SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 107-ONR
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Analysis of the Decision and Verification Stock-market Tasks using Additive Stage Process Methods
5. TYPE OF REPORT & PERIOD COVERED: Technical Report 1/1/81 - 6/30/81
6. PERFORMING ORG. REPORT NUMBER: ICS Tech. Report
7. AUTHOR(s): Murray Singer, University of Manitoba, Canada
8. CONTRACT OR GRANT NUMBER(s): N00014-78-C-0433
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Institute of Cognitive Science, University of Colroado, Campus Box 345, Boulder, CO 80309
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 157-422
11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Programs, Office of Naval Research (Code 458), Arlington, VA 22217
12. REPORT DATE: June, 1981
13. NUMBER OF PAGES: 35 pages
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Reading, comprehension, memory, decision, model processing.
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Two additive stage processing models are presented that address some of the tasks reported in Antos, Kozminsky, Bourne, and Kintsch (1981). The development and application of these models to information analysis and decision procedures demonstrates their usefulness in structuring thinking and experimental design about critical, complex, real-life sorts of decisions. Both alternative processing models presented were constructed by carefully thinking out what mental operations are necessary to perform

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) information analysis, decision and verification. Predictions are generated from the models and those predictions were compared to the data reported in Antos, et al. While some predictions of the model are supported, new experiments are needed to test more critical predictions. The intention of this work is to lay some of the theoretical groundwork for designing experiments that would provide information on the critical elements in how people understand facts in order to make decisions.

ABSTRACT

Two additive stage processing models are presented that address some of the tasks reported in Antos, Kozminsky, Bourne, and Kintsch (1981). The development and application of these models to information analysis and decision procedures demonstrates their usefulness in structuring thinking and experimental design about critical, complex, real-life sorts of decisions. Both alternative processing models presented were constructed by carefully thinking out what mental operations are necessary to perform information analysis, decision and verification. Predictions are generated from the models and those predictions were compared to the data reported in Antos, et al. While some predictions of the model are supported, new experiments are needed to test more critical predictions. The intention of this work is to lay some of the theoretical groundwork for designing experiments that would provide information on the critical elements in how people understand facts in order to make decisions.

This report presents additive stage process models which address several of the tasks used in experiments reported by Antos, Kozminsky, Bourne, and Kintsch (1981). These models make use of the principles embodied in several familiar models of sentence verification (e.g., Carpenter & Just, 1975). The theoretical character of such models has, furthermore, been carefully examined by S. Sternberg (1969). The role of the models is to specify the mental operations necessary for the performance of the tasks in question, mainly decision or verification tasks. Models of this sort generate predictions concerning the relative time that a subject should need to respond in different experimental conditions. To the extent that these predictions are confirmed by the corresponding data, the models are viewed as credible. The immediate goal is to present a few models of this sort in order to illustrate their essential principles and assumptions. It will then be possible to compare the predictions generated by the models with data already collected.

Table 1 provides an overview of all four sessions in the experiment. Table 2 shows the basic design of texts in the decision task. Table 3 shows an example of an assembled text. Finally, Table 4 shows the basic design of texts used in the value verification procedure. For further methodological details, the reader is referred to Antos, et al. (1981).

Insert Tables 1 - 4 About Here

Decision Task

As discussed on page 20 of Antos, et al., the reaction time measures RT1 and RT2 mean different things for RA and RB conditions. Because RT2 represents, in both cases, the phase at which a response is finally made, it is of greater interest in the current discussion. For the RB group, RT2 refers to the time needed to read the text and make a decision, while, for the RA group, it refers to the time needed to read the rule and make a decision. Reading the whole text involves many more mental operations than reading the rule, resulting in much higher variance (as well as magnitude) for the RB group's RT2 scores. For this reason, it is convenient to concentrate on the RA group at this time.

While terminating versus exhaustive reading is a strategic option in the rule-before condition, the rule-after subject has little choice but to read each sentence of a passage and strive to store the categories discussed, along with their values. It turns out, nevertheless, that the reader may proceed in more than one way. First, it is possible that, for each sentence, the subject stores only the category discussed explicitly, along with the value provided (e.g., capitalization--high). Second, it is possible that the reader computes the value associated with any categories correlated with the three explicit categories. While it might be argued that, ignorant of the rule, it would be useful for the RA subject to compute all correlated values, it is certainly possible to delay this until the rule is presented. Among other things, this would reduce the subject's memory load from five to three categories.

Let us examine the RT2 decision scores for the RA group. For each of the BUY, NOT BUY, and INSUFFICIENT INFORMATION problem types, response latencies differed considerably as a function of the number of explicit categories (see Table 1). If the RA group computed the value of correlated categories while reading the problem, then RT2 should not vary as a function of the number of implicit categories. For this reason, it will be assumed, for now, that RA subjects do not compute the value of correlated categories in the course of reading the problem.

Figure 1 shows a process model that addresses the RA group decision task, with the assumption that the values of correlated categories need to be computed at decision time.

Insert Figure 1 About Here

Stages 1 to 5 refer to the reading of the problem and thus do not address RT2. Let us, however, briefly consider these stages. Stage 1 is not a mental operation, but rather refers to the initialization of counters for "sentence within the passage" (_s_) and "relevant categories" (i). The model of Carpenter and Just (1975) includes such counters. Since each passage consisted of three sentences, _s_ has the possible values 0, 1, 2, and 3.

At stage 2, the sentence counter is incremented. On the first cycle, then, $s=1$. The subject then encodes the sentence to its propositional form at stage 3. Since the subject does not know the rule, values are stored, at stage 4, for each category mentioned in the passage. Finally, we arrive at stage 5, at which it is determined whether there are any more sentences to be read.

The portion of Model A that addresses the decision (and RT2) begins at stage 6, which involves the reading of the rule. The rule mentions two relevant categories, which can be referred to as R and R . At stage 7, the relevant-category counter, i, is incremented. On the first cycle, then, i equals 1.

An important assumption will be adopted with regard to the remaining processing stages. It states that the subject will proceed toward a solution in the most logical and efficient manner. Let us briefly consider two implications of this assumption. (1) Because of the nature of the conjunctive concept, the logical subject can make a NOT BUY decision upon the discovery of just a single relevant category that is negative. For this reason, the model postulates that processing may terminate as soon as a single negative relevant category is identified. (2) Similarly, processing may terminate in an INSUFFICIENT INFORMATION decision as soon as it is discovered that no information has been stored for a single relevant category.

Stage 8 is a test which asks whether a value has been stored for relevant category R . (This operation resembles one shown to play an important role in sentence verification by Singer (1981)). If no such information has been stored, it is asked at stage 9 whether R is correlated with another category. If not, it is established that no information is available for R . Control then flows to stage 12 at which a response index, initialized at BUY (Clark & Chase, 1972) is changed to "insufficient information." That response is registered at stage 13.

If it is established at stage 9 that R does have a correlated category, then the value of that category is retrieved, if possible, at stage 9'. If there is no value, control again flows to stages 12 and 13. If there is such a value, however, the value of R can be computed.

Regardless of whether the preceding operation was 8 or 9', it is asked at stage 10 whether the value of R is negative. If so, the response index may be immediately changed at stage 12 and a NOT BUY response registered. If the value is positive, test 11 asks whether there are any more relevant categories to examine. If not, then it is established that both relevant categories are positive: a BUY response is registered at stage 13. If there is another relevant category to examine, i is incremented (to 2) at stage 7. Processing continues in the way just described.

Predictions

It is easiest to illustrate the predictions of Model A within, rather than between, response types. Consider the BUY response, for example. As discussed elsewhere in this report, BUY problems differ only with respect to the number of implicit categories they include (i.e., the number of categories whose value must be inferred from a correlated category). Model A specifies identical processing for all BUY problems except in that each additional implicit category requires the execution of stages 9 and 9'. The model clearly predicts that, for BUY problems, 0-implicit RT2 values will be fastest and 2-implicit values will be slowest.

The actual mean correct decision times for BUY problems were 3.3, 5.9, and 4.8 seconds for the 0-, 1-, and 2-implicit conditions, respectively. The prediction that 0-implicit latencies would be fastest is confirmed, while the 1- versus 2-implicit prediction is reversed in the data. It should be considered, however, that the low correct proportion of .56 casts some doubt on the usefulness of the mean value for the 2-implicit condition. The reason for this is that, with high error rates, it becomes increasingly likely that the proportion correct (i.e., .56) includes numerous trials on which the subject did not perform the problem as specified by the model but, rather, guessed the correct answer (see Table 5).

Insert Tables 5-6 about here

Model A also generates a clear-cut prediction for the INSUFFICIENT INFORMATION problems. It states that decision time will be faster for the 2-missing problems than for any other type. The reason for this is that the subjects who have grasped the problem will be able to register the INSUFFICIENT INFORMATION response as soon as a single category is discovered to have no value. The processing sequence for such a problem would be 6-7-8-9-12-13 (see Figure 1). For all other INSUFFICIENT INFORMATION problem types (e.g., 1 explicit, 1 missing), there is a probability of .5 that the missing category would not be examined first, resulting in longer decision times. The actual response latencies were 3.8 sec for the 2-missing problems and 5.1 sec for all other problems.

Verification Task

Recall that in Session 4, subjects performed a verification task. After reading an antecedent passage, the subject had to judge the truth of a statement that a particular category had a specific value (e.g., capitalization was high). This task is addressed by Model B, shown in Figure 2. At stage 1, the test sentence is encoded in its propositional form. Test

2 asks whether any information was stored for the category discussed in the test sentence. if no information was stored, test 2' asks whether the test category is correlated with any other category, and if so, whether a value has been stored for the correlated category (test 2"). If either 2' or 2" fail, control flows to the response index change operation 4, and an INSUFFICIENT INFORMATION response may be registered.

Insert Figure 2 About Here

If a value is available (inverse correlations require a value inversion at stage 2""), then it may be compared with the test value at stage 3. A mismatch results in an index change to NOT BUY at stage 4, while a match leads to a BUY response at 5.

Model B generates a variety of predictions concerning the verification task response times. For categories mentioned explicitly, it states that NOT BUY responses will be longer than for BUY. This is because, relative to BUY's, NOT BUY items require the additional response index change at stage 4. The Model B also predicts that the verification of the value of an implicit (i.e., correlated) category will exceed explicit ones, as long as the subject has not computed the correlated value while reading the problem.

The only one of these predictions that can be examined, however, is the one that states that NOT BUY response times will be longer than BUY's. Collapsing across RA and RB, BUY responses (response type 1) took 2214 msec as compared with 2642 msec for NOT BUY (see Table 6). Since no verification items concerned implicit categories, the other predictions cannot be inspected.

Possible Experiments

While models A and B receive at least some support from the decision and verification data, it will be difficult to evaluate them fully without further experimentation. It would be relatively straightforward to design experiments comparable to the ones discussed here that would more incisively inspect the predictions of models such as A and B. The following are some examples.

Decision Task

In order to test his model of sentence verification, Singer (1981) used natural language materials that were as simple as possible. The rationale was that it is important, at least at a preliminary stage of an investigation, to eliminate as much variance as possible from the response latencies. For the present purposes, certain simplifications are similarly possible.

Consider the following procedure for an RA group only. The subject reads the passage one sentence at a time, instead of as a whole. Reading is self-paced, with the reader pressing a space bar to view the next sentence. Sentence reading times are recorded. As before, the passage mentions three categories, including at least one relevant and one irrelevant one. The sentences are simplified as compared with the previous materials: instead of presenting a complicated statement from which the reader deduces the value of one of the categories, the sentences simply assert such values (e.g., earnings are low). Immediately after the last sentence, the reader sees the conjunctive rule, makes a decision, and responds.

Here are a few of the predictions of model A that this experiment would address.

(1) For problems with two explicit categories, NOT BUY decisions should be much faster than BUY. This is especially true if both relevant categories were negative for the NOT BUY problems, as was the case in previous experiments. NOT BUY decisions could be signalled immediately after the discovery of the first negative category.

(2) INSUFFICIENT INFORMATION decisions are predicted to be very fast, with an operation sequence from the reading of the rule of 6-7-8-9-12-13 (see Figure 1). This sequence, of course, would be longer if the missing category was the second rather than the first one evaluated by the reader. The experiment has the merit of examining the differential predictions that the model specifies, for example, for missing-explicit versus explicit-missing problems.

(3) Perhaps the most interesting issue concerns the predictions of model A for the BUY and NOT BUY problems involving implicit categories. The predictions depend on whether or not it is believed that the RA subjects compute the value of correlated categories when they read the passage. If not, their decision would require the execution of operations 9 and 9' every time an implicit category was identified. This, in turn, would lead to the prediction of longer decision times (i.e., response time from operation 6 onward) as a function of the number of implicit categories. If, on the other hand, it is believed that correlated values are computed during initial reading, the model would predict decision times that are independent of the number of implicit categories.

The proposed experiment, in fact, provides an independent source of evidence concerning the time of the computation of correlated values: namely, the initial sentence-reading times. Computation of correlated values during initial reading should produce considerably higher reading times for sentences dealing with correlated categories.

Verification Task

A comparable experiment could be conducted to examine Model B directly. Subjects would read passages including either two or three sentences that state the values of correlated or uncorrelated categories (e.g., dividends are high). Immediately after, the reader would view a test item to verify. The subject would register responses such as TRUE, FALSE, DON'T KNOW, using appropriately labelled switches.

As was the case for the decision experiment, it would be useful, at this stage, to keep the materials as simple as possible. It would be possible, for example, to include only positive correlations to link categories. This would permit the elimination of stages 2''' and 2'''' (Figure 2). Even with this simplification, there would be numerous types of test items, including Explicit Buy and Not Buy, Implicit Buy and Not Buy, and Insufficient Information with and without a correlated category. The model specifies the operation sequence for each problem, from which sequences important ordinal predictions are derived.

General Discussion

Some general points should be made at this stage:

(1) The evaluation of Models A and B with reference to the data already collected needs to be considered with caution, both in terms of successes and failures. As has already been pointed out, none of the earlier experiments was devised to test these models. The earlier experiments included complicated stimulus materials that insured the high variability of response times. Numerous interesting predictions of the models could not be examined at all. The usefulness of such models toward understanding the tasks under consideration can only be established by means of experiments that more directly address the models themselves.

(2) The data of the proposed experiment will not be used simply to examine the ordinal predictions of the model. Rather, it will be possible to submit those data to multiple linear regression analysis, using the number of executions of each hypothetical operation as predictor variables. Such a procedure permits the determination of the proportion of variance among conditions accounted for by the model, goodness of fit, and parameter estimates (i.e., the magnitude of the duration of the component operations). This procedure is well documented (e.g., Carpenter & Just, 1975; Schustack & Sternberg, 1981).

(3) It is not asserted that these models can account for all of the features of subjects' performance in these tasks. One outcome that is apparently beyond the scope of the models is the fact, mentioned by Antos, et al. (1981, p. 21), that explicit information has more impact than implicit information on the subject's decision, when one relevant category is missing. As a second example, the models do not address the various interactions

involving the consistency variable (Tables 8 and 9 of Antos et al.). It should be noted, however, that these outcomes are somewhat reminiscent of certain semantic-relatedness effects. Rips, Shoben, and Smith (1973), for example, reported that it takes less time to verify _a robin is a bird_ than _a penguin is a bird_. Models such as that of Carpenter and Just have typically not accounted for such findings (although they could conceivably be elaborated to do so).

TABLE 1

Overview of Decision and Verification Experiment

Session 1	Session 2	Session 3	Session 4
Training: a) Learned to categorize individual stock-report statements into one of six fundamental fact categories. b) Learned to evaluate individual statements using integer scale.	Training: a) Learned to use conjoint rules b) Learned one positive and one negative correlation between fact categories. Completion task verification of report consistency.	Decision Task: Read short stock reports and apply conjunctive Decision Rule. The rule was either given before or after reading the text. Rule Before: RT1 = Read Rule RT2 = Read Text & Make Decision Rule After: RT1 = Read Text RT2 = Read Rule & Make Decision	Verification Task: Read short stock reports and verify value of probes presented after reading texts.

Table 2

Text assembled according to design specification of

Trial 2 in Table 1

Sales of large-scale data processing systems are substantial in dollar terms and are expanding modestly. Increased debt to capital ratio severely reduced Ectex cash position. Directors will meet next month and there is speculation about a stock split of 3 for 1.

Table 3

Design block example for Experiment 2. The signs "-" and "+" represent categories present in the text and their respective value range. A circle represents a category to be verified.

Trial	Sl	Er	Gr	Div	GF	Cap	Verified Category Sign	Correct Response	Trial Type
1	⊖	-				+	-	Yes	Consistent
2			+	⊖	+		+	No	Consistent
3	-	⊕				+	+	Yes	Inconsistent
4			⊕	+	-		-	No	Inconsistent
5			-	+	⊖		-	Yes	Consistent
6	+	+				⊖	+	No	Consistent
7		⊕	-	-			+	Yes	Inconsistent
8	-	-			⊕		-	No	Inconsistent
9	+		-			⊖	-	Yes	Not Inconsistent
10				⊖	+	+	+	No	Not Inconsistent
11		-	⊕		-		+	Yes	Not Inconsistent
12	⊕			+		-	-	No	Not Inconsistent

Table 4

			Completion		Verification	
GROUP	RULE BEFORE		.980 (3.85)	F(1,20)=3.7	.875 (6.89)	F(1,20)=2.7
	RULE AFTER		.953 (4.42)	[F(1,20)=1.2]	.907(6.48)	[F<1]
TRIAL BLK	FIRST		.946 (5.10)	F(2,40)=5.1*	.871 (7.74)	F(2,40)=5.1*
	SECOND		.971 (3.73)	[F(2,36)=27.8]**	.939 (6.13)	[F(2,22)=20.4]**
	THIRD		.982 (3.08)		.909 (5.34)	
SET SIZE	THREE		.968 (3.68)	F(2,40)=1.47	.919 (5.77)	F<1
	FOUR		.957 (4.09)	[F(2.40)=19.0]**	.912 (6.30)	[F(2.40)=18.6]**
	FIVE		.974 (4.67)		.889 (7.37)	
RESPONSE TYPE	NEGATIVE	(Inconsistent)	.954 (4.12)	F(2,40)=2.8	.934 (5.92)	F(2,40)=2.6
	INSUFFICIENT INFO	(Not Inconsistent)	.985 (4.11)	[F<1]	.907 (6.90)	[F(2,40)=5.5]*
	POSITIVE	(Consistent)	.960 (4.21)		.879 (6.62)	

*.01

**.0001

Note: RT averages and F-values are in square brackets. The level labels for response types enclosed in parentheses are for the verification task. Those not in parentheses are for the completion task.

TABLE 5

Mean Reaction Time in Seconds and Proportion Correct for Decision Task

Response		Rule Before			Rule After		
		RT_1	RT_2	P(c)	RT_1	RT_2	P(c)
Buy	2 Imp	3.998	18.233	.455	34.372	4.796	.564
	1 Exp/1 Imp	3.448	15.438	.673	31.492	5.866	.636
	2 Exp	2.710	15.118	.873	29.527	3.295	.800
Not Buy	2 Imp	2.506	16.516	.889	34.432	4.347	.709
	1 Exp/1 Imp	3.050	15.025	.836	31.449	5.008	.782
	2 Exp	3.975	16.961	.818	30.136	3.877	.709
Efficient Information	2 Missing	3.329	21.766	.782	33.969	3.839	.691
	1 Miss/1 Exp	3.375	20.181	.473	32.733	5.309	.491
	1 Miss/1 Imp	3.141	21.600	.673	28.602	5.046	.636

Note: Imp = Implicit

Exp = Explicit

TABLE 6

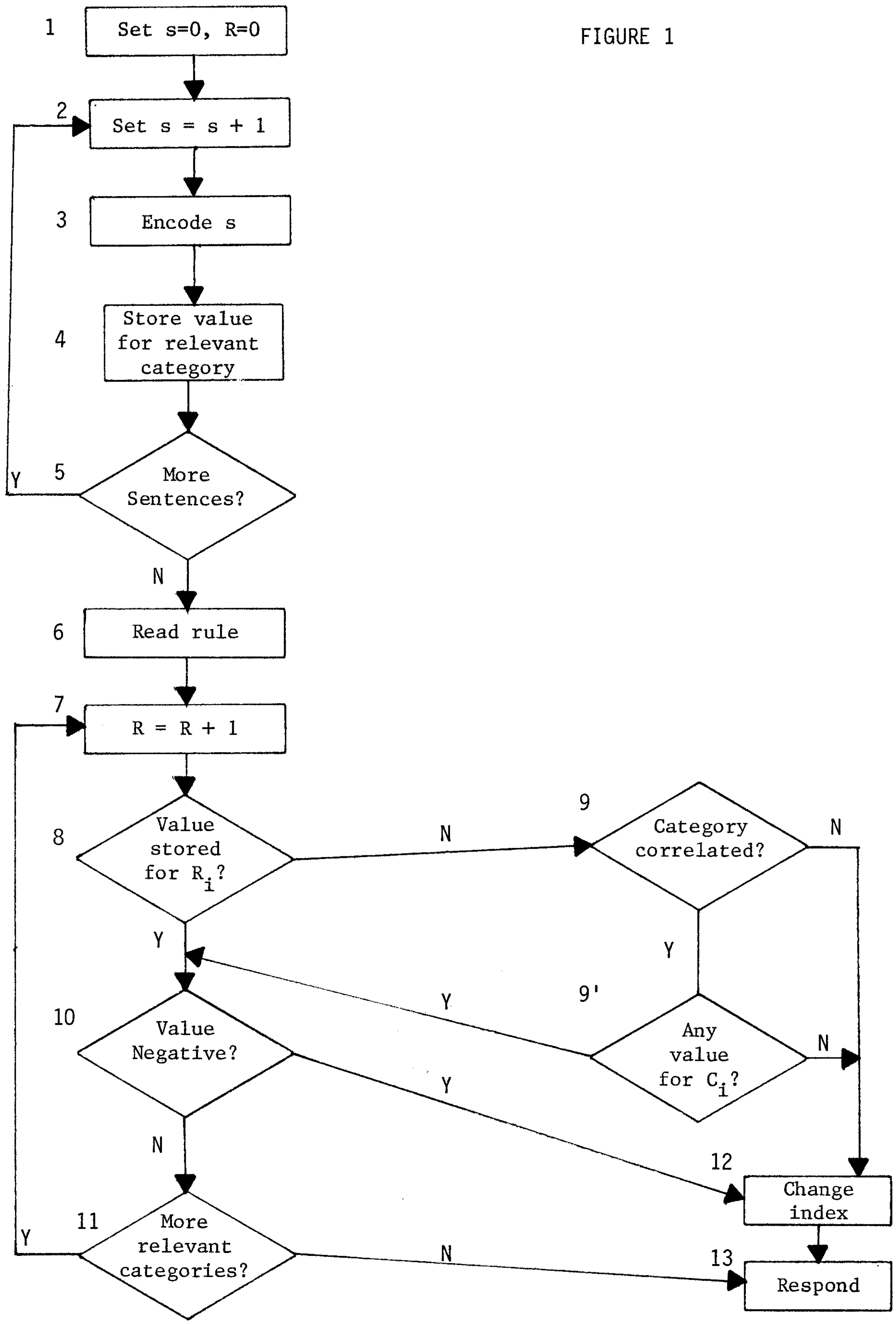
Mean Reaction Time in Seconds and Proportion Correct for Verification Task

Probe Type	Text Type	Response Type	Rule Before RT_1	Rule Before RT_2	Rule Before P(c)	Rule After RT_1	Rule After RT_2	Rule After P(c)
IN	CON	BUY	24.352	2.150	.818	31.074	2.563	.795
IN	CON	NBUY	21.018	2.578	.841	25.646	2.598	.886
IN	ICON	BUY	21.291	2.416	.864	27.829	2.007	.773
IN	ICON	NBUY	25.106	2.492	.591	30.381	3.056	.795
OUT	CON	BUY	20.380	2.055	.614	25.565	1.930	.659
OUT	CON	NBUY	20.565	2.500	.841	25.100	2.485	.773
OUT	ICON	BUY	22.133	2.118	.932	26.104	2.302	.795
OUT	ICON	NBUY	21.956	2.727	.795	27.609	2.811	.659
NC	NICON	BUY	23.816	2.238	.818	28.181	2.363	.761
NC	NICON	NBUY	22.738	2.456	.795	26.148	2.714	.830

Note: Probe Types Are: Inside or Outside a Correlated Pair or No Correlated Pair

Text Types Are: Consistent, Inconsistent, and Not Inconsistent

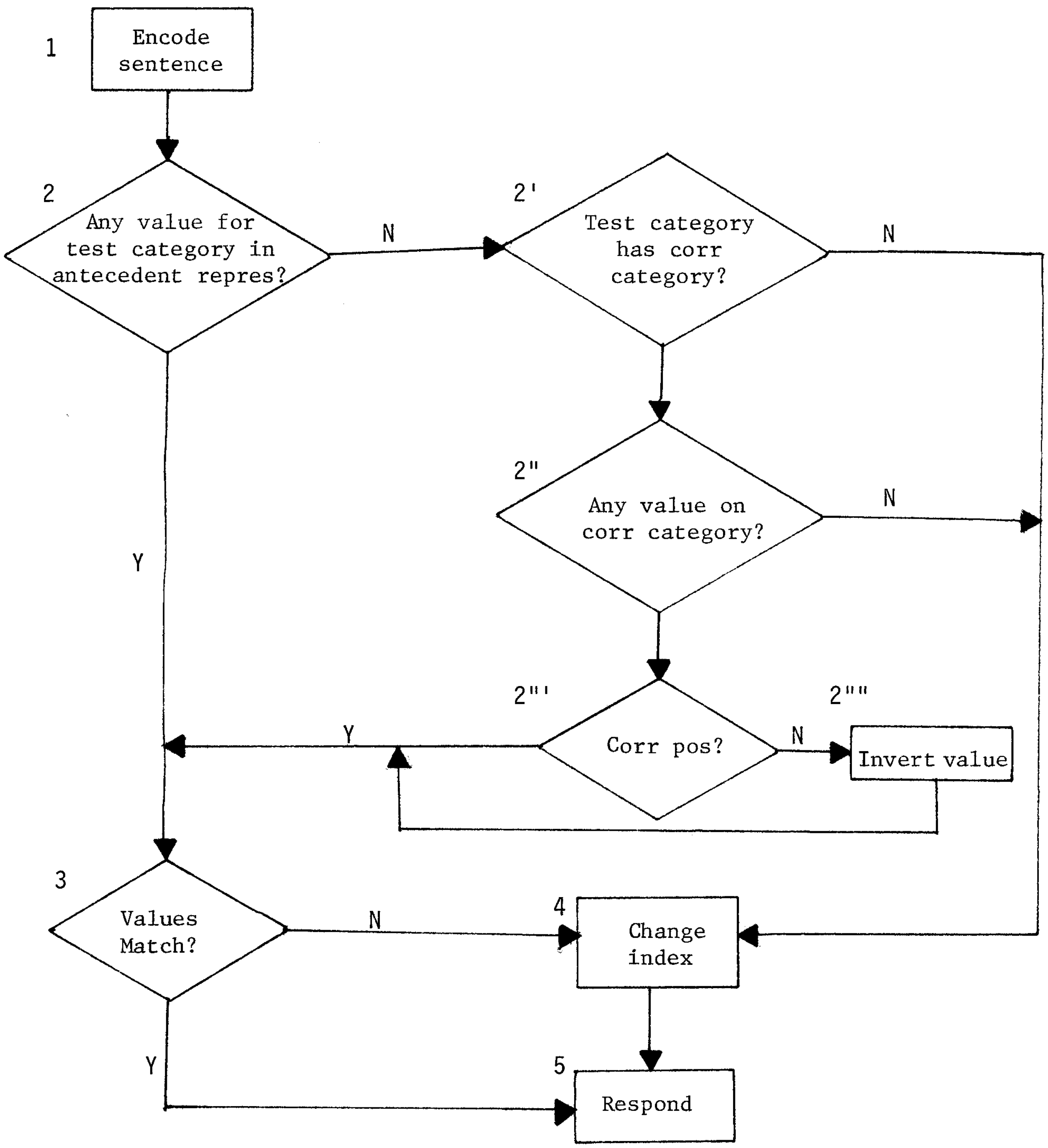
Response Types Are: Buy and Not Buy



1
Set s=0, R=0
2
Set s = s + 1
3
Encode s
4
Store value for relevant category
5
More Sentences?
Y
N
6
Read rule
7
R = R + 1
8
Value stored for R_i?
N
Y
9
Category correlated?
N
Y
9'
Any value for C_i?
Y
N
10
Value Negative?
Y
N
11
More relevant categories?
Y
N
12
Change index
13
Respond

FIGURE 1

FIGURE 2


1
Encode sentence
2
Any value for test category in antecedent repres?
N
2'
Test category has corr category?
N
2"
Any value on corr category?
N
Y
2"'
2""
Y
Corr pos?
N
Invert value
3
Values Match?
N
4
Change index
Y
5
Respond

References

Antos, S., Kozminsky, E., Bourne, L. E., Jr., & Kintsch, W. Comprehension and analysis of information in texts: IV Decision and Verification processes. Institute of Cognitive Science Technical Report #104. University of Colorado, Boulder, Colorado 80309.

Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycholinguistic model of verification. Psychological Review, 1975, 82, 45-73.

Clark, H. H., & Chase, W. G. On the process of comparing sentences against pictures. Cognitive Psychology, 1972, 3, 472-517.

Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 1-20.

Schustack, M. W., & Sternberg, R. J. Evaluation of evidence in cause inference. Journal of Experimental Psychology: General, 1981, 110, 101-120.

Singer, M. Verifying the assertions and implications of language. Journal of Verbal Learning and Verbal Behavior, 1981, 20, 46-80.

Sternberg, S. The discovery of processing stages: Extensions of Donders' method. In W. Koster (Ed.), Attention and performance II. Acta Psychologica, 1969, 30, 412-431.