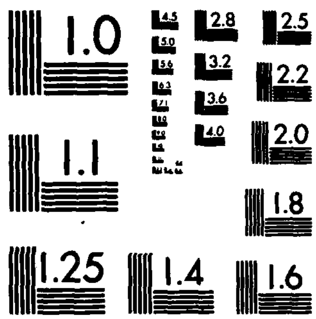
AD-A143 274 OPTIONAL STOPPING PERFORMANCE UNDER GRAPHIC AND NUMERIC 1/2
CRT FORMATTING(U) RICE UNIV HOUSTON TX DEPT OF
PSYCHOLOGY D R SCHWARTZ ET AL. JUN 84 TR-84-1
UNCLASSIFIED N00014-82-C-0001 F/G 5/10 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

RICE UNIVERSITY

Houston, Texas 77001

AD-A143 274

OPTIONAL STOPPING PERFORMANCE UNDER
GRAPHIC AND NUMERIC CRT FORMATTING

David R. Schwartz

Rice University

Technical Report #84-1

June 1984

Department of Psychology

Research Report Series

(1)

OPTIONAL STOPPING PERFORMANCE UNDER GRAPHIC AND NUMERIC CRT FORMATTING

David R. Schwartz

Rice University

Technical Report #84-1

June 1984

This research was supported by the Engineering Psychology Program, Office of Naval Research, Under ONR contract N00014-82-C-0001, Work Unit NR197-074.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 84-1	2. GOVT ACCESSION NO. AD-A143 274	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Optional Stopping Performance Under Graphic and Numeric CRT Formatting		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) David R. Schwartz and William C. Howell		8. CONTRACT OR GRANT NUMBER(s) N00014-82-C-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology Rice University Houston, TX 77251		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR197-074
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE 13. NUMBER OF PAGES 32
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)

decision making, optional stopping,CRT display formatting, cognitive processing, information purchase, time stress

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(see absract)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

Two experiments sought to determine whether, and if so how, performance varies as a function of the manner in which a progressively unfolding decision problem is displayed over time. A hurricane-tracking scenario was adapted to the "optional-stopping" paradigm, such that subjects elected to continue sampling information or to make a terminal decision (evacuate or stay) at successive points in the storm's development based upon cumulative evidence and future expectations. The display manipulation was minimal in that it applied only to non-predictive historical data (position information). Despite this conservative test, display format had a significant effect when time pressure was involved: subjects reached earlier and better terminal decisions under the analog (graphic) than the numerical format (Experiment 2). The differences reduced to nonsignificances under self-pacing (Experiment 1), although significant improvements were obtained by use of a simple aiding device (calculation of "worst case" probabilities). Results are generally consistent with Hammond's Cognitive Consistency Theory.

A-1

INTRODUCTION

Advances in computer technology provide the system designer with numerous options for the visual encoding and display of information. Even the least sophistocated machines offer a variety of formatting capabilities including both alpha-numeric and graphic presentation modes.

Considerably less advanced, however, is our understanding of how best to use these capabilities. Designers often side-step the issue by incorporating multiple modes and self-selection features into the display -- in effect, leaving the decision up to the user. While flexibility is obviously the only answer when the same information has multiple users and uses, it can prove costly when operators are under speed and/or load stress. Selection itself can take time and "mental capacity," and there is no inherent reason to believe that users are even aware of which mode produces the best results. Thus, when task requirements are fairly well defined, format decisions might better be programmed into the system design--if, in addition, the basic functions relating format to operator performance were known.

As we just noted, however, this last "if" is a rather big one. It has long been recognized that display format can have a profound, though task-dependent, influence on human performance (Hitt, 1961; Schultz, 1961; Smith, 1963; Smith & Aucella, 1983). Still for relatively few kinds of tasks could our present knowledge be considered adequate for making specific design recommendations. Notable exceptions, in addition to Hitt's early map-reading work, are recent studies on look-up tasks (Tullis, 1981, 1983, 1984) and on probabilistic judgments (Goldsmith & Schvaneveldt, 1981; Wickens & Scott, 1983). In the former, Tullis has achieved considerable success predicting keying performance and subjective ratings based on six objective display measures. In the latter, subjective weighting of cues has been shown to vary as a function of alpha-numeric vs. graphic display modes.

The fact that analog or graphic presentation sometimes produces superior performance has been attributed to "holistic" processing: people are able to integrate or interpret the information in a graphic display "at a glance" whereas they tend to deal with alpha-numerics sequentially (Goldsmith & Schvaneveldt, 1981; Wickens & Scott, 1983). In this vein, Hammond (1980) has proposed that format is one of several task features that promote either an "analytic" or "intuitive" cognitive approach to judgment.

The present research sought to extend the graphic (analog) vs. alpha-numeric comparison to another practically important task domain: the so-called optional-stopping decision problem. The principal feature of this task is that, unlike the one-shot decision problems favored in laboratory research, the operator (DM) acquires information over time and at cost, deciding when as well as how to act. In a sense, each problem involves a sequence of decisions, each stage of which requires a choice between "keep-sampling" and "terminal-action" alternatives. Cost and payoff functions are generally such that DM can reduce the uncertainty of the decision problem only at the expense of the marginal utility of his/her ultimate choice (i.e., information is costly).

The practical significance of this kind of task is that it represents a common decision situation in which time is important, a situation encountered in military, medical, political, legal and managerial contexts. A military commander or corporate CEO, for example, can be completely certain of an adeversary's intentions if he or she elects to wait until all tactical options have become futile! Though exaggerated, this example illustrates a common failing revealed in both the laboratory and in anecdotal reports from "the field:" human DM's tend to postpone action too long (i.e., to oversample) in such situations (Connolly & Gilani, 1982; Hershman & Levine, 1972; Levine, Samet, & Brahlek, 1975).

Given the potential maladaptiveness of human sampling tendencies and the suggested advantages of the analog-graphic display mode which would seem particularly salient in a progressively unfolding scenario, the optional-stopping task appeared to represent an ideal vehicle for the exploration of format effects. The aim of the present studies, therefore, was to determine whether sampling tendencies, decision quality, and/or decision efficiency vary systematically with display format. Two experiments were conducted, one under self-paced and the other under forced-paced conditions. If the two formats do encourage different cognitive strategies (e.g., "holistic" and "intuitive" vs. "sequential" and "analytic"), then one would expect time pressure to exacerbate any observed performance differences.

EXPERIMENTS

The object of both studies was simply to determine whether manipulation of display format (graphic vs. numeric mode) alters sequential decision behavior in a manner consistent with the hypothesized differences in cognitive processing. In particular, does an analog version of a developing scenario induce less sampling, faster or more efficent proccesing, and more time-sensitive decision quality than the numeric version of the same information? Experiment 1 was considered exploratory in that, despite some prior pilot work, it was the first full-blown study using a new task and methodology. Thus self-pacing was used rather than pre-selected timing, and a second variable (presence vs.absence of a decision aid) was crossed with display format in an attempt to define key task parameters for use in subsequent studies.

Experiment 2 was the immediate beneficiary of this developmental work. Forced pacing was introduced and manipulated on the basis of Experiment 1 results. In most other respects, however, the methodology for the two studies was identical. Consequently, the common features will be described first followed by a separate account of Experiments 1 and 2.

COMMON METHOD

Subjects

A total of 92 undergraduates recruited from psychology courses at Rice University participated in one of the two studies. Each served for one or two experimental sessions in exchange for course credit or $4.00 per session. The sexes were represented approximately equally in all groups; otherwise assignment was randomized. Experiment 1 called for 56 subjects; Experiment 2, for 36.

Task

Each subject was required to monitor a series of simulated hurricanes in their advance toward a heavily populated target area (city). Storm paths were described with reference to an 8 (longitude) x 7 (latitude) grid as illustrated in Figure 1. They progressed from left

---

Insert Figure 1 about here

(longitude 7) to right (longitude 0) in step-wise fashion, and subjects were required to make one of three responses after each advance: <u>wait</u> (i.e., postpone any terminal action), <u>evacuate</u> (i.e., order total abandonment of the city), or <u>stay</u> (i.e., minimize the potential losses associated with decision delay as in committing to intensified protective measures rather than retreat).

Uncertainty in the storm's path was introduced by building into each advance a .3 probability of unit latitudinal movement in either direction (thereby setting the probability of its remaining at the same latitude equal to .4). The net result of this feature was that the modal ending point of the population of storms was the target location, but the <u>a priori</u> probability of this occurrence was low ($p=.190$). For purposes of simplicity, the value components of the task were all translated into a common <u>lives lost</u> index. Waiting reduced the effectiveness of subsequent evacuation and added a small progressive loss of its own; staying carried a huge loss if the storm hit the city, but none of it did not; evacuating averted this maximum loss, but was more costly as time grew short. The specific functions used were as follows:

<u>Waiting</u> = 0 lives per advance for longitude 7-5;
20 lives per advance for longitudes 4-1.

<u>Staying</u> = 1500 lives if hit; 0 if not

<u>Evacuation</u> = 300 + [3(8 - Longitude)] lives.

Briefly, then, the subject's task was to decide when to stop gathering information and, at that point, which action to take based upon (a) the storm's current location, and (b) the subjective expectation of costs associated with the various options in that situation. The subject observed the entire course of each hurricane toward it's final (Latitude 0) destination. However, instructions emphasized that there would be a number of storms to monitor, and that the subject should attempt to minimize the loss of life over the entire session. It was explained that because hurricanes are not completely predictable, the "best" decision would not always result in a favorable outcome, and "poor" ones would sometimes win out. Since these fluctuations would even out over the course of a session, the subject was advised to concentrate on long-run consequences. To reinforce this perspective, a cumulative loss "score" was maintained and displayed continuously over the session.

The formal structure of the task, as just described, produced the expected loss pattern shown in Table 1. The values were chosen in a

---

Insert Table 1 about here

---

manner designed to avoid the "flat maximum problem" (i.e., a structure in which most of the options produce very similar marginal results), and to distribute the "correct" decision over the options in a fairly realistic fashion (i.e., early waiting followed by evacuation or staying depending on location). Whereas subjects were carefully instructed in the loss functions associated with the three options and the rule governing probability of latitude shifts (i.e., p = .3, .4, .3), they were not shown the actual values in Table 1.

The display was used primarily to indicate the present location and history of each storm. The format manipulation involved an analog representation as illustrated previously (see Figure 1) and a numeric representation as shown in Figure 2. In both cases a cumulative loss

---

Insert Figure 2 about here

---

record for the session was displayed in identical alpha-numeric form at the bottom of the CRT.

It should be noted that in the present version of the task, a Monte Carlo procedure was used to generate storm paths. Thus there was no reliable trend information in the historical "track" of a hurricane. One would, of course, expect an analog format to be most advantageous when such information exists, and we intend to explore that possibility in future research. For present purposes, however, the intent was to focus on the most rudimentary aspect of the displayed information (current position)—an aspect for which there would be no <u>inherent</u> advantage favoring the analog mode. In fact, if anything, the analog mode was considered to be at a slight disadvantage in that, based on their extra-experimental experience, subjects might be encouraged to read more into the "tracks" than the data warranted.

A final aspect of the task used in one condition of the first experiment was a <u>decision aiding feature</u>. This feature consisted of presenting the subject, at each point in each problem, with an updated calculation of the probabilities of the storm hitting the city from its current position and each remaining position that it could assume. This information was also presented numerically, but on a separate display from the position information.

Stimuli and Apparatus

Patterns of storm movement were generated randomly within the constraints noted earlier using a Monte Carlo procedure. A TRS-80 Model 2 microcomputer controlled the stimulus presentation, displayed the information (on its CRT), and recorded the performance measures (when, where, and what response was made). The program was designed such that each response resulted in a display of the option selected, an update of the storm's progress (one step to the right), and a repetition of the three response options for the new position. Following a terminal decision (stay or evacuate), the updates continued automatically at a deliberate pace until the storm's ultimate disposition was established. At that point, the session "score" was updated and, after a brief delay, the next storm was introduced.

EXPERIMENT 1

Design and Procedure

Two variables at two levels each were combined factorially in a between-group design. Display format (numeric vs. graphic) was crossed with aiding (present vs.absent) to form four groups of 14 subjects each. As noted earlier, the aiding manipulation involved making available to subjects the simple probability of the storm hitting the city from each remaining location .

Each subject tracked 50 storms in a single session. The first 10 constituted a block of practice trials. Following a short break, the 40 experimental trials proceeded in blocks of 13, 13 and 14 respectively with similar breaks introduced after the 13th and 26th trials. The entire process, including the initial explanation, required slightly more than one hour. Although generated randomly (see above), the same set of storm patterns was administered to each of the four groups, with order of appearance balanced across groups. This was done to control for the possibility of specific pattern and sequence effects.

Three principal measures were taken: (a) information sampling, or the point at which a terminal decision was reached; (b) decision accuracy, or the correspondence of DM's response to that of the EV maximization model at each choice point; and (c) latency, or the delay between the display update and the key-entered decision at each choice point. While subjects were aware that all three were being recorded, that overt calculations were prohibited, and that unnecessary delays were to be avoided, the emphasis was on decision quality rather than speed. For purposes of analysis, these measures were summarized in various ways as described in the next section.

RESULTS AND DISCUSSION

Display effects

No significant main effects or interactions attributable to the format difference were obtained on any of the measures. This is not too surprising given the minimal amount of information that was subject to format manipulation, the lack of time pressure, the emphasis on accuracy, and the performance variability typical of such tasks. As indicated earlier, the main purpose of the study was to establish a set of task conditions within which we might explore the format variable in a meaningful way; not in the format variable itself.

Despite these limitations, however, there were some indications that format is worthy of study. For example, subjects tended to sample more information (i.e., to postpone a terminal decision longer) under the alphanumeric display (a mean of 5.18 vs. 4.95 items), and—at least in the aided condition — to respond more slowly (mean RT = 521 msec vs. 466 msec). Since the two formats produced virtually identical levels of accuracy (77.2% for graphic vs. 76.8% for alphanumeric), the implication is that the alphanumeric mode induced less efficient processing. However, it benefited more consistently from aiding than did the graphic condition, which showed almost none. In the case of terminal ("evacuate" or "stay") decisions, the additional information provided by aiding increasesd response latency by an average 266 msec. for the alphanumeric display but not at all for the analog display, an interaction that approached significance, $F(1,52) = 2.88$, $p = .075$.

Taken together, these tendencies appear consistent with the "holistic-analytic" processing distinction, although by no means constituting substantial evidence for it.

Other task-induced effects. As has been reported in previous research, (e.g., Hershman & Levine, 1970; Levine et al., 1975) subjects tended to postpone a terminal decision beyond the optimal stopping point (defined as that position at which EV for "act" first exceeds that for "wait"). The mean difference between actual and optimal points of action was +.39, a difference significantly above zero: $t(55) = 3.54$, $p < .001$. Interestingly, sampling tended to increase significantly rather than decrease over trial blocks, while performance accuracy and latency both improved significantly (see Table 2).Thus, what might be construed as overcautiousness from the perspective of an ideal processor may actually have been close to optimal for the human processor: DM's

---

Insert Table 2 about here

---

appeared to make good use of the "oversampled" information. The fact that subjects sampled more and became more efficient with practice also suggests that they took the task seriously and did not succumb to boredom effects as has been reported in some earlier studies (e.g., Levine, et al., 1975). Responding thoughtlessly or carlessly simply to "get it over with" would have produced declining petormance over blocks.

The present data support an earlier finding by Levine et al. (1975) that terminal decisions are slower than sampling decisions. In the present study the difference was highly significant even without including the earlier storm positions where vitually all choices were to "wait" and were made very fast (thereby inflating the "wait" vs. "terminal" decision latency difference). A more meaningful and conservative comparison limited to the last four storm positions yielded mean latencies of 510 msec. for waiting and 712 msec. for terminal decisions, $F(1,51) = 6.93$, $p < .012$.

Finally, the presence of a decision aid raised the mean accuracy of all decisions from 75.25% to 79.50%, a difference significant at $F(1,52) = 6.88$, $p = .011$. Of course these mean values also reflect the disproportionate influence of early, easy "wait" decisions; by contrast, accuracy for the critical last four positions averaged 55.00% and 62.70% for unaided and aided conditions respectively, $F(1,52) = 6.12$, $p<.017$. While mean latency for the aided decisions in these positions averaged

122 msec longer than for the unaided ones, and the difference was principally vested in the numeric display mode, neither effect was statistically reliable (although, as noted earlier, the interaction approached significance).

The main conclusions to be drawn from Experiment 1, then, are (a) that the hurricane tracking task constitutes a promising vehicle for exploration of display format and other design variables on sequential decision behavior, (b) that aiding can improve accuracy but at a cost in time, (c) that the range of conditions over which forced pacing can meaningfully be introduced into the task extends roughly from .3 sec. to 2.0 sec. per decision, and (d) that the tendency to "oversample" or postpone terminal action in optional stopping is not necessarily maladaptive from the human perspective. In addition, despite the absence of conditions (especially time pressure) that would be expected to induce display format effects, nonsignificant trends consistent with such effects did appear. The second study, therefore, was designed to provide an explicit test of the format variable by introducing forced pacing.

EXPERIMENT 2

Design and Procedure

The chief differences between this study and Experiment 1 were that the display format variable was manipulated within subjects, forced pacing was introduced as a between-subjects variable, and (in view of its demands on processing time) the aid was not used. The design was thus a mixed model involving the factorial combination of display format (analog vs.numeric) x order of display administration (AN vs. NA) x trial blocks (3) x time stress (300, 700, and 1630 msec/update). Twelve subjects were assigned randomly to each of the three pacing groups.

The procedure was identical to that used in Experiment 1 except that more practice trials were administered (13), and more storms were included in each experimental block (29). Since the total number of trials was increased from 50 to 200, (100 trials per display mode), each subject participated in two sessions rather than one. The first included instructions, practice, and the first three experimental blocks; the second, the last three blocks plus the debriefing. All trials in a given session were performed under one of the two formats.

RESULTS AND DISCUSSION

Display Effects

Under the pressure of forced pacing, both stopping point (information sampling) and decision accuracy were influenced in the predicted ways by the display format manipulation. The pattern of these effects is best understood by considering together the analyses of overall decision accuracy, late decision accuracy (i.e. the last four longitude positions), and sampling (see Tables 3 and 4). Since timing was controlled, latency was no longer a salient index.

Insert Tables 3 and 4 about here

Looking first at the accuracy measures, it is apparent that the analog format produced consistently superior performance, particularly under the more stressful pacing conditions and in the critical later stages of each storm (i.e., the last four positions where both terminal and sampling decisions were well represented). In the overall analysis (Table 3a) where storm position (early vs.late) was included as a separate variable, it is clear that the effect was limited primarily to the later decisions. In fact, early decision accuracy was near perfect since, as noted previously, the tendency during the storm's formative stages was to postpone action, a decision that was almost always correct (95% overall). This ceiling effect did not operate in the later stages since the subject had to give serious consideration to all three options, and mean accuracy dropped to 50%. Combining insensitive early positions with sensitive later ones, the main effect of format was reduced to nonsignificance ($p=.098$); by contrast, separating early and later positions revealed highly significant ($p<.001$) position and position x format interaction effects. The format x stress interaction was also sufficiently robust to withstand the diluting influence of early positions ($p = .014$), indicating that, as expected, the analog superiority was limited to the more stressful conditions. In fact, the overall accuracy under the slowest condition was similar to that under self-pacing (Experiment 1) where it will be recalled there was no significant format effect.

A separate analysis of accuracy for later decisions (i.e., those made during the last four storm positions) revealed a highly significant format effect ($p < .001$, see Table 3b). However, despite the fact that the analog superiority again seemed more pronounced under the more stressful conditions, the format x stress interaction failed to reach significance ($p=.39$). A possible explanation is that since format was manipulated within subjects, order of presentation may have moderated the effect sufficiently to have obscured the interaction. The marginal significance of the second-order interaction ($p=.064$) supports this possibility. In any case, given the results from the overall analysis and from the self-pacing condition employed in Experiment 1, the most conservative conclusion is that analog superiority is more likely to occur when time pressure is involved than when it is not.

Turning to the sampling data, Table 4, it is apparent that, once again, the overall tendency was to oversample (see Table 4b) in this case by an average of 1.49 updates, or nearly a fourfold increase over that for Experiment 1 (i.e., 0.39). Thus, the mere exsistence of time pressure for making individual decisions, and the consequent reduction in opportunity to process the available information, appears to have caused subjects to seek more information. As in Experiment 1, the trend was for sampling to increase with experience, but in this case the effect was limited to the numeric format (the format x trial blocks interaction was highly significant). There was also a tendency for stress level to moderate the effect, but this influence was neither systematic nor easily interpretable. The safest conclusion with respect to display, therefore, is that the numerical format promotes an increase in sampling over trial blocks, but that this format effect is a relatively small modulation in a very large tendency for time stress to promote oversampling. Since overall accuracy was about 7% lower than for the comparable condition in Experiment 1, this dramatic increase in oversampling does not appear to have been very productive. Moreover, the absence of any positive relationship between the conditions that increased sampling and improved performance in the present study suggests that, unlike Experiment 1, stress-induced oversampling is not functional. Apparently people are not able to make good use of the additional information that they feel they need. In contrast, it will be recalled that they did make good use of "oversampled" information under the non-stressful self-pacing conditions of Experiment 1.

Other Task-induced effects. As may be seen in Table 3, the time-stress manipulation produced a systematic decrease in accuracy, particuarly for the later (more sensitive) decision positions (Table 3b). The main effect of this variable was highly significant in both analyses, as was its interaction with position in the overall analysis. While, as noted previously, forced pacing per se apprears to have induced a substantial increase in the tendency to oversample (0.39 vs. 1.49 updates), the specific rate of forced pacing did not yield a signficant main effect on either sampling measure (see Table 4). The update rate, did, however, amplify the format effect significantly as discussed earlier.

GENERAL CONCLUSIONS

If, as hypothesized, analog representation of an evolving decision problem encourages a more "holistic" processing strategy (or conversely, numeric representation a more "analytic" one), we would expect better performance with an analog format as time pressure exceeded some critical level. This expectation was borne out in the present findings. Under self-pacing and the lowest forced-pacing conditions, format had little effect on decision accuracy; at higher levels of time stress, a significant analog superiority emerged.

While one might argue that the obtained accuracy differences, though statistically significant, were relatively small (on the order of 5-8% in the critical later positions), it should be recognized that the experimental test was intentionally very conservative. The only information that was subject to the format manipulation was the progressive location of the storm; there was no predictability in its "track" or path. Since such information (i.e., spatial trends) would naturally favor an analog format, one can only expect the differences obtained here to become magnified under the more realistic task situation. What the present studies show, then, is that display format can affect decision performance in very subtle ways. Not only can it bring out aspects of a data set (such as trend information) that are otherwise difficult to perceive--the well-established and relatively obvious "compatibility" phenomenon--it can alter the decision maker's whole approach to information processing. In a sense it can alter his/her processing "set."

Another noteworthy finding involved sampling behavior. Consistent with previous studies, the present results again revealed a tendency for subjects to delay action beyond the optimal decision point--in effect, to buy too much information. However, these findings qualify the generalization in several important respects. First, a numeric format tends to amplify the effect. Secondly, the oversampling effect may not always be as maladaptive as it seems, particularly if one takes the decision-maker rather than an "objective" model as the point of reference. That is, the overall effectiveness of human decisions can improve beyond the point where an "ideal decision-maker" would stop sampling. Lastly, however, the present data suggest that under time pressure, the tendency toward <u>maladative</u> oversampling increases. Under the non-stressful conditions of Experiment 1, subjects seemed to derive real value from the information they received after they "should have acted," whereas under the stressful conditions of Experiment 2, their performance did not improve at all as a result of such information. Admittedly, the evidence for this last conclusion is circumstantial at best--the present work was not directed explicitly toward this issue. Nonetheless, it would appear sufficiently important, theoretically as well as practically, to warrant further investigation.

TABLE 1

The Expected Loss Association with each Response Option at the various Hurricane Positions*

Latitude	Longitude 7	6	5	4	3	2	1	Response Option
	.	.	.	77	61	40	60	Stay
1	.	.	.	97	81	60	80	Wait
	.	.	.	348	395	448	507	Evacuate
	.	.	187	178	182	175	60	Stay
2	.	.	187	192	194	194	80	Wait
	.	.	327	348	395	448	507	Evacuate
	.	272	288	310	358	400	560	Stay
3	.	265	281	308	336	373	580	Wait
	.	312	327	348	395	448	507	Evacuate
	286	307	334	371	440	550	660	Stay
4	278	298	325	360	403	507	680	Wait
	303	312	327	348	395	448	507	Evacuate

TABLE 1 (Continued)

The Expected Loss Associated with each Response Option at the various Hurricane Positions*

Latitude	Longitude							Response Option
	.	272	288	310	358	400	560	Stay
5	.	265	281	308	336	373	580	Wait
		312	327	348	395	448	507	Evacuate
	.	.	187	178	182	175	60	Stay
6	.	.	187	192	194	194	80	Wait
	.	.	327	348	395	448	507	Evacuate
	.	.	.	77	61	40	60	Stay
7	.	.	.	97	81	60	80	Wait
	.	.	.	348	395	448	507	Evacuate

*Values are rounded to the nearest unit.

Table 2

Mean Performance over Trial Blocks for the Three Principal Measures (Experiment 1)

Measure	Trial Blocks 1	2	3	Significance Test
1. Stopping point (no. of observations)	4.92	5.04	5.23	$F(2, 104) = 5.00$, $p < .009$
2. Accuracy	76%	76%	80%	$F(2, 104) = 3.87$, $p = 0.024$
3. Latency(ms)*	608	450	351	$F(2, 104) = 149.46$, $p < 0.001$

* Includes both sampling and terminal decisions over all positions. Since earlier and sampling ("wait") decisions far outnumber later and terminal decisions, and the former tend to be about 200 ms. shorter than the latter, these means are biased toward the shorter latencies (see text).

Table 3

A Summary of Major Display Effects Expressed in Percent Correct (Experiment 2)

A. OVERALL ACCURACY

Display Format	Storm Position	Stress Level (Updates/sec) .6	1.4	3.3	Mean
	Early	92	96	93	94(3)
Analog	Late	58	53	47	53
	Mean	75	75	70	73
	Early	98	95	94	96
Numeric	Late	57	45	42	47
	Mean	77	70	68	71
Mean		76	72	69	73

(1) Stress $F(2, 30) = 4.68$, $p=0.017$

(2) Format $F(1, 30) = 2.92$, $p=0.098$ (ns)

(3) Position $F(1, 30) = 1068.56$, $p<0.001$

(4) F x S $F(2, 30) = 4.92$, $p=0.014$

(5) F x P $F(1, 30) = 12.40$, $p=0.001$

(6) S x P $F(2, 30) = 5.92$, $p=0.007$

B. LATE ACCURACY

Stress Level (Updates/s)

Display

Format	Order	.6		1.4		3.3		Mean	
	AN	60		52		47		53	
Analog	NA	56	58(3)	53	53	48	47	51	53(2)
	Mean		58		53		47		53
	AN	58		51		43		53	
Numeric	NA	55	57	38	45	41	42	45	47
	Mean		57		45		42		47
Mean			57		49		45		50

(1) Format $F(1, 30) = 15.47$, $p<0.001$

(2) Stress $F(2, 30) = 11.84$, $p<0.001$

(3) F x S $F(2, 30) = 2.11$, $p=1.39$ (ns)

(4) F x S x Order $F(2, 30) = 3.00$, $p=0.064$ (ns)

Table 4

Summary of Major Sampling Effects ExpressedNumber of Updates Prior to a Terminal Decision (Experiment 2)

A. ACTUAL SAMPLING

		Stress Level (Updates/sec)			
Display Format	Trial Block	.6	1.4	3.3	Mean
	1	4.72	5.04	5.00	4.92(2)
Analog	2	4.66	5.05	5.01	4.91
	3	4.86	5.00	4.88	4.91
	Mean	4.75	5.03	4.96	4.91
	1	5.17	4.70	4.87	4.91
Numeric	2	5.49	4.91	5.09	5.16
	3	5.51	4.72	5.30	5.18
	Mean	5.39	4.78	5.08	5.08
Mean		5.07	4.90	5.02	4.79

(1) Format x Stress $F(2, 30) = 4.38$, $p=0.021$

(2) Format x Trials $F(2, 60) = 3.58$, $p=0.034$

B. OVERSAMPLING (ACTUAL-OPTIMAL)

		Stress Level (Updates/s)			
Display Format	Trial Block	.6	1.4	3.3	Mean
	1	1.02	1.17	2.15	1.45
Analog	2	.80	1.05	1.40	1.08
	3	1.82	.71	1.78	1.44
	Mean	1.22	.98	1.78	1.72
	1	1.57	1.53	1.38	1.49
Numeric	2	2.17	1.33	1.59	1.70
	3	2.11	1.74	1.48	1.78
	Mean	1.95	1.54	1.48	1.66
Mean		1.58	1.26	1.63	1.49

(1) Format $F(1, 15) = 3.39$, $p=0.085$ (ns)

(2) F x Stress $F(2, 15) = 3.08$, $p=0.076$ (ns)

(3) F x Trials $F(2, 30) = 7.45$, $p<0.003$

(4) S x T $F(4, 30) = 4.36$, $p<0.007$

(5) F x S x T $F(4, 30) = 7.22$, $p<0.001$

References

Connolly T., and Gilani, N. Information search in judgements tasks: A regression model and some preliminary findings. Organizational Behavior and Human Performance, 1982, 30, 330-350.

Goldsmith, T.E., and Schvaneveldt, R.W. The role of integral displays in decision making. In Human Factors in Computing Systems. Gaithersburg, MD: Association for Computing Machinery, 1981, 197-201.

Hammond, K.R. The integration of research in judgement and decision theory (Report No. CRJP 226). Boulder, CO.: University of Colorado, Institute of Behavioral Sciences, 1980.

Hershman, R.L., and Levine, J.R. Deviations from optimum information-purchase strategies in human decision-making. Organizational Behavior and Human Performance, 1970, 5, 313-329.

Hitt, W.D. An evaluation of five different abstract coding methods- Experiment IV. Human Factors, 1961, 3, 120-130.

Kerkar, S.P., and Howell, W.C. The effect of information display format on multiple-cue judgement. Houston, TX: Rice University, Report No. ONR-84-2, 1984.

Levine, J.M., Samet, M.G., and Brahlek, R.E. Information seeking with limitations on available information and resources. Human Factors, 1975, 17, 502-513.

Schultz, H.G. An evaluation of graphic trend displays. Human Factors, 1961, 3, 99-107.

Smith, S.L. Color coding and visual separability in information displays. Journal of Applied Psychology, 1963, 47, 358-364.

Smith, S.L., and Aucella, A.F. Design guidelines for the user interface to computer-based information systems Bedford, Massachusetts: The Mitre Corp., Report No. ESD-TR-83-122, 1983.

Tullis, T. S. An evaluation of alphameric, graphic and color information displays. Human Factors, 1981, 23, 541-550.

Tullis, T.S. (1984). Predicting the usability of alphanumeric displays: Dissertation in Psychology. Unpublished doctoral dissertation, Rice University, Houston.

Tullis, T.S. The formatting of alphanumeric diaplays: A review and analysis. Human Factors, 1983, 25, 657-682.

Wickens, C.D. and Scott, B.D. A comparison of verbal and graphical information presentation in a complex information integration decision task. Champaign, Illinois: University of Illinois, Engineering-Psychology Research Laboratory, Report No. EPL-83-1/ONR-83-1, 1983.

LIST OF FIGURES

David R. Schwartz (Optional Stopping Performance Under Graphic and Numeric CRT Formatting) recieved his M.A. degree in psychology from Rice University in 1984. He is currently fullifilling requirements for his Ph. D. at the same institution. His primary interests include decision-aiding under stress conditions, and human-computer interface issues. Until recently, Mr. Schwartz has been researching systems usability at IBM's Austiin Technical Evaluation Center, Austin, Texas.

William C. Howell is the Lynette S. Autry Professor of Psychology and Administrative Science, and Chairman of the Psychology Department at Rice University. He recieved his Ph. D. from the University of Virginia in 1958, joined the Laboratory of Aviation Psychology at the Ohio State University in 1957 (becoming its Director in 1965), and held a regular appointment on the Ohio State faculty from 1960-1968. His interests have covered a number of topics within the field of Engineering Psychology concentrating, in recent years, on judgment and decision processes.

END

FILMED

9-84

DTIC

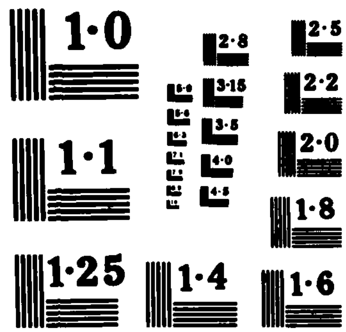
AD-A143 274 OPTIONAL STOPPING PERFORMANCE UNDER GRAPHIC AND NUMERIC CRT FORMATTING(U) RICE UNIV HOUSTON TX DEPT OF PSYCHOLOGY D R SCHWARTZ ET AL. JUN 84 TR-84-1 2/2

UNCLASSIFIED N00014-82-C-0001 F/G 5/10 NL

END

FILMED

DTIC


1.0
2.8
2.5
3.15
2.2
3.5
2.0
4.0
1.1
4.5
1.8
1.25
1.4
1.6

SUPPLEMENTARY

INFORMATION

Errata for Technical Report #84-1

Optional Stopping Performance Under Graphic and Numeric CRT Formatting

Due to clerical error, an early, unedited version of Technical Report # 84-1 (Office of Naval Research Contract # N00014-82-C-0001, Work Unit NR1978-074) was released and distributed. Upon receipt of this errata statement, Technical Report #84-1 should be considered complete. The statement includes the following:

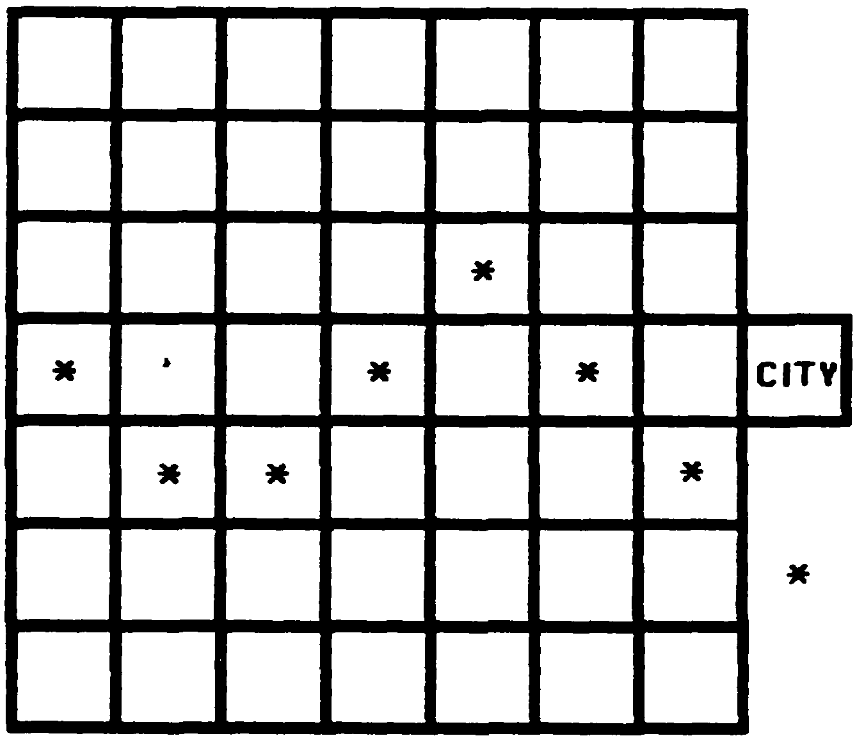
1) The evacuation cost formula on page 7 should read:

$$\text{Evacuation} = 300 + [3 \times (8 - \text{longitude})]^2 \text{ lives.}$$

2) Figures 1 and 2.

3) Tables 1, 3a, 3b, 4a, and 4b.

4) Contrary to the report's running head, David R. Schwartz is its sole author.



Trial
Losses: 80

Cum
Losses: 80

STORM'S LATITUDE	STORM'S LONGITUDE
4	7
5	6
5	5
4	4
3	3
4	2
5	1
6	0

CITY'S LATITUDE	4	0	CITY'S LONGITUDE

Trial Losses: 80

Cum Losses: 80

TABLE 1

The Expected Loss Associated with each Response Option at the Various Hurricane Positions*

Latitude	Longitude 7	6	5	4	3	2	1	Response Option
	.	.	.	77	61	40	60	Stay
1	.	.	.	97	81	60	80	Wait
	.	.	.	348	395	448	507	Evacuate
	.	.	187	178	182	175	60	Stay
2	.	.	187	192	194	194	80	Wait
	.	.	327	348	395	448	507	Evacuate
	.	272	288	310	358	400	560	Stay
3	.	265	281	308	336	373	580	Wait
	.	312	327	348	395	448	507	Evacuate
	286	307	334	371	440	550	660	Stay
4	278	298	325	360	403	507	680	Wait
	303	312	327	348	395	448	507	Evacuate

TABLE 1 (Continued)

The Expected Loss Associated with each Response Option at the Various Hurricane Positions*

Latitude	Longitude							Response Option
	7	6	5	4	3	2	1	
	.	272	288	310	358	400	560	Stay
5	.	265	281	308	336	373	580	Wait
	.	312	327	348	395	448	507	Evacuate
	.	.	187	178	182	175	60	Stay
6	.	.	187	192	194	194	80	Wait
	.	.	327	348	395	448	507	Evacuate
	.	.	.	77	61	40	60	Stay
7	.	.	.	97	81	60	80	Wait
	.	.	.	348	395	448	507	Evacuate

*Values are rounded to the nearest unit.

TABLE 3

A Summary of Major Display Effects Expressed in Percent Correct

(Experiment 2)

A. OVERALL ACCURACY

Display Format	Storm Position	Stress Level (Updates/sec) .6	1.4	3.3	Mean
	Early	92	96	93	94
Analog	Late	58	53	47	53
	Mean	75	75	70	73
	Early	98	95	94	96
Numeric	Late	57	45	42	48
	Mean	77	70	68	72
Mean		76	72	69	72

(1) Stress $F(2,30) = 4.68$, $p = 0.017$

(2) Format $F(1,30) = 2.92$, $p = 0.098$ (ns)

(3) Position $F(1,30) = 1068.56$, $p < 0.001$

(4) F x S $F(2,30) = 4.92$, $p = 0.014$

(5) F x P $F(1,30) = 12.40$, $p = 0.001$

(6) S x P $F(2,30) = 5.92$, $p = 0.007$

B. LATE ACCURACY

Display Format	Order	Stress Level (Updates/sec) .6	1.4	3.3	Mean
	AN	60	52	47	53
Analog	NA	56	53	48	52
	Mean	58	53	47	53
	AN	58	51	43	51
Numeric	NA	55	38	41	45
	Mean	57	45	42	48
Mean		57	49	45	50

(1) Format $\underline{F}(1,30) = 15.47$, $\underline{p} < 0.001$

(2) Stress $\underline{F}(2,30) = 11.84$, $\underline{p} < 0.001$

(3) F x S $\underline{F}(2,30) = 2.11$, $\underline{p} = 1.39$ (ns)

(4) F x S x Order $\underline{F}(2,30) = 3.00$, $\underline{p} = 0.064$ (ns)

TABLE 4

Summary of Major Sampling Effects Expressed as the Number of Updates Prior to a Terminal Decision (Experiment 2)

A. ACTUAL SAMPLING

Display Format	Trial Block	Stress Level (Updates/sec) .6	1.4	3.3	Mean
	1	4.72	5.04	5.00	4.92
Analog	2	4.66	5.05	5.01	4.91
	3	4.86	5.00	4.88	4.91
	Mean	4.75	5.03	4.96	4.91
	1	5.17	4.70	4.87	4.91
Numeric	2	5.49	4.91	5.09	5.16
	3	5.51	4.72	5.30	5.18
	Mean	5.39	4.78	5.08	5.08
Mean		5.07	4.90	5.02	5.00

(1) Format x Stress $F(2,30) = 4.38$, $p = 0.021$

(2) F x Trials $F(2,60) = 3.58$, $p = 0.034$

B. OVERSAMPLING (ACTUAL-OPTIMAL)

Display Format	Trial Block	Stress Level (Updates/sec) .6	1.4	3.3	Mean
	1	1.02	1.17	2.15	1.45
Analog	2	.80	1.05	1.40	1.08
	3	1.82	.71	1.78	1.44
	Mean	1.22	.98	1.78	1.33
	1	1.57	1.53	1.38	1.49
Numeric	2	2.17	1.33	1.59	1.70
	3	2.11	1.74	1.48	1.78
	Mean	1.95	1.54	1.48	1.66
Mean		1.58	1.26	1.63	1.49

(1) Format $F(1,15) = 3.39$, $p = 0.085$ (ns)

(2) F x Stress $F(2,15) = 3.08$, $p = 0.076$ (ns)

(3) F x Trials $F(2,30) = 7.45$, $p < 0.003$

(4) S x T $F(4,30) = 4.36$, $p < 0.007$

(5) F x S x T $F(4,30) = 7.22$, $p < 0.001$

END

FILMED

10-84

DTIC