THE EFFECTIVENESS OF REMEDIAL COMPUTER USE FOR MATHEMATICS IN A UNIVERSITY SETTING (BOTSWANA)

TJEERD PLOMP, JACQUELINE PILON and INGEBORG JANSSSEN REINEN
Department of Education, University of Twente, P.O. Box 217, 7500 A. E. Enschede, The Netherlands

(Received 26 July 1990; revision received 26 November 1990)

Abstract—This paper describes the evaluation of the effects of the Mathematics and Science Computer Assisted Remedial Teaching (MASCART) software on students from the Pre-Entry Science Course at the University of Botswana. A general significant improvement of basic algebra knowledge and skills could be measured, demonstrating that the combination of increased time on task and computer-assisted remedial instruction is effective. This result is relevant for this setting of a university in a less developed country, because other ways of remedial teaching, for example increased staff time, are not possible. By applying Markov chain analysis it could be proved that the improvement in the experimental group was caused by the increased time on task, and that in a situation where students are expected to attain high mastery levels, especially high ability students profit from the MASCART approach. Students like to work with MASCART because they could work individually and at their own pace.

INTRODUCTION

This paper describes the evaluation of the effects of the Mathematics and Science Computer Assisted Remedial Teaching (MASCART) software on students from the Pre-Entry Science Course (PESC) at the University of Botswana. The evaluation of the implementation of the MASCART project, which was the first application of computer assisted instruction in this university, focused on the following questions:

—Considering the conditions for successful implementation as mentioned in the literature (e.g.[1]), has the implementation of computers in the PESC Department been such that we may expect their institutional acceptance?

—Did the use of the MASCART software by students indeed result in cognitive gains (knowledge and attitudes)? And if so, can this increase be attributed to the use of the remedial software or are other factors playing a role?

The results for the first question concerning implementation and acceptance are described in [2]; here we will only deal with the effect question.

Botswana has an educational system that is still being defined. Secondary school (a typical product of the leading British role in southern Africa) lasts five years, and school leavers who wish to enter the Faculty of Science of the University of Botswana first have to go through PESC. While still in secondary school, students are selected for entering PESC by a PESC-lecturer team. This selection is based on students' performance in the Cambridge Overseas School Certificate (COSC) exams or on PESC's own selection tests, and on a recommendation from the secondary school. PESC is a remedial course, bridging the gap between secondary school and university. The course lasts for about 7 months, covering topics from mathematics, chemistry, physics, biology and language and study skills. The training in basic concepts, cognitive and practical (laboratory) skills, and methods in these subjects is considered to be more important than acquisition of pure facts. The number of students participating in the course has increased every year since the start of the course in 1975. In January 1989, about 300 students entered the PESC course.

A report on the use of computers in the University of Botswana in 1985 was the starting point for considering the introduction of computers in the PESC Department; assistance was provided by the Basic Science Program Unit (BSPU) of the Free University Amsterdam (FUA, The Netherlands). A computer laboratory, equipped with 30 microcomputers, has been available from 1989 onwards and a compulsory computer awareness course was introduced in the PESC
curriculum. The facility of a computer lab also enabled the introduction of computer assisted instruction (CAI) in mathematics and science, especially for remedial purposes. The provision of remedial mathematics and science teaching through the computer is an advantage in the PESC setting, because other methods of remediation (e.g. by the lecturer or by peer tutoring) are not possible; this would require extra manpower, which is not available.

Students entering the PESC course come from different secondary schools, which causes great variation in knowledge level among students. Although all students have passed their ordinary ('O') level exams, many do not really master the basic science and mathematical concepts. Though PESC is remedial in itself, and although the same concepts are being taught in PESC as were already taught in secondary school, with small extensions, many students still have difficulties with the subject matter. Extra individual remedial teaching via computers was expected to be a very helpful tool in this situation. Smit and Wolff[3] reported the learning difficulties of PESC students in mathematics, especially in basic algebra, caused by students' misconceptions. Traditional teaching in PESC did improve the understanding of fundamental concepts in certain areas but the formation of other concepts was still unsatisfactory. The suggestion made by Smit and Wolff, that additional instruction with the computer may change this, resulted in the development of experimental software for basic algebra: the MASCART-basic algebra project. This project also serves as a pilot, to examine the possibilities of this type of computer application for PESC-Botswana, as well as for other, similar settings within less developed countries. If this pilot should prove to be successful, the project could be extended to other subjects within PESC (biology, physics, chemistry) and/or to other countries. The main aim of the MASCART-basic algebra materials is providing remedial instruction aimed at improving the algebra knowledge and skills of students by overcoming misconceptions, which students might bring in from secondary school or from the regular PESC lessons. The MASCART software is meant to help students in remediating particular topics within basic algebra, by providing diagnostic exercises and specific feedback. Students can follow the MASCART course individually, without the help of the lecturer. The MASCART materials are complementary to the usual teaching program for mathematics, which consists of eight periods of 45 min (lectures, practicals and tutorials). In this way the MASCART software adds two periods (2 × 45 min) per week extra to the teaching program.

Although research on misconceptions is the basis for the development of the MASCART materials, this particular aspect will not be discussed here. Detailed analysis of the results concerning misconceptions was done by Smit, Oosterhout and Wolff and their reports discuss changes in misconceptions since the introduction of MASCART[4,5].

Operational research questions

There are several meta-analyses of effects of educational computer use[6–8]. Although they all conclude the existence of a positive effect of computer use on learning achievement, they sometimes show controversial results. Some relevant outcomes are:

— the positive effect on learning achievement occurs regardless of the type of CAI used, or the age range of students (here Niemiec and Walberg[8] indicated that especially younger and special populations profit from this technology);
— tutorial and drill modes seem to be more effective for low ability students than for middle or high ability students (here Roblyer et al.[7] conclude that there is no significant difference for different ability levels (pp. 79, 122);
— the use of CAI improves student attitudes towards computer use in the learning situation, while Roblyer et al.[7] also find an improvement of the attitudes of students towards the school subject.

These results are based on research situated in more developed countries. This study investigates whether the same type of results also have validity in a less developed country.

Given these results, the general research question of this study can be operationalized into more concrete evaluation questions:

1. Are there general gains in basic algebra knowledge for PESC students using MASCART, when compared to students who do not? Stated in other words: did the materials indeed prove their effect?
2. For which ability group of students (low, middle or high) is MASCART more effective?
3. What are student attitudes towards computers and do these attitudes change due to the use of computers in both the computer awareness course and MASCART?

In the next section, the design and method for the study will be presented, while the last section deals with the results.

**Research Design**

In February 1989, after the first teaching unit of PESC, the total group of 260 PESC students (40 students left during the course) were pretested. A 40-item multiple choice test on basic algebra was used, having a reliability of $\alpha = 0.80$ (Cronbach’s alpha coefficient for all items after Aiken[9]). A normal PESC teaching unit consists of about four weeks of lectures, practicals, tutorials, assignments and weekly achievement tests. Based on the results of this pretest, the students were matched into equivalent experimental and control groups. The matching took place within the different PESC groups, to keep other conditions (like kind of lectures, practicals and tutorials) as identical as possible. During unit 2 of PESC, the experimental group received extra remedial teaching, using the MASCART materials that cover unit 1. This meant spending an extra two periods (2 x 45 min) per week on basic algebra. The control group only received the regular teaching of unit 2, consisting of 8 periods of lectures, practicals and tutorials. A tutorial allows time for students to ask questions about topics discussed that week. In a practical, a student performs, if necessary with assistance of a teacher, exercises about the materials discussed in the lecture. After unit 2, the same basic algebra test (using a different sequence of questions) was administered as posttest ($\alpha = 0.80$). Afterwards, the second group received remedial teaching using MASCART. It is possible to measure the overall effect of cognitive knowledge gains by comparing the test results of both the experimental group and the control group on pretest and posttest.

Concerning the differential knowledge gains, various expectations exist:

- findings of e.g. Hasselbring[6] indicate more effect of CAI for low ability students than for middle or high ability students;
- research reviews by Roblyer et al.[7] indicate that there is no significant difference in effects for different ability groups;
- PESC lecturers, based on their teaching experience, expect that the middle ability group would gain most from MASCART.

To test these expectations, the total group of students was divided in several ways into different ability groups.

To examine the attitudes of students towards computers, two attitude questionnaires have been administered: one before students started to use computers and one after both groups finished the first MASCART course. The specific opinions of students towards the MASCART software have been examined by means of a general questionnaire.

**Instruments**

The Basic Algebra Test (BAT) used in this study is not representative of the complete basic algebra content, but was developed to cover those parts with which students have learning difficulties (according to the experience of the lecturers). Pre-piloting of the test took place in June 1988. Cronbach’s $\alpha$ proved to be 0.74, the Spearman–Brown coefficient was 0.72, which is acceptable. For the development of the attitude questionnaires, that of Sandman and Welch[10] was used. Some of the items were slightly changed to make them more suitable for the specific context, and some items were added. Cronbach’s $\alpha$ of the first attitude questionnaire is 0.73 ($N = 296$). The second questionnaire has been slightly changed: all hypothetical statements of the first questionnaire (e.g. “Learning to work with computers seems fun to me”) have been reformulated into questions (“Was it fun learning about computers?”). The $\alpha$ coefficient of the second questionnaire is 0.76 ($N = 234$). Both questionnaires consist of different subscales: motivation, enjoyment, anxiety and societal concern.

The general MASCART questionnaire was mainly the same as the one used by Smit and Wolff[3] during the pre-pilot. The students were asked to give their opinion about the quality of the feedback
and the help given by the MASCART program, the style of the program, the difficulty of the language used in the program, etc. Students could also give some qualitative remarks about the materials. The MASCART questionnaire was administered together with the second attitude questionnaire, after the MASCART sessions had finished. All instruments used in this study are described in more detail in [11].

**GENERAL RESULTS**

In this section, first the effects on student attitudes will be discussed followed by the general knowledge gains of the students. Then these knowledge gains will be examined in more detail for the three different ability groups. Also corrections for the increased time on task will be made, which enables us to make definite statements concerning the effects of the materials.

**Effects on student attitudes towards computers**

In order to investigate whether student attitudes change after using the computer for some time, the null hypothesis that no attitude changes occur is tested two-tailed with $\alpha = 0.05$. The mean attitude scores of all students (experimental and control group) are 85.29 ($N = 296$) before computer use and 85.47 ($N = 234$) after the use of computers. Clearly, the null hypothesis cannot be rejected; the attitudes appear to have stayed about the same. When considering the results for the different subscales (see Table 1), we can differentiate between these effects more clearly.

The results in Table 1 indicate a significant decrease in the subscale of motivation; the changes on the other subscales however are not significant. To mention just some exemplary items for the motivation subscale:

- Though at first 89% of the students indicates that computers are important for their future, afterwards only 78% feels this way.
- Though at first 84% of the students indicates that computers help to learn, afterwards only 73% thinks so.

Although the percentages on the posttest can still be considered as high, apparently for some students the high expectations which they had beforehand were not met completely.

The results on the MASCART questionnaire are very positive. About 97% of the students think that the software and the instructions are clear and that the feedback is very useful. Many students answer that they like to work with MASCART very much, for various reasons. Only very few students indicate that they do not like the materials, of which some indicated that they thought the materials were too easy. These results are similar to the findings of the student observations during the MASCART sessions: students seem to like to work with the materials very much and they work with much concentration.

**Knowledge gains**

The hypothesis whether students using the MASCART materials (experimental group) perform better on the posttest than students who did not use the materials (control group), was tested one-tailed (Table 2).

These results indicate that students in the experimental group indeed performed significantly better than the control group. Whether this greater improvement is caused by the quality of the MASCART materials or by the increased time on task (students of the experimental group receive about $2 \times 45$ min of remedial teaching per week extra) cannot be determined from these results. This problem will be discussed in more detail in a later section. On the other hand, if MASCART did not exist, no other form of remediation could have been possible because lecturers do not have

---

**Table 1. Mean attitude scores per subscale**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>296</td>
<td>37.16*</td>
<td>16.01</td>
<td>15.55</td>
<td>12.91</td>
</tr>
<tr>
<td>After</td>
<td>234</td>
<td>36.00*</td>
<td>16.45</td>
<td>15.99</td>
<td>12.88</td>
</tr>
</tbody>
</table>

*Significant difference ($\alpha = 0.05$).

**Table 2. Mean scores and standard deviations on pretest and posttest of the experimental and control group**

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean scores ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Pretest</td>
</tr>
<tr>
<td>Experimental</td>
<td>130</td>
<td>26.06 ± 5.89</td>
</tr>
<tr>
<td>Control</td>
<td>130</td>
<td>26.11 ± 5.89</td>
</tr>
</tbody>
</table>

*Significant difference between pre- and posttest ($\alpha = 0.05$).
The effectiveness of MASCART in a university

Pretest (exp. grp) + Pretest (contr. grp)
Posttest (exp. grp) Posttest (contr. grp)

Fig. 1. Scores on pretest and posttest per ability group.

time for this kind of activities. This leads to the conclusion that the remedial computer assisted instruction indeed results in an overall improvement of student achievement.

Effects on ability group

The next question is whether the MASCART materials have different effects for the various ability groups (low, middle and high). This question has been examined by using two different methods. For the first method, the total group of students is divided into three; the low ability group consisting of the 25% with the lowest score on the pretest, the middle ability group of the next 50% of the students and the high ability group of the top 25%. The MASCART materials were developed for the low and middle ability group especially, but the lecturers of PESC believed that the main effects would be found in the middle ability group. They expected that the entrance level of the materials may be a little too high for the low achievers. The null hypothesis to be tested is that all three ability groups will equally gain from MASCART.

Figure 1 shows the mean scores on the pre- and posttest for both experimental and control groups at the three ability levels.

The two lower lines represent the results for both groups on the pretest. The two lines are about the same, which was to be expected because the matching procedure was based on the results of this test. The upper two lines represent the results on the posttest. The general effect of MASCART is shown by the higher scores of the experimental group. The results of the control group show that the regular instruction in PESC (about other topics than basic algebra) also improves the scores on the posttest.

An analysis of variance shows that there is indeed an effect caused by the experimental treatment (the MASCART materials). The factor ability also plays a role. From these data it cannot be concluded which ability group benefited the most. Therefore, a more thorough analysis is needed.

Table 3 shows the differential scores (posttest minus pretest scores) for the three ability groups in the experimental and the control group. Variance analysis of these data shows that there is no significant interaction ($\alpha = 0.05$) between experimental treatment and ability group. This means that, no matter what program is concerned, the lowest ability group benefited the most; and, no matter what subgroup is concerned, the experimental group made more progress than the control group.

| Ability group | Group   | Low       | Middle    | High      |
|              |         |           |           |           |
|              | Experimental | 9.86 ± 4.06 (36) | 6.95 ± 3.07 (63) | 3.42 ± 2.42 (31) |
|              | Control   | 7.23 ± 4.93 (38) | 4.53 ± 4.07 (61) | 1.38 ± 2.51 (31) |
RESULTS OF THE MARKOV CHAIN ANALYSIS

In addition to the regular eight periods of instruction per week, students in the experimental group received two extra periods of remedial instruction, by working through the MASCART materials. The analysis applied so far does not allow us to determine whether the improved results of the experimental group are a result of just more time on task, or whether the quality of the MASCART materials contributed to the efficiency of the remedial activities. The approach of Chen and Plomp[12], in which they apply Markov chain analysis on repeated data on the educational achievement of a certain group of students, may shed light on this and other questions. Under the assumptions that (i) every student can reach a pre-defined mastery level if he or she spends enough time on the learning task, and (ii) the quality of the instruction given in stable, one can, by applying Markov chain analysis, calculate for each pre-defined (sub)group the expected average time to reach a pre-defined level of mastery.

The approach of Chen and Plomp[12] is summarized in Appendix 1. As indicated in the Appendix, for a homogeneous Markov chain it is assumed that the probability for students to reach a certain level is only dependent on the level the students are originally in, and does not depend on other variables such as the number of instructional units the students have attended.

Before applying Markov chain analysis, first the question will be addressed of how realistic these assumptions are for the group of students in the PESC course. As indicated before, the group of students participating in the PESC course was selected based on their performance in the Cambridge Overseas School Certificate. Having the best students of a grade cohort in the PESC course, it is reasonable to expect that this group will meet the assumptions of the mastery learning approach, namely that in principle all students must be able to master the objectives of the basic algebra course. For the same reason, we also accept the extra assumption of constant transition probabilities for the subgroups. This assumption goes one step further than the mastery learning assumption, because the latter is not referring to any property of the subgroups.

In case there is doubt about the validity of the assumption of constant transition probability, it is possible to test statistically whether the probabilities of transformation from one level to another level from time \( t \) to \( t + 1 \) is not different from the transformation probabilities from time \( t + 1 \) to \( t + 2 \). Chen and Plomp[12] mention the Log-likelihood ratio test (cf Billingsley [13]) to test the homogeneity of transition matrices. Applying this test in the experimental and control groups, we conclude that for both groups the null hypothesis is not rejected, which implies that our data fit the model underlying Markov chain analysis.

By applying the Markov chain analysis to our experimental and to our control group, differences in expected time to reach mastery level might reveal differences between the quality of instruction in both conditions, and in this way help us to separate the possible two causes of the effect of the experimental conditions (that is more time on task, and the quality of the MASCART materials).

For this purpose, the students in the experimental and control group are divided in ability groups in a different way than in the preceding section: five ability categories are defined, based on the student’s score on the pre- and posttest (40 items). Table 4 depicts this ability distribution of the pretest and posttest for both the experimental and control group.

To apply Markov chain analysis on the Basic Algebra Test (BAT), first the data on the pretest and posttest for both groups need to be presented in transition matrices, indicating how students on a certain level on the pretest score on the posttest (Table 5). From these, the inverse matrices for both the experimental and control group can be calculated (Table 6). Now the expected time for the different subgroups to reach mastery level 1 can be calculated by applying the formula of Chen.
and Plomp for Markov chain analysis on these data. The results for the subgroups in both experimental and control group are presented in Table 7.

In the experimental group, for example, the subgroup of students who scored in the pretest on level 2, needs an average of 1.19 time units to reach the defined mastery level. Similarly, in the control group the subgroup of students who are, for example, in the pretest on level 3, needs an average of 2.17 units to reach the defined mastery level. One teaching unit in the PESC course is about four weeks; but because during this period the time on task in the experimental group (MASCART) is greater than in the control group, Table 7 also contains the expected times for the experimental subgroups calculated in the time units of the control group. The data show that for the subgroups in the experimental condition in all cases the expected average time to reach the desired mastery level is smaller than for the equivalent groups in the control condition.

Because of the stochastic nature of the phenomenon we are analyzing, we cannot conclude that the experimental condition is always better than the control condition, or that, after having corrected for the differences in the value of a time unit between the experimental and the control group, the better performance of all experimental subgroups is due to the quality of the MASCART materials. For this we have to test whether the expected average time (in both the non-corrected, and the corrected situation) for the experimental subgroups is shorter than for the control group.

Chen and Plomp[12] present a procedure for testing the null hypothesis that the expected time for both the experimental group and the control group is the same, that is both groups come from the same population: that is the difference between the expected times results from stochastic observations. To test this hypothesis, the Pearson χ² procedure is applied. In this case, the contingency tables are 1*4 tables:

\[
\begin{align*}
T(\text{observed, experimental}): & \quad 1.19 \quad 1.63 \quad 1.96 \quad 2.46 \\
T(\text{observed, control}): & \quad 1.75 \quad 2.27 \quad 3.06 \quad 3.21 \\
T(\text{theoretically expected}): & \quad 1.47 \quad 1.95 \quad 2.51 \quad 2.84
\end{align*}
\]

where both \(T(\text{observed})\) are the expected times for the subgroups in the experimental condition (not corrected, see Table 7) and the control condition, and the \(T(\text{theoretically expected})\) is the mean value of the other two. The Pearson \(\chi^2\) test statistic is 13.68, while the critical value is 9.49 (\(df = 4; \alpha = 0.05\)). We can reject the null hypothesis, which means that the expected time for reaching the mastery level is significantly shorter in the experimental group than in the control group, in the case of not correcting for differences in the unit of time. The conclusion is that the experimental

<table>
<thead>
<tr>
<th>Subgroup level</th>
<th>Experimental group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.19(1.42)*</td>
<td>1.75</td>
</tr>
<tr>
<td>3</td>
<td>1.63(1.96)</td>
<td>2.17</td>
</tr>
<tr>
<td>4</td>
<td>1.96(2.35)</td>
<td>3.06</td>
</tr>
<tr>
<td>5</td>
<td>2.46(2.95)</td>
<td>3.21</td>
</tr>
</tbody>
</table>

*The unit for expected time is the amount of instruction time between pretest and posttest, for the control group 4 weeks of 4 x 2 lesson periods (Tc); for the experimental group Te is Tc plus 4 weeks of 2 periods per week MASCART Instruction; so Te = 1.2* Tc.

†The expected time for the experimental group in Tc units (multiplication factor 1.2).
condition, in which the students in a four week instructional period spend two lesson periods extra on remedial CAI, leads to significantly better results than the control condition consisting of the regular instruction. This result is consistent with the preceding section, where we tested the differences in posttest scores between both groups.

However, if we apply the test procedure on the corrected expected times for the experimental group (taking into account the longer time on task), then the test statistic is 3.99 ($df = 4$), which is not significant for $\alpha = 0.05$. This implies that if the control group could spend as much time on task as the experimental group, then the difference between conditions is no longer significant. This leads to the conclusion that the difference in achievement between both groups is predominantly a consequence of difference in time on task. One may conclude that the MASCART project has been successful in that respect, that students were brought to more time on task. The better results of the experimental group could, however, not be attributed to the quality of the MASCART materials.

Suitability of the MASCART program

Earlier we asked which group of students profit the most from the MASCART program. We concluded that all ability groups profit from the remedial computer instruction, and that, no matter what program is concerned, the lowest subgroup (the bottom 25% on the pretest) benefited the most.

The Markov chain analysis can also shed light on this question. From the matrices in Table 6 we can see that the average time for which the last two subgroups of the control group are expected to remain in their original level (1.24 and 1.37 time units) before going up to a higher level, is not much higher than that of the similar subgroups in the experimental group (1.15 and 1.07). The students of the control group originally in level 5 are even expected to spend less time (0.30) in level 4 than the similar experimental subgroup in the level 4 (0.47). Both these students are expected to spend much time in the levels 2 and 3: as we can see from the matrix, the 'control' students originally in level 4 are expected to spend 1.25 time units in level 2 and 0.56 units in level 3, while for the students originally in level 5, these expected times are 0.96 and 0.58 time units. These expected times are much higher than the expected times of the students in the experimental group in the corresponding level. This indicates that the experimental program is more suitable for the students originally in a higher level.

Chen and Plomp[12] applied their method also on a more detailed division of students in subgroups, by defining the highest level of achievement '36–40 correct' (instead of '33 up'). They show that, even if the expected time for the experimental group is corrected for the extra time on task, the time needed to reach the highest level in the experimental condition is significantly less than in the control condition. In summary, applying Markov chain analysis, it can be shown that the experimental program is more suitable for students in a higher level in a situation where a very high mastery level is considered to be important. This result is an interesting contrast to earlier findings, where we conclude that for a lower mastery level (that is, the top 25% on the pretest) the lowest ability group profited more from the experimental condition.

CONCLUSION

In examining the effects of the MASCART materials on student knowledge in basic algebra, a significant improvement is basic algebra knowledge and skills could be measured in the experimental group. We could not prove that this improvement was caused by the quality of the MASCART materials, but the data show that the increased time on task by working with the MASCART materials was useful. A more detailed analysis applying Markov chain analysis revealed that, in a situation where a very high mastery level is required, higher levels students profit more from MASCART. On the other hand, it was shown that with a lower mastery level, the lowest ability group profited the most. Our analysis nicely shows that the question of which ability group profits the most from computer-assisted instruction depends on the interaction of the mastery levels and of the ability grouping.

Given the conclusion that it is predominantly the increased time on task which caused the better results in the experimental condition, one might argue whether a cheaper way of remedial teaching
might have been to give students an extra tutorial per week, instead of the MASCART sessions. However, when considering the rather low attendance of students during regular tutorials, it is doubtful whether the same effect could have been reached using this kind of remedial teaching. Moreover, in the PESC situation the mathematics staff does not have time to give extra tutorials. The MASCART materials have been designed especially to satisfy this need and are therefore lecturer-independent. So, one might say that the computer-assisted MASCART materials have proved to be a successful means to realize remedial instruction, where at present no other possibilities for remedial instruction are easily possible.

Another effect of the MASCART materials was the effect on student motivation. Students liked to work with MASCART very much, because they could work individually and at their own pace. Whether this positive effect on students’ motivation will last when the novelty effect is gone, has yet to be studied in this situation.

These results indicate that the use of computers for remedial purposes can be very useful within the setting of a less developed country, where very often a lack of teachers and lack of time to give individual attention are important factors. In this way, the ‘internal effectiveness’—defined by Lockheed and Hanushek [14] as the ratio of learning (a non-monetary outcome of education) to specific material and non-material inputs of the system—has improved in the MASCART case. The resources available in the PESC Department (e.g. limited teachers time, computer lab, MASCART software) have been organized in such a way that the maximum feasible output is produced. Whether the use of the software also improves ‘internal efficiency’ (defined as the ratio of learning to the costs of educational inputs) or even ‘external efficiency’ (the ratio of monetary outcomes to monetary inputs) is a question which has yet to be investigated in this setting.

Acknowledgements—We would like to express our great appreciation to Chen Yu Kun for his assistance in the application of the Markov chain analysis. We also thank Kees Smit and Pieter Wolff from Free University Amsterdam for their help given during the project.

REFERENCES


APPENDIX

The Suitability of Markov Chain Analysis of Effectiveness of an Education Program

This appendix is taken from Chen and Plomp [11].

A stochastic process taking values in a set $I$ is called a homogeneous Markov chain, if the following equations hold:

$$P(X_{n+1} = j | X_n = i; A) = P(X_{n+1} = j | X_n = i) = P_0.$$ 

Here $I$ is called the *state space* and an element $i$ or $j$ of it a state; $X$ is a random variable; $n$ and $n + 1$ represent points in time; $P$ is the transition probability of a system from one state to another state; while $A$, being of the form $X_0, X_1, \ldots, X_n$, describes completely the past of our system of interest.
The meaning of the first equation is that the state of the system at the time $n + 1$ is independent of the history prior to time $n$, as long as $X_n$ is known. The meaning of the second equation is that the conditional probability does not depend on the value $n$.

We apply Markov chain analysis on a pretest–posttest design in which the pretest and the posttest consist of the same items. In this case, it is expected that, as a consequence of the instruction, nearly all students’ scores on the posttest are higher than those on the pretest. If we divide students, according to their scores on the test, in levels or classes A, B, C, D and E, and tabulate the change in the students’ scores, the table might be, for example, Table A1.

In the table the last column $n_i$ is the number of the students in the $i$th level in pretest; $x_{ij}$ is the number of the students who were in the pretest in the $i$th level, but are transferred in the posttest into the $j$th level.

Obviously, it holds:

$$x_{11} + x_{32} + x_{33} + x_{44} + x_{55} = n_i.$$

We notice that in this case all elements of the table which are above the diagonal are zero. It symbolizes the before-mentioned expectation that the students’ scores on the posttest are not lower than on the pretest. This kind of system is called a hierarchical system with no demotion. Under the name of Markov processes, a well developed technique has been developed for the analysis of this kind of system[A2].

Assume that for measuring the effectiveness of an education program a pretest and posttest are used; as mentioned above, we assume that both tests are made of the same items. Let us further assume that all students, after taking part in a program, either make some progress or remain in their original level. If we change the above table into a matrix form, the so called transition matrix $M$ might be as follow:

$$M = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
a_{21} & a_{22} & 0 & 0 & 0 \\
a_{31} & a_{32} & a_{33} & 0 & 0 \\
a_{41} & a_{42} & a_{43} & a_{44} & 0 \\
a_{51} & a_{52} & a_{53} & a_{54} & a_{55}
\end{pmatrix}$$

with $a_{ij} = x_{ij}/n_i$.

This kind of matrix, with zero elements above the diagonal, is usually referred to as a lower triangular matrix ($a_{ij} = 0$ for $i < j$). Due to the error of a measurement, in practice there might be some non-zero elements $a_{ij} > 0$ for $i < j$. But it is supposed that, compared to the other elements in the same row, the value of the non-zero element is very small so that we can treat them as zero.

The first element of the first row in the matrix is 1, which means that all students originally in the highest level stay in that level, after participation in the education program. The remaining students, being originally in various levels, will either stay in their original level or reach a higher level. If instruction continues and tests are being repeated, under the assumptions we made, then theoretically all students will reach the highest level sooner or later, and if they reach this level, they will remain there.

The pedagogical meaning of our assumption is that, if we can spend enough time on training a group of students in an education program in a proper way, then all of the students will reach the highest level in course of time. The assumption is in fact the basic ideal of mastery learning. According to mastery learning, by breaking down a complex behaviour into a chain of component behaviours and by ensuring students’ mastery of each link is the chain, it would be possible for all students to master even the most complex skills.

Now given the transition matrix and its basic property, a reasonable question is: how long, on the average, will it take for a group of students to reach the highest level?

Let us, before answering this question, suppose that the students take a pretest, then receive a unit of instruction (which may take some hours per week during several weeks), and finally take a posttest. As a result, we get the following transition matrix $M$:

$$M = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\
a_{31} & a_{32} & a_{33} & a_{44} & a_{53} \\
a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\
a_{51} & a_{52} & a_{53} & a_{54} & a_{55}
\end{pmatrix}$$

In order to make our computations more look like reality, we do not assume all of the elements above the diagonal of the matrix be zero; except for the first row (a student in the highest level, stays in that level).

Under the assumption of mastery learning, we call $t_i$, the average time for students originally in level B to reach level A; and, similarly, let $t_4$ and $t_5$ be the average times for students in level C, D, respectively, to reach level A.
Assuming that all periods of instruction have the same length of time (\( u \) is the unit time), then we have the following equations:

\[
\begin{align*}
t_1 &= u + a_{23} t_2 + a_{25} t_5 + a_{34} t_4 + a_{35} t_5 \\
t_2 &= u + a_{33} t_3 + a_{34} t_4 + a_{35} t_5 \\
t_3 &= u + a_{43} t_3 + a_{45} t_5 + a_{46} t_5 \\
t_4 &= u + a_{53} t_5 + a_{54} t_5 + a_{56} t_5
\end{align*}
\]

Let us take the first equation as an example to explain the interpretation of these equations. After one step of unit time (that is the first unit of instruction), a student will remain in level B with probability \( a_{34} \); the average time for a student to go directly from level B to level A is \( t_2 \) hence the expected time for a student to go after a first step directly from level B to level A is \( a_{34} * t_2 \). However, with probability \( a_{23} \) a student originally in level B will go to level C; then he needs \( t_3 \) to come from level C to level A; so \( a_{23} * t_3 \) is the time component in the equation for student originally in level B, who reach level A via level C as "detour". Similar explanations can be given for the terms \( a_{34} * t_4 \) and \( a_{25} * t_2 \).

The equations can be rewritten in matrix form:

\[
t = At + u
\]

where both \( t \) and \( u \) are vectors:

\[
t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix}, \quad u = \begin{bmatrix} u \\ u \\ u \\ u \end{bmatrix}
\]

and \( A \) the so-called transposition matrix:

\[
A = \begin{bmatrix}
a_{22} & a_{23} & a_{24} & a_{25} \\
a_{32} & a_{33} & a_{34} & a_{35} \\
a_{42} & a_{43} & a_{44} & a_{45} \\
a_{52} & a_{53} & a_{54} & a_{55}
\end{bmatrix}
\]

\( A \) is derived from matrix \( M \) by deleting the first row and first column, those parts of matrix \( M \) represent the state of the students who are already in the highest level (in which we are not much interested).

From matrix equation (1), we obtain:

\[
\begin{align*}
(t - At) &= u \\
(I - A)t &= u \\
t &= (I - A)^{-1} u
\end{align*}
\]

in which \( I \) is the unit matrix, and \((I - A)^{-1}\) is the inverse of matrix \( I - A \).

The equation (4) is the most important equation used to estimate the expected time for each group of students to reach the objectives of an education program. The solution of it gives the answer to our question.

### APPENDIX REFERENCES
