in the two groups.

The factor structures are so similar in the two groups of pupils, that the same test battery can be used to measure language proficiency. Thus it has not been found that the two different teaching methods under study have produced such different factor structures that they have influenced the composition of the test battery. This is a most satisfactory result and if, by generalisation, this result could be applied to other teaching methods as well, it means that the same test battery can be used in assessing different groups of pupils. Further investigation is desirable, using groups of pupils taught by more pronouncedly different methods, to see whether the results hold good.

Unfortunately, this study has not succeeded in solving the problem of how to measure skill in speaking without using oral tests. It has admittedly been shown that the free speech test can partly be replaced by a similar written test, but further experiments within this section of language proficiency are desirable.

One interesting point is that so few of the language tests show loadings on the intelligence factor. This opens the way for an interesting question, concerning the relation between ability and achievement in foreign languages during the later years of schooling. Will this connection increase or decrease during the years in the 'gymnasium' ?

Another interesting result obtained from this study concerns the question of whether by comparing factor structures, certain teaching methods can be proved to be particularly worthwhile. It has been found, for example, that the pupils who have used the UMT material seem to make more use of logical powers of deduction when solving grammatical problems. This could mean that the material's method of teaching grammatical structures, with emphasis laid on understanding, is efficient. It would be interesting to continue this line of thought in new experiments.

10. References

Ahmavaara, Y. Transformation analysis of factorial data. Helsinki: Soumalaisen Kirjallisuuden Seuran Kirjapainon Oy, . 1954.

Anvisningar till standardproven för grundskolans åk 8. Stockholm: Kungl. Skolöverstyrelsen, 1968.

Beery, A. Interrelationships between listening and other language arts areas. Elementary English, 1954, 31, 164-172.

Bloom, B.S. Testing cognitive ability and achievement. In: N. L. Gage (Ed.), Handbook of research on teaching. (4th ed.) Chicago: Rand McNally & Company, 1965. Pp 392-394.

Brooks, N. Language and language learning. New York: Harcourt, Brace & World, 1960.

Carroll, J. B. A factor analysis of two foreign language aptitude batteries. J. gen. Psychol., 1958, 59, 3-19.

Centrala prov för årskurs 2 gymnasiet. Stockholm: Skolöverstyrelsen, 1968.

Dixson, W. J. (Ed.) Biomedical Computer Programs. Berkeley & Los Angeles: University of California Press, 1967.

Engh, B. Feltyper och felfrekvenser i årskurs 9 vid fri skriftlig produktion i tyska. Pedagogisk-psykologiska problem, Nr 72, 1968.

Fruchter, B. Introduction to factor analysis. Princeton, New Jersey: D van Nostrand Company, 1954.

Hampleman, R.S. Comparison of listening and reading comprehension ability of fourth and sixth grade pupils. Dissert. Abstr. 1955, 15, 1757,

Härnqvist, K. Individuella differenser och skoldifferentiering. Stockholm: SOU, 1960.

Härnqvist, K. m.fl. F-testet. Stockholm: 1956.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

Kjellmer, G. On active versus passive proficiency in pronounciation. Moderna språk. 1968, 62, 136-143.

Lado, R. Language testing. Bristol: Longmans, 1962.

Larsen, R. P. & Feder, D. D. Common and differential factors in reading and hearing comprehension. J. educ. Psychol., 1940, 31, 241-252.

Lindell, E. Läromedelssystem för språkundervisning. Stencil, 1968.

Lindell, E. (Red.) Språkprojektet "Tyska 7" läsåret 1965-66:1. Målen för undervisningen. Pedagogisk-psykologiska problem, Nr 35, 1966.

Läroplan för fackskolan. Stockholm: SÖ-förlaget, 1965.

Läroplan för grundskolan. Stockholm: Kungl. Skolöverstyrelsen, 1962.

Läroplan för gymnasiet. Stockholm: SÖ-förlaget, 1965

Löfgren, H. Objektiva uttalsprov. Under arbete.

Löfgren, H. Mätningar av språkfärdighet i tyska: testbatteri och testdata. Testkonstruktion och testdata, Nr 4, 1969.

Löfgren, H. & Hall, P. Prediktion av hörförståelse med läsförståelsetest. Stencil, 1968.

Meier, H. Deutsche Sprachstatistik. Hildesheim: Georg Olm, 1964.

Nilsson, B. Strukturfrekvenser i tyskt tidningsspråk ("Bildzeitung"): Metodkonstruktion och några förstudier. Pedagogisk-psykologiska problem, Nr 88, 1969.

Oehler, H. Grundwortschatz Deutsch. Stuttgart: Ernst Klett, 1966.

Pfeffer, J.A. Grunddeutsch. Basic (spoken) german word list. Grundstufe. Englewood Cliffs: Prentice-Hall, 1964.

Rosengren, I. Modernt tyskt tidningsspråk. Under arbete.

Rulon, P.J. et al. A comparison of phonographic recordings with printed material in terms of knowledge gained through their use alone. Harv. educ. Rev., 1943, 13, 63-76.

Scherer, G.A.C. & Wertheimer, M. A psycholinguistic experiment in foreign-language teaching. New York: McGraw Hill, 1964.

Schwandt, E. Innehållsanalyser av läroböcker i tyska för grundskolans årskurs 8. Pedagogisk-psykologiska problem, Nr 66, 1968.

Smith, P.D. & Berger, E. An assessment of three foreign language teaching strategies utilizing three language laboratory systems. Report No. 5-0683. U.S. Department of health, education, and welfare, 1968.

Spearritt, D. Listening comprehension - a factorial analysis. Melbourne: Green & Sons, 1962.

Thorén, B. Moderna språk. Stockholm: Svenska Bokförlaget, 1960.

Tyska årskurs 7. Studiehäfte 1-8. Stockholm: SÖ-förlaget, 1968-69.

Werdelin, I. The mathematical ability. Lund: Håkan Olssons boktryckeri, 1958.

Department of
Educational and
Psychological Research
School of Education
Malmö, Sweden

Abstract card

Löfgren, H. Measuring proficiency in the German language: A study of pupils in grade 7. Didakometry (Malmö, Sweden: School of Education), No. 25, 1969.

Starting from the objective of language teaching and a theory of the relationship between language variables and integrated skills this study tries to contribute to the understanding of the foreign language ability and how to test it. It was found that the language ability can be seen as a rather self-contained unit and the correlations between the various language tests were high. However, it was possible to separate certain parts of the language ability as the factors of knowledge of vocabulary and structure pronunciation, intelligence and language fluency.

Indexed:
1. Language testing
2. Language teaching, German
3. Factor analysis, foreign language proficiency

Reference card

Löfgren, H. Measuring proficiency in the German language: A study of pupils in grade 7. Didakometry (Malmö, Swede School of Education), No. 25, 1969.