Volume II

OE
5

THE EVALUATION OF PROGRAMMED INSTRUCTION

How Much Does It Cost?

Programmed Instruction Centre for Industry,
University of Sheffield,
32, Northumberland Road,
SHEFFIELD S10 2TN.

PROGRAMMED INSTRUCTION CENTRE FOR INDUSTRY.

University of Sheffield,
Department of Psychology,
SHEFFIELD, S10 2TN.

14th December, 1967

THE EVALUATION OF PROGRAMMED INSTRUCTION

How Much Does It Cost?

Time is money, and training time is money too. Money is invested in training with the expectation of substantial rewards when the trainee becomes proficient. Undoubtedly a good investment, but a peculiar one; for in most investments we can at least say how much money is involved. But how much money is involved in training, and what are the rewards? How much does training cost? One thing at least is certain - it is amazingly difficult to put a monetary yardstick against the value of training.

There are some factors that are definitely to be included in the cost of training, the 'direct costs' such as the instructor's salary. But there are others - the 'overheads' whose inclusion, and value, depends largely on an ad hoc decision. It is such features that make the costing of any training method, including programmed instruction, virtually impossible to do precisely. Still, on the basis of the 'direct' costs, it should be possible to make a start.

It is convenient to seperate costs into two categories: Developmental costs - the cost of the production of a programme in its final form, and Variable costs - the cost of the programme in operation. The tendency will be to express these as cost per hour per student; for, although the 'absolute'

cost is useful when deciding how much a particular programme is costing a company, the concept of cost per hour per trainee is the most useful basis for comparing costs. For example, if two courses teach the same and cost the same, they can still be compared on the basis of the number of students each course will serve.

The obvious thing to do right now is to spell out the Developmental and Variable costs in some detail. But leave that for the moment. Concentrate on the question of what is the most expensive item in training. Which for example is the most expensive: Development or Variable costs?

Rummler did a survey of in-plant programming costs. The results are as in Fig.(I). (c.f. Page 3).

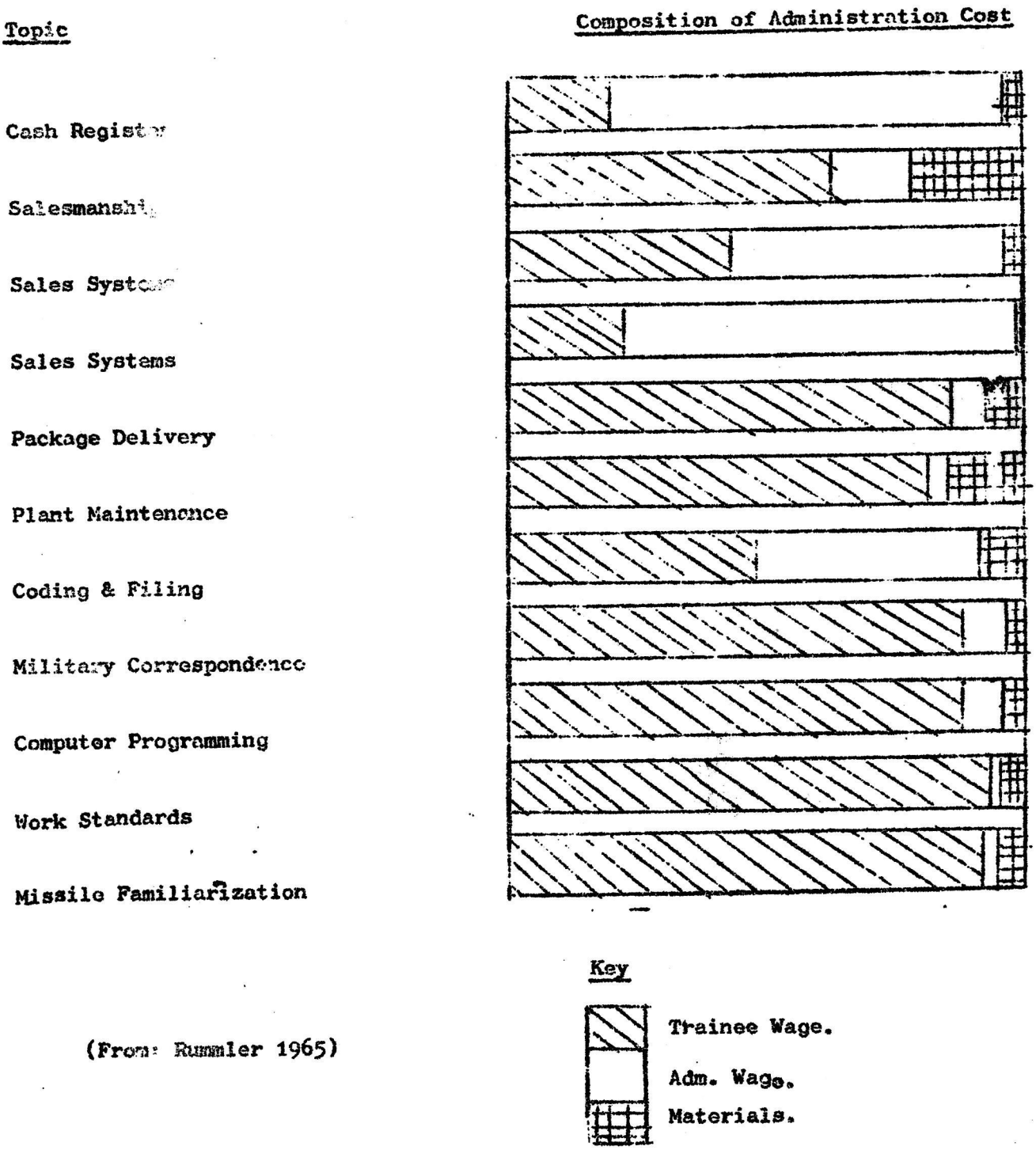
In other words, Variable or Administrative costs accounted for something like 75% of the cost per hour of training per trainee. Why should this be? Fig.(II) (c.f. Page 4), gives the answer to that one.

FIG (I) Total Cost (Development & Administration) per Hour of Training

FIRM	TOPIC	(A) DEVELOP-MENT	(B) ADMINIS-TRATION	(C) TOTAL
Department Store (1)	Operating Sales Register	$0.46	$7.12	$7.58
Department Store (2)	Beginning Salesmanship	0.39	1.98	2.38
Department Store (3)	Sales Systems	0.56	2.88	3.44
Department Store (4)	Sales Systems	1.94	6.85	8.79
Department Store (5)	Package Delivery	1.39	1.61	3.00
Oil Refinery	Plant Maintenance	1.27	6.03	7.30
Government (1)	Coding and Filing System	2.81	3.50	6.31
Government (2)	Military Coding	1.84	2.09	3.93
Equipment Manufacturer	Computer Programming	0.87	2.33	3.20
Automobile Manufacturer	Work Standards	2.56	4.87	7.43
Aero-Space	Missile Familiarization	0.66	3.29	3.95

(From: Rummler 1965)

FIG (II) Composotion of Administration Costs for Eleven In-Plant Programmes



(From: Rummler 1965)

The bare fact is that the highest single item in the entire direct cost of running a training programme is the trainee's wage. This is certainly predictable and probably applies to most methods of training. What it does mean is that if economies are to be made in the direct cost of training, the most obvious thing to look at is the most expensive item, the trainee's wage.

One way of reducing costs on the most expensive item would be to cut the trainees wage - but there is a feasable way too: Do the training faster and the faster the better. For the amount paid to the trainee depends on the time it takes to train him. The results of this can be quite amazing; for by reducing the time to train a trainee, the time a supervisor/instructor spends on training is also being reduced. And, as can be seem from Fig.(II) this is the second most expensive item in the direct variable costs. If overhead costs are added, the results are even more amazing.

For example:-

Suppose a national sales organization which hires young men fresh out of college and gives them a basic course of instruction in the product line, administrative requirements, such as reporting, expense policies, etc., and techniques of selling. The compnay operates a two-week training programme for all new salesmen. Averaging 10 students to a class, the company pays them a nominal salary of $60 per week while in training. Thereafter, they are on straight commission. During their first year, they will average earnings of $110 per week for themselves and earn $90 per week for the company.

Assume that the company runs the programme on a continuing basis, holding 25 classes per year, and that the course is taught by two experienced salesmen, selected at times from the selling ranks to serve a tour of duty in sales training. If these instructors were out selling, instead of teaching, it would be reasonable to presume that they would also earn the company at least $90 per week.

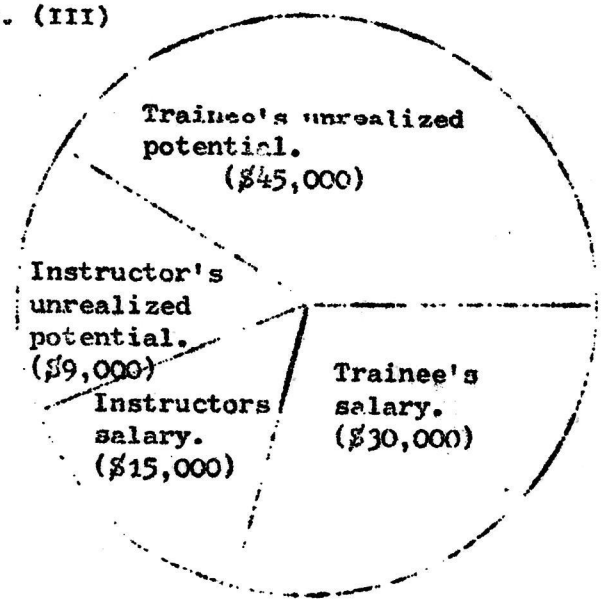
If, in such circumstances, a major evaluation and restructuring effort were to make possible equal results, through a shorter course, what savings might be expected?

On a two-week basis involving 10 days of training time, the costs are about as follows:-

For each class, $1,200 in student salaries and at least $600 in instructor salaries. Add $1,800 loss of income to the company while the students are in class and not out selling, and a similar loss of at least $360 on the time of the two instructors. This amounts to a cost of about $4,000 per class or $100,000 per year for the 25 classes. (Fig.(III) c.f. Page 7).

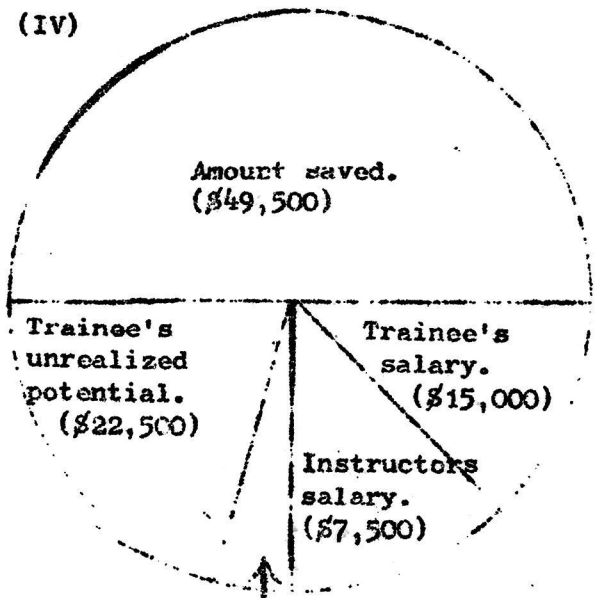
If the training time could be reduced from two weeks to one week, savings would amount to about $50,000 per year. (Fig.(IV) c.f. Page 7). A reduction in course length from 10 to 8 days saves $20,000 per year. (Fig.(V) c.f. Page 7). Over 10 years, the stakes have become quite large. (Example taken from: Lott '67).

FIG. (III)



TWO-WEEK COURSE

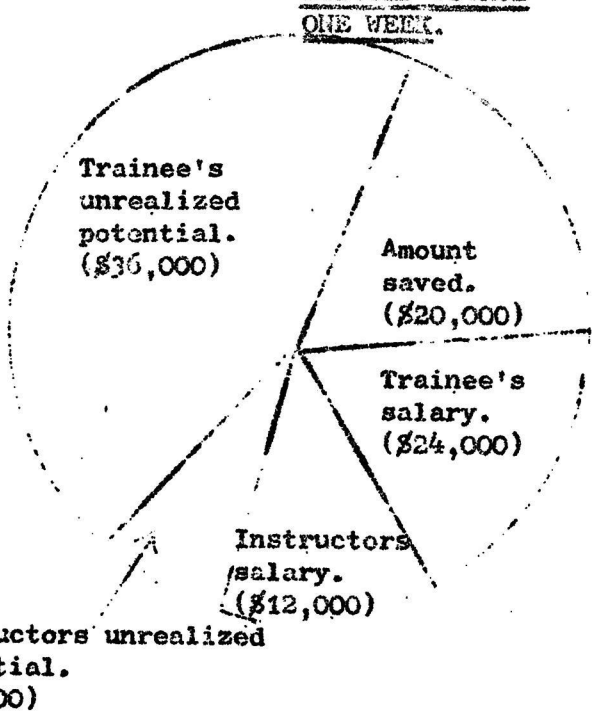
FIG. (IV)



REDUCED COURSE - ONE WEEK.

Instructors unrealized potential ($4,500)

FIG. (V)



REDUCED COURSE TO 8 WORKING DAYS.

Time is money, and training time is money too. If the example above is at all realistic, a reduction in training time of 20% can save an enormous amount of money. But how realistic is that example?

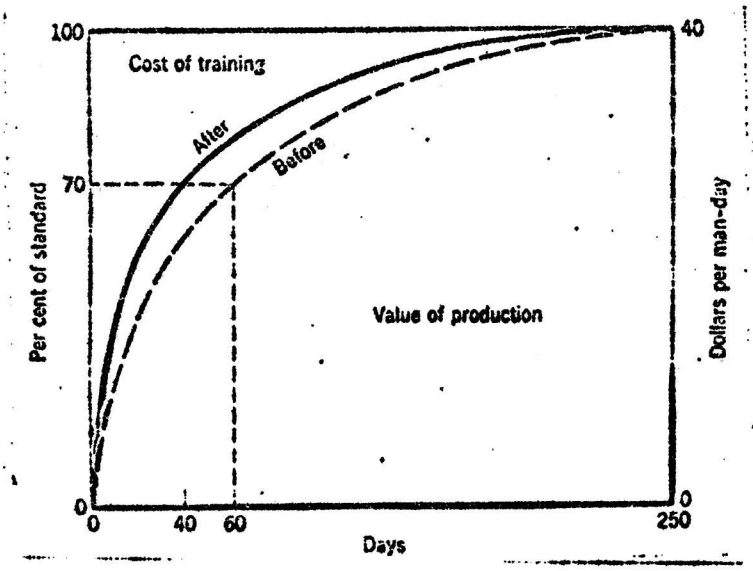
One study (Hickey 1962) suggests that the idea of such rewards for such reductions in training time is not as riduculous as it might at first appear. The study involved a prominent American manufacturer of telephone relays who employed 120 men to produce 6,000 relays every day at a manufacturing cost of $2 each; a further 120 men being employed to adjust the relays by hand at a further cost of $1 per relay.

40 relay adjusters were trained every year at a total cost of $80,000 - taking into account the actual cost of training, production lost and overheads. The method of training adopted was half-a-day's lecture/demonstration followed by having a go at adjusting simple relays under the guidanceof an experienced man in the production line.

Under this system it took about 60 working days to reach 70% of the department's standard, and a year to reach 85-100%. (See Fig.(VI) c.f. Page 9).

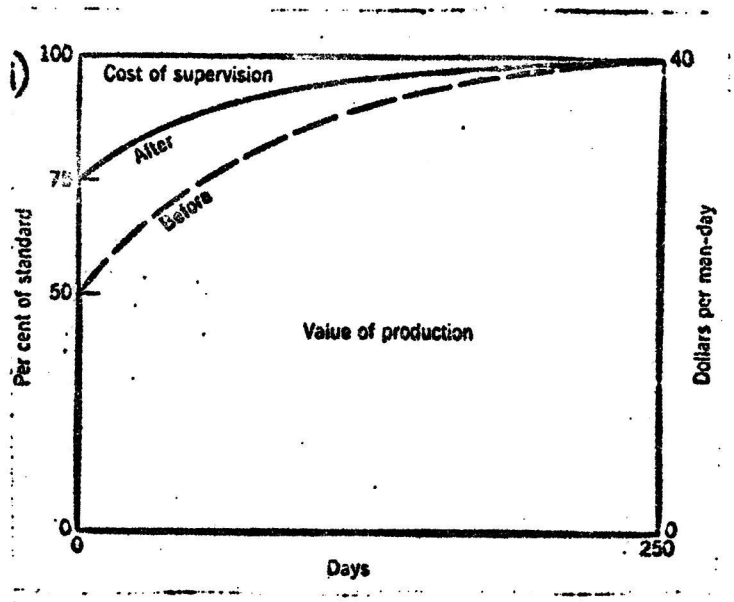
When the training was programmed, the trainee reached the 70% standard after only 40 days (as against 60) with the cost of training calculated to be $59,000 - a one third reduction in time and just about a 26% reduction in cost.

FIG. (VI) Cost of training apprentice relay adjustor before and after programmed instruction.



(From: Hickey 1962)

FIG. (VII) Cost of supervising training of relay adjustor before and after programmed instruction.



(From: Hickey 1962)

But this isn't all. Look back at Fig.(II) and it is easy to see that the second most expensive item in the administrative costs is the administrator's/supervisor's/instructor's time. And by reducing the training time for the student, the time the supervisor spends on training is also reduced. So, if the cost of supervising the trainees is added to the total cost of training the savings are even greater. (Fig.(VII). Since there was one supervisor to each trainee, and since each supervisor sacrificed about 10% of his productivity (about $1,000 per year) the total cost of supervising all the trainees was in the region of $40,000 a year. The introduction of the programme halved the supervision requirement, thereby saving $20,000. In other words, the introduction of programmed instruction resulted in a total saving of $41,000. A one third reduction on the original training costs of $120,000.

Nor is this unique. The G.P.O. claim to have made a potential saving of £70,000 a year through reducing the average training time for telephonists by five working days - by means of programmed instruction. And that £70,000 is the saving in trainees wages alone!

It is obviously illegitimate to generalize from this; it is obviously wrong to say that any programme will be faster than any other method of training. But that is not the point. The important thing is to balance the costs of programmed instruction, or any method of training, against the expected pay off. Remember this. Remember too, that although programming is by no menas cheap, the American firm mentioned above recovered the costs

of programming in one year. And the G.P.O., having so far spent £20,000 on the development of their programmes are making a saving of £70,000 per year on their old course.

So how much does programming cost?

1). Developmental Costs:

How much does it cost to write a programme? The main cost is, fairly obviously, governed by the time it takes to write a programme; but estimates tend to vary, a position that is summed up by Rowntree: 'Estimates have varied from 10 to 350 hours of programmed time to produce, in fairly final form, a programme that will take the student one hour to work through.' And anyway it is impsssible to say, in advance, how many hours of instruction will be required to teach a given topic. Nor is the position any better if the unit taken is not the time it takes to write an instructional hour's worth but the time it takes to write a frame. Estimates here tend to vary too, and anyway it's impossible to say in advance how many frames will be needed.

The outlook seems hopeless for accurately predicting the time it takes to write a programme. Why should estimates vary so much? The trouble is that there are so many variables that can affect the writing time. It may well be that the programme is written for a special purpose, (e.g. all students must get 100% on the test given immediately after the programme to see how much they have learned. Or they might be required to get 50% on a criterion test administered some time later: i.e. they might be

required to remember half of what they have learned. It may well be that the programmer has already written a programme on the relevant subject-matter, and is thus already familiar with that subject. A short programme will take proportionately longer to write, frame for frame, than a long one because of the 'start up' operations common to them both; the type of programming required will also effect the time it takes to write a programme. And the complexity and organization of the subject matter will also influence it, too; (See Fig. (VIII)).

FIG. (VIII).

Programing Conversion Data

Subject-matter level	Length of conventional course (hours)	Length of programed course (hours)	Frames prepared per hour	Preparation Time	
				Man-hours	Man-months
Difficult	20	11.66	1.5	699	4.13
	10	5.83	1.5	349	2.07
	5	2.92	1.5	175	1.04
Medium	20	11.66	2.0	525	3.10
	10	5.83	2.0	262	1.55
	5	2.92	2.0	131	0.77
Easy	20	11.66	2.5	420	2.48
	10	5.83	2.5	210	1.24
	5	2.92	2.5	105	0.62

(From: Drutsch 1962)

Because of uch factors, it is impossible to say in advance precisely how much it will cost to develop a programme; and taking into account the overhead costs as well, the outlook seems even more hopeless! If the programmer talks to the subject matter expert, say an experienced lathe operator, how much is his time worth? Is it simply a matter of 'so many

hours is such and such a proportion of a lathe operators working week' and cost the time in proportion? Or should the production lost in that time also be included in the cost? And how can you say in advance how much of the subject matter experts time will be taken up?

Still, even though it may be impossible to say in advance how much a programme will cost to write, some guidlines must be given. Figs. (IX) (X) and (XI) represent the costing of the development of programmes in an industrial and commercial concern respectively:

FIG.(IX) Cost of writing General Crane Driving programme

Cost to Write

		£ s. d.
Research	2 days	8 14 0
Preparation of format	½ day	2 3 6
Preparation of information	1 day	4 7 0
Rules (See Appendix B)	½ day	2 3 6
Frame Writing (rough)	9 days	39 3 0*
Checking, correcting, re-writing	2 days	8 14 0
Total for Programme	15 days	65 5 0**

Cost of actual writing of 110 frames = $\frac{£39\ 3\ 0^{*}}{110}$ = 7s per frame

Cost of producing final master programmes = $\frac{£65\ 3\ 0}{110}$ = 11s 9d per frame

(From: Taylor 1967)

The costing exercise, carried out on the writing of a General Crane Driving Programme at Stewarts and Lloyds, represents all the 'direct'

FIG. (X)

Showing analysis of programme costs over the three stages Investigating, Writin, Testing.

Programme title	*Telephone procedures*		*BEA consultative and negotiating machinery*		*Theory of control for supervisors*		*Management statistics parts I and II*		*Human relations at the airport*	
Description	Branching Grundytutor		Branching – text		Branching – text		Part I Branching Part II Linear Grundytutor		Branching Mulitutor	
Length	180 frames 6 hours		96 frames 2 hours		80 frames 40 minutes		725 frames 14 hours		180 frames 2 hours	
	Cost	*Time*	*Cost*	*Time*	*Cost*	*Time*	*Cost*	*Time*	*Cost*	*Time*
Investigating	£48	·1 week	£291	6 weeks	£97	2 weeks	£388	8 weeks	£97	2 weeks
% of total	1·7%	1·8%	18·8%	18·8%	11·9%	11·8%	14·8%	14·8%	7·1%	7·1%
per frame	£0·3	·01 weeks	£2·9	0·6 weeks	£1·2	·03 weeks	£0·5	·01 weeks	£0·5	·01 weeks
Writing	£2425	50 weeks	£824	17 weeks	£485	10 weeks	£1891	39 weeks	£825	17 weeks
% of total	87·7%	87·7%	53·1%	53·1%	59·6%	58·8%	72·2%	72·2%	60·7%	60·7%
per frame	£13·5	·28 weeks	£8·4	·17 weeks	£6·1	·12 weeks	£2·6	·05 weeks	£4·6	·09 weeks
Testing	£291	6 weeks	£436	9 weeks	£232	5 weeks	£339	7 weeks	£436	9 weeks
% of total	10·5%	10·5%	28·1%	28·1%	28·5%	29·4%	12·9%	13%	32·1%	32·1%
per frame	£1·6	·03 weeks	£4·4	·09 weeks	£2·9	·06 weeks	£0·5	·01 weeks	£2·4	·05 weeks
Total	£2764	57 weeks	£1552	32 weeks	£814	17 weeks	£2618	54 weeks	£1358	28 weeks
per frame	£15·4	·32 weeks	£15·8	·32 weeks	£10·2	·21 weeks	£3·6	·07 weeks	£7·5	·15 weeks

(From: Barry 1967)

FIG. (XI)

Showing analysis of programme costs over the three stages Investigating, Writing, Testing.

Programme title	*Air legislation*		*Theory of Flight*		*Fundamentals of management Parts I and II*		*The Managerial process*		*Introduction to the automatic seat reservation system*	
Description	Linear – Grundymaster and text		Branching – Grundytutor and text		Linear Part I Grundymaster Part II Text		Branching – text		Linear – text	
Length	777 frames 15 hours		580 frames 10 hours		471 frames 8 hours		146 frames 4 hours		359 frames 6 hours	
	Cost	*Time*	*Cost*	*Time*	*Cost*	*Time*	*Cost*	*Time*	*Cost*	*Time*
Investigating	£480	12 weeks	£280	7 weeks	£360	9 weeks	£280	7 weeks	£534	12 weeks
% of total	21%	22·2%	21·8%	21·9%	10·5%	11·4%	76·1%	77·7%	30·4%	30·8%
per frame	£0·5	·02 weeks	£0·5	·01 weeks	£0·8	·01 weeks	£1·9	·04 weeks	£1·5	·03 weeks
Writing	£1510	35 weeks	£960	24 weeks	£2555	59 weeks	£40	1 week	£979	22 weeks
% of total	66%	64·8%	75%	75%	74·8%	74·7%	10·9%	11·1%	55·8%	56·4%
per frame	£1·7	·04 weeks	£1·6	·04 weeks	£5·4	·12 weeks	£0·27	·01 weeks	£2·7	·05 weeks
Testing	£297	7 weeks	£40	1 week	£499	11 weeks	£48	1 week	£242	5 weeks
% of total	13%	12·9%	3·1%	3·1%	14·6%	13·9%	13%	11·1%	13·8%	12·8%
per frame	£0·3	·01 weeks	£0·1	·001 weeks	£1·1	·02 weeks	£0·32	·01 weeks	£0·7	·01 weeks
Total	£2287	54 weeks	£1230	32 weeks	£3414	79 weeks	£368	9 weeks	£1755	39 weeks
per frame	£2·9	·07 weeks	£2·2	·05 weeks	£7·3	·16 weeks	£2·5	·06 weeks	£4·9	·1 weeks

(From: Barry 1967)

staff costs of development. Only the overheads have been excluded. The figure on Page 15 represents a similar costing exercise carried out by B.E.A.

It is interesting to note that the most costly programme has a proportionately low investigating cost and that the least costly programme has proportionately the highest investigating cost. It is a bit of a schock, too, to notice the most expensive - Fundamentals of Management Part I and II at £3,414. But don't forget that if 100 students a year for 10 years receive the full 8 hours worth of instruction from that programme, that represents a cost of about 8/9d per trainee hour. And that is by far and away the most expensive programme in terms of development time. It is expensive, but don't forget to compare it with the expected pay-off.

2) Variable costs:

The following table Fig. (XII) gives some idea of the costs of programmes in operation at Stewarts and Lloyds: (c.f. Page 15)

The lefthand part of the table represents the maximum cost that might be expected: it is based on the assumption that a machine has a maximum life of one cycle - about six months use, and the complete wear out of two sets of the 17 programmes.

The righthand part represents more realistic costs; the assumption is that the machine and programmes have a life of about three years - or six cycles.

FIG. (XII)

Costing for I.T.M. Workshop Practice series on Grundymaster
October 1966 - March 1967

Number of Students	- 94
Number of Programmes	- 17 Workshop Practice Series
Average Study Time per Student for 17 programmes	- 20 hrs.
Cost of 2 sets of 17 programmes used over 6 months	- £24
Purchase price of 7 Grundymasters used - 7 x £12 10 0	= £87 10 0

Cost per programme worked inclucing full cost of machines	Cost per programme worked assuming life of machine of 3 years (6 cycles)
Total cost of machines and programmes - £111 10 0 Total programmes worked =Programmes used x No of Students=17 x 94=1598 Cost per programme = £111 10 0 / 1598 = 1s 4½d	Cost of Grundymaster for 6 months = £87 10 0 / 6 = = £14 11 8 Cost of Machines and Programmes = £38 11 8 Cost per programme= £38 11 8 / 1598 = 6d

Cost Per Hour worked on machines

Total hours worked on machines = hours per student x No. of students
= 20 x 94 = 1880 hours

Cost per hour worked on machines	Cost per hour worked on machines
= £111 10 0 / 1880 = 1s 2d	= £38 11 8 / 1880 = 5d

(From: Taylor 1967)

It should be pointed out that these costs do not include overheads or the limited amount of supervision needed or the trainees wage. What the tables mean is that, at the maximum, the cost per hour worked on the machine is 1/2d. The more realsitic cost is 6d for each hour the machine is worked. That isn't to say that all programming materials will produce that figure - and those programmes that have adjunct aids certainly won't. It's just that when estimating the cost of materials, do remember that programmes and machines do wear out! And don't forget to add the trainee's wage, the administrator's wage, and the overheads.

3) Total costs:

The total costs of programming are simply the variable costs added to the development costs. A look back at Fig. (I) gives some idea of the total costs. of programming per hour of training per apprentice. The following tables Figs. (XIII), (XIV), and (XV) give the totals per trainee hour, and also the 'absolute' cost of programming, at B.E.A. (c.f. Page 17).

There is no doubt about it - programming is expensive. £3,000 odd pounds for the most expensive one is a lot of money. And that's only the 'direct' costs!

What must be borne in mind through all the talk of the costing of training is the simple fact that training is an investment, and an investment from which we expect the rewards of skilled men. And from

FIG. (XIII)

Showing direct cost of programmes produced from October 1963 to March 1966

	£	£
Programmes written internally at B.E.A.		
Air Legislation	2,590	
Theory of Flight	1,730	
Fundamentals of Management I and II	3,030	
The Managerial Process	430	
Indtroduction to the Automatic Seat Reservation System	1,870	
Telephone Procedures	2,710	
B.E.A. Consultative and Negotiating Machinery	1,200	
Theory of Controlling for Supervisors	670	
Management Statistics I and II	3,400	
Human Relations at the Airport	1,300	
Air Conditioning in Civil Airliners I and II	1,820	
		21,650
Programmes Written Externally and Purchased by B.E.A.		
How to Type B.E.A. Correspondence	700	
How to Write a Business Letter	200	
How to Complete the International Ticket	710	
Organisational Groups	170	
Introduction to Management Statistics	140	
Airline Baggage Check-in Procedures	580	
Wage Negotiations and Collective Bargaining	50	
		2,550
Abortive Programmes		
The International Air Transport Association	390	
Attitude Training	770	
		1,160

Note: These costs do not include the purchase or hire of machines or any staff or services provided by B.E.A.

(From: Barry 1967)

FIG. (XIV) Showing outlays, student numbers in 1966 and 1967 and appropriate indices of cost per student hour.

Programme	Total cost	No. of hours instruc-tion	No. of students to date	1966 Index cost per student hour	1967 Target no. by students	1967 Index cost per student hour
Programmes writeen externally and purchased	£			£		£
Introduction to Manage-ment Statistics	140	8.0	40	0.4	70	0.2
How to Write a Business Letter	200	4.0	75	0.6	95	0.5
How to Complete the International Ticket	710	5.0	125	1.1	125	1.1
Airline Baggage Check-in Procedures I and II	580	4.0	125*	1.1	125*	1.1
How to type B.E.A. Correspondence	700	3.0	44	5.3	64	3.6
Human Relations at the Airport	1,300	2.0	25	26.0	100	6.5
TOTALS	3,630	26.0	434	34.5	579	13.0
AVERAGES		4.3		5.7		2.1
GRAND TOTALS	19,450	85.6	1,597	74.4	2,487	33.0
AVERAGES		6.1		5.3		2.3

* Part I

(From: Barry 1962)

FIG. (XV) Showing outlays, student numbers in 1966 and 1967 and appropriate indices of cost per student hour.

Programme	Total cost	No. of hours instruction	No. of students to date	1966 Index cost per student hour	1967 Target no. by students	1967 Index cost per student hour
Programmes written internally at B.E.A.	£			£		£
Managerial Process	430	4.0	210	0.5	310	0.3
Air Legislation	2,590	15.0	200	0.8	300	0.5
Fundamentals of Mamagement I and II	3,930	8.0	250*	1.9	450**	1.1
B.E.A. Consultative and Negotiating Machinery	1,200	2.0	263	2.3	413	1.4
Introduction to the Automatic Seat Reservation System	1,870	6.0	100	3.1	100	3.1
Management Statistics I and II	3,400	14.0	25	9.7	35	6.9
Theory of Control for Supervisors	670	0.6	100	10.1	200	5.0
Theory of Flight	1,730	10.0	15	11.5	100	1.7
TOTALS	15,820	59.6	1,163	39.9	1,908	20.0
AVERAGES		7.4		5.0		2.5

* Part I
**Parts I and II

(From: Barry 1967)

their skill comes the expected payoff. Before adopting a training technique, some pretty hard thinking has to be dome about costs. The idea in this paper is simply that programming may well result in the speeding up of a training course; that the cost of programming may well be justified over and over again by the results of time saved. This particular idea may not always be practicable. There is no guarantee that programming will speed up a training course.

But the general idea must always be practicable: it is simply thatt the high (-and it <u>is</u> high) initial cost of programming is not necessarily excessive. It becomes so only when the cost of programming becomes excessive in relation to the benefits that programmed instruction can provide.

References

BARRY, W.S. (1967) Programmed Instruction in B.E.A. Oxford - Pergamon.

BLAKE, C.S. (1967) A Procedure for the Initial Evaluation and Analysis of Linear Programmes. Unwin, D., Leadham, J. (Eds.) Aspects of Educational Technology. London - Methuen.

COMMUNICABLE DISEASE CENTER Training Aid for the Preparation of Instructive Communications. Atlanta - U.S. Department of Health, Education, and Welfare.

DEUTSCH, W. (1962) Programmed Learning: An Overview of Personnel and Financial Requirements. In: Margulees, S., Eigen, L.D. (Eds.) Applied Programmed Instruction. London - Wiley.

HALL, C. FLETCHER, R.N. (1967) Programmed Techniques in the G.P.O. Oxford - Pergamon.

HICKEY, A.E. (1962) Programmed Instruction in Business and Industry. Margulies, S., Eigen, L.D. (Eds.) Applied Programmed Instruction. London - Wiley. 281-293

KLAUS, D.J. LUMSDAINE, A.A. (1962) Some Economic Realities of Teaching Machine Instruction. Margulies, S., Eigen, L.D. (Eds.) Applied Programmed Instruction. London - Wiley. 198-207.

LOTT, O.C. (1967) Evaluating to Reduce Training Costs. Training Development Journal. 21. No.1. 38-41.

MARGULIES, S. (1962) Programmed Instruction: Some Economic Considerations. Margulies, S., Eigen, L.D. (Eds.) Applied Programmed Instruction. London - Wiley. 212-216.

RUMMLER, G.A. (1965) The Economies of Lean Programming. National Society for Programmed Instruction Journal. 4. No.10 8-10.

TAYLOR, E.A. (1967) The Use and Cost of Programmed Instruction at Stewarts and LLoyds. Oxford - Pergamon.

YANEY, J.P. (1966) Programming Costs and Cost Reductions. Technical Education and Industrial Training. 8 No.8.

YANEY, J.P. (1967) Choosing a Programming Method. Technical Education and Industrial Training. 9. No.6.

Volume I

06
4

THE EVALUATION OF PROGRAMMED INSTRUCTION

Does It Teach?

Programmed Instruction Centre for Industry,
University of Sheffield,
32, Northumberland Road,
Sheffield S10 2TN.

THE EVALUATION OF PROGRAMMED INSTRUCTION

a) Does It Teach?

The whole idea of asking 'Does ..... teach?' is fairly novel; it demands an assessment of a purported mode of training to see whether it really does teach, and if so, how well. There are exams, of course, such as the City and Guilds, and in some way these do reflect the effectiveness of the training undergone - but they do not reflect the effectiveness of a given piece of training, a particular lecture, for example, or a particular film. In asking 'Does ..... teach?' the methods of training are being questioned in just that way that demands an investigation of the overall training picture to identify its strengths and defects. The emphasis is being put on 'How far does this particular piece of training carry the trainees towards the training objectives?' rather than on 'Somehow, I don't know precisely how, but its something to do with his attending the apprentice school, he's got through the City and Guilds.' The change of emphasis is important; asking the question 'Does ..... teach?' each bit of training is being put on trial to see just what it contributes to the overall training picture, and once that has been done, it makes sense to ask how that picture can be improved; how training as a whole can be made more efficient.

The question 'Does programmed instruction teach?' - or its equivalent 'Do students learn from programmed instruction?' is, then, important. If it, or any training method is to be accepted, it must prove its worth, it must be demonstrated that trainees learn from it, and, even more importantly, how much and how well they learn from it.

Do trainees learn from programmed instruction? Well it's obvious that managers won't learn much, if anything, from a programme on reading, or good lathe operators much, if anything, from a programme designed to train apprentices in the fundamentals of lathemanship, however good these programmes might be. And it's equally obvious that a programme may fail to teach, not because its content is inappropriate, but because that content has not been put over in just that way that enables students to learn from it either easily or at all. The same holds true of any method of training; what it underlines is that it is impossible to say whether or not any method - including programmed instruction - teaches anything, in vacuo. For whereas it might be shown that a good programme, carefully developed, can teach certain people, it does not thereby follow that any programme will. What is possible is to determine whether or not certain people learn something from certain programmes, and how much;

and what is also possible is to determine the effects of certain definable, describable properties of those programmes. But this is not to say much about the value of the method in vacuo.

Can certain people learn something from certain programmes? There are some general considerations of relevance here; in the first case, programmed instruction isn't only teaching machines and programmed texts; programmed instruction might be summed up by the slogan that it is just what a good teacher does, only more so. The significant point is that certain features of what a good instructor does have been identified, developed, and presented by something other than a teacher. So of those programmes that do embody these features of good programmes that have been carefully developed, what is really being asked of them is 'Certain of the things that a good instructor does can be done by machine/book. But do these techniques still teach when put in this form?' And, perhaps, the surprise would rather be if it didn't teach, than if it did.

Another consideration that is relevant in asking whether certain people learn something from certain programmes is that a good programme, in its development, is amended according to whether or not students do learn from it, and according to whether or not they have difficulty in learning from it. Thus if the vast majority of the trainees make a mistake on a certain frame, which indicates that the presentation of a certain point causes them some difficulty, that frame can be altered. The idea is simple, the trainees fail to learn, so there must be something wrong with the programme; they find the presentation of a certain point difficult, so that presentation must be altered until they find that point easy to understand. The idea is simple, but crucial to the development of a good instructional programme. It would be peculiar if such a programme failed to teach those for whom it was intended.

These are general considerations however, not demonstrations of effectiveness, and the proof of the pudding is always in the eating. Do some programmes teach some people something? The question can be answered by 'How well do certain programmes teach?' for if they teach something it would at least seem feasible to measure how much that something is. So just how well do certain programmes teach certain people?

Most programmes embody what is called a 'criterion test' whose function is to sample the trainees knowledge of what the programme. is meant to have taught them. And it provides a ready means for assessing how much the trainees have learnt, for a programme can only be assessed by measuring it against what it is trying to do.

The procedure is simple: since the criterion test is intended to sample the trainees knowledge of what the programme is meant to have taught him, the test is given to the trainees before they undertake the programme to measure how much they knew initially. They are then given the programme, and then the test again. The results of this before-and-after experiment might be presented as in Fig.(1).

The same might be done for a number of other programmes, and the results put together to form a composite picture: Fig (II) represents the results of twenty seven such studies.

These results might also have been represented in several different terms - e.g. as a 'confidence ratio' or as a 'gain score'. A confidence ratio of 80/80 (e.g.) means that the programme will get 80% of the students to a final score of 80% on the criterion test; and a 'gain score' is simply the difference between the means in the test results; thus if the mean test result was 5% before the programme was administered, and 95% after the programme had been administered, the gain score would be 90%.

Does this show how well certain programmes teach certain people? Perhaps ..... but there are a few things to consider. There can be no doubt that there was improvement between tests, that at least has been shown. But what can - and should - be doubted is whether all of that improvement is due to the programme. For example, trainees will tend to get higher marks on a test, administered for them for the second time, than they did when taking it for the first time, even though no training has been undertaken in the interval between tests. Again, it might well be that a trainee cannot recall something whilst undertaking the pretest, it is easy for him to relearn it. The gain score is thus inflated beyond that gain which is solely due to the programme.

Furthermore, a lot of reliance is placed on the criterion test itself; if the measures are to be reliable it must be the case that that test really does sample what the programme teaches; Eraut points out an interesting example of the sort of failure that can occur here. (Eraut 66).

Still, although imprecise, doesn't the 'before-and-after' experiment at least give some guide? If students similar to those in the experiment were given the same programme, wouldn't they reach more or less the same level of attainment? Perhaps - but what does 'similar' mean here? Presumably, students have to be similar in those respects which are relevant to and influence their scores on the criterion test; and presumably this comes to such factors as level of initial knowledge, intelligence, personality, etc., but it is not known just what 'etc.' covers. Not only (i) does the 'before-and-after' experiment fail to give a precise measure of how much is learned from the programme, but (ii) that imprecise measure is restricted to the trainees in the experiment since it is not known precisely what features another group of trainees must have in order to be similar to the original group - and it is on the knowledge of this that the ability to generalise from one case to the other depends. Of course, it would seem that the most important respect in which the trainees must be similar is the level of initial knowledge and on the basis of this a rough and ready generalization is possible; but this is an 'informed guess' rather than a hard and fast prediction.

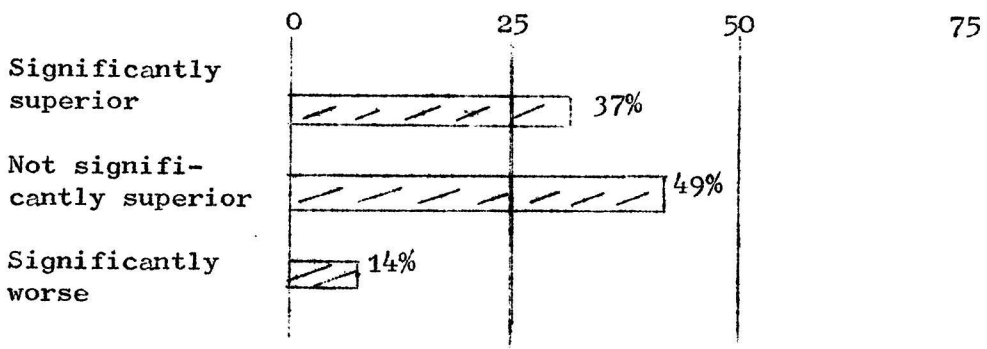
So, to the question 'How well does programmed instruction teach?' the answer given is a rather imprecise measure of how well certain trainees learnt from certain programmes; but leaving aside the imprecision, just how good is a confidence ratio e.g. of 80/80? It sounds impressive, but how good is it in fact. It is a bit like saying so many thousands of gallons of beer are drunk in England per year - the figures look impressive, but they are unfamiliar, we are not sure what to make of them.
Yet expressed in terms of something familiar - such as, on average, so many pints of beer are drunk per person per day, those figures mean something, they are familiar to us, we know what to make of them, we know whether it is a large amount or a little. The same holds true for the evaluation of programmes - a confidence ratio of 80/80 sounds impressive, but it is unfamiliar; it is not certain just how good it is. To know just how good it is, the figure must be expressed in terms of something familiar - and this is precisely what is done by the so-called 'comparison' experiment. Programmed

instruction is assessed by comparing its effects with those of what is loosely termed as 'conventional instruction', as a lecture, for example.

The basic idea is to get two groups of trainees, similar with respect to such factors as initial knowledge, intelligence, etc., one group is given 'conventional instruction' by an instructor, the other taught solely by programmed instruction. At the end, both are given the same test. The results of some one hundred and twelve such studies might be summarized as in Fig. (III). (c.f. Page 8).

Does Fig. (III) show that on the whole programmed instruction teaches at least as well as, if not better than 'conventional instruction'? No - the most it shows is that certain programmes were better than, equal to, or worse than certain instructors for certain students. No generalization can be made about the worth of programmed instruction as a whole versus 'conventional instruction' as a whole. The figures are restricted to certain programmes, certain instructors, certain students, and in no case can a generalization be made about different programmes, different instructors and different students; a good instructor will always beat a bad programme, and vice versa.

Fig. (III) RESULTS OF 112 COMPARATIVE STUDIES



(From J. Hartley, 'Research Report' - Hartley 1966)

Do the figures then show that, on the whole, certain programmes were better than or equal to certain instructors for certain students? Again, it can be doubted whether, in fact, this has been shown: it may well be that trainees worked harder at the programmes because of the effect of their novelty. It might well be that the instructor included material that was not included in the test, and was not given credit for this. And if the instructor had to keep to the same points as covered by the programme it might well be that this cramped his style, the effectiveness of his performance perhaps being decreased by such restrictions. A comparison experiment is exceedingly difficult to do precisely.

It has been stressed that the value of the comparison experiment is that it expresses something unfamiliar - the confidence ratio, the gain score, in terms of something familiar - the effects of the conventional classroom situation. The comparison experiment has no value outside this function. The point is that training is designed primarily to meet an objective; and training should be judged against how well it meets that objective. If one method of training is judged solely by comparing it with another, there is a danger of losing sight of that aim; a danger of saying 'this method is much better than that one' which may well be true, but leaves out of account the fact that neither of them may be particularly good, neither of them may give much help to trainees on their way to mastery of the subject. The standards for assessment of training techniques must be absolute, not relative. It is for this reason that although so much has been made of 'error rates' in assessing the teaching effectiveness of a programme, they are not a particularly useful measure of effect. What matters in the final count is how well the trainee does on that criterion test, i.e. on the objectives of the programme. And the error rate is only of use in assessing the effect of a programme in so far as it is related to that performance.

Does programmed instruction teach? There can be no doubt of that; students do learn from programmes, that at least has been shown; and their learning is the basic criterion for the effectiveness of any instructional method. As to how well they learn, that is difficult to measure precisely for the sort of considerations already given. It is difficult, not because of an inherent difficulty in programmed instruction, but because of the inherent difficulties of measuring how well students learn from <u>any</u> method of training. In particular, <u>the</u> criterion by which to assess any 'proof' of the effectiveness of any given method, is that the proof be <u>repeatable</u>; only then, when in a position to repeat the experiment exactly, getting exactly the same results, can one be sure that all

the variables inherent in the learning situation, have been accounted for; only then can one be sure that the measured gain is a direct result of the teaching medium, unenhanced by any variable that has been left unconsidered. Certain programmes do teach, there can be no doubt. And so do certain instructors. But it is not yet possible to say precisely how well either of them teaches.

References

A.E.R.A., A.P.A. & D.A.V.I. Joint Committee (1966) Recommendation for Reporting the Effectiveness of Programmed Instruction materials. N.S.P.I. Journal, 5, No.3, 3-9.

AUSTWICK, K. (1966) Validation of Linear Constructed - Response Program. Programmed Learning, 3, No.2. 88-93.

DODD, B.T. (1967) Programmed Instruction for Industrial Training. London: Heinemann.

ERAUT, M. (1966) The Behaviour of Analysis? N.S.P.I. Journal, 5, No.4. 6-7

HARTLEY, J. (1966) Research Report, New Education, 2, No.1

(1966) Some Guides for Evaluating Programmes. In: Assoc. for Prog. Learning, Programmes in Print. 186-192.

LUMSDAINE, A.A. (1965) Assessing the Effectiveness of Industrial Programmes. In: Teaching Machines and Programmed Learning, II Ed. Glaser, R. 267-320.

MARKLE, S.M. (1967) Empirical Testing of Programs. Programmed Instruction, The 66th Year Book of the National Society for the Study of Education. Part III. Univ. of Chicago Press. 104-110.