DOCUMENT RESUME

ED 112 893 IR 002 604

AUTHOR Grundin, Hans U.
TITLE Response Requirement and Information About Correct Responses in Programmed Instruction. Linkoping Studies in Education No. 5.
INSTITUTION Linkoping Univ. (Sweden). Dept. of Education.
PUB DATE 75
NOTE 231p.; Research Department Dissertations No. 5

EDRS PRICE MF-$0.76 HC-$12.05 Plus Postage
DESCRIPTORS Covert Response; Educational Research; Elementary Education; *Feedback; Intermode Differences; Language Instruction; *Learning Processes; Overt Response; Patterned Responses; *Programed Instruction; *Response Mode; Scoring; Teaching Methods; *Testing
IDENTIFIERS Linkoping University; Skinner (B F); Skinnerian Programed Instruction; *Sweden

ABSTRACT

To test some aspects of Skinnerian programed instruction, a 241-frame program on Swedish grammar was presented to students 10 to 12 years old. Skinnerian insistence on student-constructed overt responses and on presenting the "key answer" after each frame were contrasted with covert responses and key answer frequencies less than 100 percent. The effects of combinations of overt or covert response requirements and 100, 50 or 0 percent key answers were also studied. It was concluded that in programs of less than 100 frames, response requirement and key answer frequency have very little influence on learning; that in programs over 200 frames, overt response without key answers or covert response leads to better learning. Overt response is most effective when the program demands learning unfamiliar response terms. Covert response leads to considerably shorter working time and greater learning efficiency. The Skinnerian model is held not superior to other alternatives studied. (SK)

LINKÖPING UNIVERSITY
Department of Education
LINKÖPING SCHOOL OF EDUCATION
Research Department

Dissertations

No 5

linköping studies in education

4833

HANS U. GRUNDIN

Response Requirement and Information about Correct Responses in Programmed Instruction

ED112893

Response Requirement and Information about Correct Responses in Programmed Instruction

by

Hans U. Grundin

"There is no sharp shoreline, which marks off the land of knowledge from the ocean of not-knowledge."

A. MASLOW, The Psychology of Science.

RESPONSE REQUIREMENT AND INFORMATION ABOUT CORRECT RESPONSES IN PROGRAMMED INSTRUCTION

Inquiries about the book should be made to:

The Director, Department of Education
Linköping University
S-581 83 Linköping, Sweden.

ISBN 91-7372-012-7

Stil-Foto Tryck, Linköping, 1975.

CONTENTS

Foreword

All literature references in this book are numbered and each work is referred to by its appropriate number in parenthesis, e.g. "Anderson (59) contends...". If several references are listed, the numbers are separated by comma or ampersand, e.g. "Several authors (3, 8, 27 & 39) report...". When page numbers are given, they are separated from the reference number by colon, e.g. "The authors's discussion of this point (45:12f)...". Thus, (1:2) refers to reference no. 1, page 2, whereas (1,2) refers to references no, 1 and 2.

The appendices found in this book are given in Swedish, because I consider it very important that the exact wording of tests, instructions, program frames etc. is made available to the reader. This will, of course, be somewhat inconvenient to the reader who has no Swedish. I will, however, be happy to provide any non-Swedish-speaking reader requesting it with an English translation of these appendices.

* * *

The study presented here originated in the then National School for Educational Research in Linköping, which was also the institution where I started my work in the field of educational research. For this study as well as for a great part of my professional development I am, then, very much indebted to the National school for Educational Research, which for a ten year period, under the enthusiastic leadership of professor Eve Malmquist, constituted a highly vital centre of research and development within Swedish education. Fortunately, I have enjoyed, even after the closing down of

the National School for Educational Research, the encouragement and support of Eve Malmquist, as a colleague, supervisor and friend;

I am much indebted to Mrs. Noomi Swedberg for her conscientious typing and observant proof-reading.

To all others, no one mentioned and no one forgotten, who helped in carrying out the work reported here, I am also grateful.

Linköping, July, 1975. H. U. G.

1

Introduction

Programmed instruction (PI) has attracted considerable interest in the field of education for some twenty years now. The literature is rich in theoretical discussions and investigation results as well as in practical guidance for program construction. Several different PI models have been proposed, discussed and tested, but by and large the Skinnerian model seems to be the one most commonly accepted and adopted. In fact, many seem to believe that the rationale proposed by Skinner is the only sound basis for PI.

Variations of the Skinner model have been suggested, some of them by Skinner himself, but it is nevertheless possible to discern some basic features of what we today might call traditional PI. The instructional unit, the 'frame', in a Skinnerian program is characterized by

- the demand for overt, student-constructed response; and
- information about the correct response immediately after each unit (cf. 120 and 121).

Numerous attempts have been made to evaluate the learning effect of these features. Several alternative response modes have been tested and compared, and the differential effects of various amounts or types of information about the correct responses have also been investigated. However, so far none of the many proposed alternative PI models has been proven generally and unambiguously superior to the others. It has even been argued that such comparisons are rather futile. As one researcher claims, after having reviewed experimental results in this field:

> "I am prepared to stick out my neck and suggest that some gains so far obtained by the use of teaching machines and programmes have been primarily due to the more efficient use of pupil time." (62:435).

On the other hand, many studies have demonstrated, beyond any reasonable doubt, that the Skinnerian response mode in PI is particularly time-consuming, which may make it less efficient than other models, not demanding overt, student-constructed responses.

It has also been argued that the response mode effect in PI is dependent upon a number of different program and student characteristics. And similar relationships may well hold for variations in the information about the correct responses.

Skinner's PI model is based on a particular learning theory Whether this theory is acceptable or not is a question far beyond the scope of this study. It should be noted, though, that this learning theory has been more and more intensely challenged in recent years. Regardless of how one judges Skinner's learning theory, there are, however, practical reasons for questioning the efficiency and viability of his PI model, particularly as far as its application in school instruction is concerned.

2

Terminology

"An adequate definition permits us always to tell when a sentence containing the defined term is true..." (72:48).

"In general, definition in science stops when all descriptive terms in the definition refer either to physical objects or to some directly observable properties and relations of and among them." (72:51).

Definitions of terms are often necessary in order to ensure that propositions containing the terms in question have a clear meaning. However, such definitions are no end in themselves. Their purpose is to clarify concepts which do not already have a commonly accepted denotation, and this is achieved by relating them to concepts which do have commonly accepted denotations. Definition is also needed when the commonly accepted denotation of a term can be regarded as ambiguous or misleading.

Since PI is a comparatively new and fast developing field, it is only natural that we have not attained within it a very high level of terminological stability and precision. It seems, therefore, necessary to start this study of response mode and response information issues in PI with a discussion of certain terminological questions. The following discussion concerns only terms which, in my opinion, are used without sufficient clarity of concept. Terms not discussed in this chapter are consequently assumed to have commonly accepted - and appropriate - denotations.

2.1 PROGRAM VARIABLES AND STUDENT BEHAVIOUR

PI investigations often have the aim of studying program variables (or programming variables; see 90). Such variables are of interest because they are expected to determine or influence student behaviour and, consequently, the effect of programmed instruction.

Since student behaviour is not determined solely by program variables, it is important to make a clear distinction between the concept 'program variable', i.e. a variable program property, defined independently of student behaviour, and what might be called 'PI variable', i.e. a variable student behaviour property, which assumedly is a consequence of an interaction between program and student. It should be noted that the concept 'program' here includes all instructions given with the program.

Clearly, program variables as defined above are suitable independent variables in PI studies, since such variables can be directly manipulated. PI variables, on the other hand, are in a sense always dependent variables, as they cannot be directly manipulated by the experimentator or the programmer. With the definition of program variable proposed here, it seems necessary to choose terms for various program variables so that they do not imply any assumptions or expectations concerning student behaviour. This is especially important, when the existence of a relationship - or the degree of this relationship - between a program characteristic and student behaviour cannot be controlled.

Examples of terminology tending to obscure the difference between program characteristics and what is here called PI variables are not infrequent in the PI literature. This was pointed out several years ago by Tuel & Metfessel (51). One of the main objectives of the present discussion of PI terminology is to reveal and eliminate such sources of confusion.

2.2 RESPONSE INFORMATION

Skinnerian PI is characterized by, among other things, the student's being informed, immediately upon completion of each program unit, about the correct response, which is presented in the form of a key answer (121:99). Such presentation in full of the correct response can be labelled *response information* (cf. 56) or, more precisely, *key answer presentation*. These two terms are neutral in relation to different theories or assumptions about the effect of the presentation of the correct responses. In this respect they differ from such commonly used terms as 'confirmation of responses', 'reinforcement' and 'feedback'. The term 'response information' denotes a program variable, since the type and frequency of such information can vary independently of student characteristics and behaviour. The terms 'confirmation', 'reinforcement' and 'feedback', however, are not suitable for the denotation of program variables.

Confirmation of response implies that a correct student response has been given, otherwise only correction is possible. It also implies that the student compares his own response with the correct response, and that he is aware of the correctness of his own response. There is, of course, the possibility of confirming an incorrect response, if the student fails to notice some difference between his own response and the correct response. In any case, the occurance of response confirmation is dependent upon the student's behaviour.

Reinforcement implies, if it is taken to mean positive reinforcement, that confirmation of response has occurred and that, as a consequence of this confirmation, the probability of the student's responding correctly the next time is increased. Thus, reinforcement, even more so than confirmation, is dependent upon factors 'within' the student. It has been argued, for instance by Kay (93), that confirmation always implies reinforcement. But Kay himself emphasizes that

> "we do *not* know the conditions under which knowledge of results attains and maintains reinforcing properties" (93:30).

Feedback does not necessarily imply that the student response is correct. But it does imply some comparison between the student response and the correct response. Furthermore, as feedback in a literal sense presupposes an interaction between student behaviour and 'program behaviour', feedback is impossible in PI unless an advanced technical equipment is used. When programmed books, or simple machines are used, one can, at most, expect some kind of auto-feedback, i.e. the student may use the response information to adjust his response behaviour.

The program property *key answer presentation* can, in interaction with different kinds of student behaviour, lead to response confirmation, to reinforcement, to feedback, or to neither of these alternatives. Obviously, our terminology should reflect this fact and distinguish clearly between the objective program property, on the one hand, and its more or less probable effects in a PI situation, on the other. Basically the same point has been made by Annet:

> "It is suggested that 'reinforcement'.. should not be used arbitrarily as a substitute for other related terms as 'knowledge of results'" (61:282),

and by Lumsdaine:

> "..~~immediate~~ feedback providing confirmation or correction .. cannot be equivalated theoretically to any simple conception of reinforcement.." (100:86).

Summarizing this discussion we find that terms such as 'response information' and 'key answer presentation' are better suited than the others discussed here to denote an instructional variable, which is to be the object of experimental study. It should be noted, also, that any attempt to assess reinforcement or feedback effects in PI is meaningful only if the presentation of the correct response can be shown to have an effect.

Key answer presentation can vary as to the proportion of program units (frames) after which the key answers are given. This proportion, expressed in per cent of the total number of frames, will in the present study be labelled *key answer frequency*.

A response information schedule where key answers are given after some frames only, e.g. every second or third frame, is sometimes called 'partial reinforcement' (e.g. in 3 & 30). This term is inappropriate, since it may, wrongly, be regarded as a counterpart in PI to the much investigated partial reinforcement paradigm in animal learning experiments. An example of this misinterpretation is found in Krumboltz & Weisman (32), who avoid the term 'reinforcement' in their report title but nevertheless contend that

> "providing confirming answers on an intermittent scale may increase retention i n t h e s a m e m a n n e r (my italics) that intermittent reinforcement prolongs extinction..".

As Holland (90:91) has pointed out, this is hardly so. The crucial difference is that in traditional animal learning experiments one and t h e s a m e response is alternately reinforced and not reinforced, while in PI partial, or rather intermittent, response information means that, in a group of d i f f e r e n t responses, some are followed by key answers and some are not. Even those using the term 'partial reinforcement' in the sense criticised here sometimes show that they are aware of the inappropriateness of the term. In a report entitled "The effect of partial reinforcement" Berglund states:

> ".. it is probably inaccurate to speak of partial reinforcement in the traditional sense of the term in connection with programmed instruction" (3:2).

As a general term denoting various key answer frequencies less than 100 but greater than 0 per cent the term 'i n t e r m i t t e n t r e s p o n s e i n f o r m a t i o n' will be used in this study.

2.3 RESPONSE, RESPONSE MODE AND RESPONSE REQUIREMENT

The term 'response' is used in this study as a common label for student reactions that are answers to direct or indirect questions, or solutions to problems given, or other analogous student reactions. A response need not be observable to anyone except the student himself.

Many studies of program variables published during the 1960's are studies of what is usually called response mode. The term 'response mode' does, however, not denote a program variable; it is rather a PI variable. The corresponding program variable is the feature of the program (instruction etc.) intended to produce the desired student response mode, i.e. the r e s p o n s e r e q u i r e m e n t of the program. The distinction between response mode and response requirement may seem subtle. Where a specific response requirement produces, with great probability, the desired response mode, and where one can continually control that this is so, the distinction is in a way superfluous, and the two terms can be regarded as practically interchangeable. This may be the case when an overt, observable response is explicitly required and means for recording the response are available.

If, on the other hand, observable responses are not required, there is no immediate way of controlling the correspondence between response requirement and response mode. Also, the risk that the response is omitted is likely to increase. In this case the distinction between required and given response is important, since one and the same response requirement may lead to different response modes, with small possibility of telling which response mode each student uses. Generally speaking we can study the program variable response requirement, whereas the response mode sometimes must be regarded as a hypothetical variable. In this study, the term 'response mode', therefore denotes student behaviour, and the term 'response requirement' is used to denote program properties.

Several different terms are used in the PI literature to label different response requirements or response modes. The following terms will be discussed here:

- 'active' versus 'passive' response;
- 'overt' versus 'covert' or 'no' response;
- 'explicit' versus 'implicit' response; and
- 'student constructed' versus 'multiple choice' response.

The terms 'a c t i v e' and 'p a s s i v e' response are good examples of terminological imprecision. 'Active response' is in a sense a tautology, as a response always is some kind of action, physical or mental. From this point of view 'passive response' is an impossible term. The term 'active response' is sometimes used with the implication that the response activity in question is observable. Klaus, for instance, contends that the first rule of programming is to "require active responding" (95:43). 'Active responding' is then synonymous to the more common term 'overt responding', and consequently superfluous. On the whole, the terms 'active' and 'passive' must be regarded as inappropriate for the classification of responses in PI.

The terms 'o v e r t', 'c o v e r t' and 'n o' response obviously denote aspects of response mode. As has already been mentioned, the correspondence between response requirement and response mode can easily be checked when overt responses are required, but not when covert, non-observable responses are required. For response modes labelled 'no response' this correspondence is also difficult to check, as covert responses cannot be distinguished from 'no responses' through observation of student behaviour. Furthermore, the 'no response' mode can be a natural consequence of several different program properties: no question has been asked and no problem posed; or no instruction to answer the question has been given; or instruction has been given not to answer the question.

Empirical studies with the aim of comparing the effects of overt versus covert responses etc. are in most cases studies of the effects of the corresponding response requirements. It seems logical that the terminology used in such studies should make this clear.

The terms 'e x p l i c i t' and 'i m p l i c i t' are generally used as attributes to 'response (mode)'. Lumsdaine, for instance, says:

> "..an explicit response means a response given by the student on some specific occasion identified to him through instructions.. - - - ..i m p l i c i t responses (of a covert sort) .. are responses for which the occasion is less clearly identified, such as subvocal .. responses that occur in silent reading, in listening to a presentation, or in creative thinking" (98:486f).

Although both 'explicit' and 'implicit' stand as attribute to 'response', Lumsdaine's definition implies that 'explicit' characterizes the response requirement, whereas 'implicit' seems to characterize both response requirement and response mode. The expression "an explicit response" must obviously be taken to mean 'an explicitly required response'. Even Lumsdaine, who is obviously trying to remedy the terminological confusion, fails to distinguish between response mode and response requirement.

Response requirements are always explicit, since an 'implicit' response is not required by the program, but the result of some interaction between student and program. For practical purposes it seems reasonable to let 'explicit response requirement' denote any program feature that expressedly and unequivocally instructs the student to respond. This is also in agreement with Lumsdaine's definition (see quotation above). As Lumsdaine points out, it is essential that the dimension 'explicit - implicit' is not mixed up with the dimension 'overt - covert - no' response.

> "..the term 'implicit response' has frequently been used to embrace both explicit and implicit covert responses.. - - - Explicit responses may be either overt or covert.. - - - The further distinction to be made is that between e x p l i c i t covert responses .. and implicit responses of a covert sort." (98:487)

In principle, it is possible that a 'no response' mode is either explicit or implicit. The difference between the two modes is most easily expressed as the difference between demand for n o response, on one hand, and n o demand for response, on the other. In spite of its subtleness the distinction is of interest, mainly due to the difficulty of distinguishing, in empirical studies, non-responses from covert responses. Terminological confusion may arise here if student responses are discussed one-sidedly in terms of response m o d e, or if a proper distinction is not made between response requirement aspects and response mode aspects.

Lumsdaine (98) labels 'implicit response mode' a condition known from many studies, which consists of letting the students go through a program with the correct responses written in all the blanks - a condition often called 'reading'. The same condition is called "covert or rather n o responding" by Wittrock (55; my italics). In many studies it is impossible to tell which m o d e of responding the students have used, and the response r e q u i r e m e n t is often incompletely described. An exact rendering of the instruction given to the students is, for example, seldom found in the published reports.

In cases where the original response requirement calls for filling in blanks in the program frames, the 'reading' mode has two main variants:

- response terms are not underlined or in any other way emphasized; and
- response terms are underlined or emphasized by means of different colour, etc.

In the first variant there will probably be no responding, whereas the second would be expected to yield what Lumsdaine calls "implicit, covert responses". To label the condition 'reading' has the advantage of not presupposing anything concerning student behaviour that cannot be checked. In the terms discussed here, this condition can only be defined negatively, as the absence of any kind of explicit response requirement. Since studies of response mode/requirement mainly concern explicit response requirements, and since implicit response requirement are so very elusive, if they at all exist, the further discussion of response requirements will primarily be about explicit response requirements. The term 'response requirement' will therefore, throughout this book, mean 'explicit response requirement'.

O v e r t r e s p o n s e r e q u i r e m e n t will consequently mean an explicit demand for an observable act of responding. C o v e r t r e s p o n s e r e q u i r e m e n t will mean an explicit demand for a non-observable (mental) act of responding. N o r e s p o n s e r e q u i r e m e n t, finally, will mean the absence of an explicit demand for a response, assuming also - unless otherwise stated - that there is no response blanks in the program nor any other space designated for the recording of student responses.

2.4 PROMPTING

The term 'prompting' is often used in the PI literature. Usually it is a label for all kinds of features or cues in program frames, which have the objective of making the "emission of the correct answer more likely", to use the words of Meyer (107). The degree of prompting can vary from mere hints to maximum or complete prompting, which means that the student is told what the response should be, before he is asked to respond.

'Prompting' is often used to label the last-mentioned type of complete prompting (e.g. in 82:134). This is particularly common in studies of 'prompting versus confirmation' (e.g. 69). In order to avoid vagueness and confusion, it is suggested here that the term 'cueing' (see for instance 67: 13 & 82:134) should replace 'prompting' in its general sense, while the term 'prompting' should only be used to denote complete cueing as described above.

Like several of the terms already discussed in this chapter, the term 'prompting' has the disadvantage of being liable to imply assumptions concerning student behaviour, which may be difficult to control. In a program sequence of the prompting type (key answer presented before the student responds) it is possible that what is meant to be prompting, can in fact become confirmation, namely if the student responds - probably covertly - before the key answer is perceived (cf. 71). This vagueness can, however, be avoided, if 'prompting' is defined as the program property (key answer presentation before student responding) that is expected to have a prompting function. A recent discussion of prompting can be found in a paper by Richard Anderson (60), who uses the term in a very general sense.

2.5 RESULT VARIABLES ETC.

An evaluation of an instructional process should, in principle, take three main factors into account:
- the increase in student performance or ability, resulting from the process under study;
- the student working time needed for the process; and
- other costs of the process.

The term 'effect' is often used to label the first of these three factors, the increase in performance, while the term 'efficiency' is used in the case of simultaneous evaluation of the two first factors, e.g. performance increase per unit of student working time (cf. 101 & 118).

In the present study the terms 'effect' and 'efficiency' will be used in the above-mentioned way. Thus, 'l e a r n i n g e f f e c t' will denote an increase in student performance, which is assumed to be a result of the instructional process. Where there is no risk of confusion only 'learning' or only 'effect' will be used in the same sense. When one and the same test, or parallel tests, are used to measure performance both before and after the PI, the measured effect may of course to some extent be due to a practice effect from the first testing. This can, however, be taken into account in the definition of learning effect simply by considering the pretesting procedure as part of the instructional process studied. It should also be observed that pretest effects do not necessarily occur in PI. Hartley (89) has studied this problem rather extensively and found no such effects.

The term 'l e a r n i n g e f f i c i e n c y' - or, where there is no risk of confusion, only 'efficiency' - will in this study denote the learning effect in relation to P I t i m e. The term 'PI time' denotes the gross time spent by the student on the program, i.e. the sum of the study periods devoted wholly to the program. The term 'learning time' is avoided primarily because this term should denote the net or effective working time, and of this we generally know very little, except that it can safely be assumed to be shorter than the PI time.

Efficiency could be more strictly defined by the formula

Efficiency = Effect/PI time.

In practice it is, however, difficult to obtain reliable measures of this nature. Efficiency must, therefore, often be estimated in other ways. More will be said about this problem in section 3.3.2.4 below.

As defined here efficiency does not include any reference to any aspect of the third of the factors listed above, the costs of the process, other than student time. This could, of course, easily be remedied by changing the definition so that 'efficiency' means 'effect per working time unit and cost unit' (cf. 21:117). The difficulties of measuring efficiency defined in that way in the field of education would, however, be almost insurmountable.

In many cases one tries to measure immediate learning effect as well as retention some time after the end of the instructional process. The term 'learning effect' will be used in this study to denote 'immediate learning effect'. The effect remaining some time after the instruction has finished will be labelled, as it generally is, 'r e t e n t i o n' or 'l o n g - t e r m l e a r n i n g e f f e c t'.

3.1 OUTLINE OF THE REVIEW OF PREVIOUS RESEARCH

Research works and scientific discussions concerning response mode/requirement and use of key answers in PI using linear programs are reviewed here with an emphasis on the results of practical educational research rather than on learning research. A complete inventory of all that has been written about these topics has not been attempted. I have regarded it as sufficient to identify the major trends in the results of previous research.

Several research reviews covering the problems studied here have been published (4, 59, 61, 70, 71, 90, 98, 99, 116 & 123). Most of these reviews have been of the traditional type, that is they have consisted of short abstracts and comments of a critical, evaluative or comparative character. The review of Bernmalm (4) differs from the others in that data concerning the investigations are presented in tables, which facilitates their use for the study of problems other than those explicitly treated by the reviewer. A thorough, theoretical analysis of different aspects of response mode and so called feedback is found in Goldbeck & Briggs (84).

Most of the studies I have come across in my literature review have been discussed in one or several of the reviews listed above, and I will therefore not present here another review of the traditional type. Instead, I will try and give a summary of a more objective-analytic kind. More than 50 studies of response mode/requirement or key answer frequency in PI have been classified according to different program and student variables, and according to the results of comparisons between different experimental groups in terms of learning effect and PI time. Those classifications have been made on the basis of data presented in the reports reviewed. Summaries have been made in the form of cross-tabulations, where the unit of observation is the individual study. This procedure makes it possible to study the importance of certain program and student variables for learning effect and PI time by means of statistical description.

The value of such summaries of groups of experimental studies is limited by the following circumstances:

- Investigations with the same objective are not seldom so different in design as to diminish their comparability.
- Deficiencies in the reporting of data concerning experiments sometimes make classification difficult or even impossible.
- As all studies are given equal weight in the summaries, it is not possible to account for differences between studies in terms of subject sample size, scientific quality etc.

As regards the third factor listed above, it should be noted, though, that if two comparisons have yielded significant differences using samples of different sizes, the study using the s m a l l e r sample should be given the greater weight, since the probability of obtaining significant results is smaller the smaller the sample. This has been pointed out by Bakan (64). When the sample is large it is even possible that a statistically significant difference is practically insignificant. As Nunnally puts it:

> "..if the null hypothesis is not rejected, it is usually because the N is too small. If enough data are gathered the hypothesis will generally be rejected." (113:643)

Concerning scientific quality there exists, as far as I know, no measuring device or procedure of demonstrated reliability and validity. Nasatir & Elesh (112), for instance, have devised an instrument for that purpose, but analysing their own validation data one finds that the correlation between raters is somewhat l o w e r than could be expected if rating scores were totally randomly distributed! Many other instruments have been proposed; a review by Bartos (65) lists more than 20 items for the period 1962 - 1968. However, these instruments do not seem to have been empirically investigated as to their reliability and validity.

All three types of deficiencies listed above seem due to lack of co-ordination of educational research and to insufficient standardisation of research methods, variable definitions, methods of presenting data, etc. Research reviews of the type presented here can therefore, in addition to fulfilling their main function, contribute to the demonstration of the co-ordination and standardisation problems in educational research. The classifications of research results made here constitute compromises between what is desirable and what is reasonable in view of the available data. Deficiencies and uncertainties in the classifications will be discussed in connection with the presentation of each classification schema below.

3.2 CLASSIFICATION OF EXPERIMENTS

3.2.1 Comparison results

The results of comparisons between paired groups are usually rendered in the form of group means and t-values - or z-values - for mean differences. But some studies only report whether differences are significant or not, and whether they are positive or negative. And sometimes the signs of nonsignificant differences are not reported. In view of these circumstances the following scale seems the most reasonable basis for the classification of the results of group mean comparisons:

++ = positive significant difference
\+ = positive nonsignificant difference
0 = no difference or nonsignificant difference of unknown sign
\- = negative nonsignificant difference
-- = negative significant difference.

Concerning what is significant or nonsignificant the statements made in the reports studied have been accepted. If the result of significance testing is reported for more than one level of significance, the results at the 5 per cent level are chosen.

Comparisons between groups as to learning effect in the pro per sense (cf. definition in section 2.5 above) are rare in the studies reviewed here. As a rule comparisons are made in terms of posttest scores without correction for differen ces in initial performance. For want of more precise data i have chosen to accept posttest differences as estimates of the corresponding learning effect differences.

PI time is one of the very few variables in PI studies whic is normally measured in a quotient scale, namely number of minutes or hours. The results of comparisons between experi mental conditions can, therefore, be expressed as quotients the PI time for each experimental condition divided by the PI time for what can be regarded as the basic condition. As basic condition in the computation of such a quotient is chosen the experimental condition which logically should lead to the shortest PI time. Such PI time quotients, computed from data reported for different experiments, can be considered directly comparable.

3.2.2 Program length

The most common, and usually the only, measure of program length in the PI literature is number of program units, or frames. The length of a program unit can, of course, vary considerably between programs - and also within one and the same program. Most research reports do not permit any reliable estimate of the length or 'size' of frames in the programs used. According to Yaeger (125) it seems very difficult to find a useful measure of program frame length.

In studies of the effects of different response requirements and key answer frequencies, one of the most interesting aspects of program length is the number of responses demanded in the program. If one response per frame is required, the number of frames is a suitable measure, but in other cases this measure can be misleading. Unfortunately, the number of responses demanded in a program is rarely reported in the studies reviewed here. Therefore, the number of frames is the only generally available measure of program length.

Taking the observed variation in the number of frames per program in the studies reviewed into consideration, the following classification schema has been established for the variable program length:

Category	Definition
Long	More than 200 frames
Medium	100 - 199 frames
Short	Less than 100 frames
Unknown	Number of frames not reported

3.2.3 Students' age or educational level

Research reports usually indicate e i t h e r the age of the students o r their educational level ("college", "secondary modern", "grade 9", etc.). For the classification of experiments in this respect I have defined the following student categories:

16 years or more,
13 - 15 years,
10 - 12 years,
9 years or less,
age not reported and impossible to estimate.

Where only educational level is reported the students are considered to be of the age which is normal for the educational level in question.

3.2.4 Learning task characteristics

A classification of experimental studies according to the degree of learning task difficulty would be of great interest in this context, and so would a classification according to the occurrence of unfamiliar response terms (cf. 59: 141), i.e. what is often labelled the degree of 'response learning' (e.g. 35:504). Since sufficiently detailed program descriptions are not found in most of the reports reviewed, I have not considered it possible to construct a schema for objective classification of experiments according to such learning task characteristics. The question of possible interactions between response requirement or key answer frequency, on one hand, and different learning task characteristics, on the other, will therefore have to be discussed in more informal terms (see section 3.3.2.3. below).

3.3 STUDIES OF RESPONSE REQUIREMENT

In view of what has been said about the terminology in this field (see section 2.3 above), it seems natural to classify studies of response mode or response requirement in PI wholly on the basis of program properties. The following main types of response requirement can be distinguished:

overt, student-constructed response required (OCR);
overt, multiple-choice response required (OMR);
covert, student-constructed response required (CCR);
covert, multiple-choice response required (CMR); and
no response required (NOR).

The four first-mentioned types (OCR, OMR, CCR & CMR) may occur with or without a key answer following the student response, and they may also be combined with prompting or cueing. In practice, prompting is, however, unusual when multiple-choice answers are required.

For obvious reasons key answers or other kinds of information regarding the correct responses do not occur in connection with no-response requirement (NOR). Cueing may, however, occur in the NOR condition, when certain terms or concepts are emphasized (cf. section 2.3 above). This NOR condition comes quite close to the CCR condition combined with cueing, the only difference being that the CCR condition presupposes an explicit demand for student responses. The NOR condition is often called "reading", a term which seems most adequate when cueing does not occur.

In the literature reviewed here several experimental studies comparing two or more response requirements are reported. The number of comparisons for each pair of requirements is given in table 3.1. As this table shows, a few types of comparison dominate the literature, namely

- OCR requirement (overt constructed response) versus OMR requirement (overt multiple-choice);
- OCR requirement versus CCR requirement (covert constructed response);
- OCR requirement versus NOR requirement (no response required); and
- CCR requirement versus NOR requirement.

Table 3.1 Number of comparisons of pairs of response requirement in the literature reviewed.

	OCR	OMR	CCR	CMR	NOR
OCR	-	6	23	0	24
OMR		-	1	2	2
CCR			-	0	8
CMR				-	0
NOR					-

Note: The total sum of matrix values exceeds the number of studies reviewed, since several studies involve more than one comparison.

In the following survey only the four most frequent comparisons are discussed, i.e. OCR-OMR, OCR-CCR, OCR-NOR, and CCR-NOR.

3.3.1 Constructed versus multiple-choice response requirement

The literature reviewed contains five reports (12, 17, 20, 41, 53) of studies where student-constructed response is compared to multiple-choice response. In the study by Fry

(20) the two response requirements are compared under two different conditions of PI time, namely variable time and fixed time, which gives in all six comparisons. In all the six cases overt responses are required, i.e. the requirements compared are those here labelled OCR and OMR. The key answer frequency was 100 per cent throughout, i.e. key answers were presented to students after each frame.

Some data regarding the six comparisons are summarized in table 3.2. As the table shows, the comparison results in terms of learning effect are either neutral (3 cases) or positive, favouring the constructed response (3 cases). The positive differences are significant in two cases, but these are not fully independent, since they come from experiments made by the same researcher, using the same program (20). Comparisons as to retention have been made in five cases, and the results agree, on the whole, with the results shown in table 3.2.

Table 3.2 Data on studies comparing OCR with OMR

Ref	Effect OCR-OMR	PI time OCR/OMR	Student age	Program length	Program title
12	0	1.2	≥16	104/56[1]	Elementary Psych.
17	+	1.1	≥16	72	Symbolic Logic
20	++	1.7	13-15	--[2]	Spanish Vocabulary
20	++	1[3]	13-15	--[2]	"
41	0	--	≥16	1152	Analysis of Be
53	0	1.3	≥16	192	haviour

1) "Small" and "large" steps respectively
2) 16 items repeated "several" times
3) Same time for both conditions, preset by the experimentator

The PI time is reported in four cases, and all four PI time quotients (OCR-time divided by OMR-time) are greater than 1, i.e. the PI time increases when overt student-constructed responses are required instead of overt multiple-choice responses. Apparently, the size of this increase in PI time can vary considerably. In the studies reviewed here the PI time under the OCR condition is between 8 and 70 per cent longer than the PI time under the OMR condition (see table 3.2).

Learning efficiency is difficult to estimate on the basis of the data in table 3.2. It is obvious, however, that a demand for overt, student-constructed responses may increase PI time with up to 27 per cent - compared to the PI time using multiple-choice questions - without any significant or even noticeable increase in learning effect (53). There is, therefore, reason to suspect that OCR requirement can lead to lower learning efficiency than OMR requirement. In a complete evaluation of the relative merits of these two response requirements the fact that program production costs increase when multiple-choice items must be provided, should, however, also be taken into consideration.

The outcome of comparisons between the OCR and the OMR condition as to learning effect may also be influenced by the nature of the learning task, especially by the degree of response learning called for in the program (cf. section 3.2.1.5 above). When a program has the aim of increasing the student's response repertory, e.g. his active vocabulary, it seems likely that the demand for student-constructed responses should have a particularly good effect. It is, therefore, interesting to note that the significant positive effects of OCR as compared to OMR, reported in table 3.2, have been reached using a program with the objective of teaching Spanish vocabulary (20).

The following tentative conclusions can be drawn from this survey regarding the demand for student-constructed responses as compared to the demand for multiple-choice responses:

- Student-constructed responses lead to equal or better learning effect.
- They may in some cases lead to lower learning efficiency.
- They lead to higher effect the higher the degree of response learning called for in the program.

3.3.2 Overt or covert or no response requirement

In the literature reviewed here more than 30 reports have been found, where two or more of the response requirements OCR, CCR and NOR are compared as to their effect in PI using linear programs. One study reports 3 experiments (1), three report 2 experiments each (2, 21 & 35), and 32 studies report 1 experiment each (6-8, 11, 13, 14, 16-18, 22-25, 29-31, 33, 36-38, 40, 43, 44, 46, 47, 49-55 & 60). Since multiple-choice responses are not at all discussed in this section, the terms 'overt response' and 'covert response' are used to denote 'overt student-constructed response' and 'covert student-constructed response' respectively.

Overt versus covert response requirement is studied in 24 experiments (1, 2/2 exp:s/, 6-8, 13, 14, 17, 21/2 exp:s/, 33, 35/2 exp:s/, 36-38, 40, 46, 47, 50-52 & 60). In all these cases the response requirements have been studied in combination with 100 per cent key answer frequency.

Overt response versus no response requirement has been studied in 24 experiments (1/3 exp:s/, 6, 8, 11, 16, 18, 21/2 exp:s/, 23-25, 29-31, 43, 44, 49/2 exp:s/, 50, 53-55). In these cases, too, the OCR requirement has been studied under 100 per cent key answer frequency.

Covert response versus no response requirement has been studied in 8 experiments only (1, 6, 8, 21/2 exp:s/, 22, 31 & 50). In one of these (22) the covert response requirement is studied only in combination with 0 per cent key answer frequency, while in all the other cases the key answer frequency is 100 per cent. Studies comparing the effects of different response requirements also under key answer frequencies other than 100 per cent, are discussed further in section 3.4 below.

The results of the comparisons as to PI time, learning effect, and retention in the experiments reviewed here are summarized in t a b l e 3 . 3 . Since pretest data are usually not available, differences in posttest results are used as measures of learning effect differences. As the variance of performance increments is normally smaller than the variance of the corresponding total performance measure, this procedure will probably lead to an underestimation - rather than an overestimation - of the degrees of significance associated with the reported differences.

As table 3.3 shows, PI time data are available for about half of the experiments reviewed. The retention of learning can be estimated in just over half of the experiments, but the time elapsed between the instruction and the second posttest - intended to measure retention - varies considerably, from 2 weeks to 6 months. The distribution of comparison results concerning retention must, therefore, be interpreted with particular caution.

Table 3.3 Studies of requirements OCR, CCR and NOR: outcomes of comparisons as to PI time (PT), learning effect (E1), and retention (E2).

Comparison	Variable	++	+	0	-	--	No data
		No. of comparison outcomes					
OCR-CCR	PT	12	0	1	0	0	11
	E1	5	7	5	6	1	0
	E2	3	3	6	2	0	10
OCR-NOR	PT	9	1	1	0	0	13
	E1	7	5	6	5	1	0
	E2	4	4	2	1	2	11
CCR-NOR	PT	1	1	1	1	0	4
	E1	0	2	2	4	0	0
	E2	0	1	1	1	2	3

1) ++ or + means that the mean for the first-mentioned requirement has been higher than that for the last-mentioned.

From the summary in table 3.3 it can be concluded that overt response requirement, whether compared with covert or with no response requirement, leads to increasing PI time, but that this - often considerable - increase in PI time far from always results in a noticeable increase in learning effect. From this follows that the relationship between overt response requirement and learning efficiency is by no means unequivocal.

Regarding covert response requirement, as compared with no response requirement, one may conclude, from the summary in table 3.3, that a demand for covert responses as a rule has a rather small influence on PI time as well as on learning effect.

The perhaps most important conclusion that can be drawn from data presented in table 3.3 is, however, that any 'dogmatic' opinion concerning the general value of one specific response requirement seems untenable.

Consequently, a more detailed analysis is desirable. According to Holland (90:93ff) factors such as program length and the difficulty and relevance of the program questions should be expected to affect the result of covert responding in PI. Other researchers have pointed to the characteristics of the learning task, especially its degree of "response learning" or "associative learning" (35 & 59), and the age or intellectual maturity of the students (35) as important factors. The experiments reviewed here can, with few exceptions, be classified according to program length as well as student age.

An objective classification of program properties such as level of difficulty, relevance in relation to the criterion test, and the degree of response learning called for, is in most cases not possible on the basis of the available data. The relationship between such program properties on the one hand and learning effect on the other will, therefore, be analysed by means of a discussion of selected experiments. It should be noted here that level of program difficulty and degree of response learning are not program variables in a strict sense (see section 2.1 above), since they are to some extent dependent on student characteristics: a response term unknown to A may be well known to B, etc.

3.3.2.1 Response requirement and program length

In table 3.4 program length is cross-tabulated against the outcomes of response requirement comparisons as to learning effect.

The result of this cross-tabulation supports the hypothesis about a correlation between program length and effect of overt response requirement. In all experiments yielding a significant result favouring overt responses (++), programs with more than 100 frames have been used. When short programs - with less than 100 frames - have been used, most comparisons favour other response requirements.

Judging from table 3.4, the correlation between program length and outcome of comparisons is higher when overt response requirement (OCR) is compared with no response requirement (NOR) than when it is compared with covert response requirement (CCR). For short programs, negative outcomes are more often associated with the comparison OCR-NOR than with OCR-CCR.

Regarding the effect of OCR compared to CCR, the results of this analysis support Hollands opinion that the disadvantage of covert responding appears first when the program is

> "long enough for subjects in the covert mode to become careless..." (90:93).

It should be noted, though, that overt responding, and the increased PI time which it entails, means that a large number of short pauses occur in the students' intake of information, viz. when responses are recorded. To the extent that such pauses have a positive influence on learning, it seems reasonable that this influence should be greater the longer the program.

As regards CCR compared to NOR, the small variation between studies in program length (table 3.4) allows no conclusion concerning a correlation between program length and outcome of comparisons between these two response requirements.

A cross-tabulation corresponding to the one in table 3.4 has also been performed for the outcomes of comparisons between

Table 3.4 Cross-tabulation: Program length x outcome of comparisons between pairs of response requirements as to learning effect.

Comparison	Program length	No. of outcomes type ++	+	0	-	--	Sum
OCR-CCR	≥200	2	1	3	1	0	7
	100-199	2	1	1	1	0	5
	<100	0	3	1	3	0	7
	unknown	1	2	1	1	0	5
OCR-NOR	≥200	2	0	1	0	0	3
	100-199	4	1	2	0	0	7
	<100	1	2	3	5	1	12
	unknown	0	2	0	0	0	2
CCR-NOR	≥200	0	0	0	0	0	0
	100-199	0	0	0	2	0	2
	<100	0	2	2	2	0	6
	unknown	0	0	0	0	0	0

Table 3.5 Cross-tabulation: Program length x outcome of comparisons between pairs of response requirements as to retention.

Comparison	Program length	No. of outcomes type ++	+	0	-	--	Sum
OCR-CCR	≥200	1	1	1	2	0	5
	100-199	2	0	0	0	0	2
	<100	0	1	3	1	0	5
	unknown	0	1	1	0	0	2
OCR-NOR	≥200	1	0	0	1	0	2
	100-199	3	2	1	0	0	6
	<100	0	1	1	1	2	5
	unknown	0	2	0	0	0	2

response requirements in terms of retention of learning. The result is shown in table 3.5. There is a close correspondence between the results of comparisons as to retention and the results of comparisons as to learning effect (table 3.4). The correlation between program length and the effect of OCR appears also in the analysis of retention data. The hypothesis about the importance of program length for the effect of overt responding in PI is, on the whole, supported by the present research review.

3.3.2.2 Response requirement and student age or maturity level

The possibility of a correlation between overt response requirement (OCR), on the one hand, and student characteristics such as age, mental age, study skill, etc., on the other, has been discussed by Leith (35). He sees the overt response requirement as a means of maintaining the learner's attention and focusing it on the crucial elements of the learning task (35:504). To the extent that overt responding has such a supportive function there may well be less need for it the older or more mentally mature the students are. A definition as well as a thorough discussion of attention in the sense of the term implied here can be found in a paper by Richard Anderson (60).

The hypothesis that overt responding (plus knowledge of results) has the effect of heightening the attention of the students has also been supported in a study by Sime & Boyce (119). This study does not concern PI in a proper sense, since Sime & Boyce measured the effect of questions versus statements concerning concepts taught in a tape-recorded lecture. As they used multiple-choice questions the two response requirements compared correspond to the OMC (overt multiple-choice) and NOR (no response) conditions respectively. Their results indicate that OMC requirement heightens attention as compared to the NOR requirement.

Holland also seems to believe that overt responding primarily is a support for the student, and that, correspondingly, covert responding makes greater demands upon the study discipline of the student (90:93). If this is true, it is also reasonable to assume that overt response requirement is more important the younger the students are. Holland himself does, however, not take this possibility into consideration, but contends that overt responding should be required at all student age levels.

Rippey reviewed 14 studies comparing overt and covert response requirement in PI or other learning tasks (116). He concluded that

> "a number of .. studies indicated that overt responding produced greater achievement for students who were very young .. and who had lower intelligence quotients" (116:218)..

Rippey also concluded that the superiority of overt responding tapers off between grades 4 and 6, and that it is little evident in grade 6.

If there is a relationship between student age and effect of OCR, this should show in a cross-tabulation of student age and outcomes of response requirement comparisons. The results of such a cross-tabulation are shown in table 3.6, which, however, shows no clearcut correlation tendency. But the table demonstrates clearly that the distribution of experiments on the different age categories is skewed: most studies have used older students as subjects, particularly studies comparing OCR and NOR.

It should be noted, though, that 6 out of 8 studies of OCR involving students under 13 years of age have yielded learning effects results favouring this requirement. The data summarized here give, therefore, some support to the hypothesis that overt responding in PI is especially beneficient for the lower age groups. The lack of a more clearcut correlation in the data at hand may be due to the scarcity of experiments involving younger students. The analysis results reported here also indicate, then, what kind of experiments are most needed in the further studying of this problem.

Some of the experiments reviewed here have been designed with the specific aim of studying the relationship between student age, mental maturity, etc., on one hand, and the effect of overt response requirement, on the other.

Table 3.6 Cross-tabulation: Student age x outcome of comparisons between pairs of response requirements as to learning effect.

Comparison	Student age	No. of outcomes type ++	+	0	-	--	Sum
OCR-NOR	≥16	2	4	1	3	1	11
	13-15	1	1	3	4	1	10
	10-12	3	1	1	1	0	6
	<10	0	0	0	0	0	0
OCR-NOR	≥16	6	3	5	4	0	18
	13-15	1	0	1	1	1	4
	10-12	1	1	0	0	0	2
	<10	0	1	0	0	0	1
CCR-NOR	≥16	0	1	0	2	0	3
	13-15	0	1	1	1	0	3
	10-12	0	0	1	1	0	2
	<10	0	0	0	0	0	0

Leith & Hope (38) report significantly better learning effect with OCR than with CCR when the students' mental age is below 11 years. For students of the same chronological age with mental age above 11 years the two response requirements led to approximately equivalent learning effects. According to Leith (35), Biran, Clarke & Leith (5) found significantly better learning effect with OCR than with CCR for "low ability" students in a secondary modern school, while the more "bright" students in secondary modern and junior school learned about as much under both response requirements.

Leith & Ghuman (37) compared OCR with CCR as to learning effect at three grade levels: grades 3, 4 and 5 in grammar school, all subjects being girls. The interaction between response requirement and grade level was not significant, but mean differences pointed to a decreasing effect of overt responding with increasing student age. In grade 5 of the grammar school the difference in learning effect between the two response requirements was even significantly negative, i.e. at this age level (about 16 years) the learning effect of the program on Co-ordinate Geometry was definitely lower when overt responding was required than when covert responding was required.

In a recent study Lewis & Whitwell (96) compared the learning effect of OCR and CCR for 11-year-olds and 13-year-olds, both groups including boys of high to medium mental ability. Apparently, they found no significant interaction of age and response requirement, but mean values reported indicate that the superiority of CCR - covert response requirement - was more marked for the 11-year-olds than for the 13-year-olds. Thus, Lewis & Whitwell's study does not support the hypothesis of increasing effect of overt responding with decreasing student age. Data on high-ability and low-ability students of 11 years are, however, in better agreement with this hypothesis, since they indicate that the negative effect of OCR occurs mainly among the high-ability students. For medium ability students there was no difference between response requirements as to learning effect.

Lambert, Miller & Wiley (33) divided the grade nine students in their study of OCR and CCR into three intelligence groups: high (IQ 120-162), normal (112-119), and low (76-111). They used a fairly long program with 843 frames, but report a significant, positive effect of overt response requirement (OCR) only for students in the h i g h IQ group. Their results runs, then, counter to the hypothesis that overt responding has its greatest positive effects at low levels of intelligence.

In an unfortunately very incompletely reported study Austwick (63) found no difference in learning effect between OCR and NOR requirements when subjects were graduate students, but when he replicated the experiment with 13-year-olds in grammar school OCR led to better learning than NOR. Austwick concludes that

> "the guided instruction provided by the constructed response programme was aiding the learning process for the younger students" (63:198).

A couple of studies by Entwisle, Huggins and Phelps (16) indicate that OCR may be most beneficient - or least detrimental! - for high ability students. Their low ability students did better under a CCR condition. The programs used were apparently very difficult, so the low ability students may have been handicapped by poor understanding of the questions. Attempts to respond to questions that are too difficult may be less efficient than being told the answers right away.

Only half of the studies reviewed in the last paragraphs lend support to the hypothesis discussed here regarding a relationship between student age or maturity level, on one hand, and the effect of OCR, on the other. Also, three of the studies supporting the hypothesis have been performed within the same institution, by the same team of researchers. They should, therefore, not be considered strictly independent pieces of evidence (cf. Rosenthal's view that "ten experiments performed in a single laboratory may be worth less than the same ten experiments conducted in different laboratories", 117:153).

A study by Lockard & Sidowski (97) deserves mentioning in this context. Although it is not a PI study, it illustrates the problem discussed here. In this study the effects of OCR and CCR on the learning of nonsense syllables are compared. Subjects were students in grades 4 and 6. For subjects from grade 4 the overt response requirement (OCR) led to higher effect, although not significantly so, while for grade 6 subjects the OCR led to significantly lower learning effect than the CCR.

In a study comparing conventional programmed instruction and students' reading programs while making notes at their own discretion, Poppleton & Austwick (114) found that PI was more efficient with younger students (13 years), while reading plus note-making was more efficient with graduate students.

Summarizing the present analysis one may conclude that the hypothesis, launched primarily by Leith, that the positive effect of overt responding decreases with increasing student age or maturity, is reasonable although not unchallenged. It is probable that age as such is of no consequence, but rather some characteristics that covary with age: 'work discipline', study skills, background knowledge, etc. (cf. section 3.3.2.3 below). If that is the case, one should expect to find an interaction between student age and type of program as regards the effect of different response requirements. A response requirement that is suitable for a certain type of program and a certain age group may, for instance, be unsuitable for another program taken by the same age group. This could in turn partly explain the absence of a correlation between student age and outcome of response requirement comparisons in table 3.6.

It should be noted that the problem discussed here is part of the general problem of ATI, i.e. aptitude-treatment interactions, which has been much discussed in the recent literature. A quite extensive summary has been published by Bracht in the 1970 issue of the Review of Educational Research. The ATI discussed here need not be linear, i.e. the efficiency of a treatment need not increase or decrease monotonously with some student characteristic. On the contrary, the picture provided by the studies reviewed here indicates that there may be some optimal relationship between student ability and program difficulty, where overt constructed responding is particularly beneficient.

3.3.2.3 Response requirement and learning task characteristics

> "..explicit activity is likely to contribute to the phase of learning referred to as response learning" (35:504; note that "explicit" is used here as synonymous to "overt").
> "Clearly, an overt, constructed response should be required from a student, if he is expected to be able to emit an unfamiliar, technical term" (59:141).

As these two quotations show, both Leith (35) and Anderson (59) assume that there is a relationship between the effect of overt response requirement, on the one hand, and the degree of response learning called for in the program, on the other. A similar point, although not specifically mentioning response learning, was made by Reid already in 1964 (115:159f). It has, however, not been possible to classify the studies reviewed here objectively as to the nature of the learning tasks involved (cf. section 3.2.4 above). Also, the degree of response learning called for in a program cannot be estimated independently of student characteristics, since response terms unknown to one student may be well known to an other.

The nature of the criterion test is also important in this context. The learning effect under overt response requirement can be expected to be correlated to the amount of previously unknown response terms only if the learning of these terms is measured by the criterion test. As Holland has pointed out (90), student responding in PI will influence learning only if the responses required are relevant, i.e. relevant with regard to the criterion test. Holland and Kemp (91 & 94) also have pointed out that responses required in a program must be contingent upon the program content. Otherwise the student's responding will have no relationship to his learning of the program content. Holland and Kemp have devised a special technique and a measure, the blackout ratio, for the extent to which responses required in a program are contingent upon program content.

Most studies reviewed here are not reported in such a way as to permit judgments concerning the criterion test relevance of the required responses. The possible relationship between type of program and effect of overt response requirement may nevertheless be illustrated by means of a summary of the available data. T a b l e 3 . 7 contains a summary of some program and student data for those experiments which report significantly better learning effect, for at least some student category, under the OCR than under the CCR or NOR conditions.

As table 3.7 shows, 12 different programs have been used in the 16 experiments summarized. Four of these programs led to positive effect of OCR for some student subgroups only. Three of these programs (5, 33 & 38) have already been discussed (section 3.2.2.2 above). The fourth, a program entitled "Electronics", was used in two studies (7 & 36) and in both cases the OCR led to significantly better learning than the CCR for students with "low background knowledge" in physics, i.e. students for whom the program probably contained a large number of unfamiliar response terms.

For eight different programs, used in eleven experiments, the OCR had a significant positive effect for the whole student population. One of these program teaches spelling (35), a task very much characterized by response learning and, furthermore, a task where the writing down of responses is particularly important, since what is to be learnt is the correct writing of words. Another program addresses itself to medical students, giving "elementary information" (14: 238), which probably means that the students had little initial knowledge of the learning material. For yet another medical program of "technical, unfamiliar" nature, Tobias & Abramson (49) found a significant difference favouring the OCR condition.

Three of the programs summarized in table 3.7 treat psychological or educational subject matter. The program "Fundamentals of educational test interpretation", used in two studies by Krumbolz & Weisman (31 & 32), is said to have the objective of giving the students

> "a conceptual understanding .. of percentiles, age and grade scores, normal distribution curves, standard deviations and z-scores" (32:90).

Table 3.7 Program and student characteristics of experiments yielding significantly better learning effect under overt (OCR) than under covert (CCR) or no (NOR) response requirement.

Program (title or subject)	Student category	No. of exp	Ref.	Comparison
Fundamentals of test interpretation	Coll. stud:s	2	30,31	Both
Analysis of Behavior	"-	3	24,25,53	NOR
Programmed instruction	"-	1	43	"
Diagnosis of myocardial infarction	Medical stud:s	1	14	CCR
Medical terminology	"-	1	49	NOR
Spelling	Boys, sec.mod.	1	35	CCR
Definitions of mammals	Grade 6 stud:s	1	51	NOR
Vocabulary: difficult	Grade 5 stud:s	1	"	CCR
Electronics	Coll. stud:s w. low backgr.knowl.	2	7,36	"
Molecular interpretation of heat transfer	Sec.mod.stud:s of low ability	1	5	"
Latitude & longitude	11-year-olds, MA below 11 years	1	38	"
Sets, relations & functions	Grade 9 stud:s w. high IQ	1	33	"

This probably implies the use of a number of "unfamiliar and technical" response terms. About the program "Programed Instruction", used by Ripple (43), there is no information available in the research report. It does not seem unreasonable, however, to assume that a program with this title contains a number of response terms, unfamiliar to the students.

The program "The Analysis of Behavior" was constructed by Skinner and Holland, two of the most ardent supporters of the need to require overt responses in PI. This program was used in three studies, two by Holland himself (24 & 25) and one by Williams (53). It is worth noticing that this Skinner-Holland program, as well as the program "Programed Instruction" mentioned above, was used only in comparisons between overt (OCR) and no (NOR) response requirement. Strictly speaking, this only demonstrates that e x p l i c i t response requirements have a positive effect on learning, and not that the o v e r t n e s s of the response activity has a positive effect. An explicit response requirement, whether overt or covert, has the function of informing the student about what elements of the learning task - the program - should primarily be learnt. Such guidance may have a particularly great influence when the program is long; in a short program it may be possible to learn all there is to learn.

The observations reported here indicate clearly a correlation between the nature of the learning task and the effect of overt response requirement: the more unfamiliar response terms a program contains, the higher the positive effect of overt response requirement should be expected to be.

This conclusion is also supported by the results of other kinds of learning experiments. Eigen & Margulies (80), for instance, compared OCR and CCR when the learning material consisted of nonsense syllables at different "information levels". High information level in this case means that the association value of the syllable is low, i.e. the syllable is not similar to any existing word. The authors found that the positive effect of overt response requirement (OCR) increased with increasing information level, i.e. when response terms became more "unfamiliar" to the subjects. At the same time, the OCR had, however, the effect of decreasing the amount of "incidental learning", i.e. learning of syllables not designated as response terms.

It is quite probable, however, that the relationship between the nature of the learning task and the effect of overt response requirement is much more complicated than it has so far appeared to be. Increasing learning task difficulty should be expected to lead to an increase in the positive effect of overt response requirement. But if the degree of difficulty is very high, this positive effect may fail to appear. A study by Goldbeck & Campbell (21) lends some support to this assumption. In this study the authors used 35 items of "factual learning". The degree of item difficulty was varied by means of cueing. The OCR condition had a negative effect when items were easy, and a positive effect when they were moderately difficult. But when items were very difficult, the difference between response requirements as to learning effect was very small. The mean frequency of correct answers to the difficult items was only 21 per cent, a fact which caused Holland to remark:

> "If items cannot be answered anyway, it should make little difference whether the failure to answer is in writing or in thought." (90:96)

I agree, on the whole, with Holland here, but it should be pointed out that it may make a difference whether the "failure to answer is in writing or in thought". If by failure one means making a wrong response, the writing down of the response instead of merely thinking it may have a negative effect on learning.

The idea that overt responding may have a negative effect on learning has also been put forward by Anderson (60). In a balanced discussion of the problem of active responding he first says that

> "when a student can respond correctly to the frames .. without paying attention to critical material, attention to this material is minimized and learning suffers. However, the requirement to make overt responses might be facilitative provided that correct responding was contingent upon attention to all of the critical material" (60:355).

After a review of the Kemp & Holland study of "black-out ratio" (94), demonstrating that only cue-contingent overt responses facilitate learning, Anderson notes:

"It may be true, however, that the requirement to make overt responses which are not cue-contingent disrupts normally adaptive reading habits." (op.cit.)

It must be observed, though, that Anderson discusses a c t i v e responding, i.e. overt responding as opposed to no responding, and that he does not take into consideration the possibility of a c t i v e c o v e r t responding.

Summing up this discussion of the relationship between the nature of the learning task and the relative merit of overt, covert or no response requirement, one may safely conclude that the research reviewed indicates that no predictions regarding the effect of overt responding should be made without taking the characteristics of the learning task into consideration.

3.3.2.4 PI time and learning efficiency

Although different researchers hold different opinions about the value of overt responding in PI, there seems to be general agreement on one point: overt response requirement leads to longer PI time than covert or no response requirement. This is only natural, since the student, having decided what response to give, must devote some time to the recording of his response, that is, if he responds overtly as required. If not, the overt response requirement will not lead to longer PI time than any other response requirement.

In order to evaluate the influence of the response requirement upon learning efficiency one must know both the learning effect and the PI time. Data regarding PI time are, however, reported only in about 50 per cent of the studies reviewed here (see table 3.3 above).

For studies reporting PI time data PI time quotients have been computed (see section 3.2.1.2 above), using the PI time under the NOR or, where this is not included, the CCR condition as denominator of the quotient. The resulting PIT quotients are given in t a b l e 3 . 8 . From data in this table the following PIT quotient values can be computed:

OCR/CCR:
Mean = 1.67
Variation width: 1.36 - 2.40

OCR/NOR:
Mean = 1.78
Variation width: 1.29 - 2.32

The mean increase in PI time following OCR is, then, 67 per cent of the PI time under the CCR, and 78 per-cent of the PI time under the NOR. The lowest PIT quotient reported in any of the 13 studies in table 3.8 is 1.29, i.e. in no case did the OCR lead to an increase in PI time of less than 29 per cent, which is a considerable increase in the total working time spent on a program. Consequently, the OCR condition must lead to a considerable increase in learning effect, in order to reach the same learning efficiency as the CCR or NOR conditions.

Data available in the reports reviewed here give, as a rule, small possibilities of computing measures of learning efficiency. This is mainly due to the fact that the absolute magnitude of the learning effect usually cannot be estimated. Learning efficiency is very difficult to measure under any circumstances. Firstly, this measuring presupposes that learning is measured in an approximate quotient scale, as has been pointed out by Holland (90:102). Secondly, comparisons between quotients of learning effect through learning time presuppose a linear correlation between time and learning. This condition is not likely to be met in practice, since curves relating learning time to performance are usually negatively accelerated, i.e. a given increment produces a smaller increment in performance the longer the total learning time.

Holland's criticism of the habit of computing "efficiency index by dividing test score by learning time" (90:102) is fully justified, but the efficiency problem cannot be dismissed, as Holland seems inclined to do, by such criticism. This point was also made by Lumsdaine, who put forward critical viewpoints similar to those of Holland and ended:

> "These considerations .. do not controvert *the need to take time into account* .." (101:310; my italics).

Table 3.8 PI time (PIT), expressed as PIT quotients, under different response requirements.

Reference	Type of PIT quotient OCR/CCR	OCR/NOR	CCR/NOR
14	1.90	--	--
17	1.61	--	--
21	--	1.20	0.95
21	--	1.80	1.21
25	--	2.14	--
30	--	2.32	--
38	1.66	--	--
44	--	1.50	--
47	1.38	--	--
49	2.40	--	--
52	1.47	--	--
53	--	1.65	--
55	1.37	--	--

Out of the 13 experiments summarized in table 3.8 only one contains data on the level of learning reached by students not taking any version of the program, namely the study by Williams (53). This study reports a posttest mean of 8.8 for a control group not taking the program. If this value is substracted from the experimental group means, one gets

'learning effect' under OCR = 14.7; and
'learning effect' under CCR = 11.8.

Assuming a quotient scale, the increase in learning effect following the OCR would be 25 per cent of the effect under the CCR. But the corresponding increase in PI time is 65 per cent, according to Williams's data. Admittedly, this analysis is based on some unwarranted assumptions, but it shows nevertheless that, although the learning *effect* was highest under the OCR, it is quite possible that the learning *efficiency* was appreciably lower under this requirement than under the CCR.

If quotient scales for the measuring of learning efficiency are not available, one may attempt to judge efficiency by means of a conclusion chart based on Boolean algebra. If the relationship between the means of two experimental groups (1 and 2) in the variables effect (E) and PI time (PIT) can be expressed by < , = or > , one can draw one out of four different conclusions regarding the relationship between the two groups in terms of efficiency (EFF):

(1) EFF(1) > EFF(2) *if*
(E(1) > E(2) *and* PIT(1) < PIT(2)) *or*
(E(1) = E(2) *and* PIT(1) < PIT(2));

(2) EFF(1) = EFF(2) *if*
E(1) = E(2) *and* PIT(1) = PIT(2);

(3) EFF(1) < Eff(2) *if*
(E(1) = E(2) *and* PIT(1) > PIT(2)) *or*
(E(1) < E(2) *and* PIT(1) > PIT(2));

(4) EFF(1) ? EFF(2) /i.e. the relationship is undetermined/ *if*
(E(1) > E(2) *and* PIT(1) > PIT(2)) *or*
(E(1) < E(2) *and* PIT(2) < PIT(2)).

This conclusion chart involves no comparison as to magnitude between a difference in E and a difference in PIT. In return one must allow for the possibility of an undetermined relationship in terms of EFF, i.e. where it is impossible to judge whether EFF(1) is greater than, equal to, or smaller than EFF(2). This is expressed in conclusion no. 4 above.

If the relationships in terms of E and PIT are defined as follows:

> : significantly greater than;
= : not significantly different from; and
< : significantly smaller than;

then one has a crude but usable schema for conclusions regarding the relative efficiency of diggerent experimental conditions.

Even with this rather simple procedure one runs into difficulties when one tries to apply it to data available in the studies reviewed here. Firstly, PI time data are needed, but such data are reported only in about half of the studies (see section 3.3.2 above). Secondly, measures of learning effect (E) are needed, and such measures are, as we have seen, rare (section 3.2.1.1 above). For want of better data one must, then, assume that the relationship between group posttest means reflects the relationship in terms of effect.

The schema for conclusions regarding relative learning efficiency proposed here is applicable in 12 cases of comparison between overt (OCR) and covert (CCR) response requirement, and in 10 cases of comparison between overt and no (NOR) response requirement. The resulting conclusions are summarized in table 3.9. As this table shows, the conclusion is in 9 cases out of 12 that OCR is less efficient than CCR, and in 6 cases out of 10 that OCR is less efficient than NOR. In all the remaining cases the conclusion is that the relationship is undetermined. It is possible, then, that the OCR has been less efficient than the alternative response requirement, CCR or NOR, in all the 22 cases of comparison reviewed here.

Obviously, this evaluation of the relative learning efficiency of different response requirements gives no support whatsoever to the hypothesis that overt response requirement is more efficient than covert or no response requirement. On the contrary, one must expect efficiency to be lower when overt responses are required.

To neglect registering and reporting, in studies of response requirement, the time spent by students working on the program, diminshes in view of these findings, considerably the value of such studies, since it allows no estimate to be made of the learning efficiency of the different experimental conditions.

Table 3.9 Learning efficiency of different response requirements: Summary of conclusions according to a special schema.

(a) OCR versus CCR

Conclusion	No. of exp:s	References
1: OCR more efficient than CCR	0	--
2: OCR and CCR equally efficient	0	--
3: OCR less efficient than CCR	8	2[x)],17,21[x)],33,47,52
4: Relationship undetermined	3	14, 38,60

(b) OCR versus NOR

Conclusion	No. of exp:s	References
1: OCR more efficient than NOR	0	--
2: OCR and NOR equally efficient	0	--
3: OCR less efficient than NOR	6	21[x)],23,44,49,55
4: Relationship undetermined	4	25,31,49,53

[x)] 2 experiments

3.4 INTERACTION OF RESPONSE REQUIREMENT AND KEY ANSWER FREQUENCY

All comparisons of response requirements reviewed so far have been made in combination with 100 per cent key answer frequency for overt and covert response requirement, and, of course, 0 per cent key answers for no response requirement. It is possible, however, that there is an interaction between response requirement and key answer frequency as to the learning effect produced.

Empirical studies of such an interaction in PI have been rare: only three instances have been found in my literature review. And one of these, a study by Fiks (19), is of a very special nature, in that it involved visitors to an exhibition who volunteered to take short - 20-24 frames - programs under different conditions: OCR, CCR or NOR, and 100, 50 or 0 per cent key answer frequency. Fiks found no noticeable interaction between response requirements and key answer frequencies as to posttest results, although he did find a significant interaction between these two factors in terms of attitude to PI. This attitude interaction implied that, under the OCR condition, subjects getting 100 per cent key answers had more positive attitudes to PI than those getting 50 or 0 per cent, while there was no such attitude difference under the CCR or NOR conditions.

Moore & Smith (41) compared OCR with OMR (overt, multiple-choice response requirement) under different key answer conditions: 100 per cent key answers plus cash reward; and 0 per cent key answers and no reward. The analysis of learning effect data yielded no significant interaction between response requirement and response information.

In a recent study Lewis & Whitwell (96) reported an experiment with an, in this context, very interesting multifactorial design. This design included, among other factors, ability (high, medium), age (11 or 13 years), "reinforcement" (i.e. key answer frequency: 100, 33 or 0 per cent), and "response" (overt or covert constructed response requirement). Unfortunately, their data are somewhat unsatisfactorily reported - and data on PI time are also missing. Lewis & Whitwell claim to have found some intriguing interaction between "response" and "reinforcement", but whether one judges by their retention test data alone or by sum scores for retention, "transfer" and "generalisation" tests, there is no clear interaction between response requirement and key answer frequency in their data.

A study by Michael & Maccoby (108) deserves mentioning in this context, although it does not deal with PI. These authors studied the effect of questions in connection with the showing of instructional films; comparing OCR and CCR requirement and 100 and 0 per cent key answer frequency in a 2 x 2 design. They failed, however, to find any significant interaction between response requirement and key answer frequency in their analysis of learning effect data.

The results of this review of research in the interaction of response requirement and key answer frequency are, indeed, meagre - too meagre, perhaps, to warrant even a tentative conclusion. The studies reviewed suggest, however, that one should not expect to find any simple and clearcut relationship between key answer frequency and response requirement as regards their interaction in influencing learning effect in PI.

3.5 STUDIES OF KEY ANSWER FREQUENCY

The literature reviewed contains reports of 17 experiments concerning the effect of varying key answer frequencies (KF) in PI (3, 9, 10, 15 /2 exp/, 18, 19, 24, 30, 32, 34, 39, 41, 42, 43, 48, 49). In two cases (32 & 39) 100 per cent KF is compared with 0 per cent KF as well as with KF:s between 0 and 100 per cent, i.e. intermittent response information. In four experiments (3, 15 /2exp/ & 30) 100 per cent KF is compared only with intermittent response information. In the remaining 12 experiments only the two conditions 100 per cent KF and 0 per cent KF are compared.

3.5.1 Total elimination of key answers

The outcomes of the 16 comparisons between 100 per cent and 0 per cent KF as regards learning effect are distributed as follows (where positive sign means better effect with 100 per cent KF):

Outcome:	++	+	0	-	--
No. of studies	0	8	4	3	1

Only one study resulted in a significant difference, and this was a negative one, i.e. the learning effect was better with no key answers than with 100 per cent KF (39). The distribution of experiment outcomes lends no definite support to any hypothesis concerning the effect of key answers in PI - except, perhaps, to the hypothesis that this effect is generally small and insignificant.

Some cross-tabulations have been performed in order to illustrate the possible interactions of key answer frequency and different program or student variables. In table 3.10 is shown the result of a cross-tabulation of program length, on the one hand, and outcome of comparisons as to learning effect, on the other. (For definitions of the bases of classification see section 3.2.1 above.) The table suggests a correlation between program length and outcome of comparison between 100 and 0 per cent KF: seven out of eight positive outcomes involve programs with less than 200 frames, whereas three out of four negative outcomes involve longer programs (more than 200 frames). In the study yielding a significant negative outcome of the comparison between 100 and 0 per cent KF, the program used contained more than 1,000 frames (39).

This analysis indicates, then, that key answers, presented after the student has responded, may have some positive effect in shorter programs, but no or even negative effect in longer programs. This seems far from unreasonable. That key answers, which are properly used for the purpose of students' response control, may have a positive effect on learning seems obvious (see e.g. Anderson's discussion in 59:151). On the other hand, when students know that a key answer will always be provided immediately after the student's own responding, this may affect their way of working with the program. That is how Lublin explains her finding of a negative effect of 100 per cent KF:

> "The task .. may have been unchallenging .. due to the 100 % predictability that the right answer would immediately be supplied" (39:299).

In this situation the student may make less effort to find the right answer himself, which in its turn should be expected to affect learning negatively. If so, it is also reasonable that this tendency should be stronger the longer the program, since the temptation to get away with as little work as possible probably increases with the amount of work to be done, viz. when the program becomes longer.

Table 3.10 Comparisons as to learning effect between 100 and 0 per cent KF: Cross-tabulation of program length and outcome of comparison.

Program length	No. of ++	comparison +	outcomes 0	-	--	Sum
≥200	0	1	1	2	1	5
100-199	0	4	2	1	0	7
<100	0	3	1	0	0	4
Sum	0	8	4	3	1	16

The argument just presented is also supported by the data on PI time reported by Lublin (39). She found that the mean PI time was l o n g e r for 0 per cent KF than it was for 100 per cent KF. If key answers are actually used for the purpose of response control, this should - other things being equal - increase the total time spent on the program, i.e. PI time for 100 per cent KF should be longer than that for 0 per cent KF. If, on the other hand, key answers are simply neglected, the PI time should be the same regardless of the key answer frequency. Lublin's results must therefore imply that the students in her study did, in fact, devote m o r e t i m e to the program frames, when they knew that key answers would n o t be provided.

As in the case of studies of response requirements in PI (see section 3.3.2.4 above), it is desirable that PI time data be registered and reported in studies of key answer frequencies. In most studies reviewed here this has not been done. It is therefore not possible to judge whether Lublin's finding is typical, i.e. whether the elimination of key answers in PI usually leads to an increase in PI time.

T a b l e 3.11 contains a cross-tabulation of student age and outcome of comparisons between 100 and 0 per cent KF as regards learning effect. There is no clear tendency of a correlation between the two variables: student age and the effect of key answer frequency. It seems, therefore, reasonable to assume that hypotheses regarding the effects of key answers in PI can be formulated without regard to the age of the students involved.

Table 3.11 Comparisons as to learning effect between 100 and 0 per cent KF: Cross-tabulation of student age and outcome of comparisons.

Student age	No. of comparison outcomes*) ++	+	0	-	--	Sum
⩾16	0	4	3	0	1	8
13 - 15	0	1	1	2	0	4
10 - 12	0	3	0	1	0	4
<10	0	0	0	0	0	0
Sum	0	8	4	3	1	16

*) Positive outcome means that the effect is better for 100 per cent KF.

It has been suggested that the effect of knowledge of results upon achievement in PI interacts with student ability, and Golloday (86) even contends that "numerous studies" suggest such an interaction. Still, Golloday himself cites only one study by Eigen (79) as example of those studies and this study contains, in fact, no real support for a hypothesis concerning such an interaction. Lewis & Whitwell (96) have, however, reported a significant interaction of "reinforcement" (0, 33 or 100 per cent KF), "ability" (high or medium), and "process" (type of posttest) for 11-year-olds. Their group means indicate that learning i n c r e a s e s with increasing KF (from 0 to 33 to 100 per cent) for high ability students, but that learning d e c r e a s e s with increasing KF for medium ability students. Lewis & Whitwell provide no explanation for this interaction, and it did not appear in their corresponding analysis of data for 13-year-olds. For the time being it seems wise to regard their finding for the 11-year-olds as one of these isolated, non-generalisable research results that now and then plague us.

The outcomes discussed so far in this section concern only the immediate learning effect. For a more thorough analysis of the effects of key answers it would, of course, be desirable to have access to retention data. Such data are, however, available only in two of the studies reviewed here (43 & 96). Consequently, it has not been possible to analyse the long-term effects of key answers in PI on the basis of the data at hand.

3.5.2 Intermittent response information

The distinction between partial reinforcement - as a learning psychology concept - and intermittent response information - as a programmed instruction variable - has already been discussed (section 2.2 above). Evidently, theories concerning the effect of intermittent response information in PI cannot be based on the results of studies of partial reinforcement in animal learning, or human learning for that matter.

If effects similar to those obtained through partial reinforcement, especially increased resistance to extinction, could be obtained in PI, these effects would be at least partly negative. This point has been stated very clearly by Amsel:

> "In training of problem-solving kinds of behavior, partial reinforcement would result in ... persistance of behavior which does not solve the problem" (58:515).

Although effects similar to those of partial reinforcement may be neither probable nor desirable in PI, intermittent response information may still be worth studying. It is possible that continuous key answer presentation, i.e. 100 per cent KF, as well as total elimination of key answers, i.e. 0 per cent KF, may have a negative effect on learning. Continuous response information may make the student, who knows that a key answer will always be presented, less eager to find out the right answer himself. And total lack of key answers, on the other hand, may have the obvious negative effect of allowing the student to repeat false responses, believing that they are correct (cf. partly similar argument by Moore & Smith, 41:200).

Intermittent response information has been studied in six of the studies discussed here (3, 15/2 exp:s/, 30, 32 & 39). The comparison outcomes are summarized in t a b l e 3 . 1 2 , which also reports program length for each study.

Table 3.12 Outcomes of comparisons between continuous and intermittent response information as to learning effect.

Reference	Program length	Outcome of comparison between KF 100% and 75%	67%	50%	33%	25%	20%	10%	0%
3	140			-		+			
15a	24	-		-		-			
15b	235	+		-		-			
30	177						++	++	
32	177		-		+				+
39	1,144			--					--

Note: Positive sign means that the effect of 100% KF was better than that of the alternative KF

Table 3.12 shows that 100 per cent KF has led to lower learning effect than 75, 67 or 50 per cent KF in six cases out of seven. Compared to the effect of 33 - 10 per cent KF the 100 per cent KF has, however, led to better learning effect in four cases out of six. Apparently, intermittent response information has a positive effect, as campared to continuous response information, when less than half of the key answers are eliminated. When most of the key answers are eliminated, this seems to have a negative effect on learning.

The interpretation of the data discussed here is, however, complicated by the variations in program length. As has already been demonstrated (section 3.5.1 above), key answers seem to have no or even negative effect when programs are long, and positive effect when they are shorter. Two of the studies reviewed here may be interpreted in terms of program length. Krumboltz & Kiesler (30) found that 100 per cent KF led to better learning effect than did 20 or 10 per cent KF, and they used a 177 frame program. Lublin (39), on the other hand, used a 1144 frame program and found that 100 per cent KF led to significantly lower effect than 50 per cent KF.

In three of the six studies reviewed here PI time data are reported. If PI time for each condition is divided by PI time under 100 per cent KF, the following quotients are obtained:

Reference	PI time quotient for KF 100%	75%	50%	25%	20%	10%
3	1	--	1.08	1.16	--	--
15a	1	0.83	1.07	0.95	--	--
30	1	--	--	--	1.07	0.98

Thus, in four cases out of seven intermittent key answer presentation has increased the PI time by 7 to 16 per cent, as compared to PI time under 100 per cent KF. A similar effect of total elimination of key answers (0 % KF) was, as we have seen, reported by Lublin (39). The elimination - totally or partly - of key answers in linear programs can, consequently, in some cases be expected to induce the students to spend more time on the program. If so, positive effects on learning are a natural result of increased learning time. However, there is no clearcut relationship between the PIT quotients presented above and the corresponding outcomes in table 3.12. On the one hand, an increase in PI time of 7-8 per cent accompanying 50 per cent KF, coincides with a slight increase in learning effect. On the other hand, two instances of longer PI time when response information is intermittent (3 & 30; KF:s 25 and 20 %) are associated with lower learning effect than 100 per cent KF.

This analysis of the effects of intermittent key answer presentation in PI remains rather inconclusive, partly because experimental studies of this problem are scarce. It is noteworthy, though, that intermittent key answer presentation can have a positive effect, even if it has yet to be established under what circumstances such an effect is most likely to occur.

3.6 SUMMING UP OF THE RESEARCH REVIEW

The method employed in this review, cross tabulations on the basis of multidimensional classification of experiments, is, obviously, a valuable complement to the more traditional types of research review.

The most important problems and difficulties connected with the application of this method has already been discussed (section 3.2 above). That the method contributes to the demonstration of co-ordination and standardization problems in educational research has also been pointed out. From the review reported here one can, for example, conclude that data on PI time should always be collected and reported in studies of the effects of programmed instruction.

As regards the significance of response requirement and key answer frequency in PI, the review has primarily demonstrated that the influence of these factors as a rule can be expected to depend upon different program and student variables. Theories about the superiority of one particular response requirement can, therefore, hardly be generalised to all program types or all categories of students.

Some of the main reasons for adopting a non-dogmatic view of traditional PI features was admirably summarized by Reid already in 1964. He pointed out that Skinner-type programming was nothing new but simply improved control of learning, and continued:

> "when an interesting, meaningful subject is put into programmed form the value may lie mainly in the improvement in exposition... Formal control by making the learner write and check responses may even

occupy more time than it is worth. If he is intelligent and interested he can scarcely help formulating responses, and since the programme will make sure that his responses are very likely to be correct, he will have had his reinforcement before he starts to write, far less check. ...with dull and disconnected material, e.g. spelling ... , the programming itself may be less important and the control exerted by formal responding and checking may be essential for efficient learning." (115:159f; my emphasis)

This review also leads to the conclusion that the evaluation of different PI features must depend upon whether the learning effect or the learning efficiency, i.e. effect in relation to time spent on the program, is the primary criterion of success. From a practical instructional viewpoint it seems quite clear that the efficiency aspect cannot, without considerable loss of information, be neglected in the experimental study of PI features such as response requirement and key answer frequency. After all, the results of instruction must be evaluated with reference to the price one has had to pay in order to reach them - and student working time is one of the most important cost factors in any instructional situation.

4

The Objectives of the Present Study

The objective of the present study has been to help in providing a basis for decisions, in practical instructional situations, concerning the following issues:

- What influence, if any, has the nature of the response requirement on learning effect and efficiency in programmed instruction using linear, Skinner-type programs?
- What influence, if any, has the presentation of key answers in the same respects?
- Are the effects of response requirement and key answer frequency, if any, different at different age or grade levels?

It should be noted here, that I have considered it more important, in this context, to contribute to the basis for decisions regarding instructional practice, rather than to contribute to the formation of learning theory. This should, of course, not be taken to imply that these two aims are irreconcilable; it is merely a question of priorities.

5.

Delimitation of the Object of Study

5.1 THE INSTRUCTIONAL PROGRAM

It was considered important that the program used in the study should treat a topic included in the official curriculum of the Swedish comprehensive school. When the study was being planned such a program was constructed at the then National School for Educational Research in Linköping. The program is entitled "The main parts of the simple sentence". It treats subject matter included in the mother tongue curriculum for the middle stage (grades 4 - 6) of the comprehensive school. It is recommended that this topic be taught in grade 6, i.e. at the age of 12 years. Up to the beginning of grade 6 it would, therefore, be possible to find classes where this topic had not yet been the object of any systematic instruction. This would make it possible to use one and the same program at different grade levels.

The main objective of the program is to teach the students how to distinguish and name the two main parts of speech (subject part and predicate part) in simple sentences consisting of one head clause only, and to distinguish and name the main word (subject or predicate) within each part of the sentence.

The program is linear and the original version calls for one student constructed response per frame: either one or more missing words are to be filled in the program text, or a direct question is to be answered. In its final version the program has 241 frames. The version used in the pilot experiment of the present study was the second one. The program was revised once more on the basis of pilot study data, and the resulting, third version was used in the main experiment of this study.

It was attempted to construct a program that was easy to read and understand for students in grades 4 - 6. The readability of the final version has been measured using the most common Swedish readability index, LIX, which has been devised by Björnsson (68) and has the following, simple formula:

LIX = Mean number of words per sentence + percentage of long words;

long words being words with more than six letters. The LIX value for the program is 29, which, according to Björnsson (68), is the normal index for texts of medium difficulty for grade 4 students. Thus, students with a reading ability at grade 4 level or higher should be able to read the program without great difficulty.

Data from the readability analysis can also be used to give some idea about the length of program frames. Since this program is entirely verbal, i.e. contains no pictures or diagrams, the length of its frames can be measured in terms of words or sentences per frame. The final version of the program has on the average 20 words per frame, and between three and four sentences per frame. Sentences are, thus, very short, with an average of six to seven words per sentence.

The subject matter of the program is of a kind usually supposed to be well suited for PI, since it is clearly delimited and has a well defined structure, and since the correctness or falseness of propositions concerning the subject matter can be objectively determined. Moreover, repetition using parallel but slightly different examples can easily be built into the program. It should be pointed out, however, that this program contains relatively few terms and concepts that are new to the students, and that it aims at the teaching of identification of general structures rather than the memorizing of response terms. Thus, it belongs to the category of programs where the need for, or the advantage of overt responding can be expected to be rather small (cf. research review above, particularly section 3.3.2.3).

5.2 THE STUDENT POPULATION

As we have seen, the program chosen for the present study should be suitable for use in grades 4, 5 and 6 of the comprehensive school. A study of different PI features using this program can, consequently, include students from all three grades 4 through 6. Previous research in this area has mainly involved young adults or adolescents, viz. American college and high school students (cf. chapter 3 above). The need for additional experiments must, therefore, be considered greatest with respect to younger students.

In view of the objectives of this study the 10-12 year-olds in grades 4 - 6 are a suitable student population. Rippey (116) concluded from his review of research on overt and covert responding that the positive effect of overt responding tapers off between grades 4 and 6. Of course, Swedish students in these grades are somewhat older than their American counterparts but grades 4 - 6 still provide a reasonable age range for a study of this kind.

In terms of student population the present study has, then, been restricted to the middle stage (grades 4 - 6) of the comprehensive school, i.e. to students at the age of 10 to 12 years. The population has, however, been further delimited. Firstly, only normal class students are included, and not students in classes for mentally retarded, reading classes or other special classes. Secondly, the study concerns only students who do n o t 'master' the subject matter of the program prior to the experimental instruction; assuming that the concept 'to master' is operationally defined on the basis of a test of grammar knowledge.

In the pilot study only grade 4 and grade 5 students were included, since this study was performed toward the end of a school year, when the subject matter in question had already been treated rather thoroughly in most grade 6 classes.

5.3 PROGRAM PRESENTATION AND LAYOUT

Previous research by Malmquist & Grundin (104) had demonstrated

- that the presentation of linear Skinner-type programs by means of a simple, manually operated 'teaching machine' can be expected to result in the same amount of learning as presentation in the form of a programmed book, the so called horisontal section book; and
- that the use of this machine usually will lead to longer working time and greater risks of disturbances due to technical faults than the use of programmed books.

Other experiments comparing the effects of machine and book presentations of linear programs have, according to a review by Goldstein & Gotkin (85), yielded similar results. On the basis of a review of research Eigen (78) has recommended that programmed books be used in initial field trials of instructional programs.

In the present study it was, therefore, decided that the program should be presented in the form of a horisontal section book, which will hereafter, for the sake of convenience, be called 'book' or 'booklet'. In its final form the program consists of six booklets, each containing some 40 frames.

Book presentation implies that the much discussed risk of 'cheating' is not eliminated. (It should be noted that many manually operated 'teaching machines' are not 'cheat-proof' either.) Unless constantly supervised the student cannot, then, be stopped from checking the key answer on the following page before recording his own answer. Since constant supervising of students during PI is neither possible nor desirable in everyday school work, it was decided not to make any attempts at eliminating the risk of cheating through supervising in this study. Exact data regarding the amount of cheating that can be expected in situations similar to this one are scarce. Mills (110:130) has reported a 'cheating frequency' of less than 5 per cent in programs of moderate difficulty, but this was in a situation where the students only could cheat by copying from other students' answers.

As some students may be tempted to 'outsmart the program', if they think they have discovered an unforeseen means of doing so, the students were explicitly informed about the possibilities of 'cheating'. But at the same time they were informed that there would be a grammar test after the program, and that they could not expect to learn anything from the program unless they tried to solve the problems themselves (cf. appendix 3: Instruction to the program).

Apart from a possible tendency to 'outsmart' the teacher or the program, cheating can usually be expected to occur only occasionally. Cheating should be no problem, if only the learning task is reasonable and well adapted to the abilities of the student. Moreover, if a student really does not know the answer to a question, it seems wise to look it up rather than to make wild guesses. In other words, what according to traditional teacher and parent morality is called cheating, is in many cases a sensible and rational problem-solving method.

The mode of program presentation chosen for the present study implies, then, that the possible effect of key answers, presented to the student, might be the sum of the supposedly positive reinforcing effect of response information, on the one hand, and its 'cheat-stimulating' effect, which traditionally is supposed to be negative, on the other.

5.4 ALTERNATIVE RESPONSE REQUIREMENTS AND MODES OF RESPONSE INFORMATION

Studies of the effect of different response requirements and modes of response information in PI can include several alternative experimental conditions. The selection of alternatives for the present study has been made on the basis of the following conditions:

(1) The effect of normal Skinnerian response requirement and mode of response information shall be evaluated.

(2) Priority shall be given to alternatives to the Skinnerian model, which are likely to increase learning efficiency.

(3) Experiences from previous research in the areas shall be utilized.

(4) It shall be possible to test all combinations of response requirement and mode of response information in one program, without changing the form or content of the program frames.

(5) The number of alternatives shall be limited to what is economically and practically feasible.

5.4.1 Response requirements

The response requirement of traditional Skinnerian PI, i.e. demand for overt, student-constructed responses, must be included. (For Skinner's own description of this condition see, for example, 121:107.)

Multiple-choice responses are of theoretical interest, but a comparison between constructed and multiple-choice responses allows no evaluation of the effect of the demand for o v e r t responding.

A, from the point of view of efficiency, more interesting alternative to the OCR requirement (overt student-constructed response) is the absence of response requirement. This 'response requirement' has in several investigations been shown to be at least as efficient as the OCR requirement. The no response requirement cannot, however, be combined with different amounts of response information, and does therefore not permit the kind of multifactorial design desired here.

As alternative to the Skinnerian OCR requirement was chosen, instead, the requirement which instructs the student to think out a response but not write it down, i.e. the demand for a covert, student-constructed response - the CCR requirement. This requirement demands no change in the program, only a change in the instructions given to the students. And it can be combined with different modes of response information. Since the CCR requirement as a rule leads to considerably shorter PI time than does the OCR requirement (see section 3.3.2.4 above). it is very interesting from the point of view of learning efficiency in PI.

The following response requirements have, then, been chosen for the present experiment.

R1: an OCR requirement, i.e. the student is explicitly instructed to write down, in the response space, a response which he himself has constructed.

R2: a CCR requirement, i.e. the student is explicitly instructed to think out a response which fits into the response space, but not write it down.

Note that R 1 and R 2 label the particular variants of the response requirements investigated in this study, variants which are, in the end, defined through the wording of the program and of the instructions to the students going through the program. Thus, OCR and CCR label types of response requirement, whereas R1 and R2 label individual instances of the respective types.

5.4.2 Response information

Here too, the response information mode of Skinnerian PI must be included. This means that the student is provided, after each frame, with the correct response to that frame, i.e. the key answer frequency is 100 per cent. (For Skinner's own description of this mode see for example 121:99.) This mode of response information will hereafter be labelled F 1 .

Starting from the F1 alternative defined above, the response information can vary as to type, delay in relation to the responding, and frequency. The most common alternative to F1 in terms of type is to tell the student whether his response is right or wrong. When responses are student-constructed, this type of information is, however, impossible unless one has access to very sophisticated machinery for the evaluation of student responses prior to the presentation of response information. Obviously, this is not possible when the program is presented in book form.

Delay in the presentation of response information is also incompatible with book presentation, since a student going through a programmed book is always entirely free to determine himself the delay between his responding and his consulting the key answer - if he at all consults it. Besides, there is nothing that indicates the possibility of a gain by postponing the presentation of information that will always be presented in the end.

Key answer frequency, however, can easily be varied irrespective of the program presentation mode and whether responding is overt or covert. Previous research indicates also that variations in the key answer frequency in PI may influence learning (section 3.5 above). F1 with its 100 per cent key answer frequency is one extreme on the continuum of key answer frequencies. Possible alternatives are all frequencies below 100 per cent.

In order to judge whether the presentation of key answers at all has an influence on the learning effect of PI, or on the PI time, the effect of totally eliminating the key answers should be studied, i.e. the key answer frequency 0 per cent. Previous experiments have shown that this alternative - especially when programs are fairly long - may result in at least as good learning as the 100 per cent frequency (F1). This 0 per cent key answer frequency, hereafter labelled F0, also allows greater liberty in the layout of programmed books. The so called vertical book, where frames are read one after the other down to the bottom of the page, can, for instance, be used without a masking device. Milan & Bernath (109) have pointed this out, and claim that they have been able to construct a better program thanks to the elimination of key answers.

If the correlation between key answer frequency and learning effect in PI is linear, the two alternatives F1 and F0 would be sufficient for an evaluation of the direction and degree of this correlation. As was demonstrated in the research review (section 3.5.2 above), the possibility of a non-linear relationship between key answer frequency and learning effect should, however, not be excluded. To test the linearity of the correlation the effect must be measured at not less than three points along the key answer frequency continuum.

Since there is no really safe ground for formulating a hypothesis concerning the form of a possible non-linear correlation between key answer frequency and learning effect (cf. section 3.5.2 above), the third alternative has simply been located halfways between the extremes, i.e. at 50 per cent key answer frequency. If the 50 per cent alternative, here called F.5, is to yield a basis for estimating the linearity of the correlation in question, it would perhaps be natural to reduce the key answer frequency from 100 to 50 per cent by means of either a random elimination of half the key answers or elimination of every other key answer. Both this procedures have been used in previous studies (32 & 39), and there is no reason to expect them to differ in their influence upon the learning effect.

From the practical viewpoint of real-life instruction both these procedures of eliminating key answers seem, however, unsuitable, since they do not take account of the structure of the individual program, i.e. neither procedure can be expected to yield response information that is optimally spaced throughout the program. A variant based upon some kind of systematic variation, adapted to the individual program, between frames with and frames without key answers following them would therefore be more appropriate in a study with a practical instructional objective.

In the present study the alternative with 50 per cent key answers has the following form. The program was divided into sequences of 10 to 20 frames, which are relatively 'closed units' as to content. Within each sequence key answers are given after the first 50 per cent of the frames, but not after the remaining frames. This variant is adapted to the program used in the study, and at the same time it can be regarded as a scale point in a variable, viz. the proportion of frames, counted from the first, in each sequence that are followed by key answers.

The different experimental conditions as regards information about correct responses included in the study has, then, been

F1: key answers presented after 100 per cent of the frames in each program sequence;

F.5: key answers presented after 50 per cent of the frames in each program sequence; and

F0: key answers presented after 0 per cent of the frames in each program sequence.

Note that previously employed designations such as 100 per cent KF, 50 per cent KF, denote different types of response information modes. The designations introduced here, F1, F.5, and F0, denote only the individual instances of these types that are studied in the present experiment. Like R1 and R2 (section 5.4.1 above), these alternatives are in the end operationally defined by the material, the instructions etc. used in this study. It is especially important to distinguish between the more general concept '50 per cent key answer frequency' and F.5, since F.5 is only one out of several possible modes of response information that provides key answers after, on the average, 50 per cent of the frames.

6

The Experimental Design

The experiment has the objective of analysing the effects of the variables response requirement (R) and key answer frequency (F) in PI at different grade levels (G). Furthermore, it may be of interest to study difference between sexes (S) as to these effects. The experiment was, therefore, planned so as to allow a complete multifactorial design with

L(R) x L(F) x L(G) x L(S) cells,

where L(i) is the numbers of values of the variable i.

It has been shown (chapter 3 above) that the effect of different response requirements or key answer frequencies may be dependent upon other factors in the learning situation. An interaction between response requirement and the students' grade level is, for example, possible. The mere possibility of such interactions make multifactorial studies of the variables involved desirable. The choice of this particular multifactorial design does not imply, however, that specific hypotheses have been formulated, prior to data collecting, concerning all interactions of the first, second, and third order that can be tested in such a design. The use of a complete multifactorial design has partly the aim of creating a basis for the formulation of new hypotheses.

Most previous experiments concerning response requirement and correct response information in PI have not employed multifactorial designs. The use of very simple research designs can lead to oversimplified conclusions concerning the variables involved. This point was made already in 1961 by Lumsdaine, who emphasized that

> "the need to seek contingent rather than absolute generalizations implies the frequent use of factorially designed experiments in which the operation of one variable can be observed as it interacts with other variables." (98:500)

The question of which interactions and which sets of orthogonal comparisons will be of particular interest is further discussed in the chapter "Hypotheses" below. The number of cells in the complete experimental design is

$$2 \times 3 \times 3 \times 2 = 36,$$

since

L(R) = 2 ; i.e. R1, R2
L(F) = 3 ; i.e. F1, F.5, F0
L(G) = 3 ; i.e. grade 4, 5, 6
L(S) = 2 ; i.e. male, female.

If, for example, 8 replicates per cell is considered desirable, the minimum number of subjects that must be included in the experiment is 288.

7

Hypotheses

The principle of testing only such hypotheses as are formulated prior to the data collection will not be strictly applied in this study. My primary aim in the data analysis has been to interpret the data at hand as objectively and exhaustively as possible. A strict application of the principle mentioned above means that any new idea, arising during the course of an experiment, demands a wholly new experiment for its testing. The alternative, i.e. no strict application of this principle, means instead that the possibilities of data analysis are restricted primarily by the choice of experimental design, variables, measuring instruments, etc. Even if the data analysis in the present study does not rest on Bayesian principles, the following quotation from Edwards, Lindman & Savage is very much to the point:

> "According to the likelihood principle, data analysis stands on its own feet. The intentions of the experimenter are irrelevant to the interpretation of data once collected, though of course they are crucial to the design of experiments." (77:239)

Below are discussed the hypotheses which, in view of available knowledge, have been considered reasonable to formulate.

7.1 HYPOTHESES CONCERNING RESPONSE REQUIREMENT

7.1.1 The influence of response requirement on learning effect

The program used in this study can be considered long enough (more than 200 frames) for a positive effect of the overt response requirement (R1) to show (cf. section 3.3.2.1 above).

The subject matter programmed (elementary analysis of sentence structure) is of such a nature that the positive effect of the overt response requirement (R1) can be expected to be small or even nonexistant (cf. 3.3.2.3 above). The positive effect of R1, as compared to R2, is likely to be greatest for the students who are youngest and have least initial knowledge, i.e. the grade 4 students, and the effect is likely to be smallest for those who are oldest and have most initial knowledge, i.e. the grade 6 students (cf. section 3.3.2.2 above).

In the present experiment one can, then, expect

t h a t the learning effect and the retention is, on the whole, significantly higher under R1 (overt response requirement) than under R2 (covert response requirement); and

t h a t the positive effect of R1 compared to R2 in these respects increases when the grade level of the subjects decreases from grade 6 to grade 4.

7.1.2 The influence of response requirement on PI time

In view of the results of previous research (see especially section 3.3.2.4 above), one can expect in this study

t h a t PI time is shorter under R2 than under R1, irrespective of the variation in other design variables (key answer frequency, grade level, and sex);

t h a t the increase in PI time following R1, as compared to R2, is of the order of magnitude 50 - 100 per cent; and

t h a t this increase in PI time is smaller the higher the grade level of the subjects within the interval grade 4 to grade 6.

7.1.3 The influence of response requirement on efficiency

From the hypotheses listed above it follows

t h a t R2 (covert response requirement) on the whole leads to better learning efficiency than R1 (overt response requirement); and

t h a t this superiority of R2, compared to R1, is more marked the higher the grade level of the students within the interval grade 4 to grade 6.

7.2 HYPOTHESES CONCERNING KEY ANSWER FREQUENCY

7.2.1 The influence of key answer frequency on learning effect

When programs are long (more than 200 frames), 100 per cent key answer frequency, in combination with demand for overt responding, seems to have a predominantly negative effect compared to the effect of 0 per cent key answer frequency (cf. section 3.5.1 above). The effect of key answers when covert responses are required have rarely been studied, but these effects may well be different from the corresponding effects under overt response requirements. If one assumes, as do Leith (35) and Anderson (59), that overt responses have an attention-focusing function; this function may, in the absence of overt responding, be transferred to the key answers. The elimination of overt responding as well as key answers may, therefore, be expected to result in a negative effect, especially for the younger students, who probably need attention-focusing devices in PI most.

Regarding the influence of the key answer frequency upon the learning effect or the retention it can, then, be hypothesized

t h a t under R1 decreasing key answer frequency, from F1 to F0, results in increasing learning effect or retention; and

t h a t under R2 decreasing key answer frequency, from F1 to F0, results in decreasing learning effect or retention.

7.2.2 The influence of key answer frequency on PI time

In view of what has been established by previous research (section 3.5 above) it seems reasonable to hypothesize

t h a t PI time, at least for students working under R1 (overt response requirement), increases as key answer frequency decreases from F1 to F0.

7.2.3 The influence of key answer frequency on learning efficiency

On the basis of the hypotheses concerning learning effect and PI time listed above, one may expect

t h a t under R1 (overt response requirement) the efficiency is approximately equal for the different key answer frequencies, F1, F.5, and F0; and

t h a t under R2 (covert response requirement) the efficiency decreases when the key answer frequency decreases from F1 to F0.

7.3 HYPOTHESIS CONCERNING THE NORMAL SKINNERIAN CONDITION, R1F1

The normal condition in programmed instruction according to the Skinner model is the combination R1F1, i.e. overt response requirement and 100 per cent key answers.

If this normal condition is also, as many assume, the optimal condition, one may formulate the hypothesis

t h a t learning effect, and learning efficiency under the condition R1F1 is greater than the effect and efficiency under any other combination of R and F alternatives.

Contrary to the other hypotheses formulated here this is merely a control hypothesis, and not something the author believes to be true.

Table 7.1 Summary of hypotheses

Criterion	No.	Source of variation	Relationship between cell means
Learning effect	1:1	R	$\mu(R1) > \mu(R2)$
	2	RxG: linear component	$\mu(R1G4) - \mu(R2G4) > \mu(R1G6) - \mu(R2G6)$
	3	RxF: linear component	$\mu(F1R1) - \mu(F0R1) < \mu(F1R2) - \mu(F0R2)$
	4	R1F1 - all other RF	$\mu(R1F1) > \mu(R1F.5, R1F0, R2F1, R2F.5, R2F0)$
PI time	2:1	R	$\mu(R1) > \mu(R2)$; $1.5 \leq \mu(R1)/\mu(R2) \leq 2.0$
	2	RxG: linear component	$\mu(R1G4) - \mu(R2G4) > \mu(R1G6) - \mu(R2G6)$
	3	RxF: linear component	$\mu(F1R1) - \mu(F0R1) < \mu(F1R2) - \mu(F0R2)$
Learning efficiency	3:1	R	$\mu(R1) < \mu(R2)$
	2	RxG: linear component	$\mu(R1G4) - \mu(R2G4) > \mu(R1G6) - \mu(R2G6)$
	3	RxF: linear component	$\mu(F1R1) = \mu(F.5R1) = \mu(F0R1)$; $\mu(F1R2) > \mu(F0R2)$
	4	R1F1 - all other RF	$\mu(R1F1) > \mu(R1F.5, R1F0, R2F1, R2F.5, R2F0)$

Note: $\mu(R1)$ is the population mean for R1 after summation over all F, G and S; $\mu(R1F1)$ is the population mean for R1F1 after summation over all G and S; etc.

8

Variables: General Definitions

The design variables R, F, G, and S have already been defined in chapters 5 and 6 above. Here are presented variables which are included in the study as variates or as covariates.

8.1 VARIATES

The experiment has the objective of studying how the design variables, especially R (response requirement) and F (key answer frequency), influence PI time, learning effect, retention and learning efficiency.

These four variates (dependent variables) are defined, in general terms, as follows.

PI time:	gross time, from start to completion of last frame, for going through each program booklet, rounded off to whole minutes and summed for all six booklets
Learning effect:	increment in performance, measured by means of a special test, from pretesting before the PI to posttesting no. 1 the day after completion of the program
Retention:	ditto from pretesting to posttesting no. 2 about eight weeks after completion of the program

Learning efficiency:	learning effect in relation to PI time (a more precise measure has not been defined beforehand - see section 3.3.2.4 above).

8.2 COVARIATES OR CONTROL VARIABLES

The analysis of effect and efficiency, as defined above, presupposes the existence of a measure of the students' initial knowledge of the subject matter treated in the program. This measure can then be used either in the computation of a performance increment (posttest - pretest) or as a covariate.

A subject's value in a criterion variable will, to some extent, depend upon his intelligence. It has been contended that in PI learning depends to a very small extent - or not at all (!) - on the general intelligence of the student. This contention is, however, to the best of my knowledge not supported by unequivocal empirical evidence. A previous experiment by Malmquist & Grundin (107) showed a significant correlation between IQ of students and learning effect. In the present study this correlation can be expected to be rather high, since the learning task demands a substantial amount of logical thinking, especially in the application of a sentence analysis method, learned from certain examples, to new examples. The verbal and the logico-inductive or reasoning factors of intelligence should be particularly important.

The following covariates or control variables are, then, included in the study.

Initial knowledge:	performance, as measured by a special test, the week before the PI
Verbal intelligence:	the result on an intelligence subtest (DBA:2), administered the week before the PI

Inductive intelligence: the result on another intelligence subtest (DBA:3), administered the week before the PI.

9

The Pilot Study

9.1 OBJECTIVE

The main objective of the pilot study was to control the feasibility of the experiment as regards design, administration, etc., and to test all materials, such as program, instructions and tests. The pilot study did not have the aim of providing data for the testing of the hypotheses formulated for this study. If the statistical analysis of pilot study data has, nevertheless, been rather thorough, this is primarily because this analysis was expected to guide the choice of data analysis methods to be used in the main study.

9.2 SUBJECTS AND EXPERIMENTAL GROUPS

Subjects in the pilot study were students in the then National School for Educational Research in Linköping. Since it was carried out during a spring term, the subject matter in question had already been treated in the grade 6 classes, so the pilot study was limited to two grade levels, 4 and 5.

Within each of four experimental classes the students were divided into 6 experimental groups by means of a random procedure (lot drawing). This was done separately for boys and girls in all subgroups.

The distribution of the [illegible] students in the experimental classes on the different subgroups is shown in table 9.1.

Table 9.1 Number of students in the [illegible] experimental groups.

Group	Exp. cond.	Grade 4 M	F	M+F	M	F	M+F	M+F
1	R1F1	5	[illegible]	10	[illegible]	[illegible]	[illegible]	[illegible]
2	R1F.5	4	6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
3	R1F0	5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
4	R2F1	5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	R2F.5	5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
6	R2F0	5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Total		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

9.3 TESTS AND [illegible]

A preliminary version [illegible] this study, the [illegible] This version consisted [illegible] and 1 with 26 [illegible].

The following [illegible] subjects:

result on [illegible]
result on [illegible]
result on [illegible]
total FI [illegible]
result on [illegible] tion of the [illegible]
result on [illegible] tion of the program [illegible]

For subjects [illegible] down their [illegible] registered

number of [illegible]
booklets [illegible]

9.4 LESSON PLAN

The pilot experiment comprised 11 lessons as follows.

Lesson 1, 2:	Pretesting (DBA tests and PUSL test).
Lesson 3:	Training of the mode of working with the program (through administration of a pre-program).
Lesson 4-9:	Programmed instruction, one booklet each lesson and not more than one lesson each day.
Lesson 10:	Posttesting no. 1, the day after lesson 9.
Lesson 11:	Posttesting no. 2, six weeks after lesson 10.

9.5 IMPLEMENTATION OF THE PILOT STUDY

L e s s o n 1: The intelligence tests DBA:2 and DBA:3 were given in that order in strict accordance with the instructions in the test manual. Experimental class teachers functioned as testers, but classes had been interchanged so that no teacher tested his own students.

L e s s o n 2: The PUSL test, constructed especially for the purpose of this study (for detailed description see chapter 11 below), was administered according to its written instructions by the same testers as in lesson 1.

L e s s o n 3: Students were introduced to the mode of working with the program to which they hade been assigned. A special introductory lesson was held for each experimental condition (groups 1 through 6), so that each student got only the instructions relevant to the condition he was assigned to, and all students assigned to the same condition got the same instructions, irrespective of their grade level.

All instructions were given orally by one instructor for groups 1 - 3 and another for groups 4 - 6.

L e s s o n 4 - 9: Once the students had learnt the mode of working for each particular group, it was not considered necessary to have all six groups work in separate rooms, but only to separate those writing answers in their booklets from those not writing them. Within each grade all R1 groups were brought together in one room and all R2 groups in another. Instructors were the teachers of the four experimental classes. Each teacher supervised a group of students of the same size as a normal class.

The teachers were given detailed instructions in writing about what to say and do during each lesson. The students were given short written instructions, reminding them about the mode of work prescribed for their group. Each lesson started with a short oral introduction by the teacher/instructor.

In order to keep those students occupied, who had completed their booklets, additional tasks were given in the form of arithmetic drill exercises. The students were told to start with these exercises as soon as they had completed the program booklet. A lesson in this series (nos. 4 - 9) was considered finished when the teacher observed that all students had completed their program booklet.

L e s s o n 1 0: The day after lesson 9 the PUSL test was given once again with a slightly different instruction, by the same testers as during lessons 1 and 2. The small difference in test instruction was motivated by the fact that the students now had gone through the program and were supposed to have learnt the subject matter in question.

L e s s o n 1 1: Six weeks after lesson 10 the PUSL test was administered for the third time with the the same instruction as during lesson 10, except for a minor change in the tester's introduction due to the rather long time elapsed since the previous testing occasion.

9.6 DATA ANALYSIS PROCEDURE

9.6.1 Final set of data

Of the 116 students in the four experimental classes 115 belonged to the population defined as students in grade 4 or 5 scoring less than 30 on the PUSL test given as a pretest (cf. definition of population in 5.2 above).

In all 84 students had been present during all lessons in the experiment. The number of students with no missing data was not less than 6 and not more than 9 in any subgroup in the R x F x G design. In order to facilitate data analysis in the multifactorial design R x F x G subjects were excluded by means of a random procedure until exactly 6 subjects remained in each subgroup.

9.6.2 Analysis of variance method

Preliminary analyses of pilot study data indicated some rather marked interactions between sex of subjects and experimental conditions. Analyses of the R x F x G x S design with 24 cells were therefore desirable. As the distribution of boys and girls varied from subgroup to subgroup, the number of subjects per subgroup was not constant in this design. Equal number of subjects per cell could have been obtained only by reducing the number of subjects per cell to 2 and reducing the total number of subjects in the analyses from 72 to 48. In order to avoid this, variance analyses of the design R x F x G x S were performed by means of a method allowing unequal numbers of subjects per cell (124:381). This method involved analysis of cell means instead of cell sums, and also the computation of an approximate error mean square value on the basis of the mean square within cells, which is computed in the usual way. The formula for the approximate error mean square is

$$s_e^2 = \left(\sum_1^k (1/n_i)/k\right)\cdot(S_W/df_W) ;$$

where n_i = number of Ss in cell i,
k = number of cells,
S_w = the sum of squares within cells, and
df_w = the degrees of freedom of S_w.

Principally the same method of analysis as has been described above was employed for the purpose of covariance analyses. The only difference is that the cell means and the sum of squares within cells of the variate are adjusted for regression on the covariates included in the analysis. Adjusted means and sums of squares have been computed following this procedure:

(1) Cell means, square sums and product sums totally (T) and between cells (b) are computed as in conventional analysis of variance (see e.g. 124: chapter 15).

(2) Square and product sums within cells are computed on the basis of the relationship S(w) = S(T) - S(b). From the resulting matrix square and product sums within cells, correlations and regression coefficients within cells are computed in the usual way. The multiple correlation coefficient is then computed using the formula

$$R^2_{y.12..k} = \sum_{j=1}^{v} (r_{yj} \cdot b^*_{yj.12..k}) \; ;$$

where $b^*_{yj.12..k}$ = the partial regression coefficient for covariate j, expressed in a 'scale free' measure (124:326), and
v = the number of covariates.

(3) Adjusted cell means for the variate (Y) are computed from the formula

$$\hat{Y}_i = \bar{Y}_i - \sum_{j=1}^{v} b_{yj.12..k}(\bar{X}_{ij} - \bar{X}_i) \; ;$$

where j = a covariate,
v = the number of covariates,
i = a cell, and
b = the partial regression coefficient within cells.

(4) The sum of squares within cells of the variate is finally adjusted according to the formula

$$S_{w(adj)} = S_w(1 - R^2_{y.12..k}).$$

9.7 ANALYSIS RESULTS

9.7.1 Different modes of analysing posttest results

Four different ways of analysing posttest data by means of variance analysis have been studied. Variates and covariates included in each type of analysis are shown below.

Alt.	Variate	Covariate(s).
1	POST	--
2	POST - PRE	--
3	POST	D2
4	POST	D2, PRE

A l t e r n a t i v e n o . 1 analyses variation between cells in terms of performance level after the PI. Conclusions regarding learning effect and efficiency can only be drawn under the assumption that there were no differences between the groups as to performance before the PI. This assumption is, of course, reasonable in view of the random distribution of subjects on experimental groups.

A l t e r n a t i v e n o . 2 analyses variation between cells in terms of the increase in performance from pretesting to posttesting. Differences between groups in pretest performance are accounted for. The analysis presupposes, however, that the coefficient of regression within cells equals 1. If it does not, the estimate of the error mean square is not minimized (cf. 76:295ff). Also, the reliability of a difference between two scores can be very low, if the scores are highly intercorrelated, as has been pointed out by Guilford (87:393f) in his discussion of the Mosier formula for the reliability of a difference score (111). Unless pretest scores are close to zero, pretest and posttest scores can be expected to be rather highly correlated.

A l t e r n a t i v e n o . 3 analyses variation between cells in terms of performance adjusted for regression on one covariate. D2 is chosen simply because correlation data suggest that it is the best predictor of posttest scores (POST1 as well as POST2). This analysis, then, accounts maximally for the criterion score regression on one single covariate.

A l t e r n a t i v e n o . 4 differs from no. 3 in one respect only: the number of covariates is increased to two. The covariates D2 and PRE (pretest score) are chosen because they constitute the best set of two predictors of posttest results.

Coefficients of correlation within cells in the design R x F x G x S between the various covariates and variates are rendered in t a b l e 9 . 2 . Coefficients of regression within cells for one and two covariates respectively are rendered in t a b l e 9 . 3 .

Table 9.2 Coefficients of correlation within cells in the design R x F x G x S (df = 45).

	D2	D3	PIT	POST1	POST2
PRE	.260	.254	-.053	.541	.468
D2	--	.395	-.174	.698	.536
D3		--	-.208	.477	.340
PIT			--	-.320	-.383
POST1				--	.789

A simple measure of the relative efficiency of the different analysis methods discussed here can be obtained by means of comparing the error variances. Alternatives 2, 3 and 4 result in successively decreasing error variance compared to alternative no. 1. If error variances are expressed as percentages of the alternative no. 1 error variance, the following values are obtained.

Variable	Alt. 2	Alt. 3	Alt. 4
POST1	70	53	39
POST2	81	73	62

Table 9.3 Coefficients of regression within cells and R^2-values for POST1 and POST2, when one and two covariates respectively are included.

Covariates	Variate	b_{D2}	b_{PRE}	R^2
1	POST1	.977	--	.487
1	POST2	.676	--	.287
2	POST1	.837	.679	.626
2	POST2	.561	.558	.402

Judged by the magnitude of the error variance, alternative no. 4 seems clearly superior to all the other alternatives. It also appears reasonable that a study of the influence of certain program variables in PI should, as far as possible, eliminate variation in posttest performance caused by differences betweeen subjects in intelligence and initial knowledge; unless, of course, these variables are included as analysis of variance variables.

There has lately been a great deal of discussion of the use and misuse of covariance analysis (ANCOVA). Berglund (66) and Aiken (57) caution against indiscriminate use of ANCOVA and Elashoff treats the problem quite thoroughly in a 1969 paper (81). Aiken recommends blocking as an alternative to ANCOVA, but Cox (73), on the other hand, found that ANCOVA is clearly preferable to blocking when r - or R - is large, especially if it is greater than 0.8, which is the case in the present study. And Elashoff's conclusion is that

"violation of the assumption /underlying ANCOVA/ of homogeneity of regression ... or homogeneity of variances will be less serious if individuals have been assigned to treatments at random and the x variable has a normal distribution" (81:396).

The pilot study data will not be used for the purpose of hypothesis testing. From the point of view of design feasibility it is, however, interesting to note that ANCOVA according to alternative no. 4 (see above) has yielded a number of significant F-quotients - in spite of the small number of degrees of freedom within cells, i.e. 48. When the variate POST1 was analysed significant F-quotients were obtained for the main factors F and G, and for the interaction R x S. The POST2 analysis yielded significant F-quotients for R and G, and for the interaction R x F x S.

The analysis of learning effect and retention data from the pilot study demonstrated, then, that the design used in the study can be considered well suited for its purpose of demonstrating differences between response requirements and between key answer frequencies as to their effect on learning in PI.

9. 7.2. PI time

When students have a free choice of working tempo, time variables such as reading time or working time often have a positively skewed distribution. This usually implies a correlation between the mean and the standard deviation for different subgroups and, thus, heterogeneity of variance between groups. As a rule this can be corrected by means of a square-root or logarithm transformation of the raw score, i.e. time in minutes or seconds(cf. 76:128-131).

As figure 9.1 shows, the distribution of PI time scores in the pilot study was, however, only moderately skewed. The variance analysis could, therefore, be performed directly on the raw scores, PI time in minutes (cf. 76:132). This variance analysis yielded a number of significant F-quotients: for the main factors R, F and G, and for the interactions R x G, G x S and R x F x G. The pilot study has,

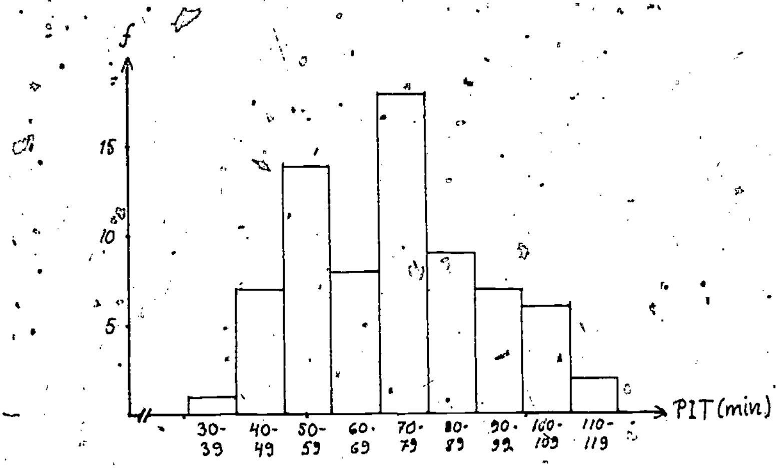


Figure 9.1 Distribution of PI time (PIT) scores for pilot study subjects (N = 72).

then, given ample additional evidence for the need to take PI time into account in studies of this kind (cf. section 3.6 above).

9.7.3 Learning efficiency

As has already been pointed out (section 2.5 above), learning efficiency measures such as performance divided by PI time are usually not acceptable. Under certain circumstances it seems, however, reasonable to try and determine learning efficiency by means of a covariance analysis, where learning effect measures are adjusted for regression on PI time.

A suitable learning effect measure (E) in this study is, according to analyses reported above (section 9.7.1)

$$E = POST1 - b_{D2}(D2 - \overline{D2}) - b_{PRE}(PRE - \overline{PRE})$$

where b is the coefficient of partial regression within cells in the analysis of variance design.

If learning efficiency (EFF) is defined as effect adjusted for regression on PI time, then

$$EFF = E - b_{PIT}(PIT - \overline{PIT}).$$

When efficiency is expressed as a function of effect and time - EFF = f(E,PIT) - the definition of efficiency (section 2.5 above) presupposes that, for a constant E, EFF increases when PIT decreases and vice versa. If

$$f(E,PIT) = E - b_{PIT}(PIT - \overline{PIT}),$$

then the b-value for PIT must be positive, i.e. the correlation between learning effect (E) and PI time (PIT) must be positive, which in its turn implies that the partial correlation between posttest score (POST1) and PI time with D2 and PRE scores held constant, is also positive.

For the pilot study subjects in grade 5 PI time was about the same for both response requirements (overt and covert). This was totally unexpected, and there was reason to believe that it was an incidental instructor effect, which could be eliminated in the main study. Computation and analysis of efficiency (EFF) data as defined above was, therefore, restricted to grade 4 subjects.

Partial regression coefficients within cells (b) were computed from the matrix of zero-order correlation coefficients. For the covariate PIT was obtained

$$b_{PIT} = +0.0704,$$

which means that the requirement stated above is satisfied.

The variance analysis of the learning efficiency variable, EFF, yielded no significant F-quotients. It is interesting to note, though, that whereas the observed effect (E) was greater for R1 than for R2 groups in grade 4, the efficiency (EFF) was greater for R2 than for R1 groups. This demonstrates clearly that the outcome of comparisons in terms of learning efficiency may be quite different from the outcomes of learning effect comparisons.

The relative learning efficiency of the different response requirements can also be estimated by means of the conclusion chart suggested in the literature review (section 3.3. 2.4 above). Due to the significant interaction R x G in the variable PIT (section 9.7.2 above) grades 4 and 5 should be treated separately. As in the previous analysis, effect (E) is posttest score (POST1) adjusted for regression on verbal intelligence (D2) and pretest score (PRE). The application of the conclusion chart yields the following results.

Grade 4: E(R1) = E(R2) and
PIT(R1) > PIT(R2), thus
EFF(R1) < EFF(R2).

Grade 5: E(R1) = E(R2) and
PIT(R1) = PIT(R2) thus
EFF(R1) = EFF(R2).

Note that '=' here means 'not significantly different from'. The conclusion would be, if conclusions were to be drawn from pilot study results, that overt response requirement (R1) leads to lower efficiency than covert response requirement in grade 4, whereas in grade 5 both requirements lead to the same efficiency.

9.7.4 Student response errors in the program

For all students in R1 groups, i.e. those who were told to write down their answers in the program booklets, the sums of incorrect and omitted responses in the booklets were computed. This variable (ERR) has a markedly skewed distribution, as can be seen in figure 9 : 2 . Square-root transformation according to the formula

$$ERT = (ERR + 0.5)^{1/2} \quad \text{(cf. 76:128ff)}$$

gives a roughly symmetrical distribution, as figure 9 . 3 shows. The analysis of the variable ERT has comprised correlation and covariance analysis. The following correlation coefficients within cells have been computed:

	PRE	ERT
D2	.148	-.696
PRE	--	-.560

This gives a squared coefficient of multiple correlation between ERT, on the one hand, and D2 and PRE, on the other, of 0.698, which means that the degree of covariation between ERT and D2 and PRE is 70 per cent.

A variance analysis of ERT cell means adjusted for the regression on D2 and PRE yielded significant F-values for the main factors F and G, and for the interaction F x G. The number of errors (ERT) increased with decreasing key answer frequency (from F1 to F0). This could be due to the fact that the students were allowed to correct their responses during their work with the program, since key answers could be utilized for the purpose of correcting student answers. This indicates that data regarding the frequency of such corrections should be included in studies of this type.

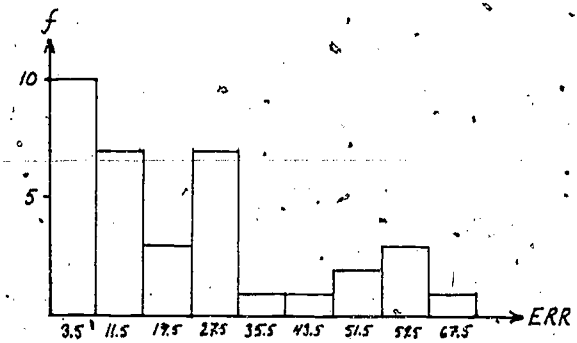


Figure 9.2 Distribution of pilot study subjects in the variable ERR (N = 36).

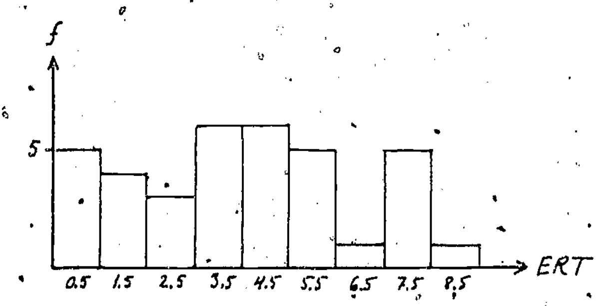


Figure 9.3 Distribution of pilot study subjects in the variable ERT (N = 36).

10

Revision of Program, Design etc. Prior to the Main Study

The pilot experiment findings indicated that the main study could be carried out with the same general design, after a final revision of the instructional program, the instructions to the students and the teachers, and the criterion test - the PUSL test. The program was revised on the basis of error response rates in the pilot experiment and suggestions from the teachers involved. The revision of instructions aimed at increasing the clarity and simplicity of expression; this revision was also performed in co-operation with the teachers involved in the pilot study.

In order to reduce as much as possible non-desirable 'instructor effects' it was decided that all instructions to the students as to how they should work with the program under the different experimental conditions, should be recorded on tape and played back via tape recorder during the introductory lesson (lesson no. 3 in the schedule). By means of brief written instructions, handed out to the students during the PI lessons (lessons no. 4 - 9) - so called reminder sheets - we tried further to diminish the need for oral instructions from the part of the teachers in the experimental groups.

The revision of the PUSL test is further described in section 11 below.

An excerpt from the final version of the program "The main parts of the simple sentence" is found in a p p e n d i x 4.

11

The PUSL Test

The sentence analysis test used in the pilot study contained five subtests with in all 60 items. This test was revised on the basis of pilot study findings. Since the test was intended to make comparisons between students in different grades possible, it was desirable that it permit a relatively great variation of scores above as well as below the expected total mean, i.e. the mean p value for all items should be close to 50 per cent.

The experimental conditions studied here probably have only marginal effects on learning, i.e. these effects will be small compared to the effect of the program as such. It was therefore important that the test should differentiate maximally in the interval around the total mean for all groups, which means that the majority of the test items should have p values fairly close to the average, 50 per cent (cf. 102: 221f).

The revised version of the PUSL test contains three subtests with 15 examples (sentences) each. In subtests A and B each example comprises one test item, whereas in subtest C it comprises two items, which gives in all 60 items. This version of the test was given to 57 students in grade 6, who had taken the program "The main parts of the simple sentence" some six months earlier - in grade 5. The distribution of p values computed from data for this sample is shown in figure 11.1. The mean p value is 52.6 per cent and 47 out of 60 items have p values between 40 and 69 per cent.

The test reliability, according to the Kuder-Richardson formula 20, computed from data for this sample of 57 students was 0.98.

The PUSL test in this version obviously had the desired characteristics as regards the distribution of item p values.

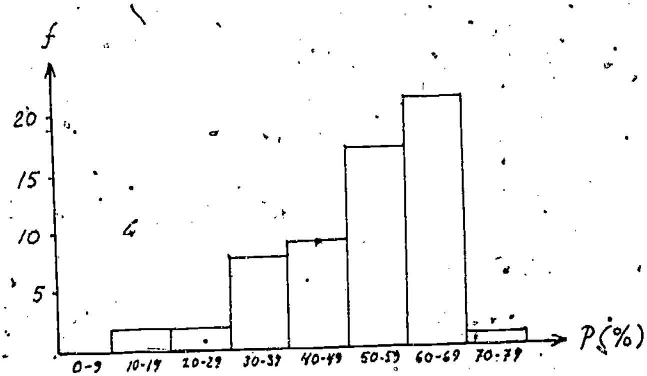


Figure 11.1 The PUSL test: distribution of item p-values (no. of items = 60; no. of subjects = 57).

And the reliability coefficient of 0.98 pointed to a very high item homogeneity. The test could, then, be considered well suited for its purpose: to be used as criterion test in the main study. One minor change was, however, made in the final version: the items were re-arranged and ranked according to p values. The testing time limit was intended to be purely administrative, but in case some students still would not have time to complete all test items, it was considered wise to order them in this way.

The reliability of the final version of the PUSL test was computed from data for the sample used in the main study. The following coefficients were obtained.

Homogeneity: r_{KR20} = 0.96 (posttest 1)

Retest reliability: r_{tt} = 0.79 (posttests 1 & 2).

The final version of the PUSL test is shown in a p p e n d i x 1 .

Instructions and time limits for the final version of the PUSL test were determined on the basis of pilot study experiences and of the try-out of the revised version. The three subtests were timed individually: six minutes for each of subtests A and B, and ten minutes for subtest C. The instructions were identical on all three test occasions (pretesting, posttesting 1 & 2), except for the introduction telling the students about the purpose of the testing etc. The complete instructions are given in appendix 2.

12

The Students Participating in the Main Study

12.1 SELECTION OF EXPERIMENTAL CLASSES

It was decided that the student sample should comprise 6 classes from each of the three grades 4, 5, and 6, making a total of 18 classes. Random selection of classes or schools was not possible, since participation in the experiment was voluntary as far as the teachers were concerned.

All teachers in grades 4 - 6 in four schools in Linköping and two in Norrköping were invited to participate. In three schools, two in Norrköping and one in Linköping, all teachers invited agreed to participate. These schools had in all 20 classes in grades 4 - 6, and 18 of them were selected as experimental classes. The total number of students in these classes was 464.

12.2 THE STUDENTS' PRIOR KNOWLEDGE OF SWEDISH GRAMMAR

All experimental class teachers declared that their students had had no systematic instruction on sentence parts or sentence analysis prior to this study. The concepts 'subject' and 'predicate' had been mentioned by the teachers in all grade 6 classes and in some grade 5 classes - mainly in connection with the teaching of English. The parts of speech 'substantive' and 'verb' had, according to the teachers, been treated rather thoroughly in all classes in grades 5 and 6, and to some extent in grade 4. Pronouns had been treated cursorily in a few classes, mainly in grade 6.

The student population of the present study was defined so as to include only normal class students in grades 4 - 6 who answered correctly less than half of the PUSL test items at the pretesting. Students with a pretest score of 30 or more were, consequently, excluded from the main study. Students who were absent from the pretesting were also excluded.

Table 12.1 shows how many students participated in the pretesting and how many scored above 30 points. The percentage of students with pretest scores of 30 or more was 0 in grade 4, 2 per cent in grade 5, and 8 per cent in grade 6. Pretest means (the variable PRE) for the three grades were as follows (students with 30 points or more excluded).

Grade 4: 3.1
Grade 5: 4.6
Grade 6: 11.3.

Obviously, prior knowledge of grammar increases with increasing grade level. Still, the grade 6 mean is sufficiently low for learning effects following the PI to be measured by means of the PUSL test even in this grade (the maximum possible score is 60).

12.3 ALLOCATION OF STUDENTS TO EXPERIMENTAL GROUPS

Within each experimental class the students were assigned by lot to six different groups. Boys and girls were allotted to groups separately, so as to obtain a fairly even sex distribution in each experimental condition (RF combination).

In order to avoid delays between the pretesting and the experimental lessons the allocation of subjects to groups was performed prior to the analysis of pretest data. Students who were absent from the pretesting or who, due to high pretest score, did not belong to the student population, were thus ~~also~~ allotted to experimental groups. The distribution on experimental groups of the 419 students who took part in the pretesting and belonged to the population was, however, relatively even, as can be seen in table 12.2.

Table 12.1 Number of students in each class with pretest score below or not below 30 points.

Grade	Class	Boys PRE <30	Boys PRE ≥30	Girls PRE <30	Girls PRE ≥30	Total PRE <30	Total PRE ≥30
4	1	10	-	5	-	15	-
	2	9	-	11	-	20	-
	3	11	-	10	-	21	-
	4	17	-	10	-	27	-
	5	13	-	14	-	27	-
	6	8		19	-	27	-
	Sum	68	-	69	-	137	-
5	1	7	1	11	1	18	2
	2	11	-	9	-	20	-
	3	14	-	15	-	29	-
	4	14	-	13	-	27	-
	5	13	-	15	-	28	-
	6	14	1	14	-	28	1
	Sum	73	2	77	1	150	3
6	1	12	1	9	1	21	2
	2	10	-	12	1	22	1
	3	14	-	11	1	25	1
	4	6	2	12	-	18	2
	5	11	2	11	2	22	4
	6	13	1	11	1	24	2
	Sum	66	6	66	6	132	12
4-6	Total	207	8	212	7	419	15

Table 12.2 Distribution of students on the 36 cells in the design R x F x G x S (N = 419).

Grade	Sex	R1 F1	R1 F.5	R1 F0	R2 F1	R2 F.5	R2 F0
G4	M	12	11	11	11	11	12
	F	11	12	12	12	12	10
G5	M	11	13	10	14	12	13
	F	14	13	14	11	12	13
G6	M	11	11	11	11	11	11
	F	12	11	10	11	11	11
Sum		71	71	68	70	69	70

13

Experimental Conditions in the Main Study

The different values of the design variables response requirement (R) and key answer frequency (F) have been defined in general terms (section 5.4 above). The experimental conditions in this study are the six possible pairs of R and F values:

R1F1 , R1F.5 , R1F0 , R2F1 , R2F.5 , R2F0.

Each condition is operationally defined by the combination of program version and instruction presented to the students in each experimental group. The program varies for different F values only (different amount of key answers), whereas the instructions to the students vary both with R and with F.

Here the experimental conditions are described by means of quotations from the taped instruction presented during the introductory lesson (no. 3), and from the 'reminder sheet' handed out to the students at the beginning of each PI lesson (nos. 4 - 9). Formulations which are specific to a particular condition are emphasized.

<u>R1 (overt response requirement), for groups 1 - 3</u>

From instruction on tape:

"In some frames there is a question and an answer space. There you shall w r i t e t h e a n s w e r t o t h e q u e s t i o n . In other frames there is something missing. There you shall f i l l i n t h e m i s s i n g w o r d o r w o r d s. .... If you discover that your answer is wrong, you are free to go back and correct it."

From 'reminder sheet':

"W r i t e a n s w e r s to all questions and f i l l i n all missing words!"

R2 (covert response requirement), for groups 4 - 6

From instruction on tape:
"In some frames there is a question and an answer space. There you shall t h i n k o u t t h e a n s w e r b u t n o t w r i t e i t d o w n . In other frames there is something missing. There you shall t h i n k o u t t h e m i s s i n g w o r d o r w o r d s b u t n o t f i l l i t i n ."

From 'reminder sheet':
"T h i n k o u t a n s w e r s to all questions and t h i n k o u t the missing words!"

F1 (100 per cent key answers), for groups 1 & 4

From instruction on tape:
"When you have finished a frame, you shall check that you have w r i t t e n t h e r i g h t a n s w e r (R1)/ t h o u g h t o u t t h e r i g h t a n - s w e r (R2). The right answer - or the missing words - are written in the left hand margin on the following page.
Maybe you think that you can look at the right answer f i r s t a n d t h e n c o p y i t (R1)/ w i t h - o u t f i r s t t h i n k i n g o u t y o u r o w n a n s w e r (R2). Of course you can, but you won't learn much. - If you want to learn this thoroughly you must first t h i n k y o u r s e l f (R2)/ t h i n k a n d a n s w e r y o u r s e l f (R1), and then check that you have w r i t t e n (R1)/ t h o u g h t o u t (R2) the right answer."

From 'reminder sheet':
"Read the right answer and check that you have w r i t - t e n (R1)/ t h o u g h t o u t (R2) the right answer before you go on to the next frame!"

F.5 (50 per cent key answers), for groups 2 & 5

From instruction on tape:
"Many times you'll find the right answer in the left hand margin on the following page. Then you shall check that you have w r i t t e n t h e r i g h t a n s w e r

(R1)/ t h o u g h t o u t t h e r i g h t a n - s w e r (R2). Sometimes the right answer is not there. Then you cannot check what you've w r i t t e n (R1)/ t h o u g h t o u t (R2). You'll surely get along fine with these questions too, if you just think carefully. - But remember always to look for the right answer, you'll often find it.'

From 'reminder sheet':

"Always look for the right answer! - If it is there, check that you've w r i t t e n (R1)/ t h o u g h t o u t (R2) the right answer, before you go on with the next frame!"

F0 (zero per cent key answers), for groups 3 & 6

From instruction on tape:

"When you have answered a question, you would perhaps like to know if your answer was correct. But t h e r i g h t a n s w e r s a r e n o t w r i t t e n i n t h e b o o k l e t . If you have read the frame and thought about the answer carefully, you'll surely know what is the right answer."

From 'reminder sheet':

" - - - - "

The instructions for the different experimental conditions are, then, paired combinations of the above R and F instructions. Other instructions are the same for all groups. The entire instruction given to group 1 (overt response requirement and 100 per cent key answers) is found in a p p e n d i x 3 .

14

Carrying Through of the Main Study

14.1 LESSON PLAN AND GROUPING OF STUDENTS

The lesson plan of the main study was the same as that of the pilot study (section 9.4 above). During testing lessons (nos. 1, 2, 10, and 11) the students all worked in their own classrooms. Testers were the experimental class teachers, forming pairs and interchanging classes within each pair. During the introductory lesson (no. 3) the students were regrouped so that only students working under the same experimental condition worked in the same room. During the PI lessons (nos. 4 - 9) the students were grouped according to response requirement, i.e. R1 groups from two classes worked in one room led by one teacher; and R2 groups from these classes in another room led by the other teacher.

It was considered important that students working under different response requirements should not influence each other. R1 demanded overt responding and should take longer time than R2 (covert responding). R1 students working together with R2 students might, therefore, be tempted to work too rapidly when they saw other students finishing their booklets very quickly.

During testing lessons, on the other hand, it was important that the situation should be, as far as possible, the same for all students regardless of what experimental condition they had worked under.

14.2 INSTRUCTIONS TO TESTERS/INSTRUCTORS

The teachers who functioned as testers or instructors in the different experimental groups were given detailed written instructions about what to do and say. At a meeting prior to the study they were informed about the design of the study and the purpose of the different instructions. The importance of following, in every detail, all instructions was strongly emphasized.

A sample of written instructions to testers/instructors is given in appendix 2, which contains instructions for

lesson 2 (pretesting lesson);
lesson 3 (introductory lesson); and
lessons 4 - 9 (PI lessons).

14.3 INSTRUCTIONS TO THE STUDENTS

During testing lessons (nos. 1, 2, 10, and 11) the students were instructed orally by the testers. When students were introduced to the experimental conditions (lesson no. 3), a brief instruction was given orally by the instructor (see appendix 2). The major part of the instruction was played back on a tape recorder. The differences between experimental conditions as regards the introductory lesson have been described in chapter 13 above. The full instruction for one of the experimental groups (no. 1: R1F1) is given in appendix 3.

During the PI lessons (nos. 4 - 9) the students worked individually after a brief oral 'starting instruction' (cf. appendix 2). Each student was also given a 'reminder sheet' where the working instructions were summarized. There was one 'reminder sheet' for each experimental condition (combination of response requirement and key answer frequency).

As in the pilot study, students who had finished their program booklets started working on arithmetic exercises. The aim of this was to avoid disturbances and irritation among students who, having finished the booklet, would otherwise have nothing to do. A PI lesson was concluded when all students in the room had finished their booklets.

15

Treatment of Main Study Data

15.1 THE RAW DATA SET

A complete set of raw data for one subject should comprise values in all the variables listed in table 15.1. The data analysis in the main study is based solely on complete individual data sets, which presupposes that the individual has been present during all eleven experimental lessons. Complete data sets are available for 312 subjects, distributed on the cells on the R x F x G x S design as shown in table 15.2. The total set of raw data in this study consists, thus, of 312 individual raw data sets as described in table 15.1.

15.2 MISSING DATA

In all 419 students took part in the pretesting and had less than 30 points on the PUSL test (cf. table 12.1), and complete data sets are - as we have seen - available for 312 students. The proportion of students with missing data is, then, about 25 per cent. This may seem a rather high figure, but it should be noted that it includes all students who were absent from any one out of eleven lessons during a period of several weeks. The absence was in all likeliness due almost entirely to normal causes, mainly temporary illness.

Pretest (PRE) and verbal intelligence (D2) scores are available for the 107 students with missing data. The means for this group have been compared with the corresponding means for the remaining 312 students, and the results are summa-

Table 15.1 Raw data set in the main study.

Variable	Definition
POST1	PUSL-test score at posttesting no. 1
POST2	Ditto at posttesting no. 2 (8 weeks later)
PIT	Total gross working time spent on program booklets 1 - 6, in minutes
PRE	PUSL-test score at pretesting
D2	DBA:2 score (verbal intelligence)
D3	DBA:3 score (inductive intelligence)
W	Total number of program responses that are omitted, erroneous or originally erroneous but corrected by the student
CW	Total number of program responses that are originally erroneous but corrected by the student
G	Grade level
R	Experimental condition in terms of response requirement
F	Ditto in terms of key answer frequency
S	Sex of student
C	No. of class to which student belongs

Table 15.2 Students with complete set of raw data: distribution on cells in the design R x F x G x S

		R1			R2			
G	S	F1	F.5	F0	F1	F.5	F0	Sum
4	M	9	9	9	6	6	10	49
	F	8	8	6	11	10	7	50
5	M	8	10	9	10	7	10	54
	F	10	11	9	10	9	12	61
6	M	10	7	8	9	8	10	52
	F	8	8	6	8	10	6	46
Sum		53	53	47	54	50	55	312

rized in table 15.3. None of the computed t-values for the mean difference in this table is significant, and most of them are close to zero. Apparently, students eliminated because of missing data do not differ markedly from the remaining sample as regards pretest (PRE) and verbal intelligence (D2) scores.

The distributions of missing data on experimental groups have also been studied. The observed frequencies are compared with the frequencies expected under the assumption of equal frequencies of students with missing data (i.e. 25.5 per cent) in all subgroups. Such analyses have been performed separately for the variables sex, grade level, and experimental condition. The results are summarized in table 15.4. All three chi^2-values in that table are very far from significant, which indicates that the variation between groups as to frequency of missing data is not greater than normal chance variation. The distribution of missing data can, consequently, be assumed to be independent of the design variables R, F, G, and S.

To sum up this discussion, it can be observed that the frequency of students with missing data may seem large, mainly due to a very strict demand for completeness of the data sets, but that there is no indication that the absence of these data has affected the differences between experimental groups in the various criterion variables.

15.3 VARIABLES GENERATED FROM RAW DATA

On the basis of the raw data (cf. table 15.1) the following variable values were computed and registered for each student.

SD2, SD3: DBA:2 and DBA:3 score respectively, expressed in a stanine scale for each grade (i.e. with each grade mean equalling 5).

PREM(C): Mean pretest score on the PUSL test for the student's class (C), computed from data for all students participating in the pretest.

Table 15.3 Comparison between eliminated and included subjects with regard to PRE and D2 means.

Variable	Grade	Elimin. Ss n	Elimin. Ss M	Elimin. Ss s	Incl. Ss n	Incl. Ss M	Incl. Ss s	t_{diff}
PRE	4	38	3.3	2.6	99	3.1	2.8	+0.4
	5	35	4.8	5.3	115	4.6	6.0	+0.2
	6	34	10.1	9.1	98	11.3	8.6	-0.7
	4-6	107	6.0	6.7	312	6.2	7.1	-0.4
D2	4	38	21.9	5.7	99	20.6	5.8	+1.1
	5	35	25.5	7.3	115	25.6	5.2	-0.1
	6	34	27.9	6.1	98	28.2	6.2	-0.3
	4-6	107	25.0	6.8	312	24.8	6.5	+0.2

Table 15.4 Distribution of subjects with missing data.

E_o = observed no. of eliminated Ss
E_e = expected no. of eliminated Ss

(a) Sex

S	E_o	E_e
M	54	52.86
F	53	54.14
Sum	107	107.00

$chi^2 = 0.05$, df=1 , N.S.

(b) Grade level

G	E_o	E_e
4	38	34.98
5	35	38.31
6	34	33.71
Sum	107	107.00

$chi^2 = 0.55$, df=2 , N.S.

(c) Experimental condition

Cond.	E_o	E_e	Cond.	E_o	E_e
R1F1	18	18.13	R2F1	16	17.88
R1F.5	18	18.36	R2F.5	19	17.62
R1F0	21	17.36	R2F0	15	17.88
Sum	57	53.62	Sum	50	53.38

$chi^2 = 1.53$
df=1 , N.S.

PREM(GC): Mean of class means (PREM(C)) for grade 6.

PREM(G): Mean pretest score for grade G, computed from data for students included in the results analysis.

On the basis of raw data and the data described here values for the 312 remaining subjects in the variables to be used in result analyses were generated as described below.

15.3.1 Pre-treatment knowledge

The variation in pre-treatment knowledge (measured by pre-test score, PRE) can be regarded as having three components:

(1) an individual component (variation between individuals in each class);

(2) a class component (variation between classes in each grade); and

(3) a grade component (variation between grades).

It is conceivable that each of these components affect test performance in a study of the present kind.

The grade component is already accounted for, since grade level (G) is a design variable in the study. The class component could be accounted for in the same manner: by including class membership as a design variable. This would, however, have led to almost insurmountable missing data problems in this case, so another method of accounting for the class component in pretreatment knowledge had to be devised.

The class component of the variable PRE reflects, among other things, differences between classes as to the nature and amount of instruction received on the subject matter covered by the test. These differences are likely to affect the students' capability of learning the tasks set by the instructional program. If, for example, two students belong to classes with different pretest means due to different amounts of teaching, it is conceivable that these students' ability to learn from the program differs, even if they have the same pretest score. Instead of simply letting the variable PRE be included in the analyses two new variables are, therefore, generated:

(1) The individual's deviation from his class mean

PREI = PRE - PREM(C)

(2) The deviation of the class mean from the mean of class means for the grade to which the class belongs

PREC = PREM(C) - PREM(GC)

Class means are computed from scores for all students participating in the pretesting, including those who scored 30 points or more and do not belong to the population. This mean is the one most appropriate for the present purpose, since it reflects the level of knowledge in the class as a whole. For the same reason the grade mean used in the above formula (2) is the mean of class means, which gives each class equal weight regardless of the number of students in it. All class means and grade means are given in table 15.5.

Table 15.5 Class means, PREM(C), and grade means, PREM(GC) in the variable PRE.

Class	G 4	G 5	G 6
1	0.9	11.0	10.0
2	2.0	3.8	15.8
3	3.2	5.8	15.2
4	5.5	4.6	15.7
5	3.2	7.4	13.4
6	3.0	6.0	11.5
Mean	3.0	6.4	13.6

Note: Class numbers are arbitrary serial numbers. There is no particular relationship between classes in different grades with the same number.

In analyses where pretest result enters as a covariate the variable PRE is, thus, replaced by the two variables PREI (individual component) and PREC (class component). The grade component is a design variable and therefore not used as a covariate:, it enters, however, into the computation of performance increment from pretest to posttest (section 15.3.5 below).

15.3.2 Intelligence measures

Intelligence scores, as measured by the tests DBA:2 and DBA:3, were intended to be used as covariates in the result analyses. The variation in intelligence scores contains, however, also a grade component. Since grade is a design variable in this study, it was desirable to eliminate the grade component from the covariates. In order to obtain this the raw scores D2 and D3 have been replaced by SD2 and SD3, i.e. the corresponding stanine scores, computed for each grade separately. Since the means of these stanine scales are 5 for all three grades, the grade component in intelligence score variation is eliminated from the variables SD2 and SD3.

15.3.3 PI time measure

As figure 15.1 shows the variable PIT has a markedly skewed distribution, which probably implies that the means and the standard deviations for subgroups covary. As suggested by Edwards (76:130), it has been attempted to eliminate the skewness by means of a logarithm transformation, viz. LPT = log PIT.

The variable LPT is, as figure 15.2 shows, approximately normally distributed.

15.3.4 Program response variables

For students in the R1 experimental groups, who wrote answers in their program booklets, the raw data set contains two response error variables: number of incorrect, originally incorrect but then corrected, or omitted answers in the booklets 1 - 6 (the variable W); number of originally incorrect answers that were spontaneously corrected by the student (the variable CW). From these two variables have been computed.

$UCW = W - CW$, i.e. number of uncorrected errors and omissions;

$PCW = CW/W$, i.e. proportion of errors corrected by the students.

The distributions of all four error variables (W, CW, UCW, and PCW) deviate from the normal distribution, as can be seen in figure 15.3.

In order to normalize the distributions the variables W, CW, and UCW were transformed according to the formula

$$x' = \log(x + 1)$$

to LW, LCW, and LUW respectively (cf. 76:130). The variable PCW, which is a proportion, was transformed as suggested by Edwards (76:131) to

$$APC = \arcsin(PCW)^{1/2}.$$

The distributions of the new variables LW, LCW, LUW, and APC are shown in figure 15.4. The three first-mentioned variables have approximately normal distributions, while the distribution of APC is more rectangular. Compared to the raw score distribution (PCW, figure 15.3d) APC is, however, much more symmetrically distributed, and it must therefore be considered more appropriate than the raw score for use in variance analyses.

15.3.5 Learning effect measures

The analysis of pilot study data indicated that learning effect in a study of this kind is best investigated by means of a covariance analysis with posttest score as variate, and pretest score and verbal intelligence score as covariates (cf. section 9.7.1 above). In the main study pretest score

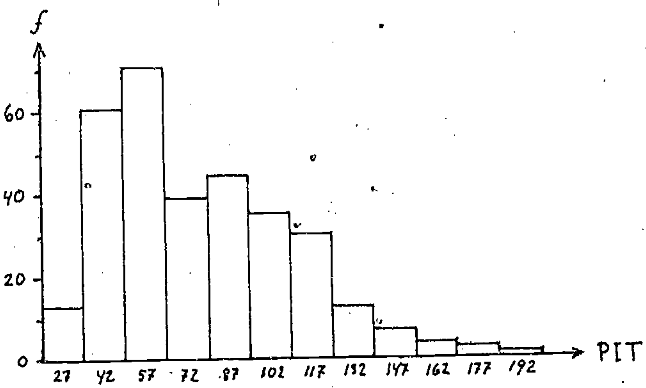


Figure 15.1 The variable PIT: distribution of scores (N = 312).

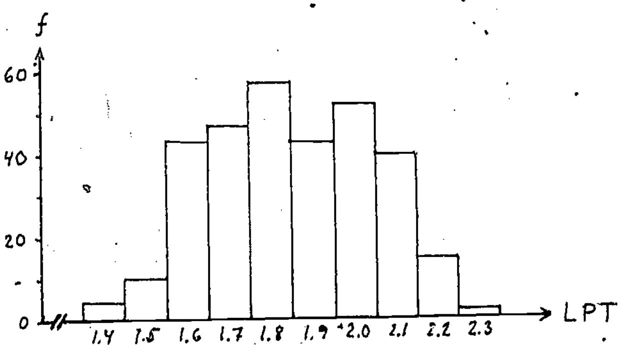


Figure 15.2 The variable LPT (= logPIT): distribution of scores (N = 312).

Figure 15.3 Raw score distributions for response error variables (N = 153)

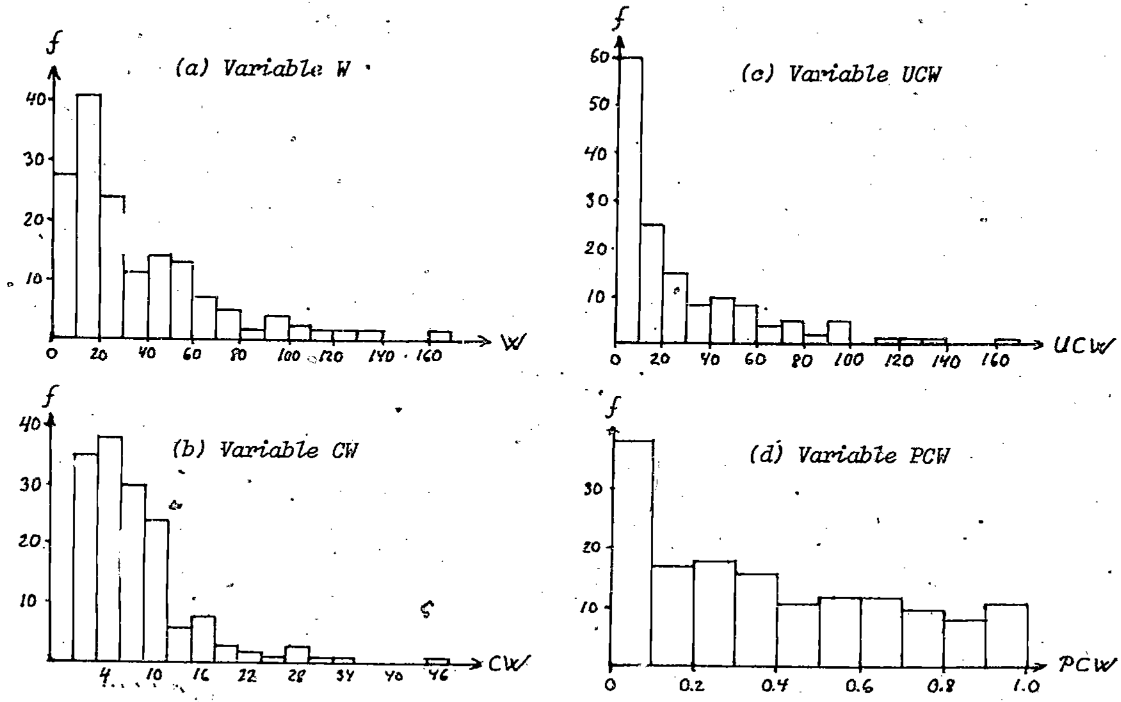
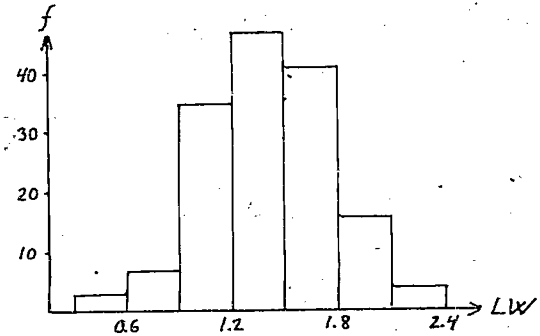
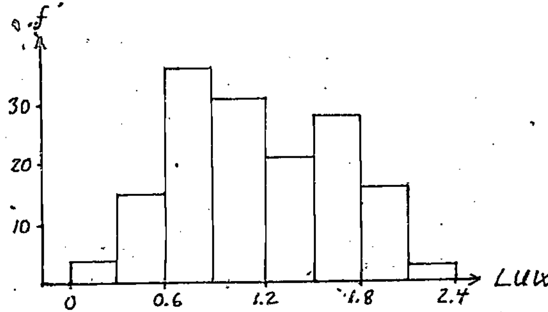


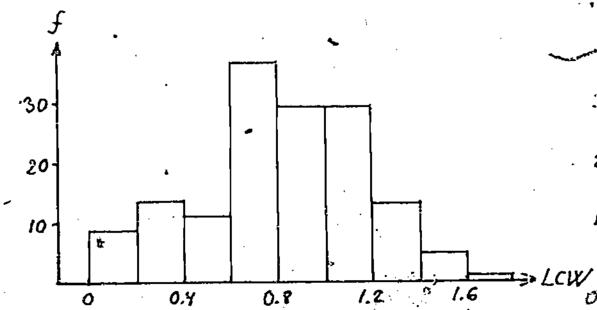
Figure 15.4 Distributions for transformed response error variables (N = 153)



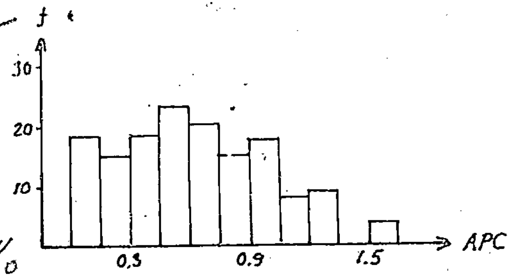
(a) Variable LW



(c) Variable LUW



(b) Variable LCW



(d) Variable APC

is replaced by two pretest score components: individual and class component (section 15.3.1 above).

Since the gain from pretest to posttest is of interest here, and not the posttest performance level in absolute terms, posttest scores are replaced by the difference between posttest score and pretest mean for the grade to which the student belongs, viz.

ED1 = POST1 - PREM(G) ; and

ED2 = POST2 - PREM(G).

The variables ED1 and ED2 are not measures of learning effect, since the individual and class components of the pretest score are not accounted for. Learning effect measures are obtained by adjusting ED1 and ED2 for the regression on the covariates PREI (individual pretest score component), PREC (class pretest score component), SD2, and SD3.

Immediate learning effect (E1) is, then, measured by the following variable

$$E1 = ED1 - \Sigma b_x(x - \overline{x}) \text{ , where } x = \text{PREI, PREC, SD2, SD3}$$

and where b_x is the partial regression coefficient within cells for covariate x. And retained effect (retention) is measured by

$$E2 = ED2 - \Sigma b_x(x - \overline{x}).$$

15.4 SUMMARY OF VARIABLES INCLUDED IN THE RESULT ANALYSES

In the analyses of data from the main study the following variables are, then, employed.

<u>Variates</u>

ED1 = POST1 - PREM(G)	Difference between individual posttest 1 score and pretest grade mean.
ED2 = POST2 - PREM(G)	Ditto for posttest 2.

LPT = logPIT	The logarithm for the individual PI time value.
E1, E2	ED1 or ED2 adjusted for regression on PREI, PREC, SD2, SD3.

<u>Covariates for all groups</u>

PREI = PRE - PREM(C)	Difference between individual pretest score and class means.
PREC = PRE - PREM(GC)	Difference between class and grade pretest means.
SD2 = stanine(D2)	Individual D2 score expressed in stanines within grade.
SD3 = stanine(D3)	Ditto for D3.
LPT	See above under variates!

<u>Covariates for R1 groups only</u>

LW = log(W + 1)	The logarithm for the number of errors and omissions in program answers (booklets 1 - 6).
LCW = log(CW + 1)	Ditto for corrected errors.
LUW = log(UCW + 1)	Ditto for uncorrected errors and omissions.
APC = $\arcsin(CW/W)^{1/2}$	The proportion of errors corrected by the student, transformed through arcsin transformation.

15.5 METHODS OF COMPUTATION AND ANALYSIS

The raw data set (section 15.1 above) and the values SD2, SD3, PREM(C), PREM(G), PREM(GC) for each subject (N = 312) were punched on punch cards. The variable transformations described above (section 15.3) were performed by computer in connection with the use of standard statistical analysis programs from the so called BMD series (cf. 75).

The covariance analyses were performed by means of a combination of automatic and manual computations. Partial regression coefficients, adjusted within cell variances, adjusted cell means etc. for the design R x F x G x S and for different sets of variates and covariates were obtained in output from the computer program BMDO4V, which employs standard statistical formulae (cf. 75). On the basis of such output variance analyses of adjusted cell means were performed manually using the method described in the report on the pilot study (section 9.6 above). In some cases the variation between groups has been analysed into a linear and a quadratic component employing formulae given by Edwards (76:chapter 10). Multiple regression coefficients have in certain cases been computed manually on the basis of zero-order correlation coefficient matrices; in those cases the Doolittle method has been employed (124:326ff).

All machine computations were performed on a GE 625/635 computer owned by Industridata AB, Solna.

16

Data Analysis Results

The word 'significant' in the discussion of main study results means throughout 'statistically significant at the 5 per cent level ($\alpha = 0.5$)', unless some other meaning is specifically stated. Test variable values are marked with one asterisk (*), if they exceed the limit for significance at the 5 per cent level, and with two asterisks (**), if they exceed the limit for significance at the 1 per cent level.

16.1 LEARNING EFFECT

Immediate learning effect is measured by E1, i.e. the ED1 score (POST1 - grade pretest mean) adjusted for regression within cells on the covariates PREI, PREC, SD2, and SD3 (cf. sections 15.3.5 and 15.4 above). The observed coefficients of partial regression within cells (b) and their corresponding t-values for testing the null hypothesis ($b = 0$) are as follows.

	PREI	PREC	SD2	SD3
b	0.48	0.22	3.27	1.26
t_b	4.0**	0.6	8.0**	3.1**

Obviously, posttest 1 results are primarily correlated with intelligence (SD2, SD3) and the individual component of the pretest score (PREI), but not with the general level of pretreatment knowledge in the class (PREC).

All cell means in the variable E1 in the design R x F x G x S are given in table 16.1.

Table 16.1 Learning effect (E1): cell means in the design. R x F x G x S.

G	S	R1 F1	R1 F.5	R1 F0	R2 F1	R2 F.5	R2 F0
4	M	9.72	13.82	13.60	9.18	2.67	8.36
	F	12.78	11.01	14.71	9.72	5.77	9.55
5	M	7.04	11.31	14.61	15.96	19.25	11.28
	F	14.37	15.49	23.53	8.92	10.93	5.37
6	M	15.21	15.31	18.43	15.22	13.05	17.11
	F	17.81	13.69	22.93	16.02	21.51	19.62
4-6	M,F	12.82	13.44	17.97	12.50	12.20	11.88

16.1.1 Variance analysis of the design R x F x G x S

The results of the variance analysis of the E1 cell means (table 16.1) are presented in *table 16.2*. The interactions R x F and R x G have been divided into a linear and a quadratic component and t-values for these components are also included in table 16.2.

On the whole, the variation in experimental conditions has not led to any marked variations in learning effect. Except for the significant R x G x S interaction, which is very difficult to interpret, there are no significant F-quotients involving R or F.

According to the hypotheses formulated in this study (section 7.1 above) some sources of variation are of particular interest, viz. R, R x F, and R x G.

Response requirement (R): The F-quotient for R is close to the significance limit (3.67 as against 3.88). The E1 mean for all R1 cells is 14.74 as compared to 12.19 for all R2 cells. This result is compatible

Table 16.2 Learning effect (E1): variance analysis of cell means in the design R x F x G x S.

Source of variation	df	s^2	F	F_α
R	1	58.47	3.67	3.88
F	2	19.17	1.20	3.04
G	2	151.38	9.50**	3.04
S	1	14.18	<1	3.88
R x F	2	28.80	1.81	3.04
R x G	2	18.21	1.14	3.04
R x S	1	28.34	1.78	3.88
F x G	4	8.15	<1	2.41
F x S	2	1.82	<1	3.04
G x S	2	6.93	<1	3.04
R x F x G	4	20.04	1.26	2.41
R x F x S	2	12.57	<1	3.04
R x G x S	2	60.44	3.79*	3.04
F x G x S	4	1.59	<1	2.41
R x F x G x S	4	2.21	<1	2.41
Error (within cells)	272	15.94	--	

Linear and quadratic components of certain interactions:

Interaction	Component	d	s_d	t_d
R x F	Linear	-34.61	19.54	-1.8
	Quadratic	23.53	33.91	+0.7
R x G	Linear	29.54	20.10	+1.5
	Quadratic	1.96	32.93	+0.1

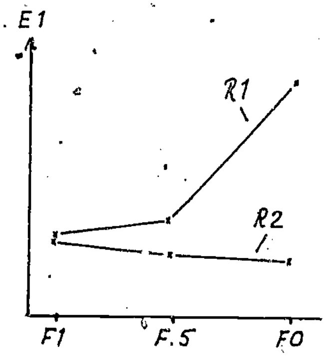


Figure 16.1

Learning effect (E1): R x F

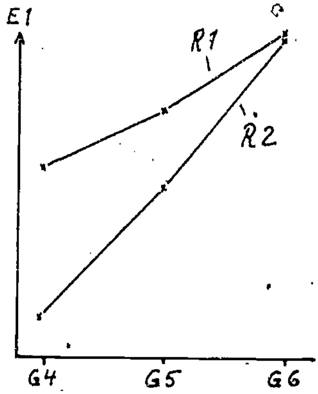


Figure 16.2

Learning effect (E1): R x G

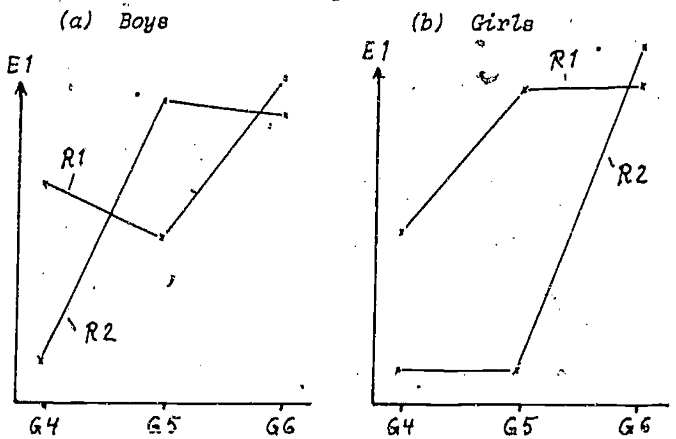


Figure 16.3

Learning effect (E1): R x G x S

with the hypothesis that the learning effect is not much greater under overt response requirement (R1) than under covert (R2).

Interaction between response requirement and key answer frequency (R x F): The interaction R x F is primarily linear, with a small quadratic omponent (table 16.2). The t-value for the linear component is fairly close to the significance limit (t = -1.8). The interaction is graphically illustrated in *figure 16.1*, which shows that decreasing key answer frequency (F) is coupled with increasing learning effect when overt responses are required (R1) but not when covert responses are required (R2). The trend of the result is, thus, in agreement with the hypothesis about differential effect of key answer frequency under different response requirements (cf. chapter 7).

Interaction between response requirement and grade level (R x G): This interaction, too, is mainly linear (table 16.2). The nature of the interaction is seen in *figure 16.2*, which shows that the difference in terms of effect (E1) between R1 and R2 groups decreases with increasing grade level (G 4, 5, and 6). In grade 6 there is practically no difference between R1 and R2 groups. Although the test variable is not significant, this result gives some support to the hypothesis that overt response requirement (R1) is most important in the lower grades and less important in the higher ones (cf. table 7.1 above).

The significant interaction R x G x S (table 16.2) implies, however, that the interaction between response requirement and grade level (R x G) is different for boys and for girls (S), as illustrated in *figure 16.3*. For grades 4 and 6 (on the basis of which the linear R x G component is computed) the outcome is roughly similar for both sexes, whereas in grade 5 there is a marked difference: for grade 5 boys R1 has led to less learning than R2, while the opposite is true for grade 5 girls. A sex difference of this kind could not be foreseen on the basis of previous research, and a satisfactory explanation cannot be presented here.

16.1.2 Grade by grade variance analysis of the R x F x S design

The R x F x G x S can be regarded as three replicates of a R x F x S design, one for each of grades 4, 5, and 6 (G). The results of a grade by grade analysis of the R x F x S design is summarized in table 16.3.

For grade 4 (G4) the F-quotient for R is significant, while other F-quotients are far from significant. The El mean for all R1 groups is 12.60 and for all R2 groups 7.54, which implies that overt response requirement (R1) in grade 4 leads to considerably higher learning effect than covert response requirement (R2). Variations in key answer frequency (F) seem to have no or little influence on learning at this grade level.

For grade 5 (G5) only the interactions R x F and R x S are significant. These are illustrated in figures 16.4 and 16.5. The R x F interaction implies that a reduction of key answer frequency from 100 (F1) to 50 (F.5) per cent has some positive effect regardless of response requirement (R), whereas removal of all key answers (F0) has a still more positive effect when overt responses are required (R1), but a clearly negative effect when covert answers are required (R2). The R x S interaction in grade 5 indicates that the overt response requirement (R1) has a - compared to the covert response requirement (R2) - negative effect on boys, but a positive effect on girls; a finding difficult to explain.

For grade 6 (G6) there are no significant F-quotients for variation between groups (table 16.3); in fact, all the F-quotients are quite small. It seems then as if variation as to response requirement (R) and key answer frequency (F) has very little influence on grade 6 students' learning from the program 'The main parts of the simple sentence'.

This grade by grade analysis of learning effect data is, in a sense, only another way of presenting the results already described of the overall analysis (section 16.1.1 above). Still, the grade by grade analysis singles out particularly clearly a couple of interesting findings:

Table 16.3 Learning effect (E1): grade by grade variance analysis of the design R x F x S.

Grade	Source of var.	df	s^2	F	F_α
4	R	1	76.96	4.83*	3.88
	F	2	10.71	<1	3.04
	S	1	3.19	<1	3.88
	R x F	2	10.24	<1	3.04
	R x S	1	1.00	<1	3.88
	F x S	2	0.70	<1	3.04
	R x F x S	2	4.66	<1	3.04
5	R	1	17.86	1.12	3.88
	F	2	7.97	<1	3.04
	S	1	0.06	<1	3.88
	R x F	2	51.92	3.26*	3.04
	R x S	1	144.91	9.09**	3.88
	F x S	2	3.26	<1	3.04
	R x F x S	2	0.19	<1	3.04
6	R	1	0.06	<1	3.88
	F	2	16.79	1.05	3.04
	S	1	24.80	1.56	3.88
	R x F	2	6.91	<1	3.04
	R x S	1	3.30	<1	3.88
	F x S	2	1.04	<1	3.04
	R x F x S	2	11.95	<1	3.04
4-6	Within cells	272	15.94	--	

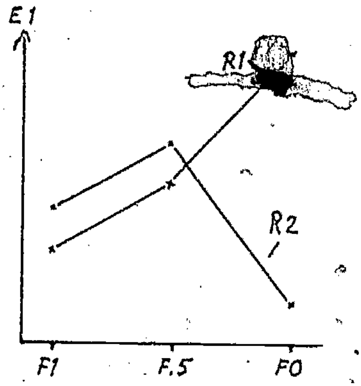


Figure 16.4
Learning effect (E1) in grade 5: R x F

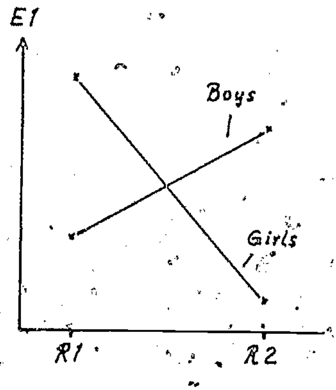


Figure 16.5
Learning effect (E1) in grade 5: R x S

- The overt response requirement (R1) has led to significantly better learning effect than the covert response requirement only at the lowest of the grade levels investigated, grade 4.

- A significant interaction between response requirement (R) and key answer frequency (F), in accordance with the hypothesis stated in chapter 7, has been demonstrated; only for grade 5, that is true, but this is the grade for which the program employed in the study was originally intended.

16.2 LONG-TERM LEARNING EFFECT

The long-term learning effect is measured by E2, i.e. the ED2 score (POST2 - grade pretest mean) adjusted for regression within cells on the covariates PREI, PREC, SD2, and SD3 (cf. sections 15.3.5 and 15.4 above). The observed coefficients of partial regression within cells (b) and their corresponding t-values for testing the null hypothesis (b = 0) are as follows.

	PREI	PREC	SD2	SD3
b	0.47	0.79	2.84	1.38
t_b	3.7**	2.1*	6.5**	3.2**

All four t-values are significant, which means that all the covariates contribute significantly to the regression equation. The coefficients for PREI, SD2, and SD3 are roughly equivalent to the corresponding coefficients obtained in the analysis of learning effect (E1; section 16.1 above), which means that the individual pretest score component (PREI) and the individual's intelligence (SD2, SD3) influence long-term learning about as much as they influence immediate learning.

The regression coefficient for PREC, however, is significant here, although it was not in the previous analysis. The class component in pre-treatment knowledge seems, then, to influence long-term learning - retention - more than it influences immediate learning. Perhaps immediate learning from the program 'suppresses' as it were the effect of previous instruction given in the class, and previous learning increases again in relative importance as the retention of learning from the program diminishes over time. It is also conceivable that learning relevant to the PUSL test, occurring within the classes after the PI, is more highly correlated with the level of pre-treatment knowledge in the class, than it is with the effect of the programmed instruction.

All cell means in the variable E2 in the design R x F x G x S are given in table 16.4.

Table 16.4 Long-term learning effect (E2): cell means in the design R x F x G x S.

G	S	R1 F1	R1 F.5	R1 F0	R2 F1	R2 F.5	R2 F0
4	M	7.86	11.25	5.51	10.88	0.41	7.43
	F	9.68	9.57	13.06	6.10	4.73	6.69
5	M	6.60	9.75	8.57	12.97	18.84	8.31
	F	9.61	12.68	20.40	6.84	8.30	4.57
6	M	12.09	18.98	11.18	11.61	17.06	11.18
	F	16.63	13.23	22.57	15.02	14.51	14.21
4-6	M,F	10.41	12.58	13.55	10.57	10.64	9.52

16.2.1 Variance analysis of the design R x F x G x S

The results of the variance analysis of the E2 cell means are summarized in table 16.5. The interactions R x F and R x G have been divided into linear and quadratic components, which are also included in table 16.5.

On the whole, the variation in experimental conditions has not entailed any great variation in long-term learning effect (E2). In that respect the results of the analysis of long-term learning correspond well with those of the immediate learning effect (E1; section 16.1 above).

The only significant F-quotient for E2 involving R or F is the R x S interaction, which is illustrated in figure 16.6. The overt response requirement (R1) - compared to the covert response requirement (R2) - seems to affect boys' long-term learning negatively or not at all, whereas girls' long-term learning is clearly affected positively. To the extent that girls in grades 4 - 6 are more mentally mature than boys in the same grades, this finding is contrary to the hypothesis that the positive effect of overt responding decreases with increasing mental maturity. It is possible, of course, that boys and girls differ in their 'mode of learning' in a PI situation. If so, the nature of their different 'modes' remains to be identified. Whatever the explanation may be, the R x S interaction found here indicates that the effect of a response requirement may depend on other variables - in this case the sex of the student.

In view of the hypotheses formulated in chapter 7 the sources of variation R, R x F, and R x G are particularly interesting. These variation sources are discussed below.

Response requirement (R): The difference between response requirements as to long-term learning (E2) points in the same direction as the corresponding difference in immediate learning (E1): the overall mean is higher for R1 than for R2 groups. The R1-R2 difference is, however, smaller for E2 than for E1, which means that it gives no support for a hypothesis that overt responding should have a more positive influence on retention than on immediate learning. Such a hypothesis would be reasonable, to the extent that overt responding is believed to 'consolidate' learning. The result of the E2 analysis is compatible, though, with the hypothesis formulated in this study, which expects the overt response requirement (R1) to lead to a - compared to the covert response requirement (R2) - rather

Table 16.5 Long-term learning effect (E2): variance analysis of cell means in the design R x F x G x S.

Source of variation	df	s^2	F	F_α
R	1	33.66	1.87	3.88
F	2	4.70	<1	3.04
G	2	171.34	9.51**	3.04
S	1	4.82	<1	3.88
R x F	2	13.13	<1	3.04
R x G	2	5.20	<1	3.04
R x S	1	93.80	5.21*	3.88
F x G	4	8.97	<1	2.41
F x S	2	30.23	1.68	3.04
G x S	2	3.26	<1	3.04
R x F x G	4	18.53	1.03	2.41
R x F x S	2	22.67	1.26	3.04
R x G x S	2	22.21	1.23	3.04
F x G x S	4	8.35	<1	2.41
R x F x G x S	4	6.54	<1	2.41
Error (within cells)	272	18.01	--	

Linear and quadratic components of certain interactions:

Interaction	Component	d	s_d	t_d
R x F	Linear	-25.10	20.77	-1.2
	Quadratic	-0.02	36.04	-0.0
R x G	Linear	14.35	21.36	+0.7
	Quadratic	10.87	34.99	+0.3

small increase in long-term learning effect.

Interaction between response requirement and key answer frequency (R x F): The R x F interaction is far from significant. The observed trend is almost totally linear (table 16.5) and very similar to the one found in the analysis of E1, only less marked. The interaction is illustrated in figure 16.7. The hypothesis that decreasing key answer frequency (F) entails increased retention when overt responses are required (R1), but decreased retention when covert responses are required (R2), is not verified. The observed cell means are, however, not at all incompatible with this hypothesis.

Interaction between response requirement and grade level (R x G): The R x G interaction is also far from significant, neither is it clearly linear (table 16.5). It is illustrated in figure 16.8, which shows that the positive effect of the overt response requirement (R1), if any, is still greatest at the lowest grade level (4). So the hypothesis that the positive effect on retention of R1 - as compared to R2 - is greater the lower the grade level gets little support from these data, even if it is not incompatible with them.

16.2.2 Grade by grade analysis of the R x F x S design

As for E1, a grade by grade variance analysis of E2 has been performed on the R x F x S design. Its results are summarized in table 16.6, which contains no significant F-quotients for grade 4 or grade 6. For grade 5, only the R x S interaction is significant, with a trend quite similar to that found in the corresponding analysis of learning effect (E1; section 16.1.2 and figure 16.5 above), as can be seen in figure 16.9. Thus, the grade by grade analysis of long-term learning effect (E2) shows the same tendencies as the corresponding analysis of immediate effect (E1), with the only notable difference that the variation between subgroup means is, on the whole, smaller.

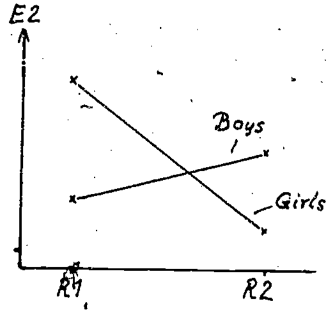


Figure 16.6

Long-term learning effect
(E2): R x S

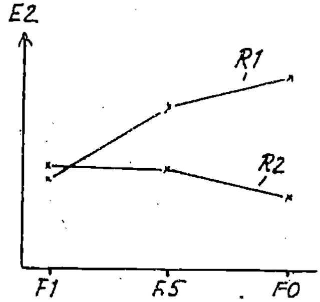


Figure 16.7

Long-term learning effect
(E2): R x F

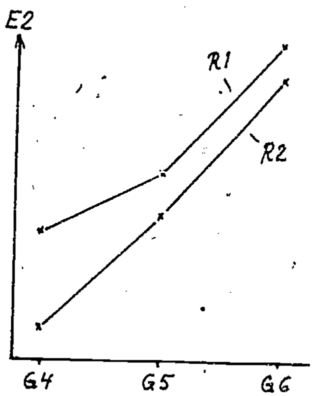


Figure 16.8

Long-term learning effect
(E2): R x G

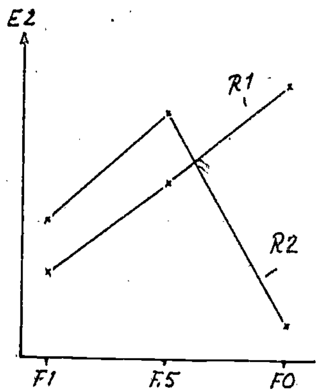


Figure 16.9

Long-term learning effect
(E2): R x F for grade 5

16.3 PI TIME

The variable LPT, i.e. the logarithm of the PI time in minutes, is used as PI time measure (cf. section 15.3.3 above). Partial regression coefficients (b) and the corresponding t-values (t_b) for the covariates PREI, PREC, SD2, and SD3 with LPT as variate have been computed with the following results:

	PREI	PREC	SD2	SD3
b	-.001	.002	-.001	-.001
t_b	-1.1	0.7	-1.1	-1.0

All the above t-values are far from significant, and the analysis of LPT has, therefore, been performed without adjustment for regression on the covariates. The cell means are given in t a b l e 1 6 . 7 , expressed in the original scale, i.e. transformed back to PI time in minutes (antilogLPT). A summary of the variance analysis of the R x F x G x S design is found in t a b l e 1 6 . 8 , which shows three significant F-quotients: for R, for R x F, and for R x G.

In view of the hypotheses regarding PI time formulated for this study (table 7.1 above), the variation sources R, R x F, and R x G - which also yielded significant F-quotients - are of particular interest.

R e s p o n s e r e q u i r e m e n t (R) : As expected the PI time varies with the response requirement. The overall mean time for R1 groups is 98.8 minutes, as compared to 49.8 minutes for R2 groups. This gives a PIT quotient of 1.98, i.e. the overt response requirement (R1) has led to twice as long PI time as the covert requirement (R2). The hypothesis that PI time is greater under R1 than under R2 is, then, strongly supported, and so is the additional hypothesis that PI time under the R1 condition is 1.5 to 2 times the PI time under the R2 condition (cf. table 7.1 above).

Table 16.7 PI time (PIT): cell means in the design R x F x G x S.

G	S	R1 F1	R1 F.5	R1 F0	R2 F1	R2 F.5	R2 F0
4	M	116.4	119.7	126.6	52.6	45.2	54.3
	F	97.2	124.2	106.7	52.5	48.6	55.3
5	M	93.4	112.5	109.3	59.1	54.2	58.2
	F	81.4	111.4	107.3	47.5	55.4	48.5
6	M	72.0	90.1	93.3	47.0	40.8	44.0
	F	84.3	79.2	78.9	44.4	48.4	44.7
4-6	M,F	89.8	104.9	102.6	50.3	48.5	50.5

Interaction between response requirement and key answer frequency (R x F): The R x F interaction is significant and the analysis of its components has yielded about equal t-values for the linear and the quadratic components, although only the latter is significant (t - 2.1; table 16.7). The interaction is graphically illustrated in two ways in *figure 16.10*: in the form of PIT means for different R and F combinations, and in the form of PIT quotients (R1/R2) for different values of F.

When overt responses are required (R1), the PI time increases with decreasing key answer frequency from 100 (F1) to 50 (F.5) per cent. When key answer frequency is further decreased to 0 per cent (F0), the PI time decreases somewhat. When covert responses are required (R2), the PI time is about the same for all key answer frequencies, only slightly shorter for F.5 than for F1 and F0 (figure 16.10a). The non-linear character of the interaction is most clearly perceived in figure 16.10b, which shows how the PIT quotient for R1/R2 varies with F. The increase in PI time entailed by the overt response requirement, expressed as percentage of PI time under the covert requirement, is 79 per cent for F1,

Table 16.8 PI time (PIT): variance analysis of cell means in the variable LPT.

Source of variation	df	s^2	F	F_α
R	1	.8001	579.80**	3.88
F	2	.0032	2.32	3.04
G	2	.0352	25.38**	3.04
S	1	.0032	2.30	3.88
R x F	2	.0054	3.91*	3.04
R x G	2	.0060	4.34*	3.04
R x S	1	.0011	<1	3.88
F x G	4	.0011	<1	2.41
F x S	2	.0022	1.56	3.04
G x S	2	.0012	<1	3.04
R x F x G	4	.0009	<1	2.41
R x G x S	2	.0010	<1	3.04
R x G x S	2	.0017	1.21	3.04
F x G x S	4	.0008	<1	2.41
R x F x G x S	4	.0017	1.19	2.41
Error (within cells)	276	.0014	--	

Linear and quadratic components of certain interactions:

Interaction,	Component	d	s_d	t_d
R x F	Linear	-.337	.182	-1.9
	Quadratic	-.663	.316	-2.1*
R x G	Linear	.499	.187	+2.7**
	Quadratic	.345	.307	+1.1

116 per cent for F.5, and 103 per cent for FO.

The hypothesis that, when overt responses are required, decreasing key answer frequency leads to increasing PI time is, then, supported by these findings. The hypothesis seems to need some modification, though, since the PI time under the R1 condition has its maximum when the key answer frequency is 50 per cent (F.5) and not when it is 0 per cent (FO). Still, the PI time under the R1FO condition is considerably longer than it is under the R1F1 condition.

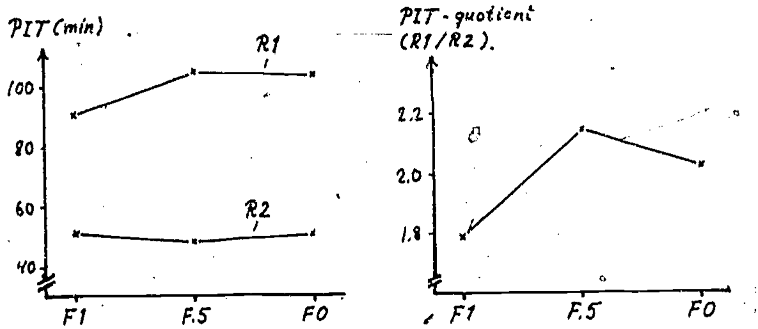
Interaction between response requirement and grade level (R x G): The interaction R x G is significant and primarily linear (table 16.7). This can be seen in figure 16.11, where it is illustrated in the form of PIT means for all R and G combinations, and in the form of PIT quotients (R1/R2) for the three grade levels (G). The difference in PI time between R1 and R2 groups decreases with increasing grade level, which means that the hypothesis formulated for this interaction (table 7.1 above) is unequivocally supported.

As figure 16.11a shows, there is a marked correlation between PI time and grade level under R1 condition, with time decreasing as grade increases. Under the R2 condition, on the other hand, PI time means for different grades are much more similar.

16.4 LEARNING EFFICIENCY

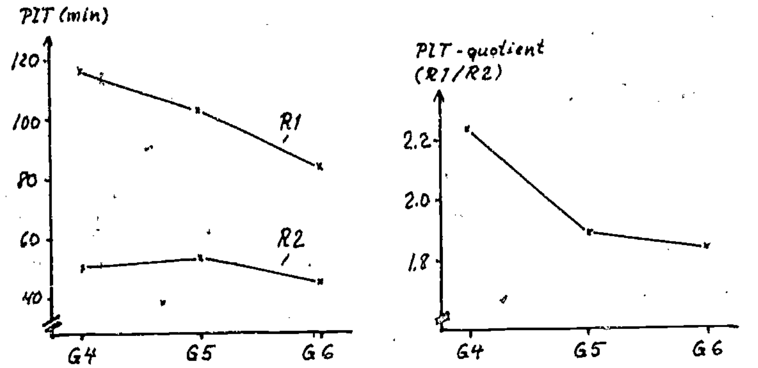
The problem of measuring learning efficiency has been discussed in different contexts (sections 2.5, 3.3.2.4, and 9.7.3) and two alternative methods of estimating learning efficiency under different experimental conditions have been suggested:

(1) The variable PI time (PIT), if necessary transformed in order to normalize its distribution, is used as covariate, which gives as efficiency measure (EFF) the effect



(a) *PIT-values for RF combinations* (b) *PIT-quotients for different F*

Figure 16.10 PI time (PIT): R x F



(a) *PIT-values for RG combinations* (b) *PIT-quotients for different G*

Figure 16.11 PI time (PIT): R x G

measure (E) adjusted for regression on PIT, i.e.

$$EFF = E - b_{PIT}(PIT - \overline{PIT})$$, where b_{PIT} is the coefficient of partial regression within cells for PIT (cf. 9.7.3 above).

(2) A conslusion chart is used for determining the relative efficiency of different experimental conditions as to learning effect and PI time (section 3.3.2.4 above).

16.4.1 Variance analysis of EFF1 : the design R x F x G x S

Of the two suggested methods for estimating efficiency the first one seems preferable when the necessary conditions implied in it are satisfied; particularly the condition that the partial regression coefficient for PIT is greater than zero (section 9.7.3 above). This method of estimating efficiency only accounts for the linear regression of learning effect on PI time, although it cannot be taken for granted that this relationship is linear (section 3.3.2.4 above). Since PIT is transformed here to LPT, which is approximately normally distributed (figure 15.2 above), the assumption of linear regression seems, however, tenable.

The efficiency measure discussed here (EFF1) has been obtained by means of adjusting the variable ED1 for regression on the previous covariates PREI, PREC, SD2, and SD3, plus the covariate LPT, i.e.

$$EFF1 = ED1 - \Sigma b_x(x - \overline{x})$$, where x = PREI, PREC, SD2, SD3, and LPT, and b_x is the partial regression coefficient within cells for x.

The following values have been computed for these partial regression coefficients and their corresponding t-values for testing the null hypothesis (b = 0).

	PREI	PREC	SD2	SD3	LPT
b	0.50	0.18	3.34	1.31	16.64
t_b	+4.2**	+0.5	+8.2**	+3.3**	+2.6**

The coefficients for the four covariates involved in the computation of the effect measure (PREI, PREC, SD2, and SD3) are very similar to those obtained in the analysis of learning effect (section 16.1 above). The regression coefficient for LPT is significantly positive, which means that the variable EFF1 is an acceptable estimate of learning efficiency in the present study. EFF1 has been analysed in the same manner as the measures of learning effect (E1) and retention (E2). The cell means in EFF1 in the design R x F x G x S are found in table 16.9, and the variance analysis of this design is summarized in table 16.10.

The variations in experimental conditions do not seem to have entailed any great differences in learning efficiency. Except for the interaction R x G x S, which is extremely difficult to interpret, there are no significant F-quotients involving the factors R or F (table 16.10). In view of the hypotheses that have been formulated concerning learning efficiency (table 7.1 above), a closer inspection of the variation sources R, R x F, and R x G is, however, needed.

Response requirement (R): The F-quotient for R is not significant, but it is fairly close to the limit value (3.39 as against 3.88). The overall EFF1 mean for R1 groups is 12.24, and for R2 groups 14.67. The observed efficiency is, then, higher when covert response are required (R2) than when overt responses are required (R1). This is in agreement with the hypothesis formulated above (table 7.1): covert response requirement leads to more efficient learning than does overt response requirement.

Interaction between response requirement and key answer frequency (R x F): The R x F interaction is predominantly linear, but it is not significant (table 16.10). The interaction is illustrated in figure 16.12, which shows that the efficiency is higher in R2 groups than in R1 groups when key answer frequency is 100 (F1) or 50 (F.5) per cent, but not when it is 0 per cent (F0). It has been hypothesized that the efficiency should be independent of key answer frequency under the R1 condition (table 7.1 above), and decrease with the key answer frequency under the R2 condition (ditto). The observed trend is not in agreement with this hypothesis, except perhaps in one respect: the difference between response requirements (R) as to efficiency is smallest for the lowest key answer frequency (F0). It seems reasonable, therefore, to reformulate the hypothesis as follows:

> The increase in efficiency entailed by R2 as compared to R1 is most marked when the key answer frequency is high, and least marked when it is zero.

I n t e r a c t i o n b e t w e e n r e s p o n s e r e q u i r e m e n t a n d g r a d e l e v e l (R x G) : This interaction is almost completely linear, but far from significant (table 16.10). It is illustrated in f i g u r e 1 6 . 1 3 , which shows that the EFF1 value is higher for R2 than for R1 groups at all grade levels. The difference between R1 and R2 groups in terms of EFF1 is, however, smallest in grade 4 and greatest in grade 6, so the result is not incompatible with the hypothesis formulated above (table 7.1; hypothesis no. 3:2).

Table 16.9 Learning efficiency (EFF1): cell means in the design R x F x G x S.

G	S	R1 F1	R1 F.5	R1 F0	R2 F1	R2 F.5	R2 F0
4	M	6.12	9.97	9.33	11.11	5.86	10.15
	F	10.46	6.99	11.66	11.79	8.45	11.30
5	M	4.69	7.85	11.39	17.13	21.13	12.57
	F	13.02	12.15	20.54	11.74	12.67	8.04
6	M	15.01	13.50	16.50	18.21	16.97	20.44
	F	16.49	12.73	21.94	19.22	24.12	23.03
4-6	M,F	10.97	10.53	15.23	14.87	14.88	14.26

Table 16.10 Learning efficiency (EFF1): variance analysis of cell means.

Source of variation	df	s^2	F	F_α
R	1	52.95	3.39	3.88
F	2	15.04	<1	3.04
G	2	234.41	15.01**	3.04
S	1	22.53	1.44	3.88
R x F	2	26.12	1.67	3.04
R x G	2	9.92	<1	3.04
R x S	1	33.56	2.15	3.88
F x G	4	6.48	<1	2.41
F x S	2	4.28	<1	3.04
G x S	2	3.96	<1	3.04
R x F x G	4	17.10	1.09	2.41
R x F x S	2	9.59	<1	3.04
R x G x S	2	51.35	3.29*	3.04
F x G x S	4	2.57	<1	2.41
R x F x G x S	4	2.15	<1	2.41
Error (within cells)	271	15.62	--	

Linear and quadratic components of certain interactions:

Interaction	Component	d	s_d	t_d
R x F	Linear	-29.24	19.34	-1.5
	Quadratic	34.58	33.56	+1.0
R x G	Linear	21.76	19.89	+1.1
	Quadratic	-2.74	32.59	-0.1

The interpretation of the R x G interaction is complicated - as was the case with the learning effect analysis (section 16.1.1 above) - by the significant interaction R x G x S, which is illustrated in figure 16.14. As this figure shows, R2 has been more efficient than R1 for both sexes in grades 4 and 6, whereas in grade 5 R2 has been most efficient with boys and R1 with girls. It has already been pointed out that such an interaction could not have been foreseen on the basis of previous research, and no satisfactory explanation of it can be presented here.

On the whole, the findings presented here are not incompatible with the hypothesis that the superiority of covert response requirement - as compared to overt - in terms of efficiency is greater the higher the grade level of the students. On the other hand, the findings do not give any strong support to the hypothesis.

16.4.2 Grade by grade analysis of the R x F x S design

The efficiency variable, EFF1, has also been subjected to a grade by grade analysis, performed on the design R x F x S. The results of this analysis are summarized in table 16.11.

In grade 4 all F quotients are less than 1, i.e. the variation between groups is small compared to that within groups. It should be remembered that the analysis of learning effect (E1) yielded a significant F quotient for the factor R in this grade, meaning that the effect was higher in R1 than in R2 groups (section 16.1.2 above). There is, however, no corresponding difference in terms of efficiency. On the contrary, the EFF1 mean for the overt response requirement groups (R1) is slightly lower than that for the covert response requirement groups (R2), viz. 9.09 as compared to 9.78. Consequently, the fact that the learning effect in grade 4 is higher under the R1 than under the R2 condition can be fully explained by the difference in working time (PI time) entailed by the two conditions.

In grade 5 the R x F and R x S interactions are sig-

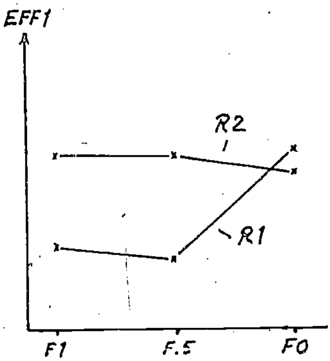


Figure 16.12

Learning efficiency (EFF1):
R x F

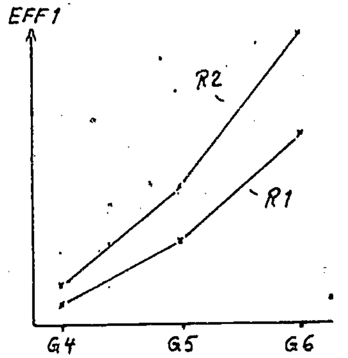


Figure 16.13

Learning efficiency (EFF1):
R x G

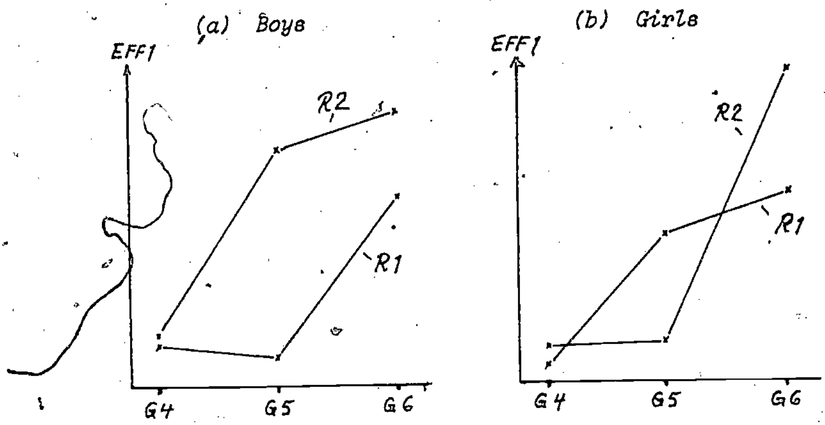


Figure 16.14 Learning efficiency (EFF1): $\overline{R}$ x G x S

Table 16.11 Learning efficiency (EFF1): grade by grade variance analysis of the design R x F x S.

Grade	Source of var.	df	s^2	F	F_α
4	R	1	1.42	<1	3.88
	F	2	8.37	<1	3.04
	S	1	5.48	<1	3.88
	R x F	2	5.19	<1	3.04
	R x S	1	0.04	<1	3.88
	F x S	2	1.94	<1	3.04
	R X F x S	2	5.70	<1	3.04
5	R	1	15.50	<1	3.88
	F	2	3.72	<1	3.04
	S	1	0.96	<1	3.88
	R x F	2	47.64	3.05*	3.04
	R x S	1	134.40	8.60**	3.88
	F x S	2	5.43	<1	3.04
	R x F x S	2	0.07	<1	3.04
6	R	1	55.86	3.58	3.88
	F	2	15.90	1.02	3.04
	S	1	24.00	1.54	3.88
	R x F	2	7.85	<1	3.04
	R x S	1	1.82	<1	3.88
	F x S	2	2.40	<1	3.04
	R x F x S	2	7.75	<1	3.04
4-6	Within cells	271	15.62	--	

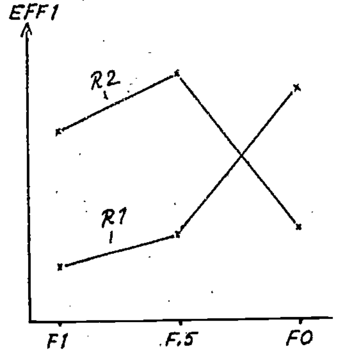


Figure 16.15 Learning efficiency (EFF1): R x F for grade 5

nificant, i.e. the same interactions that yielded significant F quotients in the E1 analysis (section 16.1.2 above). The interaction R x F, which is illustrated in f i g u r e 1 6 . 1 5 , implies that 100 and 50 per cent key answers (F1 and F.5) lead to higher efficiency under the R2 condition than under the R1 condition, whereas 0 per cent key answers (F0) has the opposite effect: higher efficiency under R1 than under R2. It is worth noting in this context that the traditional combination of maximum amount of "active responding" and information about correct answers (R1F1) as well as the combination of minimum amounts (R2F0) both seem to give lower efficiency of learning than most of the other combinations of R and F.

In g r a d e 6 there is no significant F quotient, although the value for the factor R is close to the limit value, viz. 3.58 as against 3.88. The analysis of learning effect (E1) showed practically no difference between the R1 and the R2 groups in this grade: means were 17.23 and 17.09 respectively. The observed difference in terms of efficiency (EFF1) is, however, much greater with lower efficiency in the R1 groups than in the R2 groups: EFF1 means 16.03 and 20.34 respectively. This finding is in agreement with the hypothesis of higher efficiency under covert (R2) than under overt (R1) response requirement, especially in the higher grades (cf. chapter 7.1 above).

16.4.3 Efficiency estimates based on special conclusion chart

The relative efficency of learning (EFF) under different response requirement conditions can also be evaluated by means of the conclusion chart suggested in the research review above (section 3.3.2.4 above). Since there is a tendency of interaction between response requirement (R) and grade level (G) in terms of learning effect (E1; section 16.1.1 above), this evaluation had better be made for each grade level separately.

The following relationships have been found between learning effect (E1) and PI time (PT) mean for R1 and R2 groups in the different grades.

G r a d e 4 : E1(R1) > E1(R2)
PT(R1) > PT(R2)

G r a d e 5 : E1(R1) = E1(R2)
PT(R1) > PT(R2)

G r a d e 6 : E1(R1) = E1(R2)
PT(R1) > PT(R2)

Note that = means 'not significantly different from', and > 'significantly greater than'.

From these premises the following conclusions can be drawn concerning the relative efficiency (EFF) under the R1 and R2 conditions in each grade.

G r a d e 4 : EFF(R1) ? EFF(R2)

G r a d e 5 : EFF(R1) < EFF(R2)

G r a d e 6 : EFF(R1) < EFF(R2)

Thus, efficiency of learning is apparently lower under the overt (R1) than under the covert (R2) response requirement in grades 5 and 6, whereas in grade 4 the difference between these response requirements in terms of efficiency is undetermined. These findings lend a great deal of support to the hypothesis that the R1 condition leads to lower efficiency of learning than the R2 condition. The findings are also compatible with the hypothesis that this difference between response requirements in terms of efficiency is greater the grade level of the students.

16.5 THE 'STANDARD' R1F1 CONDITION COMPARED WITH OTHER EXPERIMENTAL CONDITIONS

The standard combination, in Skinnerian PI, of overt response requirement and 100 per cent key answer frequency, i.e. the equivalent of condition R1F1 in the present study, is often assumed to provide good, if not optimal conditions for learning in most PI situations. It is not always made clear, though, if the criteria of good learning conditions refer only to the effectiveness of learning, or if they also take efficiency into account. In the present study both effectiveness and efficiency are considered important in attempts to evaluate the outcomes of PI.

As a kind of 'control' hypothesis (cf. chapter 7 above) it has been assumed in this study that the R1F1 combination of response requirement and key answer frequency is superior to all other RF combinations investigated here, both in terms of immediate (E1) and long-term (E2) learning effect and in terms of learning efficiency (EFF1). This hypothesis has been tested by means of variance analyses of the three variates (E1, E2, and EFF1) on the design (R1F1 - all other RF) x G x S, which has 2 x 3 x 2 = 12 cells.

The results of these variance analyses are summarized in tables 16.12, 16.13 and 16.14. As can be seen in tables 16.12 and 16.13, there is no significant variation associated with the factor RF, neither in E1 nor in E2. However, the overall mean for R1F1 groups is in both cases - E1 and E2 - lower than the total mean for all other RF groups: 12.82 against 13.60 in E1, and 10.41 against 11.37 in E2. Clearly, then, the standard Skinnerian

Table 16.12 The combination R1F1 compared with all other RF combinations in terms of learning effect (E1).

(a) Variance analysis results

Source of variation	df	s^2	F	F_α
RF	1	1.83	<1	3.88
G	2	43.46	4.73**	3.04
S	1	18.63	2.03	3.88
RF x G	2	4.82	<1	3.04
RF x S	1	10.14	1.10	3.88
G x S	2	0.32	<1	3.04
RF x G x S	2	5.63	<1	3.04
Error (within cells)	272	9.18	--	

Note: RF = R1F1 - all other Rf

(b) Cell means

G	S	R1F1	Other RF's
4	M	9.72	9.53
	F	12.78	10.15
5	M	7.04	14.48
	f	14.37	12.89
6	M	15.21	15.82
	F	17.81	18.75
4-6	M,F	12.82	13.60

Table 16.13 The combination R1F1 compared with all other RF combinations in terms of learning retention (E2).

(a) Variance analysis results

Source of variation	df	s^2	F	F_α
RF	1	2.77	<1	3.88
G	2	50.16	4.84**	3.04
S	1	8.55	<1	3.88
RF x G	2	4.47	<1	3.04
RF x S	1	6.18	<1	3.88
G x S	2	0.89	<1	3.04
RF x G x S	2	5.03	<1	3.04
Error (within cells)	272	10.37	--	

Note: RF = R1F1 - all other RF

(b) Cell means

G	S	R1F1	Other RF's
4	M	7.86	7.10
	F	9.68	8.03
5	M	6.60	11.69
	F	9.61	10.56
6	M	12.09	14.95
	F	16.63	15.91
4-6	M,F	10.41	11.37

Table 16.14 The combination R1F1 compared with all other RF combinations in terms of learning efficiency (EFF1).

(a) Variance analysis results

Source of variation	df	s^2	F	F_α
RF	1	37.28	4.14^{*}	3.88
G	2	75.79	8.42^{***}	3.04
S	1	18.43	2.05	3.88
RF x G	2	6.31	<1	3.04
RF x S	1	5.37	<1	3.88
G x S	2	0.03	<1	3.04
RF x G x S	2	3.85	<1	3.04
Error (within cells)	271	9.00	--	

Note: RF = R1F1 - all other RF

(b) Cell means

G	S	R1F1	Other RF's
4	M	6.12	9.28
	F	10.46	10.04
5	M	4.69	14.01
	F	10.32	13.57
6	M	15.01	17.12
	F	16.49	20.22
4-6	M,F	10.52	14.04

combination R1F1 is in no way more effective than the alternative combinations of response requirement and key answer frequency studied here.

The analysis of the efficiency variable, EFF1, has yielded a significant F quotient for RF, but no significant interaction involving RF. The efficiency mean for all R1F1 groups is 10.52 and the mean for all other RF groups 14.04. Obviously, the learning efficiency under the standard Skinnerian condition of overt response requirement (R1) and 100 per cent key answer frequency (F1) is much lower than the average efficiency of the other conditions investigated here.

This analysis gives, then, no support whatsoever for the hypothesis that the Skinnerian combination of overt response requirement and 100 per cent key answers is an optimal condition for programmed instruction employing linear programs in book form. Other conditions entail at least as good learning effect as the Skinnerian condition and much better learning efficiency.

16.6 PROGRAM RESPONSE VARIABLES

For all subjects who were instructed to respond overtly to program tasks, i.e. those in R1 groups, values have been computed in the following variables (cf. section 15.3.4 above).

- LW — the logarithm of the number of original program response errors;
- LCW — the logarithm of the number of errors corrected by the student himself;
- LUW — the logarithm of the number of remaining, uncorrected errors;
- APC — the inverse sine of the proportion of answers corrected by the student himself.

Table 16.15 Analysis of the variable LW.

(a) Variance analysis results

Source of variation	df	s^2	F	F_α
F	2	0.1623	28.47**	3.06
G	2	0.0555	9.47**	3.06
F x G	4	0.0098	1.72	2.43
Error (within cells)	144	0.0057	--	

(b) Cell means transformed back into errors in per cent of total number of responses

G	F1	F.5	F0	F1-0
4	7.1	14.8	25.3	13.9
5	4.4	10.3	17.9	9.4
6	5.7	6.0	11.3	7.3
4-6	5.6	9.7	17.3	--

Certain input data for variance, correlation and regression analyses of these variables have been obtained through the computer program BMD07D, "Description of Strata with Histograms" (75:95ff), dividing the total data set into nine groups corresponding to the cells in the F x G design. On the basis of this output, including among other things group means and within groups and subgroup correlation matrices, the different analyses reported below have been performed by means of manual computation.

Table 16.16 Analysis of the variable LCW.

(a) Variance analysis results

Source of variation	df	s^2	F	F_α
F	2	0.1552	26.21***	3.06
G	2	0.0025	<1	3.06
F x G	4	0.0038	<1	2.43
Error (within cells)	144	0.0059	--	

(b) Cell means of LCW

G	F1	F.5	F0
4-6	0.943	0.887	0.524

16.6.1 <u>Variance analyses of the design F x G</u>

Variance analyses with the design F x G have been performed for all four program response variables, LW, LCW, LUW, and APC. The results of these analyses are summarized in tables 16.15 - 18. As these tables show, all the program responses variables analysed here covary with the key answer frequency. Number of errors (LW and LUW) increases with decreasing key answer frequency, from F1 to F0, while corrections (LCW and APC) decrease. The relationships between key answer frequency, on the one hand, and the four program response variables, on the other, are illustrated in figures 16.16 and 16.17, which show approximately linear relationships for the variables LW, LUW and APC.

Table 16.17 Analysis of the variable LUW.

(a) Variance analysis results

Source of variation	df	s^2	F	F_α
F	2	0.5728	75.37**	3.06
G	2	0.1174	15.45**	3.06
F x G	4	0.0086	1.13	2.43
Error (within cells)	144	0.0076	--	

(b) Cell means transformed back into uncorrected errors in per cent of total number of responses

G	F1	F.5	F0	F1-0
4	3.0	10.7	24.0	9.3
5	1.2	6.3	16.0	5.2
6	1.4	3.1	9.4	3.5
4-6	1.7	6.0	15.3	--

Three of the variables, LW, LUW and APC, also covary significantly with the grade level (G), as is illustrated in figures 16.18 and 16.19. Not unexpectedly we find that the number of errors (LW and LUW) is smaller the higher the grade level, while the relative frequency of corrections increases with increasing grade level.

These analyses indicate that the variation in key answer frequency influences the response behaviour of the students. The finding that the number of response errors increases when the key answer frequency decreases is fully in agreement with traditional, Skinnerian PI theory. As a matter of fact, the frequency of correct program responses is some-

Table 16.18 Analysis of the variable APC.

(a) Variance analysis results

Source of variation	df	s^2	F	F_α
F	2	0.3327	79.59***	3.06
G	2	0.0255	6.09***	3.06
F x G	4	0.0017	<1	2.43
Error (within cells)	144	0.0042	--	

(b) Cell means transformed back into corrections in per cent of the number of original errors

G	F1	F.5	F0	F1-0
4	51.3	26.8	3.3	23.7
5	73.3	35.8	7.6	36.6
6	69.6	42.3	11.9	39.8
4-6	65.0	34.8	7.2	--

times regarded as a measure of program quality, in the sense that the learning effect should be greater the higher the proportion of correct student responses in the program. This interpretation of a high frequency of correct student responses in a program is, however, not supported by the results of the present study. Although decreasing key answer frequency has led to a decrease in the proportion of correct student responses, this does not seem to have entailed any marked decrease in learning effect (cf. section 16.1 above).

The students' spontaneous corrections of their response errors (the variable APC) could also be expected to influence

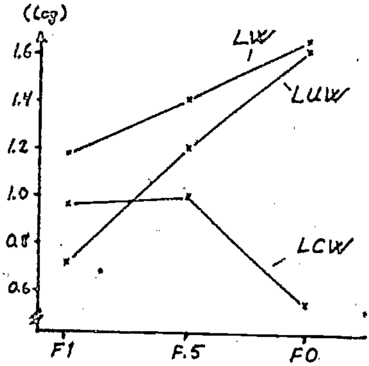


Figure 16.16

The variables LW, LUW and LCW: covariation with key answer frequency (F)

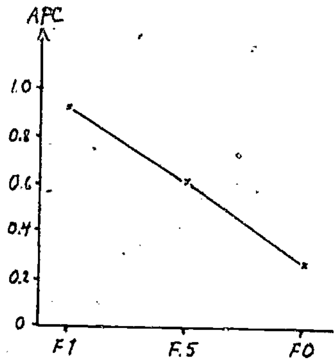


Figure 16.17

The variable APC: covariation with key answer frequency (F)

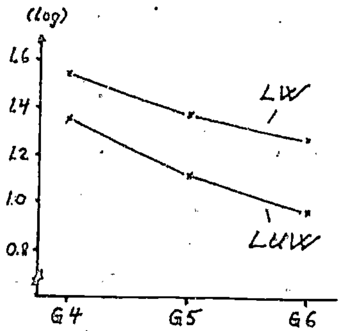


Figure 16.18

The variables LW and LUW: covariation with grade level (G)

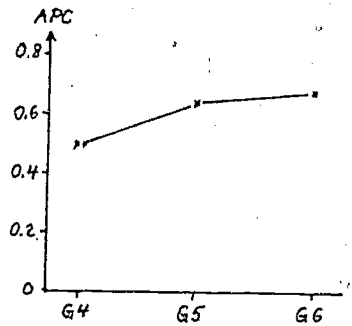


Figure 16.19

The variable APC: covariation with grade level (G) learning. Other PI conditions being equal, a higher amount of corrections should lead to a higher learning effect, since it diminishes the possibility that incorrect responses are learned and retained as correct ones. If this is so, the learning effect in this study should be expected to decrease with decreasing key answer frequency, since the proportion of corrections decreases with this variable (figure 16.17). As we have already seen (section 16.1 above) this is, however, not the case. It seems reasonable, though, to regard the fact that corrections decrease with decreasing key answer frequency as evidence that the students have used - at least to some extent - the response information provided in the form of key answers.

16.6.2 Correlation analyses

Differences between groups in the different program response variables do not, as we have seen, seem to influence learning in the ways that are usually predicted. In order to throw more light upon this question the relationships between the criteria ED1 and ED2, on the one hand, and the response variables LW, LCW, LUW and APC, on the other, have been analysed. For each of the nine cells in the design F x G the full matrix of correlation coefficients for these six variables has been computed. These coefficients are rendered in table 16.19.

The correlation coefficients in table 16.19 have been transformed to z_r (Fischer's z) and for each pair of variables the observed variance between groups in z_r has been computed according to the formula

$$s^2_{z_r} = (\sum_1^k z_r^2 - (\sum_1^k z_r)^2/k)/(k-1), \text{ where } k = \text{number of groups.}$$

Table 16.19 Zero-order correlation coefficients within cells in the matrix F x G - for R1 groups only. (Decimal points omitted.)

Cell	ED1 • ED2	ED1 • LW	ED1 • LUW	ED1 • LCW	ED1 • APC	ED2 • LW	ED2 • LUW	ED2 • LCW	ED2 • APC	LW • LUW	LW • LCW	LW • APC	LUW • LCW	LUW • APC	LCW • APC
G4 F1	894	-081	-452	213	409	022	-220	129	226	485	723	297	-194	-667	857
F.5	805	-767	-796	-221	736	-784	-788	-282	750	980	391	-835	239	-910	032
F0	579	-855	-840	-064	333	-736	-759	406	788	999	-223	-635	-262	-673	831
G5 F1	853	-317	-685	-054	582	-339	-593	-121	419	690	899	-077	332	-743	340
F.5	614	-713	-747	-081	633	-766	-768	-172	586	975	275	-685	103	-823	456
F0	847	-926	-936	476	822	-792	-807	449	734	996	-361	-778	-427	-827	844
G6 F1	812	-046	-059	-007	069	-091	-115	-015	136	-111	943	648	-371	-782	851
F.5	918	-712	-733	-397	082	-667	-743	-311	168	836	821	223	388	-337	710
F0	799	-781	-801	-034	297	-704	-744	-071	335	987	290	-162	179	-313	735
$\bar{z}_r$	1.14			-.02				.01	.55				-.00		
$\bar{r}$	.81**			-.02				.01	.50**				-.00		
$s^2_{z_r}/\sigma^2_{z_r}$	1.24	4.02**	2.92**	0.95	1.96*	2.53*	1.92*	1.16	1.85	22.87**	7.51**	6.38**	1.47	2.09*	2.87*

Note: $z_r = 0.5 \ln((1 + r)/(1 - r))$, i.e. Fischer's z.

$s^2_{z_r} = (k\sum_1^k z_r^2 - (\sum_1^k z_r)^2)/k(k - 1)$, where k = number of cells, i.e. 9.

$\sigma^2_{z_r} = (\sum_1^k (1/n_i - 3))/k$, i.e. 0.0729, where n_i = number of Ss in cell i.

The expected variance for the nine groups, assuming a random normal distribution of z_r, is

$$\sigma^2_{z_r} = \sum_1^k x(1/(n_i - 3)/k = =.0729,$$

where n_i = number of subjects in group i.

The observed variation between groups for z_r has been tested using the test variable

$s^2_{z_r} / \sigma^2_{z_r}$, which has a chi^2/f-distribution (91:44).

The observed values of this variance quotient are found in the bottom row of table 16.19. The variance quotient is non-significant for six of the variable pairs: ED1xED2, ED1xLCW, ED2xLUW, ED2xLCW, ED2xAPC and LUWxLCW. For these variable pairs the variation between groups in the observed correlation coefficients is, then, not greater than could be expected if coefficients were randomly sampled from a population with a mean equalling the observed mean. These observed mean correlations are also given in table 16.19, where one can see that three of these values (mean r) are not significantly different from zero, namely ED1xLCW, ED2x LCW and LUWxLCW.

Apparently, the variable LCW (errors corrected by the student) is uncorrelated with ED1 and ED2, i.e. the number of program response errors corrected by the student does not seem to influence the learning effect. The variable LCW may indicate that the available response information (key answers) has been utilized by the students, but it does not contribute to the explanation of variations in learning effect. The variable LCW is, therefore, excluded from the subsequent analyses.

The variables LW and LUW (original and remaining, uncorrected program response errors respectively) correlate highly with each other in F.5 and F0 groups. Coefficients between 0.84 and 1.00 indicate that the two variables are practically identical in these experimental groups. Within F1 groups the correlation between these two variables is only moderately high, and the correlations with the effect variables ED1 and ED2 are higher for LUW than for LW. In analyses of relationships between program response variables and post-

Table 16.20 Partial correlations within cells in the design F x G between APC and ED1 with LUW held constant.

Group		APC•ED1	APC•ED2
G4	F1	.161	.109
	F.5	.047	.129
	F0	-.579	.576
G5	F1	.149	-.401
	F.5	.046	-.127
	F0	.243	.201
G6	F1	.037	.074
	F.5	-.257	-.131
	F0	.082	.161
$s^2_{z_r}/\sigma^2_{z_r}$		0.99	1.11
$\bar{r}$		-.017	.072

Note:

Concerning z_r and $s^2_{z_r}$ see note in table 16.19.

$\sigma^2_{z_r} = \sum_{i=1}^{k}(1/(n_i - 4))/k$,

where k = number of cells and

n_i = number of Ss in cell i.

test results the variable LUW seems, then, to be a more adequate measure than the variable LW, i.e. the remaining, uncorrected errors made by the student provide a better measure than the original errors.

The two program response variables having the highest correlations with posttest results (ED1 and ED2) are consequently LUW, number of uncorrected errors, and APC, proportion of errors corrected by the student. These two variables are also fairly highly correlated with each other: coefficients for the nine groups vary between -0.31 and -0.91 (table 16.19). Since LUW on the whole shows a higher correlation with ED1 and ED2 than does APC, the partial correlation coefficients for APC and ED1 - or ED2 - with LUW kept constant have been computed. These coefficients are given in table 16.20.

The variance quotient for Fischer's z, i.e. observed variance divided by expected variance, is 0.99 for the partial correlation APCxED1 and 1.11 for APCxED2 (table 16.20). Neither of these values is significant, which means that the variation between experimental groups in partial correlation for these variables is not greater than could be expected if coefficients were randomly sampled from a population with a mean equalling the observed mean. The mean r values computed, -0.02 for APCxED1 and 0.07 for APCxED2, are both nonsignificantly different from zero. The conclusion is, then, that there is no partial correlation between APC, on the one hand, and ED1 or ED2, on the other, when LUW is kept constant. Of the four program response variables studied here only one, the number of uncorrected errors (LUW), seems consequently to have any particular value for the prediction of posttest results.

The correlation between LUW and ED1 and ED2 respectively varies, as table 16.19 shows, from one experimental group to another. The variation between groups in Fischer's ž for these two variable pairs has been analysed by means of variance analyses on the design F x G. The results are summarized in tables 16.21 and 16.22. As these tables show, the variance quotient (observed through expected variance) is significant for the factor F in both cases, but not for G or for the interaction F x G. The correlation between the number of uncorrected errors made by the students (LUW) and posttest results (ED1 or ED2) varies, then, with the key answer frequency (F), in that this correlation is higher the lower the key answer frequency (cf. r mean values in tables 16.21c and 16.22c).

In order to clarify further what factors influence the variable LUW under different conditions as to key answer frequency, the standardized coefficients of the partial regression of LUW upon the covariates in the previous analyses, namely PREI, PREC, SD2, SD3 and LPT, have been computed for F1, F.5 and F0 groups respectively. These partial regression coefficients (R) are found in table 16.23. The coefficients in this table must, of course, be interpreted with particular caution in view of the low number of degrees of freedom involved. They give, however, some interesting indications about the nature of the influence pretreatment knowledge (PREI, PREC), intelligence (SD2, SD3) and PI time (LPT) have upon the number of program response errors made by students.

Table 16.21 Variation between groups in correlation between uncorrected errors (LUW) and posttest result (ED1).

(a) z_r*-values for R1 groups in the design F x G*

G	F1	F.5	F0	F1-0
4	-.487	-1.088	-1.221	-.932
5	-.839	-.966	-1.705	-1.170
6	-.059	-.935	-1.102	-.699
4-6	-.462	-.996	-1.343	--

(b) Variance analysis

Source of var.	df	s^2	s^2/σ^2
F	2	0.5910	8.11**
G	2	0.1666	2.29
F x G	4	0.0473	0.65

(c) $\bar{r}$*-values*

	F1	F.5	F0
$\bar{r}$	-.430	-.759	-.880

Table 16.22 Variation between groups in correlation between uncorrected errors (LUW) and posttest 2 results (ED2).

(a) z_r-values for R1 groups in the design F x G

G	F1	F.5	F0	F1-0
4	-.224	-1.066	-.994	-.761
5	-.682	-1.015	-1.119	-.939
6	-.116	-.957	-.959	-.677
4-6	-.341	-1.013	-1.024	--

(b) Variance analysis

Source of var.	df	s^2	s^2/σ^2
F	2	0.4593	6.30**
G	2	0.0534	0.73
F x G	4	0.0235	0.32

(c) $\bar{r}$-values

	F1	F.5	F0
$\bar{r}$	-.328	-.767	-.772

Table 16.23 The regression of LUW on PREI, PREC, SD2, SD3 and LPT for different key answer frequencies (F; for R1 groups only).

F	b^*_{PREI}	b^*_{PREC}	b^*_{SD2}	b^*_{SD3}	b^*_{LPT}	R^2
F1	-.080	.122	-.386	-.089	-.507	.309
F.5	-.146	.083	-.684	-.260	.004	.760
F0	.035	.166	-.575	-.390	-.160	.617

Note: b^*_i is the standardized (scale free) partial regression coefficient for variable i with the other four variables held constant.

$$R^2 = \sum_{i=1}^{k} b^*_i r_{yi} \text{ , where } i = \text{PREI,PREC,SD2,SD3,LPT}$$
$$y = \text{LUW}$$
$$k = \text{number of covariates}$$

When the key answer frequency is 100 per cent (F1), LUW obviously depends largely upon other variables than those investigated here, since these only account for 31 per cent of the LUW variance. The highest partial correlation - for F1 groups - is that between LUW and LPT (with a coefficient of -0.51), which implies that the response errors in the program depend to a marked extent on the PI time. The correlation is negative, so the shorter the time devoted to the program the more numerous the errors made. It seems therefore reasonable to assume that the errors made in the program and not corrected by the students, when the key answer frequency is 100 per cent, are primarily mistakes caused by neglect rather than errors caused by lack of knowledge. If this is so, one should not expect these errors - or mistakes - to correlate highly with the posttest results (cf. tables 16.21-22).

When the key answer frequency is 50 or 0 per cent (F.5 or F0), the amount of uncorrected errors in the program (LUW) depends to a much greater extent on the five covariates studied here. The percentage of LUW variance explained by the covariates is 76 for F.5 and 62 for FO, which should be compared to the figure 31 per cent for F1. All covariates do not, however, influence the number of errors left uncorrected by the students to the same extent. The PI time (LPT) seems to be of very little importance in this context (the coefficient is 0.04 for F.5 and -0.16 for FO). But the uncorrected errors (LUW) are fairly highly correlated with intelligence (SD2, SD3), particularly with verbal intelligence (SD2, with the coefficient equalling -0.68 for F.5 and -0.58 for FO).

Summarizing the findings presented here concerning the various program response variables we note that

- the covariation between the amount of corrections of program response errors and the key answer frequency provides evidence that the key answers have been used by the students for the purpose of response control;

- the correlation between program response errors and posttest result is highly dependent upon the key answer frequency, in that this correlation is high only when the key answers have been partly or wholly removed from the program;

- response errors made in the program are of a different type depending on the key answer frequency, particularly in the sense that errors made under the condition of 100 per cent key answers are likely to be mistakes (caused by neglect), which is not the case under the conditions of 50 or 0 per cent key answers; and

- the responses a student is expected to give in the program 'The main parts of the simple sentence' must on the whole be considered criterion relevant, since the correlation between response errors and posttest result is high, r = -0.88, for those groups who have not been able to correct any responses with the aid of key answers, i.e. the FO groups.

16.7 SUMMARY OF THE RESULTS OF THE MAIN STUDY

16.7.1 The importance of the response requirement

The findings of the main study reported in this chapter agree, in general, fairly well with the hypotheses formulated concerning the influence of the response requirements R1 and R2 upon learning effect, PI time and learning efficiency (hypotheses 1:1-3 in table 7.1 above). It has been found

- that learning effect (E1) and retention (E2) were somewhat greater under the overt response requirement (R1) than under the covert (R2), and that this difference was larger in the lowest grade (G4) than in the highest (G6);
- that PI time (PIT) was about twice as long under the overt (R1) as under the covert (R2) response requirement, and that this difference was greater in grade 4 than in grade 6;
- that learning efficiency (EFF1) was lower when overt responses were required (R1) than when covert responses were required (R2), and that this difference was greater in the highest grade than in the lowest.

It is quite clear, then, that this study gives no support to the theory that overt response requirement is generally superior to covert response requirement in programmed instruction using linear, Skinner-type programs. This is particularly obvious if one looks at the results for groups with 100 per cent key answers, i.e. the normal Skinnerian condition, separately. In these groups both learning effect and retention were about the same for the overt as for the covert response requirement (E1 means 12.88 and 12.50 and E2 means 10.41 and 10.57 for R1 and R2 groups respectively). The learning efficiency under 100 per cent key answer frequency was, however, much lower for the overt (R1) than for the covert (R2) response requirement (EFF1 means 10.97 and 14.87, which are significantly different from each other.

These findings are, on the other hand, in good agreement with the theory that the influence of the response requirements R1 and R2 in programmed instruction is dependent upon a number of program and student variables. In particular the findings support the hypotheses

- that the overt response requirement (R1) has a very small positive effect when used with a program requiring a small amount of response learning (cf. 3.3.2.3, 5.1 and 7.1 above); and

- that for a given program the overt response requirement (R1) has less positive influence upon learning the older and the better equipped as to background knowledge, study skills etc. the students are (cf. 3.3.2.2 and 7.1 above).

This discussion of the importance of the response requirement in programmed instruction is based on the assumption that any response required in a program is relevant with regard to the criterion test. This issue is further discussed below (section 16.7.4).

16.7.2 The importance of the key answers

The analyses of data from the main study reported here give some support to the hypotheses formulated regarding the influence of key answer frequency upon the various criterion variables (cf. 7.2 above). It has been found

- that learning effect (E1) and retention (E2) under the overt response requirement (R1) increased, as expected, with decreasing key answer frequency - from 100 to 0 per cent - while both effect measures (E1 and E2) seemed unaffected by the variation in key answer frequency when covert responses were required (R2);

- that the PI time (PIT) increased, as expected, with decreasing key answer frequency - from 100 to 0 per cent - when overt response was required (R1), whereas this variable (PIT) was not affected by key answer frequency variations when covert responses were required (R2);

- that the learning efficiency (EFF1) instead of remaining unchanged under the overt response requirement (R1) when key answer frequency decreased - as was hypothesized (section 7.2 above) - seemed to increase, ~~while the~~ efficiency (EFF1) under the covert response requirement (R2) remained uninfluenced by the decrease in key answer frequency - instead of decreasing as hypothesized (section 7.2 above).

In combination with a demand for overt responses the providing of key answers seems to have a predominantly negative effect on learning from the program used in the present study. And this is true in spite of the fact that the key answers apparently have been used to a large extent for the purpose of control and correction (cf. section 16.5.1 and tables 16.17-18 above).

A negative effect of the availability of key answers upon learning from a program seems explainable in view of the fact that 100 per cent key answers has led to shorter working time (PI time) than 50 or 0 per cent key answers, although the increased amount of corrections under the 100 per cent key answers condition must have consumed a certain amount of working time.

One can conclude, therefore, that the availability of 100 per cent key answers has caused the students, who were required to give overt responses, to devote less time to the content of the program frames. In other words, the effective learning time has been shorter the more key answers have been provided in the program.

When covert responses were required (R2), the variation in key answer frequency seems to have had little influence on learning effect or working time. The most plausible explanation for this finding is probably that the students have not utilized the information provided through the key answers. It seems reasonable that the need for control of the responses is felt less strongly by the student when the responses are not written down or otherwise registered.

16.7.3 The combination of overt response requirement and 100 per cent key answers

The hypothesis that the Skinnerian combination of overt response requirement (R1) and 100 per cent key answer frequency (F1) is superior in terms of learning effect to other conditions studied here (section 7.3 above) has not received any support by the results of the main study. In fact, both learning effect (E1) and retention (E2) under the Skinnerian condition (R1F1) were somewhat lower than the average effect and retention under the other five conditions. And the learning efficiency (EFF1) under the Skinnerian condition of overt response requirement and 100 per cent key answers was significantly lower than the average efficiency under the other conditions.

That the Skinnerian model of programmed instruction, involving overt response requirement and 100 per cent key answers, should be generally superior to alternative forms of response requirement and response information in programmed instruction seems, then, very unlikely - particularly if the efficiency criterion is taken into account.

16.7.4 Students' errors and corrections in overt responding groups

The number of program response errors made by the students increased and the amount of corrections of such errors decreased with decreasing key answer frequency. This seems to imply that the available key answers (whether 100 or 50 per cent) have been utilized by the students for the purpose of control and subsequent correction of errors. Key answers may also have been used for the purpose of a form of 'self-prompting', which is often called cheating, i.e. the key answer is read before the student has written down, or perhaps even thought out, an answer of his own.

The correlation between response errors in the program and the posttest result was high only for groups with 50 or 0 per cent key answers ($r = -0.77$). When 100 per cent key answers were provided, this correlation was very moderate (r = -0.33). This obviously means that in programmed instruction of this kind- with the program presented in book form - the learning effect cannot be predicted from the number of program response errors made by the students, if key answers have been given after each frame.

The very high correlation (r = -0.88) between program response errors made by the students and posttest result for groups who were provided no key answers (cf. table 16.21 above) can be regarded as indirect evidence of the relevance of the program responses with reference to the criterion test (the PUSL test). Holland (90), for example, has made the point that the lack of positive effect of overt responding in programmed instruction is often attributable to the lack of program response relevance. This is a very good point, and it should therefore be observed that the lack of positive effect following overt response requirement - as compared to covert - in the present study can hardly be due to such a lack of response relevance.

17

A Synthesis Attempt

There seem to be three main theories regarding the importance of overt response requirement in programmed instruction using linear, Skinner-type programs. The first theory, represented by, among others, Holland, contends that overt response requirement has a general positive influence upon learning (90:93ff), and that

> "studies failing to reveal this effect are largely explained by failures in meeting the rationale for overt responding and by poor experimental procedures." (90:101f).

It is recognised that overt responding is time-consuming, but the solution is believed to be found in the development of more expedient modes for the registering of students' responses (90:102).

The second type of theory is represented by Anderson (59). He accepts in principle Holland's theory, but he modifies it in admitting that the positive effect of overt responding may be greater or smaller due to other factors - primarily the degree of "response learning" (59:139ff). The learning efficiency problem entailed by the fact that overt responding is time-consuming is, in fact, not at all discussed by Anderson, although his analysis of the role of overt responding is otherwise quite thorough.

The third type of theory, represented by, among others, Leith, does not ascribe any generally positive effect upon learning to students' overt responding. This kind of theory maintains that the effect of overt responding - or demands for overt responding - can be positive or negative or nonexistent, due to a number of other factors in the total situation of programmed instruction. Factors which are believed to be particularly important are the degree of response learning, the student's age or maturity level and the student's background knowledge (35). In his discussions of this problem, Leith indicates that he is aware of the learning efficiency problem involved, but his empirical studies contain no serious attempt at efficiency analysis.

The results of my own studies and the results of the research reviewed in chapter 3 above are not consistent with Holland's theory concerning overt responding, unless, of course, his demands on program design and program length are interpreted so literally as to become almost absurd. If his dictum that, in order to demonstrate the superiority of overt responding, "programs must be long enough for subjects in the covert condition to become careless" (90:93) is very strictly interpreted, then Holland's theory is, at least in this respect, a priori confirmed: if overt responding does not show positive effects the program just isn't long enough! It also deserves mentioning here that Holland discusses in considerable detail the reasons why overt responding should be preferred to covert, but in his own experimental studies (24 & 25) he only compares overt response requirement with no response requirement - not with covert response requirement. This means that he has not really attempted to verify empirically his own theory.

According to Holland it can also be argued that a text must be demonstrated to be programmed, before studies of response requirement etc. are carried out. This implies that the responses required of the student in the text must be contingent upon the greater part of the text. If this is not so, most of the text could be, to use Holland's term, "blacked out" without this affecting the student's possibility of responding correctly. What Holland calls "blackout ratio" (91 & 94) has not been computed for the program used in the present study, but since questions in the program concern grammatical examples, it seems obvious that the responses required are contingent upon the text: no correct responses could be given without the student having studied the examples given in the text.

As regards the two other types of theory discussed here it seems somewhat more difficult to judge their respective merit. The difference between the two lies mainly in the fact that one - Anderson's - does not allow for the possibility that overt response requirement has a negative effect on learning, which the other - Leith's - does. (In a later essay on this subject Anderson (60) has recognised the possibility that overt responding can have a negative effect on learning.) The evaluation of these two theories is further complicated by the fact that they are, to a considerable extent, based on different sets of experimental data, since the British studies on which Leith (35) relies heavily are not at all discussed in the review by Anderson (59).

As far as the learning effect under different response requirements is concerned, the results of my own experiment seem consistent with Anderson's as well as with Leith's theory regarding the effect of overt responding. Some of the studies presented by Leith (35), and also certain other results (cf. section 3.3.2 above), indicate, however, that the effect of overt response requirement can be negative compared to the effect of no response requirement or covert response requirement. A study which does not deal with programmed instruction in the usual sense, but which is still of interest in this context has been presented by McGuire (106). The learning of tools' names by means of an audio-visual device was studied, using two different speeds of stimulus presentation. With the higher presentation speed the covert response requirement led to higher learning than the overt. Apparently, to force the student to give overt responses may interfere with the learning process, if the student is still uncertain of what is the correct response.

It seems, then, as though the hypothesis that overt response requirement can have a negative effect on the learning effect in programmed instruction, should not be discarded.

Markle has proposed in a thought-provoking essay that

> "in most instances the important part of a response is what takes place before the overt behavior" (105:193).

This can be developed into an interesting theory: the covert response activity is the important part of the student's responding in programmed instruction, and to require overt responses will have an effect insofar as it affects the probability - positively or negatively - of the occurence of the covert response activity. Such a theory is not at all inconsistent with Holland's thesis:

> "The reason for public (i.e. overt; my comment) answer is not theoretical but practical ... private (i.e.

covert; my comment) answering often may change to private omissions" (90:93).

A great deal of evidence shows that the effect of covert response requirement in programmed instruction can vary with a number of program and student variables. Which are, then, the most important of these variables?

T h e r e l e v a n c e o f p r o g r a m r e s p o n s e s is obviously one of them. Unless there is a definite connection between program responses, on the one hand, and the content to be learned, on the other, the effect of overt - or any - responding cannot be expected to influence the measured outcome of the learning process.

P r o g r a m l e n g t h is, as this study has shown, in all likeliness one of these factors, although there still remains the task of defining an acceptable measure of program length.

The a g e and the i n t e l l i g e n c e l e v e l o f t h e s t u d e n t also belongs to the group of variables on which the effect of overt responding depends. The relation between the knowledge, aptitude etc. of the student, on the one hand, and the program design, on the other, is likely to be of great importance. This relation is expressed in student-related program variables, here called PI variables, such as program difficulty and the degree of response learning required in a program. The interaction between student age and the effect of overt response requirement is in that case merely a secondary effect: a consequence of the covariation of age with knowledge and aptitude, variables which in their turn affect the relative difficulty of a program. In this sense we are dealing here with an instance of aptitude-treatment interaction of a fairly precise nature.

Concerning the importance of key answers in programmed instruction the different positions do not seem to be very clearly formulated. Anderson summarizes:

> "..while it is generally known that KCR /knowledge of correct responses/ facilitates learning, there is good reason to doubt that it functions primarily as a reinforcer... KCR may be chiefly ... a source of correc-

tive feedback" (59:151).

The reinforcing effect of key answers has been even more strongly questioned by Davies:

> "Frequent and apparent success ... need not by itself be intrinsically reinforcing; the socalled programming principle that mere progress through a programme is inherently reinforcing is not only poor learning theory - it is an abandonment of common sense." (74:206)

Various reasons why the lack of key answers in a program should, under certain circumstances, entail better learning have been presented by Lublin (39:299ff). Generally, the evaluation of the effect of key answers in programmed instruction is complicated by the fact that you have to reckon with different types of effect: feedback, reinforcement, negative effect on attitudes or motivation; possibilities of cheating etc. But our measures of learning can only estimate the net effect of these different influences, some of which may even cancel each other. In some cases even this net effect of key answer presentation on learning is difficult to assess.

The research review presented here (section 3.5.1 above) has demonstrated that the effect of key answers combined with a demand for overt responses is dependent upon the length of the instructional program. The nature of this interaction is such that 100 per cent key answers influences learning positively in shorter programs (less than 200 frames), but has no or even negative effect in longer programs. The results of the present study are consistent with this finding. It is also noteworthy that the time spent by students on the program (PI time) is longer when no key answers are provided than it is when the key answer frequency is 100 per cent both in Lublin's study (39) and in my own (section 16.3 above). This seems to indicate that the presentation of key answers after each frame made students less inclined to reflect upon the tasks given in the frames.

Anderson, commenting upon Lublin's finding, contends that there is a simple explanation of this negative effect of 100 per cent key answers:

> "The procedure and program format ... may have permitted the students ... to copy the correct answers into the blanks" (59:151).

Such an explanation is, of course, always plausible when programs are presented in book form, as in both Lublin's and my study, or by means of a simple, not 'cheat-proof' machine. My analysis of errors made by the students in the program (section 16.5.1 above) has shown, however, that the key answers to a great extent have been used for the purpose of what Anderson calls "corrective feedback". In the groups given 100 per cent key answers on the average 65 per cent of the errors were spontaneously corrected by the students, whereas the corresponding percentage of error corrections in groups provided with no key answers was as low as 7 per cent! This finding makes it most unlikely that the students just copied the key answers. After all, if key answers are copied incorrectly, what information could the student use to correct those errors? It is also worth noting that the teachers participating in the study did not observe such copying, except on a few occasions.

Lublin's explanation of the negative effect of continuous key answer information seems, therefore, more tenable than that of Anderson. Lublin maintains that programmed instruction with 100 per cent key answer frequency makes the task

> "so unchallenging, or so uninteresting (due to the 100 per cent predictability that the right answer would immediately be supplied), that they /the students/ were not attending to the concepts being taught in the frames." (39:300)

The absence of key answers, on the other hand, made

> "the subjects study each frame more carefully" (39:299).

The research review presented in the present study covers no studies of the influence of key answers when covert responses are required. A study by Kanner & Sulzer (92) is, however, of interest in this context, although it does not concern programmed instruction directly. These authors found that the removal of key answers in a learning situation where covert responding was required affected learning negatively, but not significantly so.

The conclusions which can be drawn here concerning the effect of key answers under the condition of covert response requirement must, consequently, be based mainly on the findings of my own experiment. These findings indicate that the variation in the key answer frequency under the covert response requirement (R2) affects neither learning effect nor PI time, which means that learning efficiency is not affected either. The most reasonable explanation of this finding is that the key answers simply have not been used for the purpose of control to any great extent - or perhaps not at all.

The synthesis attempted here results, then, in the following major conclusions concerning programmed instruction, using linear, Skinner-type programs presented in book-form:

- When programs are short, less than 100 frames, variations in response requirement and key answer frequency have little influence on the learning effect. Consequently, the least time-consuming PI conditions can then be expected to lead to the highest learning efficiency.

- When programs are fairly long, more than 200 frames, overt response requirement with no key answers or covert response requirement - with or without key answers - can be expected to entail better learning than any other combination of response requirement and key answer frequency studied here. Which alternative is the most efficient will then depend almost exclusively on the 'cost' of overt responding in terms of time spent by the student on the program.

- Overt response requirement has its greatest positive effect when the program demands, for a given group of students, a high degree of learning of unfamiliar response terms or otherwise has a high information level, i.e. is difficult to master.

- Covert response requirement leads to considerably shorter PI time than overt response requirement, if student-constructed responses are demanded. And normally, the learning effect is almost as great under covert as under overt response requirement. If efficiency of learning is the primary criterion, covert response requirement can, therefore, be expected to lead to results better than - or at least as good as - those reached with overt response requirement.

- The Skinnerian combination of a demand for overt responses, constructed by the student, and presenting a key answer after every student response cannot in any respect be considered generally, i.e. with regard to various program types and student categories, superior to other possible combinations of response requirement and key answer frequency.

18

Practical Educational Consequences of the Results

Programmed instruction of the Skinnerian type was originally developed on the basis of one particular learning theory. The acceptability and usefulness of this theory can, no doubt, be questioned. And it has been questioned both with regard to its explanation value and its agreement with available experimental data, or rather lack thereof, and for its ideological implications as a theory of behaviour control and shaping. This is not the place to reopen this discussion. My point here is instead that even if the Skinnerian theory could be accepted, it is doubtful whether laws or principles-of practical school instruction can be deducted from its principles of learning. The possibilities - and also the desirability - of strict control of various aspects of learner motivation in a real-life school situation are rarely - if ever - comparable to that of a learning laboratory situation.

This problem has been touched upon earlier in this study, in chapter 2. It has been emphasized that program characteristics should not be mixed up and confused with PI functions, which are dependent upon student behaviour. The program characteristic 'presentation of key answers after each frame', for example, cannot be said to be unequivocally related to one particular type of student behaviour, whereas the term 'reinforcement' undeniably implies that something is happening to - or in the student, although this 'something' may not be easily observable.

Decisions concerning the design of instructional programs - as well as those concerning other instructional materials - must primarily be dealing with program variables, i.e. they deal with programmer behaviour rather than with student behaviour. Naturally, the wisdom of the decision is ultimately judged with reference to some kind of student behaviour. But this means only that the behaviour of the student is the dependent variable by means of which we study the effect of the independent variable: the programmer behaviour.

The reinforcement problem provides a good illustration of this point. Immediate and frequent success in working on the learning tasks is believed to further learning in that it functions as a reinforcer of the desired student behaviour. The presentation of key answers after the student has responded to a program frame is assumed to make the student experience success and, thus, to reinforce his learning of that particular response. This is, for example, Skinner's own position (121:96ff).

A precise and definite relationship between a particular program characteristic - the presentation of key answers - and a student behaviour - experiencing success - is simply taken for granted. As a consequence, the positive effect of the key answer presentation is also taken for granted - and even believed to follow by logical necessity from the learning theory in question. To deny - or even doubt - the importance of key answers in programmed instruction may then easily be construed as questioning the importance of reinforcement to learning, which is quite a different matter altogether. The confusion is, of course, not diminished if, as often happens, the presentation of key answers in an instructional program is labelled 're-inforcement'.

The presentation of key answers in a program can, as was pointed out in chapter 17 above, influence the student's behaviour in several different ways. It may, for example, diminish the student's desire to make an effort to solve the problems presented to him, which will, in all likeliness, affect his learning negatively. To question the positive effect of a program characteristic - the key answer presentation - does, therefore, not necessarily mean that one denies the positive effect of the student behaviour supposed to be correlated to the program characteristic, namely the student's observing that his answer was correct. It simply means that the program characteristic and the student behaviour are not believed to be highly correlated.

The response requirement problem can be looked upon in a similar vein. The principle that 'learning is doing' is not necessarily questioned, because one questions whether one, in programmed instruction, should demand overt, student-constructed responses of the type and to the extent presupposed in the Skinnerian PI model. Besides stimulating, with a higher or lower degree of probability, some specific response behaviour (in this case the construction and writing down of a response), the response requirement may well influence the student's behaviour in other ways - which could be desirable or undesirable. In prolonging the time the student has to spend on the program, the response requirement may, for example, have a negative effect upon the student's motivation to work with the program.

This investigation has studied some alternative response requirements and key answer frequencies in programmed instruction using linear programs. On the basis of both previous research and my own experiment I have drawn certain conclusions, which were presented in chapter 17 above. I want to emphasize that the present study does not pretend to have demonstrated in exactly what ways these response requirements or key answer presentations influence the students' behaviour and their learning. What has been studied here is primarily the net effect of different program characteristics, as this effect appears in measures of learning effect, time spent on the program (PI time) and learning efficiency.

This net effect of a given set of program characteristics is, however, of very great practical importance. And it is a primary objective of this study to try and improve the basis for practical educational decisions concerning programmed instruction. Even if our understanding of the 'mechanisms' of different program designs is much less than complete, we will probably want to make a number of decisions regarding what designs to use in different situations, when we want students to learn by programmed instruction. Before making any such decision we should consider very carefully all the available data concerning the probable net effect on learning and learning efficiency of each of the alternative designs among which we want to choose.

The conclusions to which the present investigation has led, and which have been presented in chapter 17, should not be regarded as definite in any way - conclusions rarely should! - since the relevant experience and research results are not unanimous and conclusive. These conclusions are, however, in my opinion more reasonable and likely to be true than most other possible conclusions, they are also more consistent with the available empirical data. These conclusions should, therefore, be regarded as a summary of the basis for decisions regarding response requirement and key answer frequency in programmed instruction which educational research can offer at present.

The practical consequences of the conclusions drawn from the findings of this study may not be altogether self-evident. It should, therefore, be useful to try and spell them out. I see the following consequences as the most important ones:

- The application of principles of learning theory does not automatically lead to the optimal design of programmed instruction as regards response requirement and key answer presentation.

- The efficiency criterion, i.e. primarily the cost of learning in terms of student time, must be considered in the choice of programmed instruction model, particularly when it comes to choosing the type of response requirement.

- To accept and apply generally, i.e. for all program and student categories, the Skinnerian demand for overt, student-constructed answers and 100 per cent key answers is not defensible, especially not in view of the learning efficiency criterion.

- The less time-consuming covert response requirement is particularly suitable for programs which, for a given group of students, call primarily for associating a previously learned response repertoire - or a limited number of new response terms - with new concepts, phenomena etc., i.e. in programs entailing only a small amount of response learning.

- If overt responses are required, the key answer frequency should preferably be less than 100 per cent, in particular when the program is fairly long, i.e. has more than 200 frames or the equivalent thereof.

- As a consumer of programs to be presented in book form or by means of simple machines one should not accept data regarding errors made - or not made - by students going through the program as measure of the program's potential learning effect, if these error data have been recorded in a situation where 100 per cent key answers were presented to the student.

- Finally, it should be observed that other programmed instruction designs than the traditional Skinnerian design may make it feasible to construct more flexible types of program, which are more easily adapted to the different learning tasks involved in various school subjects.

REFERENCES

A. *Sources of data on studies included in the research review, chapter 3*

(BMN indicates that the study is reviewed by Bernmalm - no. 4 below - and that data are taken from him.)

1 Alter, M & Silvermann, R E, "The response in programmed instruction". *Journal of Programmed Instruction 1*(1962):1, 55-78.

2 Ashbaugh, W H, "Effect on achievement of written responses to programmed learning material for students of differing academic ability". *Psychological Reports 14*(1964), 780-782.

3 Berglund, G W, "The effect of partial reinforcement in programmed instruction". *The Institute of Education, Uppsala University, Report No. 39,* 1968 (mimeo).

4 Bernmalm, S, Programmerad undervisning: En granskning och sammanställning av amerikanska forskningsrapporter. Pedagogiska institutionen vid Göteborgs universitet, Mars 1965 (mimeo).

5 Biran, L A, Clarke, R & Leith, G O M, "A comparison of two modes of responding to a linear programme containing new vocabulary and concepts". *Technical Report, National Centre for Programmed Learning,* 1965 (mimeo).

6 Buckland, P R, "The response in a linear program: its mode and importance". *Programmed Learning & Educational Technology 4*(1967), 47-51.

7 Buckle, C F, "Relations between attainment of component skills, retention and transfer". *Research Reports on Programmed Learning No. 24, National Centre for Programmed Learning,* July 1967 (mimeo).

8 Campbell, V N, *Adjusting Self-Instruction Programs to Individual Differences: Studies of Cueing, Responding and Bypassing.* San Mateo, Cal.: A.I.R., 1961.

9 Campeau, P L, *Level of Anxiety and Presence or Absence of Feedback in Programed Instruction.* Palo Alto, Cal.: A.I.R., February 1965.

10 Campeau, P L, "Test anxiety and feedback in programmed instruction". *Journal of Educational Psychology* 59(1968), 159-163.

11 Case, H W & Roe, A, *Basic Properties of an Automated Teaching System*. Report No. 64-29, Los Angeles: Dept. of Engineering, University of California, June 1964.

12 Coulson, J E & Silberman, H F, "Effects of three variables in a teaching machine". *Journal of Educational Psychology* 51(1960), 135-143.

13 Christ, R L, "Overt versus covert responding and retention by sixth-grade students". *Journal of Educational Psychology* 57(1966), 99-101.

14 Cummings, A & Goldstein, L S, "The effect of covert and overt responding on two kinds of learning tasks". *Technical Report 620919, Center for Programmed Instruction*, 1962. (BMN)

15 Ekman, B, "Effekter av olika facitsvarsschemata vid programmerad undervisning". *Pedagogiska institutionen, Lärarhögskolan i Uppsala; Rapport nr 24*, 1971 (mimeo).

16 Entwisle, D R, Huggins, W H & Phelps, F W, Jr, "Response mode in technical programs". *American Educational Research Journal* 5(1968), 403-411.

17 Evans, J L, Glaser, R & Homme, L E, "An investigation of 'teaching machine' variables using learning programs in symbolic logic". *Journal of Educational Research* 55(1962), 433-452.

18 Feldhusen, J F & Birt, A, "A study of nine methods of presentation of programmed learning material". *Journal of Educational Research* 55(1962), 461-466.

19 Fiks, A J, "Some treatment and population variables in programed instruction". *Journal of Educational Psychology* 55(1964), 152-158.

20 Fry, E B, "A study of teaching machine response modes". Lumsdaine, A A & Glaser, R (Eds), *Teaching Machines and Programmed Learning: A Source Book*. Washington, D.C.: N.E.A., 1960. Pp 469-474.

21 Goldbeck, R A & Campbell, V N, "The effect of response mode and response difficulty on programed learning". *Journal of Educational Psychology* 53(1962), 110-118.

22 Hamilton, N R, "Effects of logical versus random sequencing of items in an autoinstructional program under two conditions of covert response". *Journal of Educational Psychology 55*(1964), 258-266.
23 Hartman, T F, Morrison, B A & Carlson, M E, "Active responding in programmed learning materials". *IBM Research Reports RC-694*, 1962. (BMN)
24 Holland, J G, Program design and use. Paper read at the A.P.A., Chicago, September 1960. (Rel. by Silberman in DeCecco, J P (Ed), Educational Technology, 1964, 160f).
25 Holland, J G, "Response contingencies in teaching machine programs". *Journal of Programmed Instruction* 3(1965)1.
26 Hough, J B & Revsin, B, "Programed learning at the college level: A study of several factors influencing learning". *Phi Delta Kappan 44*(1963), 286-291.
27 Karis, C & Kent, A, "Effect of variation in confirmation procedures on program performance and retention of programmed materials". *Journal of Educational Research 65*(1972), 475-480.
28 Keislar, E R & McNeil, J D, "A comparison of two response modes in an autoinstructional program with children in the primary grades". *Journal of Educational Psychology 53*(1963), 127-131.
29 Krumboltz, J D, The nature and importance of the required response in programmed instruction. Paper presented to the A.E.R.A., Chicago, February 1963. (Rel. by Schramm, see no. 45 below.)
30 Krumboltz, J D & Kiesler, C A, "The partial reinforcement paradigm and programed instruction". *Journal of Programmed Instruction 3*(1965), 9-14.
31 Krumboltz, J D & Weisman, R G, "The effect of overt versus covert responding to programed instruction on immediate and delayed retention". *Journal of Educational Psychology 53*(1962), 89-92.
32 Krumboltz, J D & Weisman, R G, "The effect of intermittent confirmation in programed instruction". *Journal of Educational Psychology 53*(1962), 250-253.
33 Lambert, P, Miller, D M & Wiley, D E, "Experimental folklore and experimentation: The study of programmed learning in the Wauwatosa Public Schools". *Journal of Educational Research 55*(1962), 485-494.

34 Leedham, J, "A summary of research and development with programmes and machines in Leicestershire since 1960". Unwin, D & Leedham, J (Eds), *Aspects of Educational Technology*.... London: Methuen & Co., 1967. Pp 97-105.

35 Leith, G O M, "The role of overt responding in programmed learning". Unwin, D & Leedham, J (Eds), *Aspects of Educational Technology*.... London: Methuen & Co., 1967. Pp 503-513.

36 Leith, G O M & Buckle, C F, "Mode of response and non-specific background knowledge". *Technical Report, National Centre for Programmed Learning,* 1965 (mimeo).

37 Leith, G O M & Ghuman, A S, "The effects of prompting and confirmation on two methods of responding to a self-instructional programme on co-ordinate geometry". *Research Notes on Programmed Learning, No. 6, National Centre for Programmed Learning,* March 1966 (mimeo).

38 Leith, G O M & Hope, B, "A comparison of two modes of responding to linear and branching programmes". *Educational Review 17*(1965), 223-233.

39 Lublin, S C, "Reinforcement schedules, scolastic aptitude, autonomy need and achievement in a programed course". *Journal of Educational Psychology 56*(1965), 295-302.

40 McCrystal, T J & Jacobs, T O, The Effect of Programed Instruction Response Conditions on Acquisition and Retention. Technical Report 66-20, Alexandria, Va.: Human Resources Office, George Washington University, December 1966. (Abstract in AV-Communication Review 15(1967), 223.)

41 Moore, J W & Smith, W I, "A comparison of several types of 'immediate reinforcement'". Smith, W I & Moore, J W (Eds), *Programed Learning*. Princeton: van Nostrand, 1962. Pp 192-201.

42 Moore, J W & Smith, W I, "Knowledge of results in self-teaching spelling". Smith, W I & Moore, J W (Eds), *Programed Learning.* Princeton: van Nostrand, 1962. Pp 150-162.

43 Ripple, R E, "Comparisons of the effectiveness of programed text with three other methods of presentation". *Psychological Reports 12*(1963), 227-237. (BMN)

44 Roe, A & al., Automated Teaching Methods Using Linear Programs. Report No. 60-105. Los Angeles: Dept. of Engineering, University of California, December 1960. (BMN)
45 Schramm, W, *The Research on Programed Instruction: An Annotated Bibliography*. US Dept. of Health, Education and Welfare: Office of Education, Washington, D.C., 1964.
46 Stake, R E & Sjogren, D D, Activity Level and Learning Effectiveness. Lincoln, Nebraska: University of Nebraska, March 1964. (Abstract in AV-Communication Review 12(1964), 493f.)
47 Stolurow, L M & Walker, C C, "A comparison of overt and covert response in programmed learning". *Journal of Educational Research 55*(1962), 421-429.
48 Stukat, K G, "Ett experiment rörande kännedom om resultat och stegstorlek i ett språkläreprogram". *Pedagogisk forskning* 1965:2, 93-102.
49 Tobias, S & Abramson, T, "Interaction among anxiety, stress, response mode and familiarity of subject matter on achievement from programmed instruction". *Journal of Educational Psychology 62*(1971) 357-364.
50 Tobias, S & Weiner, M, "Effect of response mode on immediate and delayed recall from programmed materials". *Journal of Programmed Instruction 2*(1963) 9-13.
51 Tuel, J K & Metfessel, N S, "An investigation of overt versus covert responding in programmed instruction". *Californian Journal of Educational Research 16*(1965), 29-33.
52 Widlake, P, "The effects of mode of response on learning". *Educational Review 16*(1964), 120-129.
53 Williams, J P, "Comparison of several response modes in a review program". *Journal of Educational Psychology 54*(1963), 253-260.
54 Williams, J P & Levy, E I, "Retention of introductory and review programs as a function of response mode". *American Educational Research Journal 1* (1964), 211-218.
55 Wittrock, M C, "Response mode in the programming of kinetic molecular theory concepts". *Journal of Educational Psychology 54*(1963), 89-93.

B. Other references

56 Ahlström, K G & Amcoff, S. "Two ways of giving response information with learning machines of multiple-choice type". *Institute of Education, Uppsala University, Report No. 31,* 1967 (mimeo).

57. Aiken, L R, Jr, "Attitudes towards mathematics". *Review of Educational Research 40*(1970), 551-596.

58 Amsel, A, "Error response and reinforcement schedules in self-instructional devices". Lumsdaine, A A & Glaser, R (Eds), *Teaching Machines and Programmed Learning: A Source Book.* Washington, D.C.: N.E.A., 1960. P. 506-516.

59 Anderson, R C, "Educational psychology". *Annual Review of Psychology 18*(1967), 129-164.

60 Anderson, R C, "Control of student mediating processes during verbal learning and instruction". *Review of Educational Research 40*(1970), 349-369.

61 Annet, J, "The role of knowledge of results in learning: A survey". DeCecco, J P (Ed), *Educational Technology.* New York: Holt, Rinehart & Winston, 1964. Pp 279-286.

62 Annet, J, "Research strategies for a technology of teaching". Unwin, D & Leedham, J (Eds), *Aspects of Educational Technology...* London: Methuen & Co., 1967. Pp 429-437.

63 Austwick, K, "Automation in schools". Austwick, K (Ed), *Teaching Machines and Programming.* Oxford: Pergamon Press, 1964. Pp 174-201.

64 Bakan, D, "The test of significance in psychological research". *Psychological Bulletin 66*(1966), 423-437.

65 Bartos, B B, "Review of instruments developed to be used in the evaluation of the adequacy of reported research". *Phi Delta Kappan, Occasional Paper No. 2,* February 1969.

66 Berglund, GW, "Kovariansanalys och kvasi-experiment". *Pedagogiska rapporter nr 11, Umeå Universitet, Beteendevet. inst., avd. för pedagogik,* 1970 (mimeo).

67 Bjerstedt, Å, "Att konstruera undervisningsprogram: en studieguide". *Kompendier från pedagogisk-psykologiska institutione, Lärarhögskolan i Malmö, Nr 4,* 1967 (mimeo).

68 Björnsson, C H, *Läsbarhet.* Stockholm 1968.

69 Briggs, L J, "Prompting and confirmation conditions for three learning tasks employing the Subject-Matter-Trainer". Lumsdaine, A A (Ed), *Student Response in Programmed Instruction: A Symposium*. Washington, D.C.: N.A.S./N.R.C., 1961. Pp 375-388.
70 Briggs, L J & Angell, D, "Programed instruction in science and mathematics". *Review of Educational Research 34*(1964), 354-373.
71 Briggs, L J & Hamilton, N R, "Meaningful learning and retention: Practice and feedback variables". *Review of Educational Research 34*(1964), 545-558.
72 Brodbeck, M, "Logic and scientific methods in research on teaching". Gage, N L (Ed), *Handbook of Research on Teaching*. Chicago: Rand McNally & Co., 1965. Pp 44-93.
73 Cox, D R, "The use of a concomitant variable in selecting an experimental design". *Biometrika 44*(1957). 150-158.
74 Davies, I K, "Mathetics - a functional approach". Unwin, D & Leedham, J (Eds), *Aspects of Educational Technology....* London: Methuen & Co., 1967. Pp 205-216.
75 Dixon, W J (Ed), *BMD - Biomedical Computer Programs*. Health Science Computing Facility, UCLA, Revised edition, September 1965. (Also later editions.)
76 Edwards, A L, *Experimental Design in Psychological Research*. New York: Holt, Rinehart & Winston, 1960.
77 Edwards, W, Lindman H & Savage, L J, "Bayesian statistical inference for psychological research". *Psychological Review 70*(1963), 193-242.
78 Eigen, L D, "Some problems in field testing programs for teaching machines". *Journal of Educational Sociology 34*(1961), 372-376.
79 Eigen, L D, "Comparison of three modes of presenting a programmed instruction sequence". *Journal of Educational Research 55*(1962), 453-460.
80 Eigen, L D & Margulies, S, "Response characteristics as a function of information level". *Journal of Programmed Instruction 2*(1963), 45-54.
81 Elashoff, J D, "Analysis of covariance: a delicate instrument". *American Educational Research Journal 6*(1969), 383-401.
82 Fry, E B, *Teaching Machines and Programmed Instruction: An Introduction*. New York: McGraw-Hill, 1963.

83 Geis, G L & al., "The role of the printed answer in programmed instruction". *NSPI Journal 9*(1970)6, 8-18.

84 Goldbeck, R A & Briggs, L J, "An analysis of response mode and feedback factors in automated instruction". *American Institute of Research, Technical Report,* 1960. No. AIR-328-60-IR-133.

85 Goldstein, L & Gotkin, L G, "A review of research: Teaching machines versus programed textbooks as presentation modes". *Journal of Programmed Instruction 1*(1962), 29-36.

86 Golloday, F L, "Problems in the econometric analysis of educational technology". *OECD Conference on Policies for Educational Growth: Education and Distribution of Income, Some Exploratory Forays.* OECD Doc. DAS/EID/70.3, Paris, 20th May, 1970.

87 Guilford, J P, *Psychometric Methods.* New York: McGraw-Hill, 1954.

88 Hald, A, *Statistical Tables and Formulas.* New York: John Wiley & Sons, 2nd printing, 1958.

89 Hartley, J, "Factors affecting the efficiency of learning from programed instruction". *AV-Communication Review 19*(1971), 133-148.

90 Holland, J G, "Research on programming variables". Glaser, R (Ed), *Teaching Machines and Programmed Learning II.* Washington, D.C.: N.E.A., 1965. Pp 66-117.

91 Holland, J G & Kemp, F D, "A measure of programing in teaching-machine material". *Journal of Educational Psychology 56*(1965), 264-269.

92 Kanner, J H & Sulzer, R L, "Overt and covert rehearsal of fifty per cent versus hundred per cent of the material in film learning". Lumsdaine, A A (Ed), *Student Response in Programmed Instruction: A Symposium.* Washington, D.C.: N.A.S./N.R.C., 1961. Pp 427-443.

93 Kay, H, "General introduction to teaching machine procedures". Austwick, K (Ed), *Teaching Machines and Programming.* Oxford: Pergamon Press, 1964. Pp 1-42.

94 Kemp, F D & Holland, J G, "Blackout ratio and overt responses in programed instruction". *Journal of Educational Psychology 57*(1966), 109-114.

95 Klaus, D J, "The art of auto-instructional programing". DeCecco, J P (Ed), *Educational Technology.* New York: Holt, Rinehart & Winston, 1964. Pp 35-50.

96 Lewis, D G & Whitwell, M N, "The effects of reinforcement and response upon programmed learning". *Programmed Learning and Educational Technology 8* (1971), 186-195.

97 Lockard, J & Sidowski, J B, "Learning in fourth and sixth grades as a function of sensory mode of stimulus presentation and overt or covert practice". *Journal of Educational Psychology 52* (1961), 262-265.

98 Lumsdaine, A A, "Some conclusions concerning student response and a science of instruction". Lumsdaine, A A (Ed), *Student Response in Programmed Instruction: A Symposium*. Washington, D.C.: N.A.S./N.R.C., 1961. Pp 471-500.

99 Lumsdaine, A A, "Instruments and media of instruction". Gage, N L (Ed), *Handbook of Research on Teaching*. Chicago: Rand McNally & Co., 1963. Pp 583-682.

100 Lumsdaine, A A, "Some theoretical and practical problems in programmed instruction". DeCecco, J P (Ed), *Educational Technology*. New York: Holt, Rinehart & Winston, 1964. Pp 77-91.

101 Lumsdaine, A A, "Assessing the effectiveness of instructional programs". Glaser, R (Ed), *Teaching Machines and Programed Learning II*. Washington, D.C.: N.E.A., 1965. Pp 267-320.

102 Magnusson, D. *Testteori*. Halmstad 1961.

103 Malmquist, E, "Undervisningsmaskiner och deras användning". *Skola och samhälle - Pedagogiskt forum* 1961:6, 175-187.

104 Malmquist, E & Grundin, H U, "Olika former för bjudning av programmerat undervisningsmaterial". *Forskningsrapporter från Statens Försökskola i Linköping, nr 8*. Falköping 1968.

105 Markle, D G, "Controlling behavior changer's behavior". *AV-Communication Review 16*(1968), 188-203.

106 McGuire, W J, "Audience participation and audio-visual instruction: Overt-covert responding and rate of presentation". Lumsdaine, A A (Ed), *Student Response in Programmed Instruction: A Symposium*. Washington, D.C.: N.A.S./N.R.C., 1961, Chapter 27.

107 Meyer, S R, "Report on the initial test of a junior high-school vocabulary program". Lumsdaine, A A & Glaser, R (Eds), *Teaching Machines and Programmed Instruction: A Source Book*. Washington, D.C.:

N.E.A., 1960. Pp 229-246.
108 Michael, D N & Maccoby, N, "Factors influencing the effects of students participation on verbal learning from films; Motivating versus practice effects, 'feedback', and overt versus covert responding". Lumsdaine, A A (Ed), *Student Response in Programmed Instruction: A Symposium*. Washington, D.C.: N.A.S./N.R.C., 1961. Chapter 18.
109 Milan, N & Bernath, P, "The use of programmed learning in grammar for 11-year-old pupils". Unwin, D & Leedham, J (Eds), *Aspects of Educational Technology....* London: Methuen & Co., 1967. Pp 125-131.
110 Mills, G H, "The analysis of response errors in the evaluation of technical teaching programs". *Programmed Learning and Educational Technology* 6(1969), 121-132.
111 Mosier, C I, "Symposium: The need and means of crossvalidation. I: Problems and designs of crossvalidation". *Educational and Psychological Measurement* 11(1951), 5-11.
112 Nasatir, E D & Elesh, D, "The measurement of quality in educational research...". Lazarsfeld, P F & Sieber, S D, *The Organization of Educational Research. Project Memorandum No. 5.* Bureau of Applied Social Research, Columbia University, September 1965 (mimeo).
113 Nunnally, J, "The place of statistics in psychology". *Educational and Psychological Measurement 20* (1960), 641-650.
114 Poppleton, P K & Austwick, K, "A comparison of programmed learning and note making at two age levels". *British Journal of Educational Psychology 34* (1964), 43-50.
115 Reid, R L, "Linear programming and learning". Austwick, K (Ed), *Teaching Machines and Programming*. Oxford: Pergamon Press, 1964. Pp 136-173.
116 Rippey, R M, "The Gunther model: four dimensions of research on instruction". *The Elementary School Journal 69*(1969), 215-223.
117 Rosenthal, R, *Experimental Effects in Behavioral Research*. New York: Appleton-Century-Crofts, 1966.
118 Silberman, H F, "Self-teaching devices and programmed materials". *Review of Educational Research 32* (1962), 179-193.

119 Sime, M & Boyce, G, "Overt responses, knowledge of results and learning". *Programmed Learning and Educational Technology 6*(1969), 12-19.
120 Skinner, B F, "Teaching machines". Lumsdaine, A A & Glaser, R (Eds), *Teaching Machines and Programmed Learning: A Source Book*. Washington, D.C.: N.E.A., 1960. Pp 137-158.
121 Skinner, B F, "Why we need teaching machines". DeCecco, J P (Ed), *Educational Technology*. New York: Holt, Rinehart & Winston, 1964. Pp 92-112.
122 Stukát, K G, "Ett försök med programmering av svensk språklära". *Rapporter från pedagogiska institutionen vid Göteborgs universitet*, januari 1963 (mimeo).
123 Tobias, S, "Review of the response mode issue". *Review of Educational Research 43*(1973), 193-204.
124 Walker, H M & Lev, J, *Statistical Inference*. New York: Holt, Rinehart & Winston, 1953.
125 Yaeger, B W, "A consideration of some objective properties of linear programs". Ofiesh, G D & Meierhenry, W C (Eds), *Trends in Programed Instruction*. Washington, D.C.: N.E.A., 1964. Pp 182-184.

Statens Försöksskola
Linköping

Bilaga 1 (s 1)
Exp PUSL67
Prov

Namn: ______________________________ Klass: __________

Skola: __________________________ Datum: _______ Grupp: ____

A. I var och en av meningarna nr 1 - 15 (nedan) är ett ord understruket. Du skall svara på två frågor om det understrukna ordet!

a) Är det understrukna ordet subjekt eller predikat i meningen? Svara genom att sätta kryss i en av rutorna till vänster om meningen!

b) Till vilken ordklass hör det understrukna ordet: substantiv, pronomen eller verb? Svara genom att sätta kryss i en av rutorna till höger om meningen!

a) ordet är: subjekt	predikat		b) ordet hör till ordklassen: substantiv	pronomen	verb
☐	☐	1. Harald springer fort.	☐	☐	☐
☐	☐	2. Räven smet in i skogen.	☐	☐	☐
☐	☐	3. Pojkarna mötte en hund.	☐	☐	☐
☐	☐	4. Den hade gått vilse	☐	☐	☐
☐	☐	5. Tåget kommer om en kvart.	☐	☐	☐
☐	☐	6. Det här huset byggdes förra året.	☐	☐	☐
☐	☐	7. Hon är 11 år gammal.	☐	☐	☐
☐	☐	8. Ni måste komma i tid.	☐	☐	☐
☐	☐	9. Bo kunde inte sin läxa i dag.	☐	☐	☐
☐	☐	10. I går kunde han inte heller sin läxa.	☐	☐	☐
☐	☐	11. Snart kommer tåget.	☐	☐	☐
☐	☐	12. Stugan låg långt inne i skogen.	☐	☐	☐
☐	☐	13. Nu åker jag hem.	☐	☐	☐
☐	☐	14. Flickorna hade väldigt roligt.	☐	☐	☐
☐	☐	15. Lisas bästa vän heter Klas.	☐	☐	☐

VÄND INTE BLAD FÖRRÄN DU BLIR TILLSAGD !

B. I var och en av meningarna nr 16 - 30 (nedan) skall du stryka under hela subjektsdelen med ett streck och hela predikatsdelen med två streck!

16. Hunden skäller.
17. En flicka ramlade.
18. Den svarta katten försvann.
19. Svenssons bil är mycket gammal.
20. Karin tittade ut genom fönstret och hörde inte på.
21. Några pojkar lekte i skogen.
22. Två av dem gjorde upp eld.
23. De tänkte grilla korv.
24. En gammal torr buske fattade eld.
25. Elden spred sig snabbt.
26. En av pojkarna sprang efter hjälp.
27. Snart kom brandkaren.
28. Både brandsoldaterna och pojkarna satte igång med släckningen.
29. Folk som bodde i närheten hjälpte också till.
30. Till slut släcktes skogsbranden.

VÄND INTE BLAD FÖRRÄN DU BLIR TILLSAGD!

C. I var och en av meningarna nr 31 – 45 (nedan) skall du göra följande:

Skriv ett P mitt ovanför predikatet (ett ord) och ett S mitt ovanför subjektet (ett ord)!

Stryk sedan under hela predikatsdelen med två streck och hela subjektsdelen med ett streck!

31. En vit katt sprang förbi.
32. Den där hunden skäller aldrig.
33. En röd bil står på gården.
34. Den lilla pojken vinkade ivrigt.
35. Månen lyser om natten.
36. En flicka från Jönköping vann tävlingen.
37. Några elever pratade hela tiden.
38. Det sista tåget går snart.
39. Barn under 6 år åker gratis.
40. Alla båtarna i hamnen seglade ut.
41. Där borta kommer en bil.
42. Om dagen sover alla fladdermöss.
43. Den här meningen är svår.
44. Zebran är ett randigt djur.
45. Den har svarta och vita ränder.

SLUT PÅ PROVET

Statens Försöksskola, Linköping

INSTRUKTION FÖR PROVLEDARE/ARBETSLEDARE VID EXP PUSL67

(pl: provledare, al: arbetsledare)

Lektion 1: DBA-testning

Pl lägger före lektionens start ut provblankett DBA:2 (motsatser) på alla elevplatser. (Se till att blanketterna läggs med rätt sida upp.)

När eleverna intagit sina platser säger pl:
"Låt papperet, som ni har fått, ligga på bänken och lyssna på mig. Den här klassen ska under dom närmaste veckorna vara med i ett försök. Det försöket går ut på att pröva olika sätt att lära sig saker på egen hand. Innan vi startar försöket vill vi veta hur duktiga ni är på att lösa vissa sorters uppgifter. Därför ska ni i dag få gå igenom tre olika prov - två prov den här timmen och ett prov en annan timme.
Jag kommer att tala om precis hur ni ska göra - när ni ska börja och sluta, när ni ska vända blad osv. Gör som jag säger - men gör ingenting annat! Ni ska arbeta snabbt. Men slarva inte. Och var och en måste arbeta helt på egen hand utan att kika på sina kamrater. ---
Nu ska ni skriva namn, födelsedatum, skola och årskurs nedtill på det papper, som ni har fått."
När alla har fyllt i dessa saker, fortsätter pl:
"Ni har nu framför er anvisningarna till ett prov som heter 'Motsatser'. Följ med i anvisningarna, när jag går igenom dem."
Pl läser anvisningarna högt. Uppehåll göres, så att eleverna kan fylla i svaren till övningsexemplen. Därefter kontrolleras svaren till övningsexemplen genom frågor till någon eller några elever. Se till att alla elever får klart för sig vad som är rätt svar på resp. övningsexempel.

Facit: Övn.ex. DBA:2	uppg.	2	3	4	5
	svar	3	1	4	2

Därefter frågar pl:
"Är det någon som inte förstått, hur ni ska lösa såna här uppgifter?"
Eventuella frågor besvaras i enlighet med anvisningarna (t.ex.: "Se efter vilket av de fyra numrerade orden som är motsatsen till det första ordet. Skriv numret på detta ord, 1, 2, 3, eller 4 i svarsrutan.") Sedan säger pl:
"Nu ska vi gå över till själva provet. Det består av samma slags uppgifter som övningsexemplen. --- Vänd blad. --- Börja."
Efter e x a k t 1 0 m i n u t e r säger pl:
"Tiden slut. Lägg ifrån er pennorna."
Pl låter samla in blanketterna till DBA:2 och dela ut blanketter till DBA:3 (Bokstavsgrupper). När detta är klart säger pl:
"Ni har nu framför er anvisningarna till ett prov som heter 'Bokstavsgrupper'. Följ med i anvisningarna, när jag går igenom dem."
Pl läser anvisningarna högt. Uppehåll görs, så att eleverna kan fylla i svaren till övningsexemplen. Därefter kontrolleras svaren till övningsexemplen genom frågor till någon eller några elever. Se till att alla elever får klart för sig, vad som är rätt svar på resp. övningsexempel.

Facit: Övn.ex. DBA:3	uppg.	2	3	4	5	6	7
	svar	1	1	3	2	1	3

Därefter frågar pl:
" Är det någon som inte har förstått, hur ni ska lösa såna här uppgifter?"
Eventuella frågor besvaras i enlighet med anvisningarna (t ex: "Det gäller att se efter, vilken bokstavsgrupp, som skiljer sig från dom övriga. Numret på den gruppen ska man skriva i svarsrutan"). Sedan säger pl:
"Nu ska vi gå över till själva provet. Det består av samma slags uppgifter som övningsexemplen. --- Vänd blad. --- Börja."

Efter exakt 15 minuter säger pl:
"Tiden slut. Lägg ifrån er pennorna."
Därefter samlas provblanketterna in.

Lektion 2: Kunskapsprov i språklära (förtest)

Pl delar ut provhäftena (märkta "Exp PUSL67, Prov" i övre högra hörnet) till eleverna och låter dessa fylla i namn, klass, skola och datum överst på första sidan i häftet (i utrymmet för gruppnummer - "Grupp: ...." - skall ingenting skrivas.). När alla eleverna har gjort detta säger pl:

"Lägg ifrån er pennorna och lyssna noga. --- Ni brukar ju ofta få prov för att ni ska få visa vad ni har lärt er i skolan. - Men det här är inte något sånt prov. Det här häftet innehåller också uppgifter på sånt som ni ännu inte har lärt er i skolan. -- Ni ska få lära er dom sakerna längre fram. Men nu vill vi veta, om det är några av er, som kan det ändå -- fast ni inte har fått lära er det i skolan. -- Ni ska alltså inte bli ledsna, om det är något ni inte förstår alls. Gör så gott ni kan, så blir det nog bra. -- De uppgifter ni absolut inte kan lösa, får ni hoppa över. --
Nu ska jag tala om hur ni ska göra.
Det är tre blad i häftet. Vi ska arbeta med ett blad i taget och ni får aldrig vända blad förrän jag säger till. --
På första sidan har vi A. Överst på sidan står det hur man ska lösa uppgifterna (pl håller upp häftet och pekar). --
Nu läser vi dom anvisningarna. Följ med i häftet allihop."

Delprov A:

Pl läser ordagrant upp anvisningarna till delprov A ur häftet och säger sedan:
"Kom ihåg alltså, att ni ska sätta kryss på två ställen vid varje mening: dels i en av de båda rutorna till vänster (håll upp häftet och peka), dels i en av de tre rutorna till höger (peka). Ta linjalen till hjälp, så ni håller er på rätt rad."
Pl visar eleverna hur de ska lägga linjalen tvärs över sidan, strax under resp. rad med mening och tillhörande svarsrutor. Sedan säger pl:
"Ta upp pennorna. - Om ni glömmer bort, hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja."
Efter exakt 6 minuter[1] säger pl:
Lägg ifrån er pennorna och vänd blad. Här har vi prov B. -- Nu läser vi anvisningarna överst på sidan. - Följ med i häftet allihop."

Delprov B:

Pl läser ordagrant upp anvisningarna till delprov B ur häftet och säger sedan:
"Ni ska alltså stryka under subjektsdelen med ett streck och predikatsdelen med två streck. Och ni ska göra så med varje mening. -- Använd linjal, när ni stryker under. -- Ta upp pennorna. -- Om ni glömmer bort, hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja."
Efter exakt 6 minuter[1] säger pl:
"Lägg ifrån er pennorna och vänd blad. --"

---

1) Om samtliga elever uppenbart slutar arbeta med provet, får pl avbryta provet 30 sek efter det han konstaterat, att så är fallet. I sådant fall skall pl också notera, hur mycket provtiden därigenom förkortats.

Delprov C:

Pl säger vidare:

"Här har vi prov C. -- Nu läser vi anvisningarna överst på sidan. - Följ med i häftet allihop."

Pl läser ordagrant upp anvisningarna till delprov C ur häftet och säger sedan:

"Kom nu ihåg att göra allt det som står i anvisningarna. -- Skriv P ovanför predikatet och S ovanför subjektet. Stryk sen under predikatsdelen med två streck och subjektsdelen med ett streck. - Och gör så med varje mening. -- Använd linjal när ni stryker under. -- Ta upp pennorna. -- Om ni glömmer bort, hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja."

Efter e x a k t 1 0 m i n u t e r[1] säger pl:

"Lägg ifrån er pennorna. Nu är provet slut. -- Kontrollera nu att ni fyllt i namn, klass, skola och datum på första sidan. Låt sedan provhäftet ligga på bänken."

Pl samlar in provhäftena, såväl använda som ev. oanvända, och behåller dem i sitt förvar, tills han personligen kan överlämna materialet till försöksledaren.

Lektion 3: Inskolning i arbetssättet för resp grupper (1-6)

Före lektionen skall al

1) kontrollera att material erhållits för rätt grupp (jfr arbetsordning och utrustn.anvisn.)
2) kontrollera att bandspelare finns på plats och är i användbart skick (Det kan också vara lämpligt att al i förväg har lyssnat på instruktionsljudbandet.)
3) skicka den egna klassens elever - utrustade med gruppnummerlappar - till resp salar (jfr arbetsordning samt förteckn. över gruppindelning).

När eleverna anlänt till salen vid lektionens början skall al kontrollera

1) att endast elever med rätt gruppnummer finns i salen
2) att det har anlänt elever från samtliga de klasser, som enl. arbetsordningen skall samordnas under lektion 3.

När alla elever har den material som behövs för lektionen på sina bänkar (jfr utrustningsanvisn.) låter al dem fylla i namn, klass, skola och datum på första sidan i arbetshäftet.

Sedan säger al:

"I dag ska vi börja ett försök, där ni ska få arbeta helt på egen hand med en rad arbetshäften. Ni ska först få lära er hur ni ska arbeta. Men det är inte jag som ska tala om det för er. - Alla anvisningar finns på det här ljudbandet. När jag sätter igång bandspelaren, ska ni följa med noga allihop i arbetshäftet och göra precis som det sägs. Om ni har något att fråga om, så får ni vänta tills vi har spelat igenom bandet. -- Är alla klara? -- Då sätter vi igång."

Al startar bandspelaren.

Om al blir t v u n g e n att svara på någon fråga eller på annat sätt ingripa, så skall bandspelaren stoppas tills saken är uppklarad.

Bandet slutar med uppmaningen: "Nu får ni fortsätta själva med det här häftet". När bandet är slut, får al besvara frågor om arbetssättet i enlighet med instruktionerna på ljudbandet och den s.k. kom-ihåg-lappen. Al:s svar bör om möjligt inskränkas till uppmaningen "Läs på kom-ihåg-lappen" eller uppläsning

---

1) Om s a m t l i g a elever u p p e n b a r t slutar arbeta med provet, får pl avbryta provet 3 0 s e k efter det han konstaterat, att så är fallet. I sådant fall skall pl också notera, hur mycket provtiden därigenom förkortats.

av vad som står på den lappen. Arbetstiden behöver inte noteras under denna lektion. När alla elever är färdiga med häftet, avbryts verksamheten (även arbetet med räkneuppgifterna). Därefter i n s a m l a s a l l m a t e r i a l.
OBS! I häftet "Förprogram" (version 1 och 2) förekommer det, att det rätta svaret lyser igenom, så att det kan läsas utan att man vänder blad. Denna brist har eliminerats i huvudprogrammets häften. Om någon elev under lektion 3 påpekar att svaren lyser igenom, så skall al upplysa hela gruppen om att det inte kommer att vara så i de följande häftena.

Lektion 4-9: Genomgång av resp. häften i programmet "Den enkla meningens huvuddelar"

Före varje lektion skall al
1) kontrollera att material erhållits för rätt grupper (jfr arbetsordning och utrustn.anvisn.)
2) skicka iväg de elever i den egna klassen, som skall till en annan sal - utrustade med gruppnummerlappar (jfr arbetsordning).

När alla elever fått den material som behövs för lektionen på sina bänkar (jfr utrustn.anvisn.), låter al dem fylla i namn, klass, skola och datum på arbetshäftets första sida samt namn och klass på bladet med räkneuppgifter. Sedan säger al:

"I dag ska ni arbeta med det här häftet. Var och en ska arbeta precis på det sätt han eller hon har fått lära sig. För säkerhets skull har ni fått var sin kom-ihåg-lapp. På den står det hur ni ska arbeta. Läs allihop tyst igenom kom-ihåg-lappen. (Paus för läsningen.) Ni får gärna läsa på lappen om igen medan ni arbetar. Om ni ändå m å s t e fråga om något, så räck upp handen och vänta tyst och stilla, tills jag kommer till er. När ni är färdiga med hela häftet fortsätter ni med räkneuppgifterna. Kom ihåg en sak: Det här är inget prov och absolut ingen hastighetstävling. Det är inte viktigt att bli färdig så fort som möjligt, men det är viktigt att arbeta ordentligt och noggrant. Nu får ni börja."

Al a n t e c k n a r k l o c k s l a g (avr. t. hel minut), när arbetet p å b ö r j a t s . (Utrymme härför finns på bladet "Utrustningsanvisningar"). Under arbetets gång får al endast ge vägledning ang. arbetssättet. Detta sker genom hänvisning till kom-ihåg-lappen samt ev. uppläsning för enskild elev av vad som står där. (Kom ihåg att alla elever inte har samma instruktion.) Al går runt och a n t e c k n a r k l o c k s l a g (avr. t. hel minut), när varje elev a v s l u t a t arbetet med häftet. När alla elever är klara med arbetshäftet, kan verksamheten avbrytas. Al låter de elever, som börjat med räkneuppgifter, markera på uppgiftsbladet, hur långt de hunnit. Därefter i n s a m l a s a l l m a t e r i a l . Bladen med räkneuppgifter skall al behålla och utdela på nytt nästa försökslektion (t o m lektion 9). Eleverna får alltså nytt blad med räkneuppgifter först när de gjort färdigt uppgifterna på det förra. OBS! Al måste givetvis se till, att varje elev får häfte resp. kom-ihåg-lapp för den grupp som eleven tillhör.

Lektion 10: Kunskapsprov i språklära (eftertest 1)

Pl delar ut provhäftena till eleverna och låter dessa fylla i namn, klass, skola, datum och gruppnummer överst på första sidan i häftet. När alla elever har gjort detta säger pl:

"Lägg ifrån er pennorna och hör upp. --- I dag ska ni få visa vad ni har lärt er av de häften ni har arbetat med de senaste veckorna. Det här provhäftet består av tre blad. Vi ska arbeta med ett blad i taget och ni får aldrig vända blad förrän jag säger till. På första sidan har vi prov A. Upptill på sidan står det hur man ska lösa uppgifterna (pl håller upp häftet och pekar). -- Nu läser vi dom anvisningarna. Följ med i häftet allihop."

Delprov A:

Pl läser ordagrant upp anvisningarna till delprov A ur häftet och säger sedan:
"Kom ihåg alltså, att ni ska sätta kryss på två ställen vid varje mening: dels i en av de båda rutorna till vänster (pl håller upp häftet och pekar), dels i en av de tre rutorna till höger (pl pekar). --- Ta linjalen till hjälp, så ni håller er på rätt rad."

Pl visar eleverna hur de ska lägga linjalen tvärs över sidan, strax under resp. rad med mening och tillhörande svarsrutor. Sedan säger pl:
"Ta upp pennorna. - Om ni glömmer bort, hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja"

Efter e x a k t 6 m i n u t e r säger pl:
"Lägg ifrån er pennorna och vänd blad. Här har vi prov B. -- Nu läser vi anvisningarna överst på sidan. Följ med i häftet allihop."

Delprov B:

Pl läser ordagrant upp anvisningarna till delprov B ur häftet och säger sedan:
"Ni ska alltså stryka under subjektsdelen med ett streck och predikatsdelen med två streck. Och ni ska göra så med varje mening. -- Använd linjal, när ni stryker under. -- Ta upp pennorna. -- Om ni glömmer bort, hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja."

Efter e x a k t 6 m i n u t e r säger pl:
"Lägg ifrån er pennorna och vänd blad.-- Här har vi prov C. -- Nu läser vi anvisningarna överst på sidan. Följ med i häftet allihop."

Delprov C:

Pl läser ordagrant upp anvisningarna till delprov C ur häftet och säger sedan:
"Kom nu ihåg att göra allt det som står i anvisningarna. -- Skriv P ovanför predikatet och S ovanför subjektet: Stryk sen under predikatsdelen med två streck och subjektsdelen med ett streck. -- Och gör så med varje mening. -- Använd linjal, när ni stryker under. -- Ta upp pennorna. -- Om ni glömmer bort hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja.

Efter e x a k t 10 m i n u t e r säger pl:
"Lägg ifrån er pennorna. Nu är provet slut. -- Kontrollera nu att ni fyllt i namn, klass, skola, datum och gruppnummer på första sidan. Låt sedan provhäftet ligga på bänken."

Pl samlar in provhäftena, såväl använda som ev. oanvända, och ser sedan till, att all material återlämnas till försöksledaren.

Lektion 11: Kunskapsprov i språklära (eftertest 2)

Pl delar ut provhäftena till eleverna och låter dessa fylla i namn, klass, skola, och datum samt gruppnummer (om uppgiften är tillgänglig). När alla elever gjort detta säger pl:
"Lägg ifrån er pennorna och hör upp. -- Ni kommer säkert ihåg, att ni för några veckor sedan gick igenom en del arbetshäften i språklära. De handlade mest om subjekt och predikat och liknande saker. -- Nu ska vi se vad ni minns av det ni lärde er då. Det här provhäftet består av tre blad. Vi ska arbeta med ett blad i taget och ni får aldrig vända blad förrän jag säger till. På första sidan har vi prov A. Upptill på sidan står det hur man ska lösa uppgifterna (pl håller upp häftet och pekar). -- Nu läser vi dom anvisningarna. Följ med i häftet allihop."

Delprov A:

Pl läser ordagrant upp anvisningarna till delprov A ur häftet och säger sedan:
"Kom ihåg alltså, att ni ska sätta kryss på två ställen vid varje mening: dels i en av de båda rutorna till vänster (pl håller upp häftet och pekar),

dels i en av de tre rutorna till höger (pl pekar). -- Ta linjalen till hjälp, så ni håller er på rätt rad."

Pl visar eleverna hur de ska lägga linjalen tvärs över sidan, strax under resp. rad med mening och tillhörande svarsrutor. Sedan säger pl:

"Ta upp pennorna. - Om ni glömmer bort, hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja."

Efter e x a k t 6 m i n u t e r säger pl:

"Lägg ifrån er pennorna och vänd blad. Här har vi prov B. -- Nu läser vi anvisningarna överst på sidan. Följ med i häftet allihop."

Delprov B:

Pl läser ordagrant upp anvisningarna till delprov B ur häftet och säger sedan:

"Ni ska alltså stryka under subjektsdelen med ett streck och predikatsdelen med två streck. Och ni ska göra så med varje mening. -- Använd linjal, när ni stryker under. -- Ta upp pennorna. -- Om ni glömmer bort, hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja."

Efter e x a k t 6 m i n u t e r säger pl:

"Lägg ifrån er pennorna och vänd blad. -- Här har vi prov C. -- Nu läser vi anvisningarna överst på sidan. Följ med i häftet allihop."

Delprov:

Pl läser ordagrant upp anvisningarna till delprov C ur häftet och säger sedan:

"Kom nu ihåg att göra allt det som står i anvisningarna. -- Skriv P ovanför predikatet och S ovanför subjektet. Stryk sen under predikatsdelen med två streck och subjektsdelen med ett streck. -- Och gör så med varje mening. -- Använd linjal, när ni stryker under. -- Ta upp pennorna. -- Om ni glömmer bort hur ni skulle göra, så läs anvisningarna överst på sidan en gång till. -- Nu får ni börja."

Efter e x a k t 10 m i n u t e r säger pl:

"Lägg ifrån er pennorna. Nu är provet slut. -- Kontrollera nu att ni fyllt i namn, klass, skola, datum och gruppnummer på första sidan. Låt sedan provhäftet ligga på bänken."

Pl samlar in provhäftena, såväl använda som ev. oanvända, och ser sedan till, att all material återlämnas till försöksledaren.

Statens Försöksskola
Linköping

Bilaga 3 (s 1)
Exp PUSL67
Instr. förprogram

INSTRUKTION VID BJUDNING AV FÖRPROGRAM I EXP PUSL67

G r u p p 1 (krav på yttre svar, 100 % facitsvar)

Lägg ifrån er pennorna och låt häftena ligga på bänken. -- Lyssna nu noga. --

Ni ska under de närmaste veckorna få pröva på en ny typ av arbetsövningar i språklära. Det är arbetshäften, som ni ska få arbeta med helt på egen hand. Alla elever som använder de här häftena får samma arbetsuppgifter. Men olika grupper av elever ska arbeta på olika sätt med uppgifterna. Sen får vi se vilken grupp som lyckas bäst.

Ni som sitter här tillhör allihop samma grupp - grupp 1. Ni ska allihop arbeta på samma sätt. I dag ska ni få lära er hur ni ska arbeta. Titta nu på det arbetshäfte som ni har fått. På första sidan har ni fyllt i namn och en del andra saker. Nu vänder ni blad och tittar på nästa sida. -- Varje sida i häftet har två uppgiftsrutor, precis som den här sidan. Ni ska börja med översta rutan på första sidan. Sen ska ni fortsätta med översta rutan på nästa sida, osv till sista sidan i häftet. När ni har gått igenom den översta rutan på varje sida, då börjar ni från början av häftet igen. Men då med den nedersta rutan. Så fortsätter ni med den nedersta rutan på varje sida till slutet av häftet.

Ni ska alltid börja med att läsa noga allt som står i rutan. I en del rutor finns en fråga och plats för ett svar. Där ska ni skriva svaret på frågan. I andra rutor är det något som fattas. Där ska ni fylla i det ord eller dom ord som fattas. När ni är färdiga med en ruta, ska ni kontrollera att ni har gjort rätt. Det rätta svaret - eller det som fattades - står i marginalen till vänster på nästa sida. Nu tycker ni kanske att man kan titta på det rätta svaret först och sen skriva av det. -- Visst kan man det, men då lär man sig nog inte mycket. -- Ska man lära sig det här ordentligt, måste man först tänka efter och svara själv och sen kontrollera att man gjort rätt. --

Nu ska vi gå igenom en del av häftet tillsammans, så ni säkert förstår hur ni ska arbeta. Ta pennan i handen och följ med i häftet hela tiden. -- Läs texten tyst för er själva samtidigt som jag läser upp den. --
Vi läser ruta 1. Där står: "Substantiv -- Titta noga på ordet. Se efter hur det stavas. -- Hur många bokstäver är det i ordet substantiv? Svar: ..." Ja, hur många är det? Räkna dom och skriv svaret. -- Nu vänder vi blad. -- Har ni svarat rätt? Se efter i marginalen till vänster. -- Där står t i o . Det betyder att ordet substantiv har tio bokstäver. --
Vi läser ruta 2. Där står: "I en bok finns massor av ord. Det finns olika slags ord, olika ordklasser. Ett slags ord kallas substantiv. Substantiv är namn på en ..." Ja, där fattas ett ord. Vilket ord fattas på sista raden? -- Tänk efter vad det är och skriv dit det ordet. -- Vänd blad. Se efter i marginalen vilket ord det var som fattades i förra rutan. -- Om ni gjort fel, så får ni gärna gå tillbaka till förra rutan och ändra så att det blir rätt. --
Vi läser ruta 3. Där står: "Det är bara en särskild sorts ord som kallas substantiv. Ordet penna är ett substantiv, men ordet skriva är inte något substantiv. Vilket ord är ett substantiv, penna eller skriva? Svar: ..." -- Tänk efter vilket det är och skriv svaret. Vänd blad. -- Har ni svarat rätt? Se efter i marginalen till vänster. -- Där står "penna". Ordet penna är alltså ett substantiv. --
Vi läser ruta 4. Där står: "Ordet tvål är ett substantiv, men ordet tvätta är inte något substantiv. Vilket ord är ett substantiv, tvål eller tvätta? Svar:.." -- Tänk efter vilket det är och skriv svaret. -- Vänd blad. Se efter i marginalen om ni har svarat rätt. -- Där står "tvål". Rätt svar är alltså "tvål". --

Vi läser ruta 5. Där står: "Ordet boll är ett substantiv, men ordet kasta är inte något substantiv. Vilket ord är ett substantiv, boll eller kasta? Svar:..." -- Tänk efter vilket det är och skriv svaret. -- Vänd blad. Se efter om ni har svarat rätt. -- Svaret är "boll". Det står i marginalen till vänster. --
Vi läser ruta 6. Där står: "Här är två ord: Penna, skriva. Den ena ordet är namn på en sak, som du ofta håller i. Det andra ordet är inte namn på en sak, det är något man kan göra. Vilket ord är namn på en sak? Svar: ..." Tänk efter vilket det är och skriv svaret. -- Vänd blad. -- Har ni svarat rätt? -- Se efter i marginalen. -- Svaret är "penna". --
Vi läser ruta 7. Där står: "Här kommer två ord: kasta boll. Det ena ordet är namn på en sak, som man kan leka med. Det andra ordet är inte namn på en sak, det är ju något man gör. Vilket ord är namn på en sak? Svar: ..." -- Tänk efter vilket det är och skriv svaret. -- Vänd blad. -- Har ni svarat rätt? -- Se efter i marginalen till vänster. --
Vi läser ruta 8. Där står: "Här kommer två ord: bok, läsa. Det ena ordet är namn på en sak, men det är inte det andra. Vilket ord är namn på en sak? Svar: ..." -- Tänk efter vilket det är och skriv svaret. -- Vänd blad. -- Se efter om ni har svarat rätt. --
Vi läser ruta 9. Där står: "Alla ord, som är namn på saker, kallas substantiv. Här kommer två ord: spade, gräva. Vilket av orden är ett substantiv? Svar: ..." -- Tänk efter vilket det är och skriv svaret. -- Vänd blad. -- Har ni svarat rätt? -- Se efter i marginalen. --
Vi läser ruta 10. Där står: "Alla namn på saker är substantiv. Här kommer två ord: tvätta, tvål. Vilket av dem är ett substantiv? Svar: ..." -- Tänk efter vilket det är och skriv svaret. -- Vänd blad och se efter om ni har svarat rätt. -- Det rätta svaret står i marginalen till vänster.
Lägg ifrån er pennorna. -- Nu vet ni säkert allihop hur ni ska arbeta med såna här häften. Ni ska strax få fortsätta på egen hand med det här häftet. För säkerhets skull har var och en fått en kom-ihåg-lapp. Där står det vad ni särskilt ska tänka på, när ni arbetar. Läs igenom den lappen om ni är osäkra. Ni får läsa den när ni vill och hur ofta ni vill.
Om ni upptäcker att ni svarat fel på någon fråga, så får ni gärna ändra ert svar, så det blir rätt.
När ni kommer till översta rutan på sista sidan, kom då ihåg att gå tillbaka till början av häftet och fortsätta med den nedersta rutan på varje sida.
När ni är färdiga med det här häftet, tar ni itu med de räkneuppgifter som ni har fått. --
Nu får ni fortsätta själva med det här häftet."

SLUT PÅ BANDET

EXCERPT FROM THE INSTRUCTIONAL PROGRAM

Statens Försöksskola
Linköping

Exp. PUSL67
Arbetshäfte

DEN ENKLA MENINGENS HUVUDDELAR

V e r s i o n 1 (för grupp nr 1 och 4)

H ä f t e 1

Namn: -- Klass: ------------

Skola: -- Datum: ------------

Arbetet med detta häfte avslutat klockan --------------

Note 1: In order to prevent reading of key answers through the semi-transparent paper, every sheet of paper in the program had a dot screen printed on its back. This dot screen is not included in the program excerpt given here.

Note 2: The original program was printed in A51 format. To save space two program pages are reprinted on each page in this appendix. Frame numbers in the upper right-hand corner of each frame indicate the order in which the frames should be read.

Kom ihåg att du skall gå igenom uppgifterna i <u>nummerordning</u>.
Börja överst på nästa sida ⟶

---

xxxxxxxxxxxx xxxxxxxxxxxx	22 Den del, som berättar något om subjektsdelen, kallas <u>p r e d i k a t s d e l</u>. P R E - D I - K A T S - D E L (tonvikt på KAT) Se efter hur ordet stavas och skriv det här: ____________________
subjektsdel	12 PELLE SPRINGER. Denna mening handlar om PELLE. Ordet PELLE är alltså ____________________ .
brinner	34 JOHAN DISKAR. Predikatsdelen <u>berättar</u> något om subjektsdelen. Vilket ord i meningen ovan är <u>predikatsdel</u>? <u>Svar:</u> ____________________ .

xxxxxx	När vi talar, använder vi många olika ord. Varje ord betyder något. Av orden bildar vi meningar. DU SKRIVER. De två orden här ovan bildar en ____________________.
predikatsdel	23 PELLE SPRINGER. Det ord som berättar något om PELLE kallas p r e d i k a t s d e l. Vilket ord är p r e d i k a t s d e l ? Svar: ____________________
subjektsdel	13 HUNDEN SKÄLLDE. Denna mening handlar inte om katten eller fågeln. Den handlar om ____________________.
diskar	35 BLOMMORNA VISSNAR. Predikatsdelen berättar något om subjektsdelen. Vilket ord är predikatsdel? Svar: ____________________.

mening xxxxxx	FAGELN FLYGER. Den här meningen handlar inte om hunden eller hästen. Den handlar om ________.	2
springer	FAGELN FLYGER. Det ord som berättar något om FAGELN kallas predikatsdel. Vilket ord är predikatsdel? Svar: ________.	24
hunden	HUNDEN SKÄLLDE. Ordet ________ är subjektsdel.	14
vissnar	BARNEN SKRATTAR. Barnen är subjektsdel. SKRATTAR är ________.	36

	3
fågeln	FAGELN ...... FAGELN FLYGER är en mening, men ordet FAGELN ensamt är inte en ________.
	25
flyger	HUNDEN SKÄLLDE. Predikatsdelen berättar något om subjektsdelen. Vilket ord är predikatsdel i meningen ovan? Svar: ________.
	15
hunden	LISA SKRATTADE. Den här meningen handlar inte om PELLE. Den handlar om ________.
	37
predikatsdel	BARNEN SKRATTAR. SKRATTAR är predikatsdel: BARNEN är ________.

mening	FAGELN FLYGER Ordet FAGELN säger vad vi talar om. För att det skall bli en mening måste vi berätta något om fågeln. Vilket ord i meningen berättar något om FAGELN? Svar: ______________
skällde	26 VATTNET FRYSER Ordet FRYSER berättar något om VATTNET. Ordet FRYSER är alltså ______________
Lisa	16 LISA SKRATTADE. Meningen handlar om LISA. Ordet LISA är ______________
subjektsdel	38 En mening har två delar S U B J E K T S D E L och P R E D I K A T S D E L. Predikatsdelen berättar något om ______________

flyger	FAGELN FLYGER. Meningen handlar om FAGELN. Den berättar också något om FAGELN. Den berättar att den ______________.	5
predikatsdel	PELLE SPRINGER. Meningen handlar om PELLE (subjekts-del). Ordet SPRINGER berättar något om subjektsdelen. Ordet SPRINGER är alltså ______________.	27
subjektsdel	FAGELN FLYGER. Vilket ord är subjektsdel? (Tänk efter vad meningen handlar om.) Svar: ______________.	17
subjektsdelen	En mening har två delar: s u b j e k t s d e l och p r e d i k a t s d e l. Vilken del säger vad meningen handlar om? Svar: ______________.	39

flyger	6 FAGELN FLYGER. Dessa två ord bildar en ____________________.
predikatsdel	28 BARNEN SKRATTAR. Meningen handlar om BARNEN (subjektsdel). Vilket ord i meningen ovan berättar något om subjektsdelen? Svar: ____________________.
fåglar	18 EVA SPRANG. Vilket ord är subjektsdel? (Tänk efter vem meningen handlar om.) Svar: ____________________.
subjektsdelen	40 En mening har två delar: s u b j e k t s d e l och p r e d i k a t s d e l. Vilken del berättar något om subjektsdelen? ______________________________.

mening	FÅGELN FLYGER. Den här meningen handlar om ______________.	7
skrattar	BARNEN SKRATTAR. Ordet SKRATTAR berättar något om BARNEN (subjektsdelen). Ordet SKRATTAR är alltså ______________ ______________.	29
Eva xxxxxxxxxxx xxxxxxxxxxxxxxx	1. POJKEN SPRANG FORT. 2. FLICKAN SPRANG FORTARE. <u>Vem</u> handlar mening nr 1 om? <u>Svar</u>: ______________. <u>Vem</u> handlar mening nr 2 om? <u>Svar</u>: ______________.	19
predikatsdelen	En mening har vanligen två delar. Vi kallar den ena delen <u>subjektsdel</u> och den andra delen ______________.	41

fågeln	Det ord, som säger vad meningen handlar om, kallas s u b j e k t s - d e l. PELLE SPRINGER. Ordet PELLE är subjektsdel, eftersom meningen handlar om ______________.
	30
predikatsdel	LAMPAN LYSER. Denna mening handlar om LAMPAN Meningen berättar att lampan LYSER. Ordet LYSER är ______________.
	20
1. Pojken 2. Flickan	POJKEN SPRINGER FORT. Vilket ord är subjektsdel i meningen? Svar: ______________.
	42
xxxxxxxxxxxxxxxx xxxxxxxxxxxx predikatsdel	En mening har vanligen två delar. De kallas ______________ och ______________.

227

Pelle	S u b j e k t s d e l är ett långt ord. Vi delar upp det i stavelser: SUB - JEKTS - DEL (tonvikt på JEKT). Se efter hur ordet stavas och skriv det här: ______________________.
predikatsdel	31 ISEN SMÄLTER. Denna mening handlar om <u>ISEN</u>. Meningen <u>berättar</u> att isen SMÄLTER. Ordet SMÄLTER är ______________.
pojken	21 FLICKAN SPRINGER FORTARE. Vilket ord är <u>subjektsdel</u> i meningen? <u>Svar</u>: ______________.
subjektsdel och predikatsdel	42a S L U T P Å H Ä F T E 1.

subjektsdel	Subjektsdelen säger vad meningen handlar om. FÅGELN FLYGER. Denna mening handlar om FÅGELN. Ordet FÅGELN är alltså _______________.
	32
predikatsdel	PELLE SPRINGER. Predikatsdelen berättar något om subjektsdelen. Vilket ord i meningen ovan är predikatsdel? Svar: _______________.
	21a
flickan	Fortsätt nedtill på första sidan i häftet.

subjektsdel	Subjektsdelen säger vad meningen handlar om. VATTNET FRYSER. Denna mening handlar om VATTNET. Ordet VATTNET är alltså ______________.	11
springer	HUSET BRINNER. Predikatsdelen berättar något om subjektsdelen. Vilket ord i meningen ovan är predikatsdel? Svar: ______________.	33

LINKÖPING STUDIES IN EDUCATION

Dissertations

[1.] HOLM, ROLF, Massed and Distributed Criterion Information. Studies of Verbal Learning. 1973. ISBN 91-7222-024-4.

2. TROPE, BARBRO, Individual Discovery and External Direction. Studies of the Effects of Different Methods of Instruction. 1974. ISBN 91-7222-078-3.

3. LINDGREN HOOKER, BERNICE, Educational Flow Models. With Applications to Arab Statistical Data. 1974. ISBN 91-7222-091-0.

4. HOLM, ULLA, Criterion Information and Principle Information. Studies of Verbal Learning. 1975. ISBN 91-7222-108-9.

5. GRUNDIN, HANS U., Response Requirement and Information about Correct Responses in Programmed Instruction. 1975. ISBN 91-7372-012-7.

ISBN 91-7372-012-7