

Reprinted from

The IEA Study of Written Composition I: The International Writing Tasks and Scoring Scales

Edited by

T P GORMAN

National Foundation for Educational Research
in England and Wales, UK

A C PURVES

State University of New York at Albany, USA

R E DEGENHART

University of Jyväskylä, Finland



Published for the International Association
for the Evaluation of Educational Achievement by

PERGAMON PRESS

OXFORD · NEW YORK · BEIJING · FRANKFURT
SÃO PAULO · SYDNEY · TOKYO · TORONTO

1988

1

Origins of the International Study of Writing

Sauli Takala

Introduction

The International Study of Written Composition was planned and carried out under the aegis of the International Association for the Evaluation of Educational Achievement (IEA), a cooperative research organization which has been conducting international surveys for almost 30 years. Since the late 1950s a number of educational researchers and research institutions have been working on an empirically oriented comparative research program. A small feasibility study was carried out by the IEA in 1959-62, followed by the First Mathematics Study in the mid-1960s, with 12 countries/school systems participating. The ambitious Six Subject Survey was conducted in 1970-71 and included studies of science (19 countries), reading comprehension (15), literature (10), French as a foreign language (8), English as a foreign language (10), and civic education (10). The organization has recently completed a second mathematics study and a study of classroom environment. Current projects include a second science study and studies of preprimary education and computers in education. Thus the International Study of Written Composition, the subject of this report, has benefited from the expertise and experience developed during earlier studies already completed and reported, as well as from other currently active studies.

The writing study examines the teaching and learning of written composition in the schools of 14 countries: Chile, England, Finland, Federal Republic of Germany (Hamburg), Hungary, Indonesia, Italy, Netherlands, New Zealand, Nigeria, Sweden, Thailand, the United States, and Wales. Several other countries/school systems participated in various phases of the study, but for a number of reasons were unable to complete the full range of activities, which included curriculum analysis, pilot testing, main testing, and data analysis.

This volume is the first of three planned international reports on the results of the study. This first report will describe how the writing tasks used in the study were selected, formulated, and scored. The second volume will report the curriculum and teaching practices in the participating countries or subnational independent school systems. The third publication will report on the results concerning student writing performance.

Volume I is presented in two distinct parts. In Part I the design of the study (purposes, populations, sampling) is briefly presented, followed by an exposition of the domain of school writing and the development of the writing assignments. The specification of the writing tasks is then described. The scoring system developed for this study is presented and the results obtained in establishing the scales (benchmark compositions) are reported. Part II of the report describes the writing assignments and their scoring schemes in greater detail, with examples of each to illustrate the system.

Why Study Written Composition?

Composition, perhaps more than reading, has become a focal point for critics of schools. Writing, for obvious reasons, is one of the most visible products of education, and incorrect usage and spelling have been taken to be signs of a personal scholastic failure, and an alleged widespread deterioration of writing ability an indication of inadequacies in whole school systems.

A growing awareness of the importance of literacy, and more specifically of writing, is probably a concomitant of the growing importance of schooling and education. These are realized through the medium of "texts," after an initial stage of schooling. In an increasingly more complex society and world it is inconceivable that spoken language could effectively handle all communication needs. Written text has several features which recommend writing as an effective mode of communication in a number of situations (cf. Perera 1984, Takala 1982, Vachek 1973).

In view of the importance of writing in society and in the educational system, it is not surprising that some countries/school systems have begun to assess systematically the efficacy of the teaching and learning of writing (e.g., National Assessment of Educational Progress in the United States; Assessment of Performance Unit in England and Wales). However, large-scale assessment became an area of interest only in the late 1970s.

The study reported here was mounted to accommodate the internationally strengthened interest in the assessment of writing. In August 1980 the General Assembly of the International Association for the Evaluation of Educational Achievement (IEA) approved a study of written composition. This reflected the recognition by the IEA of the central place that the study

of the mother tongue (sometimes referred to, with or without a shift in meaning, as first language, native language, language of instruction) occupies in the school curriculum. Introducing students to written language, and thus promoting literacy, has traditionally been perhaps the principal task of the school. However, while learning to read has always been emphasized, writing has tended to receive somewhat less attention. Recently there has been a movement toward a greater balance between the two forms of literacy.

In addition to the interest in assessment, the IEA study seeks to set learning in the context of the cultural framework, curricular emphases, and teaching practices, rather than just to ascertain the level of achievement. It also aims to make it possible for each participating country or educational system to assess its relative strengths and weaknesses in writing instruction against an international background. Consequently, as in other IEA studies, important components in the name of the project, the International Study of Written Composition, are *international* and *study*. The *international* aspect has already been discussed. The word *study* is meant to convey that this is not a mere survey of the status of learning, but a project based on specific research questions.

The Context of the Study

The Cultural Context

One of the most important considerations the project had to deal with was that writing seems to differ, for example, from mathematics and science, in that the criteria of what is the correct, or at least, a good, response may vary somewhat from culture to culture. In this respect the writing study resembles the earlier literature and civics studies more than any other earlier or on-going IEA studies. It would be presumptuous to maintain that there is only one single correct response to a typical writing assignment. A characteristic typical of all composition tasks, not only in mother tongue instruction but in many other subjects as well, is that there are several acceptable approaches and several acceptable products.

There tends to be a fair degree of agreement within certain cultures, which are sometimes called interpretive communities, but cultures may differ in terms of

- (a) what functions of writing are emphasized in school;
- (b) what patterns of organization (style, rhetoric) are preferred and rewarded;
- (c) what topics are appropriate to write about;
- (d) what approaches to writing are considered appropriate (e.g., personal vs. impersonal, serious vs. humorous);

- (e) what forms of task instructions are appropriate (e.g., a simple title vs. detailed prompting);
- (f) what amount of time students should be given to write in response to an assignment;
- (g) what the appropriate criteria are for rating compositions.

This essential cultural relativism is at the same time one of the major problems of the project and one of the most interesting and challenging aspects of the study (see Takala and Vähäpassi 1986).

The Context of Earlier Comparative Work on Mother Tongue Teaching

In contrast to second and foreign language teaching, there has been surprisingly little systematic international cooperation in mother tongue teaching. There are, for instance, several journals for L2 researchers and teachers which have an international authorship and readership. The International Association of Applied Linguistics (AILA), which was founded in the mid-1960s, has been wholly dominated by the L2 research community. The first session to be systematically devoted to issues of mother tongue teaching within the AILA world congresses was held in Brussels in 1984.

The mother tongue teaching profession appears to believe that, since mother tongues differ, it is not possible to benefit very much from the experience of other countries. The general term "mother tongue" does not seem even to be used in some countries. Yet, in spite of the fact that relatively little has been done internationally in the area of mother tongue teaching, some progress can be reported.

In the series of *European Curriculum Studies*, a volume was produced on mother tongue teaching in upper secondary schools in 18 member countries (Marshall 1972). This report was a useful source for the construction of the IEA curriculum questionnaire with which detailed information on curricular emphases was collected. Unesco has also sponsored useful work in this area. Recently, the International Mother Tongue Network has been created, with headquarters in Enschede, the Netherlands, and published its first comparative volumes on the curriculum in several European countries. In the 1960s there was intensive and interesting research done in Great Britain in primary education and in the teaching of the mother tongue. This attracted attention world-wide, and led to concrete cooperation between British and American educators.

One somewhat surprising observation made during the IEA project is that the history of mother tongue education in general, and of the teaching of written composition in particular, is rather inadequately documented. In order to get a better idea of this historical context of written composition

teaching, each national center was asked to interview a number of experts and produce a brief national case study. This has proved a useful exercise both nationally and internationally.

The Context of Current Research on Writing

Another contextual feature that the project needed to take into account was that the IEA written composition study was mounted at a time when there was a growing interest in the study of writing. There are several strands in current research on writing.

One line of research focuses on the role and impact of writing. The role of literacy in cognitive functioning and in societal development has been the object of theoretical and empirical studies (see, e.g., Bruner 1972, Gaur 1984, Olson 1977, Ong 1982, Scribner and Cole 1981, Vygotsky 1978).

Another research strand is interested in the functional relationships between speech and writing. More recent is the interest in exploring the relationships between reading and writing (e.g., Perera 1984, Rubin 1980, Takala 1982, Tannen 1982).

A third prominent research paradigm is part of the cognitive psychology movement, and explores the cognitive processes relating to writing. This includes, for example, protocol analysis and computer simulation (Flower and Hayes 1980, Hayes and Flower 1983).

A fourth line of research focuses on the product of writing (text, written discourse). Different text structures (genres) are studied to see how discourse is organized (e.g., the study of story grammar, narratology, argumentation patterns). Related to this is cross-cultural rhetoric, that is, a study of culturally preferred modes of discourse organization (e.g., Rumelhart 1975, Kaplan 1966, Kaplan 1983).

A fifth research strand concerns the readers of written texts, especially how readers interact with and respond to texts. Related to this is, of course, the vast amount of literature on reading comprehension (e.g., Purves and Ripperre 1968, Purves 1984).

A sixth research paradigm has a more pedagogical orientation. It looks at what writing is taught in schools, how the curriculum is organized, and what teaching practices seem to be effective (e.g., Hilllocks 1986, Westdorp 1982).

A seventh research direction focuses on problems related to different procedures used in the rating of compositions: their reliability and validity and rating behavior in general (e.g., Cooper and Odell 1977).

Related to the above seven areas of research, an interest in assessing actual student performance in writing has emerged. In the United States the National Assessment of Educational Progress (NAEP) has conducted

periodic assessments of writing since 1969. The Assessment Performance Unit (APU) has also conducted ten surveys of writing performance in England and Wales. Australia conducted a study of basic literacy and numeracy in the mid-1970s. In Canada assessments were carried out in the provinces of Ontario and British Columbia in the 1970s, and several countries such as the Netherlands and New Zealand, as well as states within countries like the USA, have begun conducting assessments.

However, when the IEA study was started in 1980, most of the participating centers had never carried out a large-scale empirical survey of writing in their school systems. For this reason, it was decided that a high priority in the IEA study should be to produce a good account of the teaching of written composition and good national assessments of writing performance.

With regard to the seven research areas mentioned earlier, it seems evident that, in addition to its basic assessment function, the study can contribute most to the pedagogical study of writing, to cross-cultural rhetorics, and to the study of rating procedures.

Design of the Study

The following discussion is only a brief summary of the design of the study. A more detailed account of the design, sampling, instruments, and analysis will be presented in the second and third international report volumes.

Purpose

Given the cultural and research context outlined in the preceding discussion, the IEA International Study of Written Composition was designed to accomplish the following tasks:

1. to contribute to the conceptualization of the domain of writing and particularly the domain of school-based written composition;
2. to develop an internationally appropriate set of writing tasks and a system for assessing composition which is applicable across countries, school systems, and languages;
3. to describe recent development and the current state of instruction in written composition in the participating countries/school systems; and
4. to identify factors which explain differences and patterns in the performance of written composition and other outcomes, with particular attention to cultural background, curriculum, and teaching practices.

Population and Samples

The study includes three populations: Population A was defined as students at or near the end of primary education and the self-contained classroom. Population B consists of students at or near the end of compulsory education, that is, students who are in the last year of the shortest secondary program and those in longer programs who have completed the same number of years of schooling whether or not they have finished their program. Population C is composed of students at or near the end of the academic secondary school.

The recommended minimum sample sizes were 50 randomly sampled classes for Populations A and C and 100 for Population B. In most countries/school systems a two-stage sampling design was used which involved random sampling of schools within strata and then random selection of mother tongue classes within schools. The desired tolerance level for estimates of variable means has been specified as 5-7 percent of the standard deviation. To achieve tolerance levels smaller than 5 percent requires samples whose sizes would increase project costs unreasonably, either by requiring that cluster sampling be abandoned in favor of simple random sampling or by requiring that the number of students/clusters included be increased substantially. On the other hand, to permit the tolerance level to go much beyond 7 percent leads to sampling errors that are too large to allow much confidence in the estimates obtained. The principles of sampling are described in detail in the subsequent publications.

TABLE 1 *The populations tested in each participating country/school system*

Country	Population A	Population B	Population C
Chile		×	
England		×	
Finland	×	×	
Hamburg (FRG)			×
Hungary		×	
Indonesia	×		×
Italy	×	×	
Netherlands		×	×
New Zealand	×	×	
Nigeria		×	
Sweden	×	×	×
Thailand			×
United States	×	×	×
Wales		×	

Independent and Dependent Variables

To fulfill the aims set for the study, information on a large set of independent variables was gathered. The major constructs underlying the

independent variables of the study and their presumed interrelationships, as well as their relationships with the dependent variable, are illustrated in figure 1.

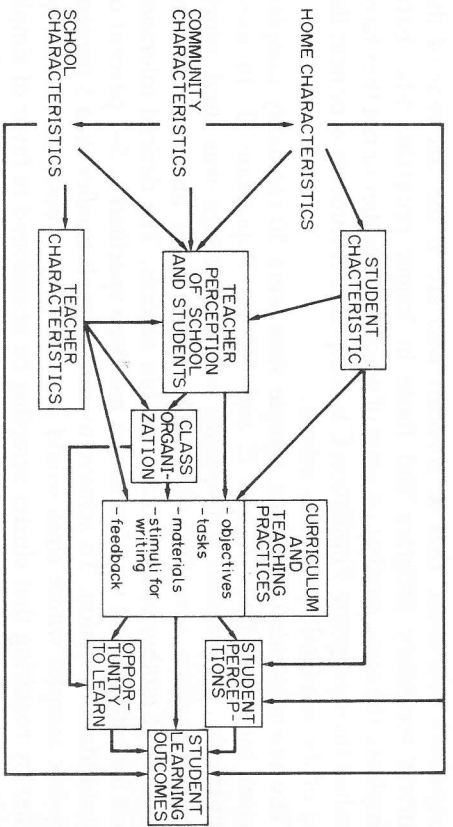


Fig 1

The figure suggests that student performance in written composition results from an interrelationship of factors, including what might be thought of as “frame” or contextual factors in the community and the school system including such specific factors as the examination system in composition and the type of training of mother tongue teachers. These various contextual factors influence what goes on in the school and particularly the teaching practices of the individual classroom. These in turn, together with the students’ perceptions of the curriculum and instruction, affect the opportunity to learn and therefore the achievement of the students.

Data on the independent variables were collected by means of a National Context Questionnaire, National Case Studies, Expert Interview Schedule, Curriculum Questionnaire, School Questionnaire, Teacher Questionnaire, and Student Questionnaire. These instruments will be described in detail in the subsequent publications.

Since the purpose of this report is to give a detailed account of how the dependent variable, writing performance, was managed, Part I is devoted to the problems that were faced during the study in obtaining scores pertaining to writing ability. Figure 2 shows how the approach was conceptualized. The figure shows that the study consisted of three phases: test development, test performance, and analysis and reporting. Clearly the three phases interact with one another and the proposed system of analysis and reporting influenced and were influenced by the test

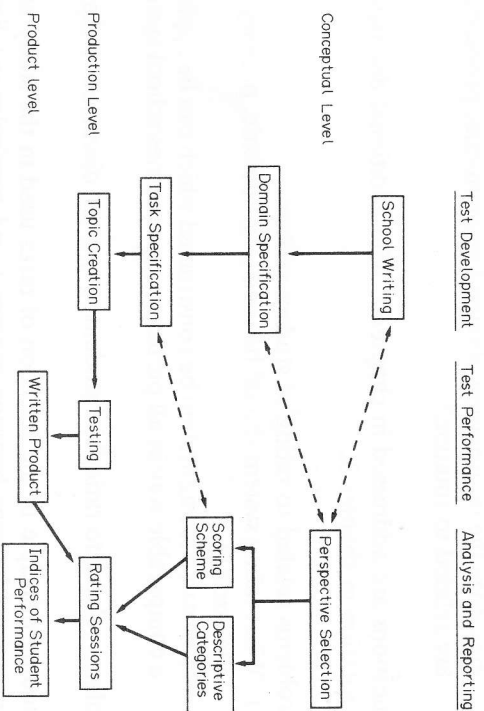


Fig 2

development and test performance phases. Test development involves domain specification and, from that, task specification and topic creation. At that point there can be testing. Similarly the analysis and reporting phase begins with a conceptualization of the perspective to be taken on the results, including the extent to which it should be descriptive or judgmental. From that selection comes a scheme for scoring or describing, and from actual scoring, a set of indices of performance. The following chapters describe the conceptual work in the development of the topics and the broad indices of student performance.

The main problems addressed were:

- I. Problems related to the construction of writing tasks:
 - (a) What is the total domain of writing and, more particularly, the school-based domain of written composition?
 - (b) What is the appropriate sample from the total domain for the students concerned? What sub-domains should be included in the set of writing tasks?
 - (c) What is the appropriate system for specifying the tasks?
 - (d) How should the actual writing tasks be formulated?
- II. Problems related to the allocation of writing tasks:
 - (e) What kind of tasks should be presented to each of the three student populations?

- (f) How should populations be linked through common tasks?
- (g) How many tasks should/can each student be asked to respond to?
- (h) If task rotation is necessary, how should it be done so as to maximize the information obtained and minimize problems that are related to rotation?

These questions are addressed in chapter 2, with additional discussion of the task rotation in chapter 3.

III. Problems related to rating of student scripts:

- (i) What rating system should be used (e.g., holistic, primary trait, analytical)?
- (j) How can a rating system be constructed which can be applied in a comparable way in all participating countries/school systems?

The problems related to rating the student scripts are discussed in chapters 3 and 4.

The guiding principle in the selection of tasks used in the study was to obtain an optimal balance between construct and curricular validity. To maximize construct validity a considerable effort was made to define the domain of writing (see chapter 2). Selection from the domain so that a high degree of match between teaching and tasks is guaranteed was guided by curriculum analysis and by an analysis of examinations.

Since educational systems offer instruction in several tasks, and since students are taught several different kinds of writing in order to help them become competent and flexible writers, it was decided to sample student writing across tasks to cover the domain well. Getting several writing samples from each student was considered necessary also in order to be able to study the structure of writing ability. Data on the dependent variables were collected by means of nine different task types, containing 14 different task versions. There were three compulsory core tasks for each population plus one or two international optional tasks. Each National Center could also develop national optional tasks. Thus each student wrote on at least three assignments representing different cells of the domain.

Limitations of testing time made some task rotation necessary. Within the constraints of being able to link students within populations and across populations, task rotation was based on the following principles: (a) each student should write on as many different types of tasks (different cells of the domain) as feasible, and (b) rotated tasks should take approximately the same amount of writing time.

In spite of task rotation (see appendix A), populations were linked so that in most cases (12 out of 14) the same task was common to two populations (A and B, or B and C). One task was the same for all three

populations (argumentative/persuasive task). Two tasks were meant for Population A only.

Similarly, students were linked so that there was one common task within all three populations; that is, all students wrote on one common task and the others were randomly rotated in class. This common task was a narrative/story for Population A, a letter of advice for Population B, and an argumentative/persuasive composition for Population C.

Within each population, and also partly across populations, the task order was standardized so as to control the possible task order effect.

The second part of this volume illustrates the results of the task scheme and scale construction as they were to be applied to the compositions written by the students in the various countries and educational systems in the study. *In no way are the examples intended to illustrate the results of the actual scoring sessions.*

Management of the Study

As in all IEA studies, the Written Composition study has a complex management structure. General policy decisions are made at the annual meeting of the International Project Council (IPC), which is composed of the General Assembly representative from each country participating in the study. Alan C Purves (USA) is the Chair of the IPC. More specific planning and detailed development of the project is the responsibility of the International Steering Committee, chaired by Anneli Vähäpassi of Finland. Steering Committee members are Thomas P Gorman (England), Judit Kádár-Fülöp (Hungary), Pietro Lucisano (Italy), Hildo Wessdorp (Netherlands, until 1985), Pai Obanya (Nigeria), Eva L Baker (USA), and, as special advisor to the committee, Raimo Kontinen (Finland). The full-time management of the study during the stages of development and testing (1981-86) was supervised by the International Coordinator, Sauli Takala (Finland). He was assisted by R Elaine Degenhart (USA), who became the International Coordinator in 1986. The International Study Committee (ISC) is made up of the National Research Coordinators (NRCS) who are responsible for the actual management and conduct of the study within each country. The NRCS established the National Centers and selected National Committees to assist them in the collection and preparation of the national data. Several intensive meetings of the ISC were held, during which pilot test results were examined and discussed, scoring schemes were tested and revised, and the training of raters was finalized.

The International Coordinating Center was located at the Curriculum Laboratory, University of Illinois at Urbana-Champaign, from January 1981 to August 1985. In 1985 the coordination of the project was moved to the Institute for Educational Research (KITL), University of Jyväskylä,

Finland, where the analysis of the international data has been carried out by the Data Manager, Kari Törmäkangas (Finland).

Summary

This chapter has outlined the origins of the International Study of Written Composition by describing its context and management, and by presenting very briefly its overall design including major constructs, tested populations and sampling, and task rotation. Chapter 2 will turn to the main theme of this book, namely the specification of the domain of the dependent variable, school-based writing, and the development of the writing tasks themselves.