Applied Ergonomics 1980, **11**.1, 2–6

Electronic calculators: which notation is the better?

S.J. Agate and C.G. Drury

Department of Industrial Engineering, State University of New York at Buffalo

Tests of an Algebraic Notation Calculator and a Reverse Polish Notation Calculator showed the latter to be superior in terms of calculation speed, particularly for subjects with a technical background. The differences measured were shown not to be due to differences in calculation speed of the calculators nor to differences in dexterity between the subjects.

Introduction

The advent of the electronic calculator has had a tremendous effect on the working life of many scientists and technologists and is now beginning to affect the lives of ordinary consumers. The change from the cams and gears of electromagnetic machines to integrated circuits has brought with it large changes in the industry manufacturing these devices, (Anon, 1976), but more importantly it has forced each manufacturer to opt for a particular logic routine. This is the set of rules governing the sequencing of input to the machine by the operator.

The two logic routines in general use are Algebraic Notation (AN) and Reverse Polish Notation (RPN). In the former, keys are pressed in the order of evaluating simple expressions, thus for 2 + 4 = 6 we have:

[2] [+] [4] [=]

and the answer 6 appears on the display.

For Reverse Polish Notation, named after the Polish mathematician Lukasiewicz, the operator follows the operands, thus:

[2] [ENTER] [4] [+]

and again the answer 6 appears on the display. (For calculator dimensions and layout see Figs. 1 and 2.)

Advertisements for calculators with Algebraic Notation stress the naturalness of left-to-right entry of operands and operators but usually do not point out that parentheses may be needed in complex calculations for correct results. An advertisement appearing in the May 1976 issue of *Datamation* states: "Texas Instruments chose Algebraic Notation because it is the most ***natural*** to use with easy left-to-right entry." The rival RPN advertisements claim that a calculation can be started from any point and that it is the most efficient method known for evaluating arithmetic expressions. Engineers at the Hewlett-Packard Co, manufacturers of the HP-35 (RPN) calculator, state: "First, as Reverse Polish Notation is scanned from left to right, every operator that is encountered may be executed immediately. This is in contrast to notation with parentheses where execution of operators must be delayed.' (Whitney, Rode and Tung, 1972.)

It is surprising that the two rival sides of the pocket calculator industry should not have published experimental data to support their claims. A literature search revealed two articles in *Consumer Reports* (1975a and b) suggesting that RPN "takes some getting used to" and that a simple rather than a scientific calculator should be used for simple calculations such as checkbook balancing. Another article published by the Consumers' Association (*"Which?"*, 1973), where 20 different calculators were tested for various operating functions that a person should look at

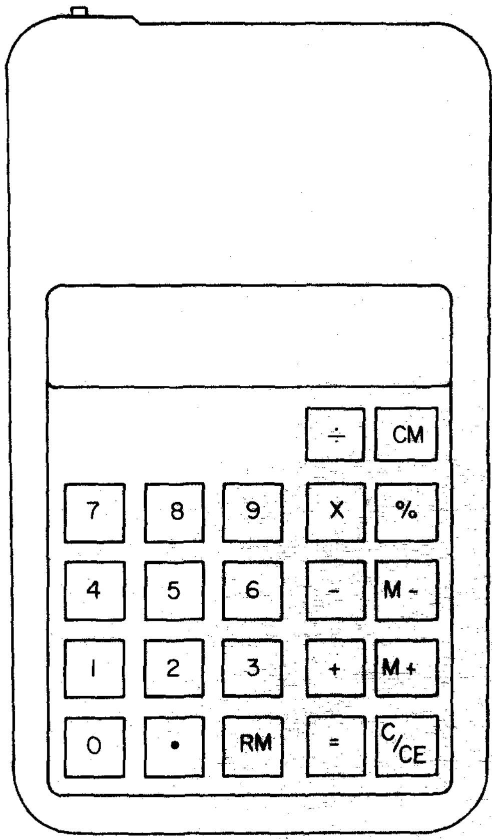

Fig. 1 AN (approximately ⅔rds actual size)

0003–6870/80/01 0002–05 $ 02.00

prior to buying a calculator, found that after doing sums on tested calculators there was little to choose between the AN and RPN calculator once an operator got used to them.

The only study remotely relevant (Anon, undated) is one by a manufacturer of RPN machines showing them to be 5 to 10 times faster than a slide rule. No data are given to support this rather obvious finding.

It was thus felt that a controlled study involving selected groups of industrial users on AN and RPN calculators would provide valuable information to those wishing to purchase equipment fitted to human capabilities and limitations.

Method

While RPN calculators tend to be aimed more at the scientific end of the market and AN at the simple end, this is certainly not a universal finding, with examples of both logic routines appearing throughout the market. It is thus necessary to test both machines on both groups of potential users to obtain any generality of findings. Accordingly, two major groups of subjects were chosen from a large industry in the vicinity of Buffalo, New York. Both groups consisted of 10 subjects and all had daily experience of using various calculators. However, one group was recruited from the accounting department and used calculators mainly for fairly simple calculations. This was labelled the Non-Technical Group. The second group, called the Technical Group, was recruited from the engineering staff and used calculators for more complex calculations.

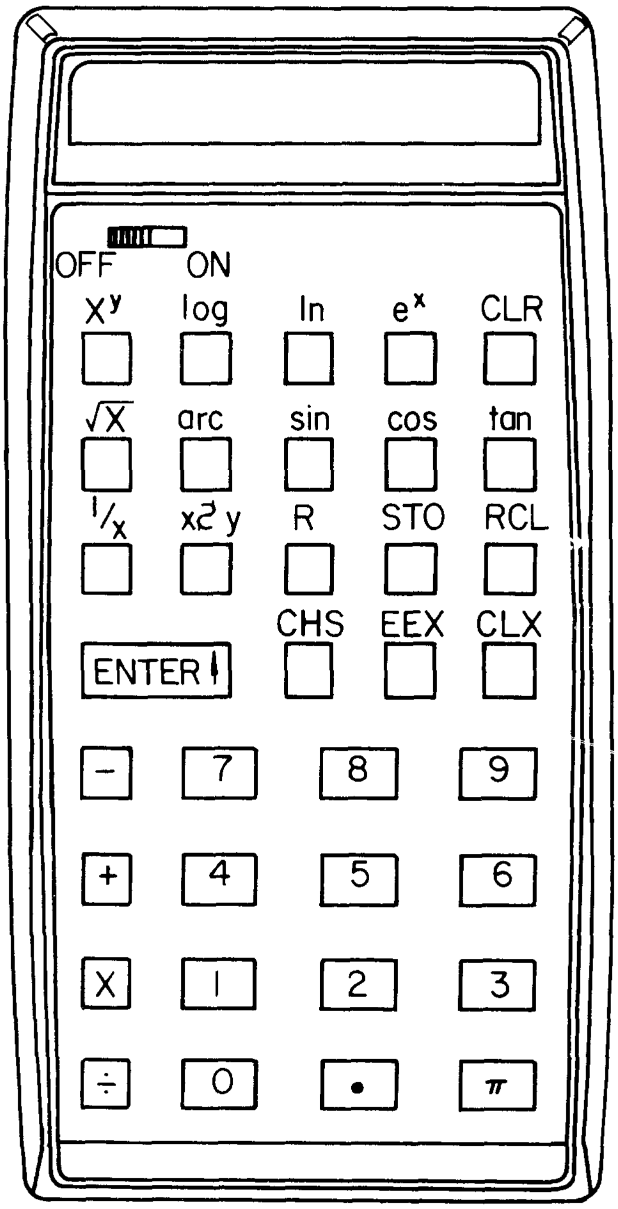

Fig. 2 RPN actual size

It is difficult to find two calculators exactly matched for all factors except their logic routines. The two selected for this study are labelled AN and RPN. The RPN calculator had a wider array of scientific functions (eg, Log, Sin, Cos, etc). The use of two specific calculators causes a number of variables, such as keyboard layout, to be confounded with logic systems. To check, at least in a rough and ready way, operational differences between the two calculators in terms of mere mechanics of manipulation, ten subjects (five associated with each calculator) performed an addition of three numbers of three digits each. The results showed no difference in mean times between the two calculators ($t = 0{\cdot}48$, 18 df, $p > 0{\cdot}10$)

A more obvious confounding factor was the different memory systems employed by the two calculators. Both had an addressable memory to store intermediate results and both, of course, had the working storage corresponding to the display. However, the RPN calculator had an implied 'stack' wherein intermediate results of sequential calculations were stored automatically and retrieved automatically when the appropriate operator button was pushed. This 'stack' had three locations in addition to the display.

Because of the transfer of training effect that exists when a subject operates one type of calculator and then operates another type, all subjects used only one calculator throughout the experiment. The four groups each consisting of five subjects were employed as follows:

Technical, AN
Non-Technical, AN
Technical, RPN
Non-Technical, RPN

Each group was given training with feedback on their assigned calculator before the experiment.

The experiment itself consisted of measuring the times and errors of each subject on ten questions which made up a 'basic' test. The questions ranged from the very simple such as:

$$(3{\cdot}3 + 4{\cdot}5)\ (5{\cdot}2 + 6{\cdot}1)\ (7{\cdot}3 + 8{\cdot}4) = \underline{\qquad\qquad}$$

to the more complex such as:

$$3{\cdot}14 \times 435 \times 3{\cdot}5((1{\cdot}24)^4 - (0{\cdot}8)^4)\,/2 = \underline{\qquad\qquad}$$

All four groups of subjects performed this basic test.

In addition, the extra functions available on the RPN calculator were utilised by having both RPN groups (ie, the Technical and Non-Technical) also perform a 'scientific' test. This consisted of 10 items, ranging from questions such as

$$e^{3{\cdot}8} = \underline{\qquad\qquad}$$

to more complex questions such as

$$\frac{\text{Sin}\ (30{\cdot}5) \times \text{Cos}\ (150)}{\text{Tan}\ (26{:}6)} = \underline{\qquad\qquad}$$

All questions were selected from the manuals supplied with the calculators.

Table 1: Summary of analyses of variance for the Purdue pegboard tests.

Source of variance	df	Right hand mss	F	p	Left hand mss	F	p	Both hands mss	F	p	Assembly test mss	F	p
Between groups	3	2·53	1·58	>0·10	0·18	0·09	>0·10	0·05	0·006	>0·10	4·93	0·25	>0·10
Within groups	16	1·60			2·05			8·03			19·88		

None of the between group differences is significant at p = 0·10

To check that the groups were well matched in the basic motor abilities (ie, finger dexterity) required in calculator manipulation, all subjects were given the four tests of manual dexterity as defined in the manual of the Purdue Pegboard Test (Salvendy and Seymour, 1973).

Results

The Purdue Pegboard Test scores (on the four tests defined in the manual) were analysed using a one-way analysis of variance. The results are summarised in Table 1, where it can be seen, for each Purdue Pegboard Test, that the four groups had very similar mean scores.

This result, together with the result of the three digit addition test given earlier, makes it implausible to assume (were any differences in performance to arise between the groups) that these differences could be attributed to inherent differences in physical manipulations required by the calculators, or to inherent differences in manipulative ability between the four subject groups.

The main results of interest are the speed and accuracy scores on basic tests. Table 2 shows the summary of analysis of variance carried out on scores of accuracy (number correct out of 10 questions) and speed (total duration in minutes) in the basic test.

A similar pattern of results is shown for both analyses. Statistical significance is reached at the $p < 0{\cdot}001$ level for differences in mean speeds between calculators. Differences in mean number of errors between calculators is only significant at $p < 0{\cdot}05$. For both speed and error scores the RPN calculator is better than the AN. Fig. 3 shows mean speed and mean percentage error scores on each calculator for each group of subjects.

Overall the RPN calculator groups took 67% of the time required by the AN calculator groups to complete the basic test. The error rate of the RPN groups was about half that of the AN groups.

Although the interaction between calculators and groups is not significant for either speed or accuracy (see Table 2) there is a tendency for the Technical Group to benefit more from the RPN than the Non-Technical Group for both performance measures (see Fig. 3).

The next question is whether, for the RPN calculator, the scientific test changes the ordering of the Technical and Non-Technical groups. Speed and accuracy scores for the basic and scientific tests were analysed for the groups using the RPN calculator. The results are shown in Table 3.

These findings on the RPN calculator show that the differences between groups were relatively small on the basic and scientific tests when performance was measured in terms of errors. The difference between groups was large for the scientific test, but not for the basic test, when the performance measure was speed.

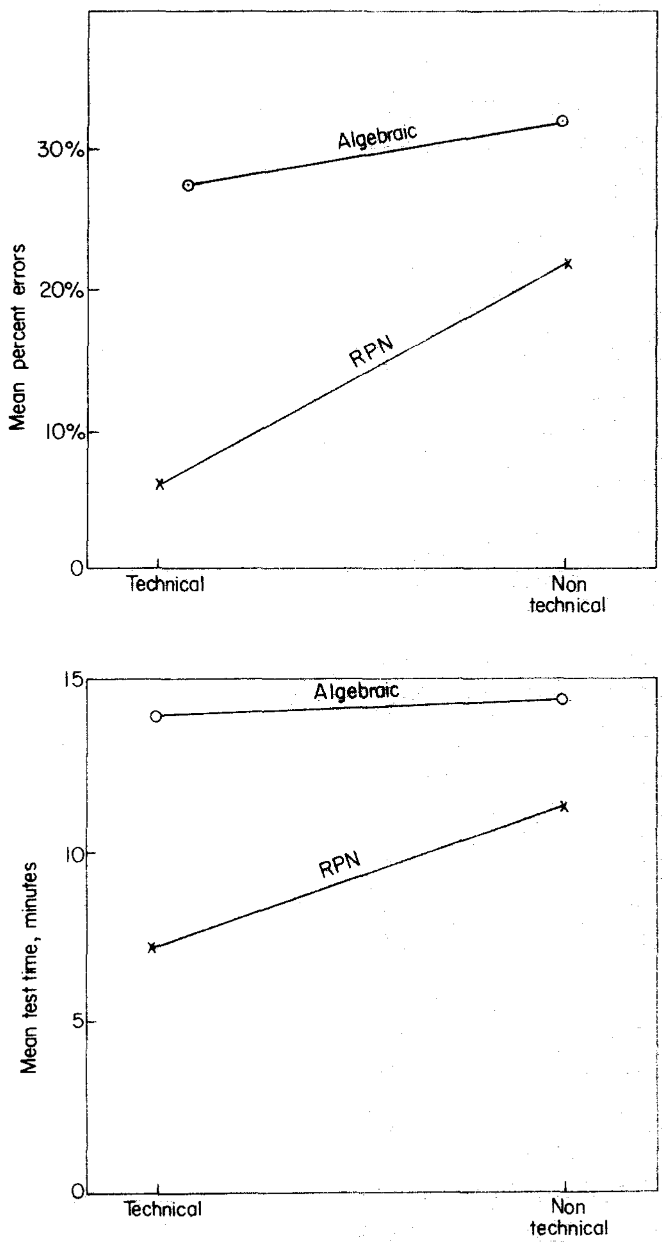

Fig. 3 Error and time scores on the basic test

Table 2: Summary of analyses of variance for the basic test

		Speed			Accuracy		
Source of variance	df	mss	F	p	mss	F	p
Between calculators (C)	1	109·28	9·75	$< 0·001$	12·80	5·01	$< 0·05$
Between groups (G)	1	23·87	2·20		5·00	1·92	
C x G	1	16·73	1·54		1·80	0·69	
Within groups	16	10·86			2·60		

Table 3: Summary of analyses of variance for the RPN calculator

		Speed			Accuracy		
Source of variance	df	mss	F	p	mss	F	p
Between tests (T)	1	1·78	0·38	–	0·45	0·41	–
Between groups (G)	1	200·38	55·94	$< 0·001$	6·05	5·62	$< 0·05$
T x G	1	22·37	6·25	$< 0·05$	1·25	1·16	–
Within groups	16	3·58			1·08		

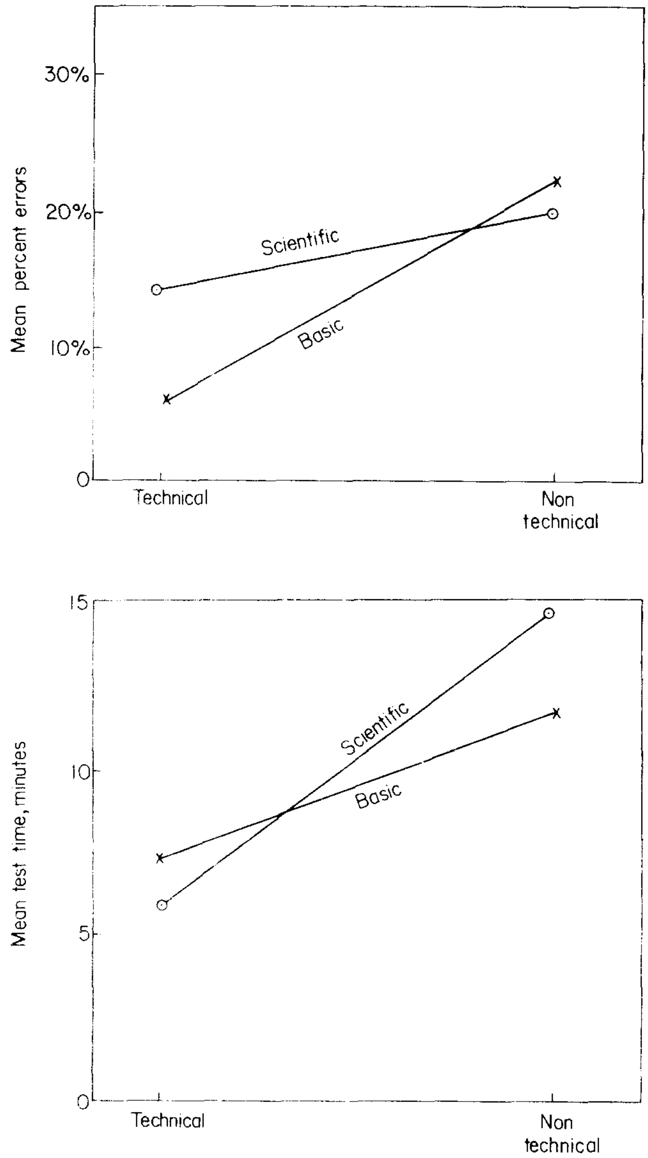

Fig. 4 Error and time scores on the RPN calculator

The Technical and Non-Technical groups are significantly different ($p < 0·001$ for speed and $p < 0·05$ for accuracy) with the Technical Group faster and more accurate on both the basic and scientific tests. There is no significant difference between the tests, a somewhat irrelevant finding since there was no special attempt to make the tests comparable in terms of speed or inherent difficulty. There is a significant interaction ($p < 0·05$) between tests and groups for speed with the two groups being more widely separated on the scientific test than on the basic test. However, as shown in Fig. 4 this could be due to a speed/accuracy trade-off as the direction of the interaction is different for speed and accuracy scores.

Discussion and conclusions

The difference observed between the AN and the RPN calculations showed a clear superiority in speed and accuracy in favour of the RPN calculator, statistically significant even with the modest sample sizes used. If this difference is to be attributed specifically to logic system differences then other potential confounding factors need to be considered and eliminated. There are broadly two groups of confounding factors, those pertaining to the subjects and those pertaining to the calculators.

The subjects could be unrepresentative, but similar results were obtained on two groups with widely different histories of calculator use, so that it seems an unlikely explanation. There is still the possibility that most of those tested were regular RPN users, but this too is unlikely in view of the smaller numbers of RPN calculators sold in the USA. The differences between subject groups in the main AN/RPN comparison were not due to manual dexterity differences, which were very similar for all groups. Finally, the experimental design eliminated unwanted transfer-of-training effects by using independent groups on each logic system.

The calculators themselves were of course different in ways other than the logic systems used but these differences were small enough not to affect manipulation time on the simplest test possible – addition. Perhaps the major confounding variable was the presence of an automatic memory stack in the RPN calculator. This stack is inherent in the use of all RPN calculators and its effect independent of the logic system could only be measured with the aid of a purpose-built experimental calculator.

Thus, the observed superiority of the RPN calculator is unlikely to be due to subject differences and represents a true difference in favour of that calculator. This difference may be due to the automatic stack or the logic system itself but, from a practical point of view, when you buy one, you buy the other.

In terms of practical choice between calculators, it would appear that RPN is faster and more accurate overall but particularly for technical users and particularly for more complex problems.

Statistical significance has been demonstrated; practical difference still remains. A test of many brands and designs of calculators, of many degrees of technical sophistication of user and of many training levels of user would be needed to establish a practical finding under all circumstances. In many instances, the choice of calculator rightly may be more influenced by user-specific technical features than by type of logic. But the results remain that under controlled and reasonably representative conditions there *is* a difference in ease of use between the logic routines.

Acknowledgements

The authors wish to thank the Department of Industrial Engineering, State University of New York at Buffalo and the Western Electric Co, Inc, for support in performing and publishing this study.

References

Anon
Undated A capability study on the HP-35 pocket calculator, Hewlett-Packard, Inc, Cupertino, California.

Anon
1976 *Business Week*, June 14, A calculator loser seeks a winning formula.

Consumer Reports
1975a *Consumer Reports,* **40,** No 9, September. Calculators.

Consumer Reports
1975b *Consumer Reports,***40,** No 12, December. Calculators.

Which?
1973 Small Electronic Calculators. Consumers' Association December.

Datamation
1976 *Datamation,* 22.25, May. Advertisement.

Salvendy, G., and Seymour, W.D.
1973 Prediction and development of industrial work performance. Wiley, New York. pp 245–258.

Whitney, T.M., Rode, F., and Tung, C.C.
1972 *Hewlett-Packard Journal,* June. The 'Powerful pocketful': an electronic calculator challenges the slide rule.