Instructional Science 11 (1983) 357–372
Elsevier Scientific Publishing Company, Amsterdam – Printed in The Netherlands

HOW WE SHOULD CONTROL TIME ON TASK – OR SHOULD WE?

P. R. J. SIMONS
Department of Instructional Psychology, University of Tilburg, the Netherlands

ABSTRACT

Typical experimental studies of learning aids will control in various ways for the time spent on the learning task. This paper reviews such methods and compares the possible confounding influences of indirect time effects in experiments with broad or restrictive time limits. Indirect time effects refer to changes in time spent on a text that are caused by learning aids. They are to be discerned from direct time effects that refer to the time the reading of the learning aids take. In particular, we give an overview of the adequacy of various time limits for answering different research questions. It is argued that experiments without time limits are potentially more likely to bring out performance effects. In the final part of this paper we give some illustrations of the different effects and methods, based on a series of six experiments into the influence of concrete analogies on learning. Those experiments show that indirect time effects are important and that performance results depend on the way one controls reading time.

Introduction

In many experiments in instructional psychology, experimental groups that study learning material with extra aids (for example, advance organizers, objectives, questions, overviews, analogies) are compared with control groups that study the learning material without these aids. The present article discusses methodological problems with regard to time-on-task in experiments in which effects of extra study aids are evaluated.

1. Overview of Time-on-Task Control

Extra learning aids often require extra study time. For instance, reading an advance organizer may take several minutes. Therefore, performance effects will often be contaminated by time effects. Several solutions have been proposed in

0020-4277/83/0000-0000/$03.00

the literature to circumvent or prevent these contaminations. This section describes five such time-on-task control methods. Section 3.2 considers further possibilities, not described before.

The first method, proposed by Ausubel (1968), will be called the *neutralization method.* Here, the extra time the experimental subjects need for the extra learning aids is neutralized by letting the control subjects perform a neutral task during a certain length of time and then for another fixed length of time, studying the learning material (Fig. 1a). Typically, this neutral task consists of an historical or biographical introduction of the same word length as the learning aids, but there is no reason why other possibilities may not be thought of.

The second method will be called the *post-organizer method.* This method was invented by Mayer (1976) and is restricted to structuring aids. Experimentally, a structuring aid is studied before the material to be learnt, while a control group first learns the learning material and afterwards the organizer (Fig. 1b). Fixed amounts of time are allowed to study both the learning material and the aids.

The third method is known as the *efficiency method* (Peeck, 1970). Control subjects are allowed as much time to study the learning material by itself as is available to the experimental subjects for both the learning material *and* the extra learning aids (Fig. 1c). In this way it is possible to evaluate the efficiency of the learning aids.

Fourth, we have what might be called the *effectiveness method.* This does not control for the extra time which subjects are allowed to spend on the extra learning aids (Fig. 1d). With this method it is possible to evaluate whether extra learning aids are effective at all. The time problem is more or less ignored. The method has been criticized by Carver (1972) and Peeck (1970), but defended by Rothkopf (1974).

Finally, we have the *unlimited time method* (Fig. 1e), where no time restrictions are imposed on the learners, either for the extra learning aids, or for the learning material itself. According to Faw and Waller (1976), many researchers use this method without even estimating the study time spent. Consequently, performance effects and time effects are contaminated. When, however, learning times are being measured, efficiency scores may be calculated by dividing the performance of a subject by the time spent.

Within each of these five methods some variability exists. For instance, in methods 2 and 5 some investigators measured the times devoted to the extra aids, others did not. A further important difference concerns massed vs. distributed practice: some investigators presented the aids in the same session as the text, others presented aids and texts in different sessions (cf. Peeck, 1977). Section 3.2 considers further possibilities.

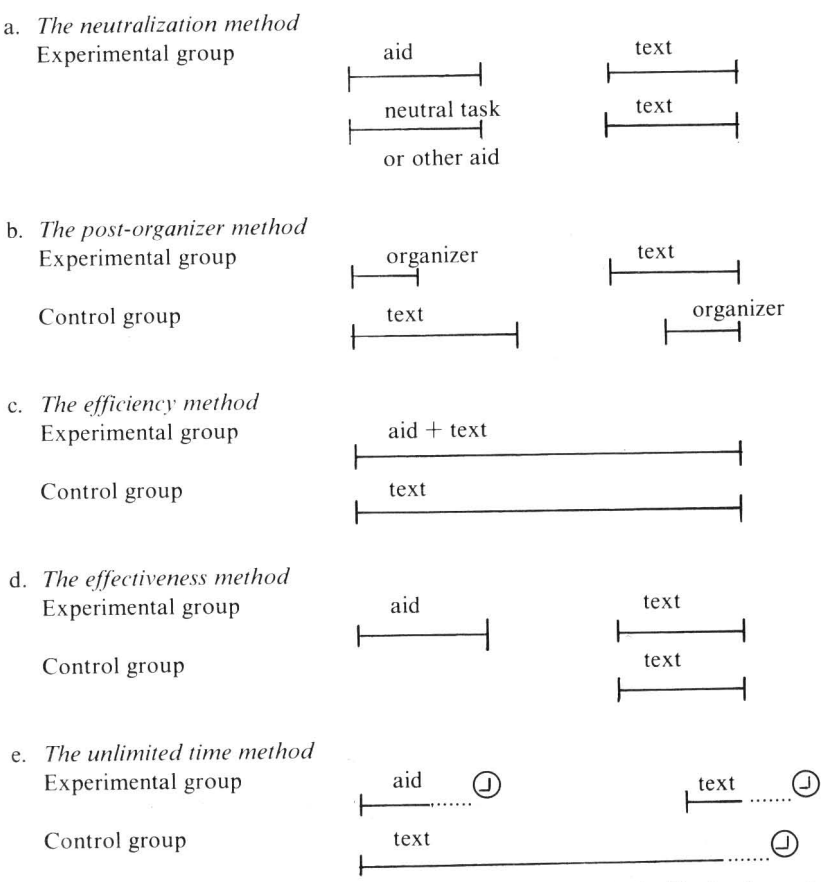

Fig. 1. Schematic representation of the five time control methods. The horizontal axes represent time-axes. The vertical dashes represent the beginning and the end of a time period or session. Dots refer to the absence of a time limit. Above the time-axes it is indicated what subjects should do in the time period or session. The little clocks indicate time measuring.

2. Some General Objections to the Different Methods

The *neutralization method* has had several critics. For example, Faw and Waller (1976) argued that it was strange to control study time through a neutral task: why should we compare performances of subjects performing potentially useful activities with performances of subjects doing useless things? They argued that any learning aid effect "found" in such a design means nothing more than that extra learning aids are better than bad, useless tasks. Though their point seems to us to be right, there are, however, circumstances where the use of the neutralization method is defensible: for example, to simulate classroom use ("normal ecological practice"), where specific kinds of introductions are com-

mon (as is for instance the case with historical or biographical introductions); and again, when different kinds of introductions or aids are compared with each other. Another objection to the neutralization method was put forward by Jaspers (1977) and by Mayer (1979 a,b). Although investigators can equalize the number of words in the learning aid and the neutral task, they cannot equalize the number of ideas to be learned. The extra learning aid concerns the ideas in the text, while in the neutral task novel ideas or novel concepts are presented. Control subjects, therefore have to learn – in the same amount of time – more ideas than the experimental subjects (because they do not know the neutral task information is unimportant). Thus any difference here between the performance of experimental and control groups may be artefactual, arising because control subjects learned a number of ideas that the experimental subjects did not learn (and which were considered unimportant by the investigator).

We did not find any objections to the *post-organizer method* in the literature (perhaps because it is still too new). But two limitations of this method are rarely made explicit. The first is that this method is restricted to structuring aids (advance organizers, overviews). The second point concerns generalizability. The post-organizer method may only lead to conclusions like the following: advance organizers work better when they are studied before a text than afterwards. Although this second limitation applies in general, there are circumstances where the method may be useful: when an investigator wants to show that a structuring aid fulfils certain functions or wants to find support for a theory. Mayer (1979 a,b), for instance, showed that organizers function as an assimilative context by showing that they only are effective when they are presented in advance of learning material and not when they are presented afterwards.

The third method, the *efficiency method*, is seen by some (e.g. Faw and Waller, 1976) as being a solution to the study time problem. However, this method also has its drawbacks, as shown in a study by Peeck (1977) on the effects of advance organizers. There, two control groups were employed. Subjects in one control group studied a neutral task; subjects in the other control group were allowed ten minutes extra for the text proper. No significantly different results were obtained for these two control groups (the average scores were even somewhat higher for the neutral task group). Subjects in the "ten minutes extra" control group reported that they stopped studying long before the final time signal was given. Obviously, the subjects did not use the extra time allowed, i.e., their *functional* study time was less than the *nominal* study time. This seems to pose serious problems for designs based on the efficiency method (see section 3.2).

The main objections to the *effectiveness method* concern the contamination of performance and time effects: it is impossible to determine whether differences in results between experimental and control groups are caused by experimental manipulation as such or by the extra time it requires (Peeck, 1970, 1977; Faw and

Waller, 1976). Carver (1972) argued that time is the most important explanatory variable in experimental research (on questions in texts). Rothkopf (1974) however, defended the effectiveness method by stating that, in education, only the total amount learned counts and that the time increase because of extra learning aids is so little that it can be ignored. Faw and Waller (1976) reacted as follows: "... nevertheless in the long haul techniques which increase learning simply at the expense of additional time are more of a disservice than an aid to students, since the totality of useful (or interesting) knowledge to be acquired is so extensive and the time available to accomplish the task is so limited" (p. 712).

Many investigators who used the *unlimited time method* failed to measure (or report) study times (Faw and Waller, 1976). Then, a confusion of performance and time occurs.

When learning times are measured, an investigator is confronted with the problem of how to analyze his data. Should the investigator divide performances by learning times (efficiency measures, see Faw and Waller (1976)), or control statistically for time differences (for instance by way of analysis of covariance or regression analysis), or analyze the two dependent variables separately? None of the three methods of analysis is without difficulties and often interpretation problems will result. The main objection to the unlimited time method was first noted in a different context by Underwood (1964): when subjects are allowed to study a limited amount of material during an unlimited time, one would expect near perfect results. Such was always the case in early laboratory studies on paired associate learning. Learning time should be the sole dependent variable in these cases. Underwood objected to the inclination of text investigators to use the unlimited time method with performances instead of time as dependent variable.

3. Some New Objections to Setting Time Limits

The literature on time control (e.g., Faw and Waller, 1976) stresses direct time effects (these are effects related to the extra time extra aids require). In this section we discuss some indirect time effects. *Indirect time effects* refer to changes in time spent *on the text proper* caused by learning aids. For instance, subjects might spend more time on a text after studying an advance organizer, because they try to relate the information in the text to the advance organizer, or subjects might spend less time on a text after studying a concrete analogy because they sooner understand the text proper. Here, the time to read the additional materials counts as a direct effect; this time may change with how the materials are used, illustrating the potential difficulty of separating direct and indirect effects.

One might object that any indirect time effects are controlled in most

experiments because experimental and control groups are allotted equal amounts of time. In this section we show that this objection is wrong.

3.1. BROAD TIME LIMITS

First, it is necessary to distinguish nominal and effective study time (cf. Peeck, 1977). Nominal time refers to the time allotted by the investigator, while effective or functional time refers to the time actually spent by a subject. Because, in nearly all experiments in which extra learning aids are investigated, rather broad time limits are imposed, *effective times will be shorter than the nominal time for most of the subjects.* This was unwittingly shown in an unpublished experiment we carried out in 1978. In this experiment eighth-grade pupils were to learn a huge amount of information on electricity. The subjects first had to do a pre-test (consisting of post-test items) so that they would know what kind of performance was expected. Then three groups had to study these items of information about electricity in three different ways. Afterwards they did several post-tests. The subjects in the three conditions were allowed seventy minutes to study. Unexpectedly, hardly any subjects used up all of this time. Most subjects stopped after about forty-five minutes and claimed (when asked) that they had mastered the subject matter. Their results on the post-tests, however, were far from good. Why would subjects stop studying before the nominal time is over? Peeck (1977) provided a partial answer: ". . . subjects will spend a certain period of time on the text until some satiation point or some subjective level of familiarity is reached; this level probably corresponds to the anticipated degree of mastery required . . ." (p. 206). Thus, Peeck attributes the stop-behavior of his subjects to an internal mastery criterion. Other reasons why subjects may stop prematurely are in our view: a) because they think they have studied long enough, (i.e., an internal time criterion); b) because they think they saw the text often enough, (i.e., an internal rehearsal criterion); c) because they are tired, no longer interested, or want to give up.

Secondly, differences between nominal and effective learning times may differ under diferent experimental conditions. In other words, when nominal and effective time differ, indirect time effects may occur. This was also shown in our above-mentioned experiment. Subjects in one condition spent significantly more time than subjects in the other two conditions, (in spite of the equal time limits). Differences in time-spending caused by independent variables have been shown by others: Wolters (1975), for instance, showed that subjects spent more time on a text when they got pre-questions than when no questions were available, and Rickards and August (1975) and Rickards and Dener (1979) showed likewise that some forms of underlining used by subjects lead to longer effective learning times than others.

Why would these differences in time-spending occur? The answer to this

question is not easy to give. All we can do is suggest some possibilities: a) when subjects study until they reach a personal mastery criterion, some independent variables may influence the time needed to satisfy this criterion, for instance, a text may be understood sooner because of a concrete analogy or because it has a list of objectives which shows readers where to look; b) when subjects study until they satisfy a personal "rehearsal" type criterion, such as reading a text once or twice, some independent variables will influence the time needed for one "rehearsal," for instance, because they need to perform extra activities such as underlining or comparing the text organizer; c) when subjects study until they satisfy a personal time criterion, some independent variables may influence the time criterion, for instance, subjects may be willing to spend more time on two different activities, such as organizer plus text, than on only one activity, such as text only (cf. Peeck, 1977).

Although it is as yet impossible to state which personal criteria subjects use, two things seem clear: a) independent variables may shorten or lengthen the time needed to satisfy personal criteria; and b) independent variables may weaken or strengthen subjects' personal criteria. Because of such effects, effective learning times in different experimental conditions may differ even though the nominal times are equal.

The indirect time effects pose serious methodological problems for designs in which broad time limits are imposed on the learners. When significant performance differences are found in such designs, these may be caused by a lengthening of effective learning times. When, for instance, questions cause subjects to use a greater part of the nominal study time, performance effects obtained are contaminated by these indirect time effects.

But also the absence of performance differences in broad time limit designs may be caused by indirect time effects. When learning aids cause experimental subjects to use a smaller part of the nominal study time than control subjects do, learning aids may fail to bring about significant performance effects. When, for instance, objectives make it easier to learn from a text, subjects may satisfy their personal mastery criteria sooner than when they do not have these objectives and may stop sooner, so performance differences between an experimental and a control group may fail to appear.

3.2. RESTRICTIVE TIME CONDITIONS

What would happen if we could set restricted time limits so that the effective time equalled the nominal time? Other things being equal, indirect time effects could not then occur. But other problems emerge. One problem is that the character of the studying could be affected in each condition. Further, some subjects might not read or learn all of the learning material. Another problem concerns the choice of a time limit. In advance, one is never certain that a limit

question is not easy to give. All we can do is suggest some possibilities: a) when subjects study until they reach a personal mastery criterion, some independent variables may influence the time needed to satisfy this criterion, for instance, a text may be understood sooner because of a concrete analogy or because it has a list of objectives which shows readers where to look; b) when subjects study until they satisfy a personal "rehearsal" type criterion, such as reading a text once or twice, some independent variables will influence the time needed for one "rehearsal," for instance, because they need to perform extra activities such as underlining or comparing the text organizer; c) when subjects study until they satisfy a personal time criterion, some independent variables may influence the time criterion, for instance, subjects may be willing to spend more time on two different activities, such as organizer plus text, than on only one activity, such as text only (cf. Peeck, 1977).

Although it is as yet impossible to state which personal criteria subjects use, two things seem clear: a) independent variables may shorten or lengthen the time needed to satisfy personal criteria; and b) independent variables may weaken or strengthen subjects' personal criteria. Because of such effects, effective learning times in different experimental conditions may differ even though the nominal times are equal.

The indirect time effects pose serious methodological problems for designs in which broad time limits are imposed on the learners. When significant performance differences are found in such designs, these may be caused by a lengthening of effective learning times. When, for instance, questions cause subjects to use a greater part of the nominal study time, performance effects obtained are contaminated by these indirect time effects.

But also the absence of performance differences in broad time limit designs may be caused by indirect time effects. When learning aids cause experimental subjects to use a smaller part of the nominal study time than control subjects do, learning aids may fail to bring about significant performance effects. When, for instance, objectives make it easier to learn from a text, subjects may satisfy their personal mastery criteria sooner than when they do not have these objectives and may stop sooner, so performance differences between an experimental and a control group may fail to appear.

3.2. RESTRICTIVE TIME CONDITIONS

What would happen if we could set restricted time limits so that the effective time equalled the nominal time? Other things being equal, indirect time effects could not then occur. But other problems emerge. One problem is that the character of the studying could be affected in each condition. Further, some subjects might not read or learn all of the learning material. Another problem concerns the choice of a time limit. In advance, one is never certain that a limit

will be such that no subjects stop learning before the stop signal. Only by means of a pilot study can we estimate suitable time limits. The most important problem however, concerns generalizability to other time limits. In many cases time limits are arbitrarily chosen. But results obtained under a certain time restriction are not generalizable to other time restrictions. With shorter limits, performance differences may disappear. Under less restrictive time conditions, performance differences may turn up that were absent under more restrictive conditions. Because the choice of a certain time limit may thus determine the kinds of results one may get, this choice should be well founded and considered. Time limits, for instance, may be based upon knowledge of learning times spent spontaneously by subjects learning a text only or upon time allowed for the subject matter (in classroom practice).

The last problem is related to the previous one. Sometimes extra aids may only be effective if extra time is available to make a different kind of processing possible. Any improvement in performance as a result of this processing may be greater than the extra time costs. In this case learning aids might be called efficient. This kind of efficiency cannot be detected with a design with short time limits.

3.3. NO TIME LIMITS

Because both under broad and restrictive time limits all kinds of interpretation problems arise, time control by imposing time limits upon learners often seems to be a kind of "would-be control." There is, in my view, hardly any case for the time limits approach and a strong case against it. The main arguments against it were discussed in the last two sections: its ecological validity (i.e., the imposition of time requirements makes it impossible to generalize to those natural settings where time limits are unlikely (but see Malone et al., 1979)); significant performance effects may be confounded with unnoticed time lengthening; insignificant performance differences may be caused by unnoticed time shortening; there is no basis on which the choice of a particular time limit might be based and yet results may depend on the limit chosen; sometimes learning aids are only effective when there is extra time available. Therefore, in my view, under many circumstances *no time limits should be imposed upon learners.* When control of variables is not possible, as is the case with the time variable, these variables should be measured. As we saw in section two, one problem with any design with broad or no time limits concerns the stop criteria which subjects use. Apparently subjects do not continue to study until they have perfect mastery of the subject matter, since sufficient variance remains. But why do they stop prematurely? Do they stop under different conditions because of the same criteria? Because these uncertainties may cause error, it is, in my view, necessary to control the stop criteria of subjects explicitly by way of instructions to them.

Five different instructions on stop criteria seem conceivable: 1) a mastery criterion, 2) a "rehearsal" criterion, 3) an ecological time criterion, 4) an ecological rehearsal criterion and 5) an ecological mastery criterion. Examples of these instructions are presented below

1. *Mastery criterion*: Read this text until you know it so well that you will be able to answer 80% of the questions correctly.
2. *Rehearsal criterion*: Read this text just once without looking back.
3. *Ecological time criterion*: Read this text for the same time period as you would spend on such a text when doing your homework.
4. *Ecological rehearsal criterion*: Read this text as often as you would if you were doing your homework.
5. *Ecological mastery criterion*: Read this text until you know it as well as you normally learn comparable texts.

4. Different Questions, Different Methods

I have argued that, under many circumstances, no time limits should be imposed. But the choice of a study time control method (or measurement method) should depend on the kind of research question posed. Some methods are especially suited to answer specific questions and are not suitable for other questions. Moreover, some of the above-mentioned problems only occur when one wants answers to certain questions. The following six questions typify those normally of interest:

a) Are the learning aids effective? Do the aids lead to better performances?
b) Do the learning aids function as was theoretically assumed in advance?
c) What influence do learning aids have on reading times?
d) Do learning aids lead to better performances when a certain nominal time is given?
e) Do the learning aids make for efficient learning: Is there, in addition to a quantitative effect (because of differences in time spent), also a qualitative effect (because of differences in learning processes)?
f) Is one learning aid more effective than another?

To four of these questions a, c, e, and f, the unlimited time method provides the best approach, because of the arguments put forward in section 3: a) in terms of effectiveness; c) for the influence on reading time; e) for efficiency and qualitative effects; and f) for comparison of different aids. Though the unlimited time method seems to be the best approach to these four questions, some other methods might be used too.

For example, question (a) might also be answered by means of the effectiveness method. There are here two methodological difficulties to take into

account in the design and analysis. Firstly, an indirect lengthening of reading time may cause possible performance differences. This possibility could be checked by measuring the effective time spent on the task. Secondly, when no performance differences are significant, this may be so because an indirect lengthening of reading time might form a necessary condition for effectiveness to show itself, or because an indirect shortening effect prohibited performance differences from occurring. In those cases, no significant differences appear while the reading aid is an effective one.

Question (e) might also be answered by means of the efficiency method. In fact, Faw and Waller (1976) advised this method to shed light on this question. But if one uses the efficiency method, one should realize that several interpretation problems may occur. The most important one is that results depend on the time limit chosen. Efficient reading aids may not bring about significant differences in the efficiency method because the particular time limit was too short. A solution to this problem is to be found in a design with two or more time limits. In that case one may see how the time limits influence results. Another problem occurs when one uses time limits of an intermediate value. Efficiency indexes as proposed by Faw and Waller (1976) are in that case unreliable. For some of the students the time spent is restricted by the time limit, for others it is not.

We see no alternative to the unlimited time method in the case of question (c). How could one measure reading times reliably when there is a time limit?

As to question (f) the neutralization method and the efficiency method form acceptable alternatives to the unlimited time method. One should, however, take into account the possible confounding effects discussed above in relation to question (a).

So far, four questions have been discussed. For all of them the unlimited time method has been considered the best method, but some other methods could be used, if appropriate measures are taken. For the remaining two questions, (b) and (d), the unlimited time method is not the best choice. For question (b), the post-organizer method seems to be the better choice, because only with this method can we answer the question of whether or not organizers function as assimilatory contexts in which new information is absorbed. Perhaps, however, in this method, too, time limits should be abandoned because of the problems discussed in section 3.

For question (d) only the efficiency method is suitable, since in the question itself a time limit is assumed.

These points are summarized in Table I, which also indicates why certain methods cannot help answer questions. These reasons have not been discussed in this section, but they have been implicit in previous sections.

TABLE I.

Different questions, different methods

Question	Method				
	Neutralization	Post-organizer	Efficiency	Effectiveness	Unlimited time
a. Effectiveness	1	1	6		
b. Theoretical functions	1		6	8	
c. Reading times	5	5	5	5	
d. Efficiency, given a certain time	1	1		3	7
e. Efficiency, qualitative effects	1	1		4	
f. Comparison of different aids		1		2	

1. With this method one may only conclude that a good reading aid is better than a bad one or that an advance organizer does not work when you present it afterwards. (See section 2).
2. With this question this method becomes the same as the neutralization method.
3. The (nominal) study times are not the same whereas the question assumes equal times.
4. Direct time effects are not controlled for. Time effects can not be measured because of the time control.
5. Time is controlled for, thus it cannot be measured.
6. Direct and indirect lengthening of reading time is not possible because of time control, while one is – with this question – only interested in performance effects.
7. The question assumes a specific amount of time, thus time control is necessary (see this section).
8. With this question this method becomes the same as the post-organizer method.

5. Some Experimental Illustrations

In this section, direct and indirect reading time effects and their operation in different control methods will be illustrated. The illustrations stem from a series of six experiments on the influence of the presentation of concrete analogies on learning performance and reading times (Simons, 1980, 1982). Details are given here only of the first four of these experiments, since they are the most relevant to this paper.

The first three experiments were in many ways comparable in that the same subject matter and the same kinds of subjects were employed. In two of these, the unlimited time method was used, in the third one the efficiency method. Thus the results of these experiments may be compared and conclusions may be formulated as to the different reading time control methods. The other three experiments were less comparable, in that not only the reading time method but also the

subject matter and the kinds of subjects differed. In what follows, only those results will be discussed that are relevant to the present discussion on reading time control and measurement. In each of the experiments 80 to 100 subjects participated. Subjects were from secondary school (3 experiments), elementary schools (1 experiment) and college students. All experiments consisted of 3 or 4 sessions of 2 hours each. During the first session a pre-test was given consisting of items from the post-test. Furthermore, subjects were given other tests. In the next session(s) subjects read texts of approximately 20 pages on various topics, being assigned randomly to one of two or three conditions. In one (or two) condition(s) texts were extended with common concrete analogies (for instance electricity concepts were compared to water flow concepts). The other condition was a "text only" control condition. In three of the experiments the unlimited time method was used (with a rehearsal criterion, see section 3.3). Subjects were instructed to read the text once (registering their reading time), and afterwards to attempt a comprehension test. In two experiments, they were told to read the text a second time, again registering their reading time. In one experiment the different readings were spread over three sessions, in the other one the readings took place within one session. In the fourth experiment the efficiency method was used and in the fifth a variant of the neutralization method. In the final experiment a variant of the effectiveness method was used. In that experiment, lessons were presented orally to elementary school children, so, instead of reading times, instruction times were registered. In all experiments a final session was devoted to a long-term retention test.

In the first experiment (unlimited time method, secondary-school children learning science), both significant performance effects and significant direct and indirect time effects were evident. Thus subjects learned more because of the concrete analogies but this learning took more time, both because of a direct effect (reading of the extra material) and because of indirect effects. These indirect effects were complex: subjects required significantly more time to read the text as such (apart from the direct effect) for the first time in the experimental condition, ($F(1,59) = 33.06$; $p < 0.1$) but during the second and third reading this indirect lengthening effect was compensated for somewhat (non-significant) by an indirect shortening (see Fig. 2). When indirect reading time effects were, by means of regression analysis, partialled out of performance variance, a long-term retention effect remained significant ($F(1,52) = 8.16$; $p < 0.01$), whereas three other main effects were no longer significant.

The second experiment was a replication of the first, but now subjects learned the texts only twice instead of three times. Again significant performance and direct time effects were found (see Fig. 3) [1]. As in the first experiment, subjects required more time for the first reading; this difference, however, was not significant ($F(1,81) = 2.53$; $p < 0.10$, (one-tailed)). As in the first experiment, subjects also required less time for their second reading but now this difference

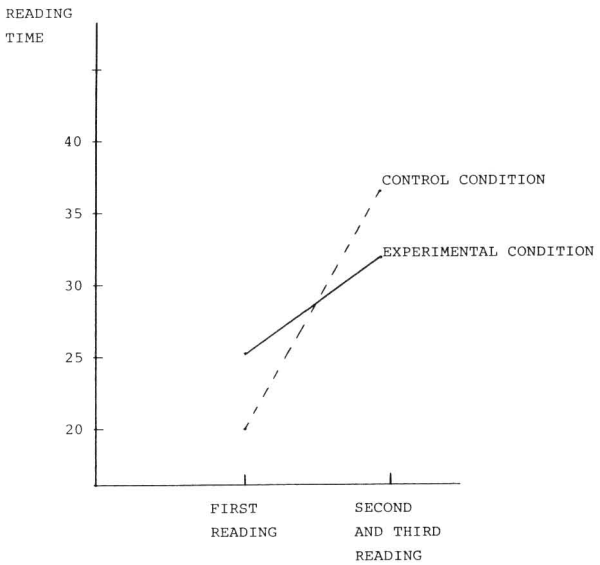

Fig. 2. Average reading times in the first experiment.

was significant ($F(1,81) = 13.51$; $p < 0.01$; (two-tailed)). Partialisation of indirect time variance out of performance variance did not alter any of the performance effects.

Thus far, only partialisations of indirect time effects out of performance variance were discussed; partialisations of direct effects were not considered.

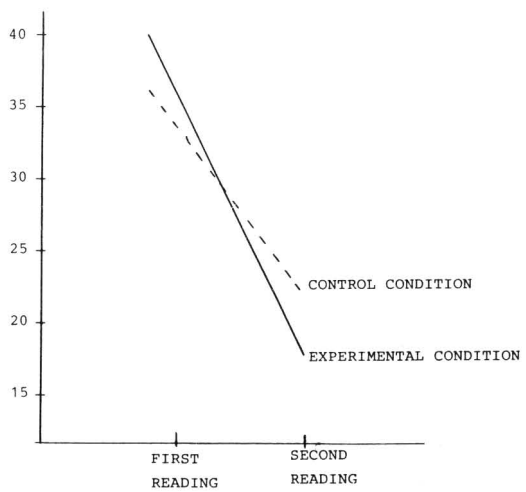

Fig. 3. Average reading times in the second experiment.

Partialisation of direct time variance, however, was not possible due to the fact that all subjects had the same time score for the analogies since these reading times were not measured separately. We are not able to see what performance effects remain when both direct and indirect time effects are partialled out of performance variance. Instead, one may look at these data in another (less satisfactory) way. Increase in time spent due to the analogies (direct + indirect) may be compared with the increase in performance. In the first experiment 40% more time was used to achieve 20% better results. In the second experiment, however, 13.5% more time brought 13.7%, 1.2%, 17.5% and 9.3% better results on an immediate post-test, a retention test, a transfer test, and a relation test respectively. Thus, in the first experiment the relation between extra time investment and improvement in results was less favourable for the analogy condition than in the second experiment.

In the third experiment the same kinds of materials and subjects were employed as in the first two, the reading time control, however, being different (efficiency method). Subjects learning with and without analogies were allotted the same amounts of time. The results of this experiment indicated no differences between the two conditions on an immediate post-test ($F(1,67) = 0.89$; n.s.) nor on a retention test ($F(1,67) = 0.00$; n.s.). Another unexpected but important result of this experiment was that subjects did not use all of the time allowed. Experimental subjects, on the average used 86.4% of the time and control subjects only 62.5% ($F(1,66) = 59.9$; $p < 0.01$).

Thus, there was an indirect lengthening of reading time due to the analogies. In spite of this, no performance differences appeared.

How is it to be explained that these differences failed to appear in this experiment while they did appear in the previous ones? The most likely explanation seems to be that, for about half of the experimental subjects, the learning process was stopped because of the time limit. If these experimental subjects could have studied as long as the subjects did in the previous experiments, significant performance differences could have been expected.

An overall conclusion from these three experiments is that concrete analogies of the type we used are effective reading aids for secondary-school children. This effectiveness, however, only showed up in our data when two conditions were met:

a) the subjects must have time to read the analogies, and

b) the subjects must have time to compare the concepts to be learned with the more concrete concepts of the reading aid.

Furthermore learning with analogies does not seem to be more efficient than learning without them, but not less efficient either. We found some indications that concrete analogies may lead to qualitatively different results when time conditions permit them to show up (transfer test, relation test).

In the fourth experiment the unlimited time method was again used.

College students ($N = 110$) read a text on computer programming just once. Two experimental and one control conditions were employed. In one experimental condition the analogies were again those in common use in this area, but were taken from different domains (e.g. games, car driving, typewriting etc.): in the other one all analogies were taken from the same domain (a post office). In the control conditions no analogies were presented. The analogies covered about 400 words, the text about 1,600 words.

The results on time-spending indicated that subjects in the experimental conditions spent somewhat less time on reading the text plus the analogies than the control subjects did on the text alone (28.3 and 28.3 minutes versus 29.7 minutes); this difference, however, was not significant. Yet, this should be considered a remarkable result; experimental subjects read about 400 words more than control subjects in about the same time. Obviously, direct and/or indirect lengthening effects were compensated for by an indirect shortening effect.

6. Conclusions

In the first sections of this article – within a theoretical framework – several time effects were described that might confound results of experiments on effects of extra reading aids. Several time control methods were described and problems of these methods were reviewed. Furthermore, new problems caused by indirect time effects were discussed. The conclusion was drawn that it is better not to control reading time but instead to use the unlimited time method. Exceptions were made for two specific research questions. Also, measures were discussed that could be taken if one should want to impose time limits. The last section presented results of experiments that illustrate several of the theoretical points made before. Thus, indirect effects were shown to operate: both lengthening and shortening occurred. Furthermore, indirect effects influenced the kinds of results one may find in an experiment. Thus in the third experiment no performance differences showed up because the chosen time limit was too short to permit direct and indirect lengthening of reading time (shown to be necessary in the first experiments). Likewise in the fourth experiment an indirect shortening of reading time probably decreased differences in performance.

Notes

1 Lines are ascending in Fig. 1 and descending in Fig. 2, because of the difference in number of rehearsals (three times in Fig. 2 and twice in Fig. 3).

References

Ausubel, D. P. (1968). *Educational Psychology: A Cognitive View*. New York: Holt, Rinehart & Winston.

Carver, R. P. (1972). "A critical review of mathemagenic behaviours and the effect of questions upon the retention of prose materials," *Journal of Reading Behaviour* 4: 93–119.

Faw, H. W. and Waller, T. G. (1976). "Mathemagenic behaviours and efficiency in learning from prose," *Review of Educational Research* 46: 691–720.

Jaspers, J. P. (1977). "Learning of an instruction text without organizers, with advance organizers and with post-organizers" (in Dutch). Unpublished report, University of Amsterdam.

Malone, T. W., Macken, E. and Suppes, P. (1979). "Toward optional allocation of instructional resources: dividing computer-assisted instruction time among students," *Instructional Science* 8 (2): 107–120.

Mayer, R. E. (1976). "Some conditions of meaningful learning of computer programming: Advance organizers and subject control of frame sequencing," *Journal of Educational Psychology* 68: 143–150.

Mayer, R. E. (1979a). "Can advance organizers influence meaningful learning?" *Review of Educational Research* 49: 371–383.

Mayer, R. E. (1979b). "Twenty years of research on advance organizers: assimilation theory is still the best predictor of results," *Instructional Science* 8 (2): 133–167.

Peeck, J. (1970). "Effects of prequestions on delayed retention of prose material," *Journal of Educational Psychology* 61: 241–246.

Peeck, J. (1977). "Preinstructional strategies and extra reading time in learning from text," *Tijdschrift voor Onderwijsresearch* 2: 202–207.

Rickards, J. P. and August, G. J. (1975). "Generative underlining strategies in prose recall," *Journal of Educational Psychology* 67: 860–865.

Rickards, J. P. and Dener, P. R. (1979). "Depressive effects of underlining and adjunct questions on children's recall of text," *Instructional Science* 8 (1): 81–90.

Rothkopf, E. Z. (1974). "Barbarism and mathemagenic activities: Comments on criticism by Carver," *Journal of Reading Behaviour* 6: 3–8.

Simons, P. R. J. (1980). "By way of comparison: research into the influence of metaphors on learning" (in Dutch). Doctoral Dissertation, Tilburg University.

Simons, P. R. J. (1982). "Concrete analogies as aids in learning from text," in: A. Flammer and W. Kintsch (eds.), *Discourse Processing*. Amsterdam: North-Holland Publishing Company, (in press.).

Underwood, B. J. (1964). "Degree of learning and the measurement of forgetting," *Journal of Verbal Learning and Verbal Behaviour* 3: 112–129.

Wolters, L. A. M. M. (1975). "Questions about questions" (in Dutch), Unpublished Report, University of Nijmegen.