N82-13721 D56

INTERRUPTION AS A TEST OF THE USER-COMPUTER INTERFACE

By John G. Kreifeldt and Mary E. McCarthy*

Dept. of Engineering Design
Tufts University
Medford, MA. 02155

17th Annual Conference on Manual Control

SUMMARY

In any practical interaction with a computer, the user is required to formulate, phrase and enter the problem for solution through a series of steps. One very common and important characteristic of this procedure is that the operator may be interrupted at any point and required to attend to other tasks before resuming. Such interruption could have pronounced effects on the time and errors made while completing the computer task.

In order to study the effects different logic systems might have on interrupted operation, an Algebraic calculator and a Reverse Polish Notation calculator were compared when trained users were interrupted during problem entry. The RPN calculator showed markedly superior resistance to interruption effects compared to the AN calculator although no significant differences were found when the users were not interrupted.

Causes and possible remedies for interruption effects are speculated. It is proposed that because interruption is such a common occurrence, it be incorporated into comparative evaluation tests of different logic system and control/display system and that interruption resistance be adopted as a specific design criteria for such design.

INTRODUCTION

Designing the user-computer interface has been a decided challenge and will continue to assume growing importance. One of the vital aspects concerns interfacing the human user's cognitive abilities and characteristics with those of the computer, i.e. its logic program.

*The work reported here is based on the unpublished M. S. thesis of Mrs. Mary McCarthy.

Grace (1970.) suggested that a new relationship should be regarded as existing within a man-machine logic system, and that the logic component introduces a new set of considerations for the human factor discipline.

Man-computer cooperation in problem solving places emphasis on the design of the console devices for efficient communication of information between man and the computer. This mainly involves the computer display and the console manipulatory devices such as the switches, light pens, and plugs. The display and console manipulatory devices serve not only to transmit information between the human to the computer but also, in varing degrees to record and establish memory of this action and its consequence. (Zeigler and Sheridan, 1965.) The form and content of the information conveyed by these devices must be such as to aid the human in his thought processes, and thus make it easier for the operator to perform multitasks more efficiently.

In 1977, Durding, Becker and Gould studied the human memory skills as to how they aid or hinder data processing. They considered that memory organization depends on the task condition (computer memory) and if this were optimized then the human memory would be used more efficiently. However, at best, in the real world situation, the human mind seldom is without distraction, as for example unexpected interruptions. Enstrom and Rouse in 1977, did considerable work in analyzing how the human's allocation of attention will affect the input-output of the human relative to the computer.

It is obvious that the display and devices of the data processing machines immediately effect the transfer of information, yet there is a more fundamental area where man and machine meet, where the reasoning powers of man interact with the very essence of the machine, its logic system.

It is clear that this logic interfacing is less than ideal. However, one of the most significant challange is to specify useful design principles and evaluation criteria for logic design in addition to its functional characteristics in a manner similar to specifying physical characteristics such as button placement, size, etc., relating to proper computer hardware design for human use.

Common (and common sense) criteria for evaluating alternative logic design at the user level ordinarily include the number of key strokes required, number of errors made, and the time required to enter the problem and obtain the results.

Card (1979) found no difference between RPN (Reverse Polish Notation) and AN (Algebraic Notation) calculators on the basis of the time required to enter and solve problems of various degrees of complexity. However, a recent study (Agate & Drury, 1980) found significant differences in percentage errors and average completion times between AN and RPN with RPN showing the more favorable performance for both measures. On the other hand, a different study (Kasprzyk, Drury, Bialas, 1979) found smaller significant differences between AN and RPN with RPN still superior.

One of the testing principles in evaluating alternative designs that do not show clear performance differentiation is to apply increasing amounts of user or environmental stress, simulating conditions which could reasonably arise, and compare the alternative design performances. The design maintaining a higher level of performance over the wider range of stress could reasonably be judged the superior design. The assumption holds that the user is taxed less by the more stress-tolerant design, which would seem to be a sensible design criteria (Chapanis, 1959).

This principle of user stress was applied to a performance re-evaluation of an RPN and an AN hand calculator. The purpose of the experiment was not to evaluate particular calculators per-se but rather to devise a test methodology for evaluating different logic system designs in a practical, user-oriented manner. The calculators were used because they provided existing (and presumably optimized) examples of alternative logic designs for solving identical problems.

The "stress" applied to the user simulated real world conditions often encountered in which the problem solving task is interrupted part way through for some reason, with a subsequent diversion of attention. After the interruption, the operator is required to resume the task. Examples of this type of interruption are familiar to pilots, controllers, and others in a multi-task environment. A more homely example might be that of a housewife programming a microwave oven, being interrupted by a child or other distractions and then resuming the programming. In fact, anyone using a hand calculator has commonly experienced being interrupted during a lengthy series of key strokes with subsequent consequences.

The issue raised by the interruption is its impact on task completion. It is a reasonable assertion that a superior design would, among other virtues, show the least difference on the primary entry task between interrupted and uninterrupted performances. Common effects of an interruption include causing the user to loose the place, and/or to start over.

Thus a "good" logic design should, in principle, permit the user to resume the task after interruptions with negligible effects on subsequent performance. A corollary principle is that the user should be able to reprogram at any point making maximum use of previously programmed material. An example of this could be found among pilots programming an on-board navigational computer through a series of waypoints with associated altitudes, speeds and times. If interrupted during this task, the pilot should be able to resume as though no interruption occurred and/or, if necessary, reprogram at any time to meet new requirements. A clearly inferior logic design would require the pilot to begin again from the beginning effectively wasting any time previously spent on programming material which remains a subject of the existing problem.

There are at least two major components to a logic system for human use. One component is the logic system which on hand calculators manifests itself as the manner in which a problem must be phrased for entry (e.g. Reverse Polish Notation or Algebraic Notation). The other component is the console unit of input and display systems whose purposes include verification, feedback, and resultant display.

One could assume that the logic system should be compatible with the user's own "logic system" or cognitive characteristics of problem phrasing, whether learned or innate, while the console system should at least aid during the process of problem entry particularly if information processing is required (Ziegler, Sheridan, 1965). Except for very novice users, problems are never entered from a prewritten list of key strokes (except possible for lengthy programs entered into a programmable calculator). Thus since the hand calculator user works directly from mind to calculator, errors in solution (barring key stroke mistrikes) as well as the time taken to enter a problem (sequence of key strokes) could be attributable to mismatches between the human and machine logic systems and inadequacy of the display system. Any distraction interrupting the user during the entry process would put stress on the logic interface and on the adequacy of the console system.

The objective of our experiment was to test the effects of interruption during the calculation task on errors and completion time for AN and RPN logic systems as examples of different logic systems designed for the same purposes. The purposes of the experiment were to examine task interruption - a common occurrence - as a potential sensitive probe of performance with different logic systems and to shed further light on potential differences between AN and RPN calculators in a simulated multi-task environment. Based on the equivocal findings in performance between the two types of calculators when the user is not interrupted, the superior design would be expected to show the lesser difference between interrupted and noninterrupted performances.

METHODOLOGY

The Hewlett Packard Model 31E and the Texas Instrument Model TI57 hand calculators were chosen as examples embodying the RPN and the AN logic, respectively. Both calculators are examples of scientific calculators equipped to solve problems of considerable formulation complexity.

Four problems of different prima facie complexities were formulated to span a range of types that might be encountered. The problems were:

(1) $$\frac{5 \times (4 + (3 \times (2 - 6)))}{\frac{6 \times \sqrt{22}}{17} + \sin 37^{\circ}} = (-17.7206)$$

(2)

254
762
321
854
952
1859
2
400
37
369
824
592
333
1
(7460)

(3) $3 \times (4^{(2 - \sqrt[4]{7})}) = (4.7000434)$

(4) $\dfrac{1.2 \times 10^{16} + 3.45 \times 10^{14}}{(4 \div 5^{2}) \times 7 + (3 \times 5^{\cos 60^{\circ}})} = (1.577 \times 10^{15})$

The answers in parentheses were, of course, not presented to the subjects. At this point, the reader is encouraged to imagine which of the above problems, if any, might reveal the largest performance differences between the two calculators.

In order to contrast RPN with AN, the first problem is shown in Figure 1 as it would be entered into each calculator.

$$\frac{5 \times (4 + (3 \times (2 - 6)))}{\dfrac{6 \times \sqrt{22}}{17} + \sin 37^{\circ}} = -17.7206$$

REVERSE POLISH NOTATION			ALGEBRAIC NOTATION		
KEY	DISPLAY	COMMENT	KEY	DISPLAY	COMMENT
2	2.		5	5.	
ENTER	2.00		x	5.	Mult. pending
6	6.		(	5.	
-	-4.00	result of Subt.	4	4.	Add pending
3	3.		+	4.	Add pending
x	-12.00	result of Mult.	(	4.	
4	4.		3	3.	
+	-8.00	result of Add	x	3.	Mult. pending
5	5.		(	3.	
x	-40.00	result of Mult.	2	2.	
ENTER	-40.00		-	2.	Subt. pending
6	6.		6	6.	
ENTER	6.00		)	-4.	result of Subt.
22	22.		)	-12.	result of Mult.
$\sqrt{x}$	4.690	SQ. RT.	)	-8.	result of Add
x	28.14	result of Mult.	=	-40.	result of Mult.
17	17.		÷(	-40.	defines divisor
÷	1.66	result of Div.	6	6.	
37	37.		x	6.	Mult. pending

REVERSE POLISH NOTATION

KEY	DISPLAY	COMMENT
SIN	0.6018	results:SIN
+	2.26	result of Add
÷	-17.7206	result of final Division
	-17.7206	ANSWER

ALGEBRAIC NOTATION

KEY	DISPLAY	COMMENT
22	22.	
$\sqrt{x}$	4.69	results of SQ.RT.
÷	28.14	results of Mult.
17	17.	
+	1.66	result of Div.
37	37.	
SIN	0.60	result of SIN
)	2.26	
=	-17.7206	results of final Division

Figure 1. A Test Problem as It Would be Entered in Reverse Polish Notation and in Algebraic Notation.

The test procedure required two groups of 24 subjects each, with one group for each calculator. Each group was further divided into two subgroups of 12 subjects each; one subgroup to solve the four problems without interruption and the other subgroup to solve each problem with a standardized 1 minute interruption beginning 12 seconds into each problem. During the 1 minute interruption, subjects were asked to write the multiplication table of nines, eights, etc. After 1 minute elapsed, subjects were asked to stop multiplying and resume solving the test problems. All subjects had paper and pencil available at all times.

A short explanation of the experiment was read to each subject and questions answered. The subject was given the test calculator, paper, pencil and the four problems, each on a separate paper placed face down. Subject was instructed to turn over one test paper on GO and proceed to solve the problem. The Interruption subjects were interrupted 12 seconds into the problem by "INTERRUPT" to perform the multiplications for one minute. At RESUME, these subjects returned to the problem signalling its completion by saying "FINISHED". The noninterrupted subjects proceeded from GO to FINISHED without interruption.

Four times were recorded by stopwatch. T_0 = GO, T_1 = first key pressed, T_2 = first key pressed after RESUME ^{0}, T_3 = FINISHED. For the noninterrupted groups, only T_0, T_1, T_3 were recorded. Problem answers were recorded in all cases.

The 48 subjects had each been using a calculator for more than five years, most frequently on a daily basis. The subject population comprised undergraduate and graduate engineering students (60%) and faculty members (40%) of Engineering, Mathematics and Physics and one engineer from Hewlett-Packard. There were 27 males and 21 females ranging in age from 19 to 35 years. Subjects for each group (RPN, AN) were chosen from those owning and using a calculator with the respective notation system although all considered the particular test calculator as somewhat unfamiliar from the one they owned. Subjects were allowed to practice with the test calculator. Approximately 40% of the RPN subjects and 25% of the

AN subjects were faculty members although the faculty members reported using their calculators less frequently than did the nonfaculty subjects.

RESULTS

In comparing the two types of calculators the following six measures were computed.

(1) T_U - Total (uninterrupted) time to perform the problem

(2) T_I - Total time to perform the problem when interrupted (interruption time of 1 minute removed)

(3) T_N - Net effect on solution due to the interruption = $\bar{T}_I - \bar{T}_U$

(4) T_1 - Time to press first key

(5) T_2' - Time to press first key after the interruption ceases (T_2 - time at RESUME)

(6) E - Incorrect solutions.

Table 1 presents average and standard deviation values of the first three measures for each of the four problems as performed on the Reverse Polish (RPN) and Algebraic (AN) calculators. The fractional and unbiased percentage differences in solution times of each calculator are also shown.

The measures are computed for the 12 subjects in each of the four test sections.

TABLE 1

Interrupted An Uninterrupted Problem Entry Times For A Reverse Polish and An Algebraic Notation Hand Calculator.

	Avg (sec) / Stnd Dev. (sec)	Problem 1	Problem 2	Problem 3	Problem 4
Reverse Polish Notation	$\bar{T}_{I_{RP}}$	84.40 / 47.54	57.45 / 32.01	72.40 / 46.41	111.45 / 44.21
	$\bar{T}_{U_{RP}}$	77.58 / 34.06	45.67 / 25.32	69.85 / 38.02	110.35 / 23.13
	$\bar{T}_{N_{RP}}$	6.65 / 16.90	12.05 / 11.78	2.55 / 14.41	1.10 / 17.12
Algebraic Notation	$\bar{T}_{I_A}$	117.15 / 60.20	89.60 / 31.60	83.35 / 39.31	122.20 / 37.54
	$\bar{T}_{U_A}$	88.15 / 39.04	45.10 / 21.90	69.95 / 40.74	90.85 / 26.10
	$\bar{T}_{N_A}$	29.00 / 20.74	44.50 / 11.05	13.40 / 13.24	31.35 / 13.37
$\frac{\bar{T}_{N_A} - \bar{T}_{N_{RP}}}{\bar{T}_{N_{RP}}}$		3.36	2.69	4.25	27.55
$200 \times \frac{\bar{T}_{N_A} - \bar{T}_{N_{RP}}}{\bar{T}_{N_A} + \bar{T}_{N_{RP}}}$		125 %	114 %	136 %	186 %

Figure 1 presents just T_U, $\bar{T}_I$ and $\bar{T}_N$ from Table 1 for both calculators. The test method allows only one sample of T_N to be computer for each problem in each group. However, the variance of T_N was computer as the sum of the two independent sample mean ($\bar{T}_U$, $\bar{T}_I$) variances and the corresponding standard deviation is shown in Table 1 for each T_N value.

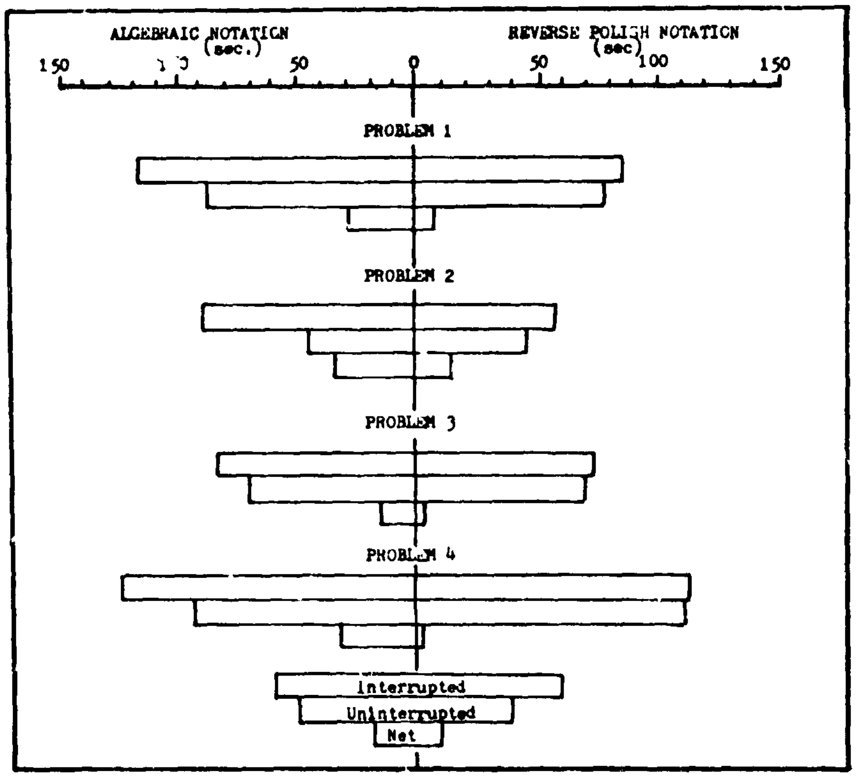

Figure 1. Interrupted and Uninterrupted Problem Entry Times for Reverse Polish and Algebraic Notation.

Both Table 1 and Figure 1 show that there was a consistently greater lengthening of solution time (T_N) on each problem for the Algebraic Notation compared to the Reverse Polish calculator when their operators were interrupted at the task. This difference was statistically confirmed by a student t test using the T_N value for each problem as sample data points and comparing the four values from the RPN with those of the AN calculator. ($t = 5.037$, $p < .02$). Table 1 shows that this statistically significant finding also represents a large unbiased percentage difference of at least 114% for the two calculators.

A similar student t test showed that the uninterrupted times did not differ significantly between the two calculators which is in agreement with several previous findings but contrary to Agate & Drury. Although the interrupted solution times (T_I) were consistenly longer on each problem for the Algebraic compared to the Reverse Polish Notation calculator, this difference was not significant with the exception of problem number 2 ($t = 2.47$, $p < .05$).

The data were also reduced to obtain the time required before pressing the first key at problem start (T_1) and again following the end of the interruption (T_2'). The latencies are shown in Table 2.

TABLE 2

First Key Press Latency Times At Problem Start And Following the Interruption.

		PROBLEM				
		1	2	3	4	AVERAGE
START	$\bar{T}_{RP_1}$	8.15	4.2	6.47	7.45	6.57
	$\bar{T}_{A_1}$	7.15	3.6	8.65	6.18	6.40
INTERRUPTED	$\bar{T}'_{RP_2}$	4 80	4.35	5.05	5.29	4.82
	$\bar{T}'_{A_2}$	3.45	2.20	4.95	5.55	4.04

The only consistent finding among these latencies is that resuming the entry task after interruption required less latency than when first beginning it and the addition problem required the least latency of the four problems. There was no significant difference between the initial or between the post interruption latencies for the two calculators. It might have been assumed that the greater lengthening of entry time under interruption for AN compared to RPN would correspond to a similarly greater post interruption latency. However, this is not the case. If anything, the AN latencies appear a bit shorter (not significantly) rather than longer compared to those for RPN.

Incorrect solutions were also compared for the two calculators on each problem interrupted and uninterrupted and the results are set out in Table 3.

Although problem two showed the fewest incorrect solutions out of 12 (except for the AN - interrupted) only the Reverse Polish calculator - uninterrupted) showed a statistically significant difference between the problems ($x^2 = 7.33$, $p = < .06$).

TABLE 3

Number of Incorrect Solutions (Out of 12 Possible/Problem) For The RPN And AN Calculators Working With And Without Interruption.

		PROBLEM				
		1	2	3	4	AVERAGE
INTERRUPT	RPN	4	2	5	8	4.75
	AN	4	5	9	8	6.5
NON INTERRUPT	RPN	4	1	4	9	4.5
	AN	7	1	7	7	5.5

DISCUSSION

Previous experiments were equivocal whether Reverse Polish Notation had an advantage over Algebraic Notations regardless of problem complexity. The no difference findings were reconfirmed in the present experiment when sugjects were not interrupted during problem entry. When subjects were interrupted, as might naturally happen in real life the problem entry time increased for both notation systems (calculators) over that required under no interruption. However, the average increase for the Algebraic Notation calculator was more than twice as great as that for Reverse Polish Notation for the addition problem at the least and twenty-seven times as great at the most, or one of the complex problems. Regardless of the problem, the increase in entry time was longer for the AN calculator. Coincidently, although not statistically significant, more incorrect solutions occurred with the AN calculator as well.

Although it is tempting to attribute the difference in entry times to the nature of logic or notation it is difficult to support this conclusion directly. If one or the other notation system is per se more compatible with the subjects' own mode of problem organization, difference might be expected in the respective times taken to press the first key for each calculator either at the start of the problem or resuming after interruption. Presumably, these times might reflect the amount of "mismatch" between human and machine "logic" systems. However, these respective times were nearly identical for each calculator. (Again, and although not statistically significant, the addition problem required the least amount of this latency time which was also consistently less following the interruption than when just beginning the problem.) The negligible differences between key latencies, and between uninterrupted solution times suggests that the user can adapt to either notation system.

It is, in fact, difficult to find the mechanism of the entry time lengthening (T_N) in the interrupted mode. Neither the recording nor observational methods were sufficiently detailed enough to monitor each of the individual key entry times. Examination of such a record might have suggested plausible mechanisms for these lengthenings. At present, and for lack of more detailed data, it could be assumed that any lengthening (T_N) is manifested as a uniform slowing of entry, for whatever reason, after interruption. It may be that problem recapitulation is necessary after the interruption and that the AN system places more memory stress on the user resulting in a longer time to execute the problem. With the increased memory burden adding to the task difficulty, resources invested in doing the entry task can do less, decreasing user efficiency (Novon and Gopher, 1979).

However, any mismatch between the human logic system, and the AN (or RPN) calculators thus can also be sought in key board and/or display differences.

A potential explanation of the solution time differences under interruption involves display rather than key board differences. The simple addition problem of [5 + 2 = (7)] as entered and displayed on each calculator is instructive for this purpose as shown in Figure 2.

RPN (HP)			AN (TI)		
Key	Display	Comment	Key	Display	Comment
5	5.		5	5.	
ENTER	5.00	'.00' shows that the number has been entered	+	5.	no indication to show if + has been keyed.
2	2.		2	2.	nothing to indicate 2 is not a resultant.
+	7.00	'.00' indicates 7 is not an entry but must be a resultant	=	7.	does not show if 7 is an entry or an answer.

Figure 2. Sample Entries of the Same Problem on HP(RPN) and TI(AN) Calculators.

As presently constructed, the HP (RPN) calculator display differentiates between resultants and entries while the TI (AN) calculator does not. Of course, if the entry had exactly as many decimal places as the displayed resultant, this difference would no longer hold. A second display difference appears in the familiar left-to-right concatenation of digits as they are entered on the HP mimicking the normal mode of writing as opposed to the left shift of the entire digit string as a new numeral is entered on the TI calculator.

The HP display features may be a better aid to "place holding" during interruption than those of the TI calculator permitting a more rapid recovery as the problem is resumed.

At present it is not unequivocably possible to attribute the large percentage differences in interrupted solution times to notational or display differences or both for the two calculators.

However, from a practical point of view, the HP calculator (RPN notation and display) is superior in terms of the shorter time taken to complete a problem when interrupted - a common real life occurrence - and would be recommended where many such computations must be made or where time is a critical factor.

Since task interruption is a fact of life, and since it also degrades performance, it is reasonable that the logic/display design be as performance resistant to the interruption as possible. It is suggested that on computer systems without printers, the display show simultaneously at all times the last datum entry and operation as well as the current resultant and that a review key permit sequential viewing of the previous few datum + operation pairs as though a key stroke list had been internally compiled for later display. Although no definite suggestion can be made for the notational system itself, it is possible that some hybred of the RPN and AN may be more compatible with the user's own modes of problem phrasing and recall.

Further studies will be made of the notational and display features unconfounded in search of the "ideal" calculator which at this point could be said to produce identical interrupted and uninterrupted solution times regardless of problem complexity in addition to other desirable features. For design and evaluation purposes, it is suggested that interrupted solution times serve as one of the performance criteria.

REFERENCES

(1) Agate, S., Drury, C., Electronic Calculators - Which Notation is Better? (1980) Applied Ergonomics 11 (1), March.

(2) Card, Stuart, K., A Method for Calculating Performance Times for Users of Interactive Computing Systems. (1979) Proc. 1979 International Conference on Cybernetics and Society. Denver.

(3) Chapanis, A., Research Techniques in Human Engineering (1959) Johns Hopkins Press.

(4) Durding, B., Becker, C., Gould, J., Data Organization (1977). Journal of the Human Factors Society, 19 (1), 1-14.

(5) Enstrom, K., Rouse, W., Real Time Determination of How a Human Has Allocated His Attention Between Conrol and Monitoring Control. (1977) IEEE Transactions on Systems, Man and Cybernetics, SMC-7 No. 3.

(6) Grace, G. L., Preface to Human Factors Journal Special Issue on Information Processing Systems. (1970) Journal of the Human Factors Society. 12(2), 161-164.

(7) Kasprzyk, D., Drury, C., Bialas, W., Human Behavior and Performance in Calculator Use With Algebraic and Reverse Polish Notation, (1979), Ergonomics, 22(9) 1011-1019.

(8) Novon, D., Groper, D. Interpretation of Task Difficulty in Terms of Resources, Efficiency, Load, Demand, Cost, Composition. (1979) Technion Institute, Haifa, Israel.