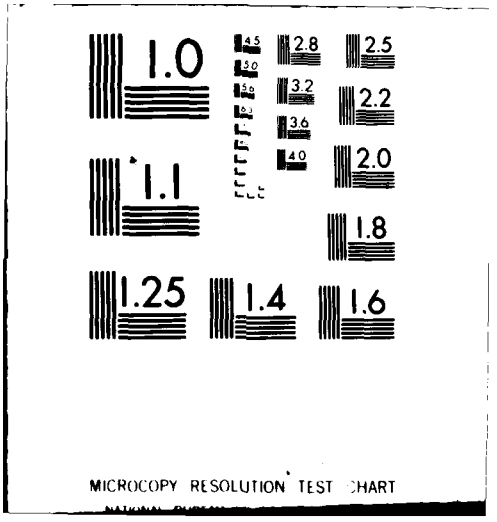
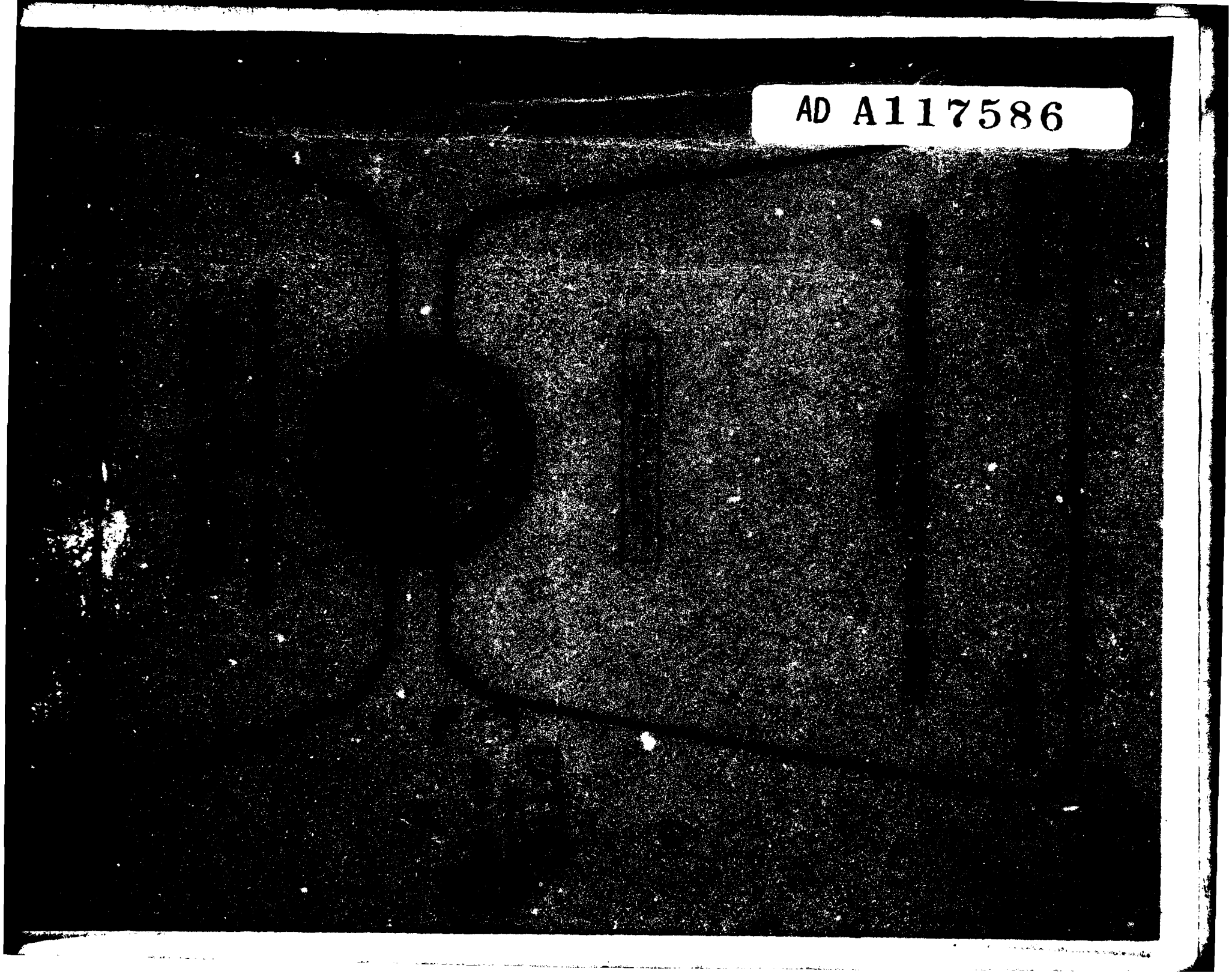
AD-A117 586 NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN D--ETC F/G 5/10
HUMAN MEMORY LIMITATIONS IN MULTI-OBJECT TRACKING.(U)
JUN 82 F L GREITZER, R T KELLY, R L HERSHMAN

UNCLASSIFIED NPRDC-TR-82-48 NL

END
DATE FILMED
9-82
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART


AD A117586

NPRDC TR 82-48

June 1982

HUMAN MEMORY LIMITATIONS IN MULTI-OBJECT TRACKING

Frank L. Greitzer
Richard T. Kelly
Ramon L. Hershman

Reviewed by
Robert E. Blanchard

DTIC
ELECTE
JUL 28 1982
S D
H

Released by
James F. Kelly, Jr.
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPRDC TR 82-48	2. GOVT ACCESSION NO. AD-A117586	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) HUMAN MEMORY LIMITATIONS IN MULTI-OBJECT TRACKING		5. TYPE OF REPORT & PERIOD COVERED Interim Report June 1980-Nov 1981
		6. PERFORMING ORG. REPORT NUMBER ~~17-81-16~~
7. AUTHOR(s) Frank L. Greitzer Richard T. Kelly Ramon L. Hershman		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS SF57-525-001-026-03.01 PE 6 5
11. CONTROLLING OFFICE NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		12. REPORT DATE June 1982
		13. NUMBER OF PAGES 39
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)

Human performance measurement
Visual memory
Mathematical models of memory
get tracking
Hum memory limitations
Visual tracking

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

Basic performance data were obtained on the effect of critical task variables in unaided multi-object tracking behavior. Six observers viewed computer-generated displays in which five, seven, or nine objects represented targets that moved in random linear trajectories at one of two speeds. Displayed positions were updated six times at

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

intervals of 5, 8, 13, or 18 seconds, and no track history was provided. The task for the observer was to monitor the trajectories and then predict the next position of each object.

Results showed that the unaided observer can keep track of up to about seven moving objects. Performance improved as the interval between updates was increased to about 13 seconds. These variables interact in their effects on tracking performance and may be traded off in a complex manner.

A family of mathematical models of human memory that focus on the encoding, learning, and rehearsal processes of the observer was developed. Two of the models' predictions were consistent with the data observed and those reported in the psychological literature.

The analysis of human memory and information processing limitations should be extended to more complex operational tasks to support system designers with quantitative estimates of operator performance.

Accession For
NTIS GRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

DTIC COPY INSPECTED 2

FOREWORD

This research was conducted in support of Project SF57-001-022-03.01 (Human Factors in Command and Control) under the sponsorship of the Naval Sea Systems Command. Additional funding was provided by the Naval Electronics Systems Command under Project PE 62731N. The objective of this project is to enhance the effectiveness of command and control systems through improved design of the human-computer interface. In particular, the project seeks to identify the limits of the human ability to assimilate tactical information quickly and accurately.

The experiment reported here grew out of discussions with Dr. Joel Lawson of the Naval Electronics Systems Command. Results are intended for use by command and control system designers.

JAMES F. KELLY, JR.
Commanding Officer

JAMES J. REGAN
Technical Director

PRECEDING PAGE BLANK-NOT FILMED

SUMMARY

Problem

An important command and control (C^2) task is the concurrent tracking of multiple targets on a radar or computer-generated display. As the number of targets increases, the operator's ability to retain critical information can be expected to diminish. Understanding human limitations in this type of task is essential to ensure the design of C^2 systems that suit the user's capabilities.

Objectives

The objectives of this research were:

1. To obtain basic performance data on the effect of critical task variables on multitarget tracking behavior.

2. To develop and assess models that quantify human performance limits in a basic multitarget tracking task.

Method

Six civilian psychologists served as observers in five sessions in a multi-object tracking task. Each session consisted of 12 trials that were defined by the factorial combination of two independent variables: the number of objects to be tracked (5, 7, or 9) and the interval (5, 8, 13, or 18 seconds) between the displayed updates of their positions. In each trial, objects on random linear trajectories moved at one of two constant speeds. After the initial display, positions were updated six times at one of the four inter-update intervals. No track history was provided. The observer's task was to monitor the trajectories and then predict the next (eighth) position of each object.

Results

1. Tracking accuracy was high (although imperfect) with five objects, but decreased as the number of objects increased. Monitoring of nine objects clearly exceeded the unaided observer's processing capacity.

2. Tracking performance improved as the interval between display updates increased. Presumably, the longer intervals allowed more time for rehearsal of the objects' trajectories.

3. Longer display times compensated for the greater processing load associated with more objects. For example, performance in the 18-second/9-object condition was comparable to that in the 5-second/7-object condition.

4. There was evidence that observers grouped sets of objects into chunks on the basis of their trajectories and proximity. Such "chunking" effectively reduced the memory load and improved tracking performance.

5. A family of models of human memory that focus on the encoding, learning, and rehearsal processes of the observer were developed. The models make quantitative predictions of performance based on the number of objects, the interval between updates, and the number of updates displayed. Two of the models' predictions are consistent with the obtained data and those reported in the psychological literature.

PRECEDING PAGE BLANK-NOT FILMED

Conclusions

1. The unaided human observer can keep track of about seven moving objects. Performance improves as the interval between updates is increased up to about 13 seconds and, presumably, as more updates are displayed. These variables interact in their effects on tracking performance and may be traded off in a complex manner.

2. Multi-object tracking performance is accounted for by two mathematical models that express human memory limitations in terms of encoding, learning, and rehearsal processes. Evidence supports the unconventional notion of a variable rehearsal capacity in short-term memory.

Recommendation

The analysis of human memory and information-processing limitations should be extended to more complex operational tasks to support system designers with quantitative estimates of operator performance.

CONTENTS

	Page
INTRODUCTION	1
Problem	1
Background	1
Human Information-processing Limitations	1
Hypothesized Performance Relationships	2
Objectives	3
METHOD	3
Observers	3
Apparatus	3
Tracking Task	3
Performance Measurement	5
Design	5
RESULTS AND DISCUSSION	5
Effects of Major Variables	5
Learning Strategies	8
MODELS OF MULTIOBJECT TRACKING	10
Information-processing Framework	10
Mathematical Models	12
Development	12
Assessment	19
CONCLUSIONS	23
RECOMMENDATION	23
REFERENCES	25
APPENDIX--DERIVATIONS OF THE PREDICTIONS FOR THE FOUR MODELS	A-0
DISTRIBUTION LIST	

LIST OF TABLES

Page

1. Analysis of Variance on the Effects of Test Sessions, Number of Objects, and Inter-update Interval 6
2. Analysis of Variance on the Effects of Clustering in the Nine-object Displays 9
3. Number of Objects Tracked for the Four Models on Each Update 19
4. Parameter Estimates and Prediction Errors for the Four Models 20

LIST OF FIGURES

1. Hypothesized performance surface for multi-object tracking 2
2. Trajectories in a representative nine-object display 4
3. Tracking accuracy for the five test sessions 7
4. Tracking accuracy as a function of number of objects and inter-update interval 7
5. Tracking accuracy for nine objects as a function of the number of clusters in the display with inter-update interval as a parameter 9
6. Schematic model of information-processing framework. 11
7. Illustration of learning states for nine objects before, during, and after a sequence of K updates 15
8. Required rehearsal time function proposed in assumption R2 16
9. Example of rehearsal process during an update 17
10. Alternative assumptions for Models I-IV 18
11. Observed data and predictions of the four models 21

INTRODUCTION

Problem

An important command and control (C^2) task is the concurrent tracking of multiple targets on a radar or computer-generated display. As the number of targets increases, the operator's ability to retain critical information can be expected to diminish. Understanding the human limitations in this type of task is essential to ensure that C^2 systems are designed to suit the user's capabilities.

Background

Most C^2 functions in the combat direction center rely heavily on humans to process information and make decisions. One such function requires that operators keep track of moving targets displayed on a radar screen. For example, tracking is a critical requirement in the Navy Tactical Data System (NTDS) for positions such as the detector-tracker, the air intercept controller, and the tactical action officer.

These personnel interact with hardware and software subsystems (by hooking targets, entering data, etc.) and thereby contribute to the overall knowledge base of the system. Reasonable answers to such questions as "What is this base of knowledge?" or "What does the system know?" might be obtained from the computer's data base on the various tracks of interest and their attributes (type of platform, position, course, speed, etc.).

The focus here, however, is on limits in operators' processing and on what they know, particularly on what humans can remember about the tactical environment when limited to their own resources. For this purpose, consider a primitive tracking problem without any external aids or computer support to provide target symbology, identify track numbers, etc. and operators who have been freed from disruptive auxiliary tasks, such as communications. In this environment, the single aspect of tactical knowledge that is of concern is the unaided ability to remember the movements of multiple targets. When left to their own processing resources, operators must first integrate positional data over time to acquire course and speed information. They then must devote considerable cognitive effort to retain this information for a future report. Investigation of these processes should reveal basic human perceptual and memory limitations that are relevant to performance and man-machine design in more general environments.

Human Information-processing Limitations

Clearly, there will be no track information to retain, unless target movement is first observed and detected by the operator. These operations involve one class of human capabilities and limitations; namely, attention and perception. It is well known that, if target positions are updated at sufficiently rapid rates (less than one per second), the requisite integration of successive displays is readily achieved by the visual system (e.g., Pollack, 1972, 1974). This was dramatically demonstrated earlier by White's (1956) "time compression" technique. He increased the update rate for a radar display from 0.1 per second (an extremely difficult tracking task for experts) to 24 per second (a very easy task even for novices). In these experiments, the observer reported immediately on new information extracted from the display. There was no need to maintain contact and no requirement to remember and report target movements later.

As the requirement to retain and report track history is imposed, however, a second class of human capabilities and limitations becomes most relevant; namely, human

memory. The incoming information persists in the human visual system for only a brief period (about 1 second) before fading away. During this time, the object's trajectory must be encoded; that is, recognized and stored in the memory system to enable further processing. Since encoding requires about 0.3 to 0.5 second (Mackworth, 1962; Posner & Boies, 1971), all of the perceived trajectories cannot always be encoded at once. Thus, the encoding operation may continue for several display updates--until the limit of short-term memory is reached.

Numerous experiments have demonstrated that the unaided human can handle about 7 ± 2 items at the same time (e.g., Miller, 1956). The best performance by an unaided operator should then be limited to between five and nine concurrent objects. To retain this information, the operator must periodically rehearse it, as has been demonstrated with visual and verbal material (e.g., Murray & Newman, 1973). Moreover, if rehearsal is inhibited or disrupted, performance degrades quickly (Peterson & Peterson, 1959).

Hypothesized Performance Relationships

The research on perceptual and memory limitations just cited leads to the hypothesized performance surface shown in Figure 1. The surface relates tracking performance (i.e., memory for object trajectories) to three variables: (1) the number of objects to be tracked, (2) the number of updates that are displayed, and (3) the time between updates or inter-update interval (IUI).

Generally, the hypothesis is that performance improves as more time is available for processing objects. Either a longer IUI or more display updates can serve this purpose. The more potent variable seems to be display updates, because each new update provides an additional learning opportunity and refreshes the information already in memory.

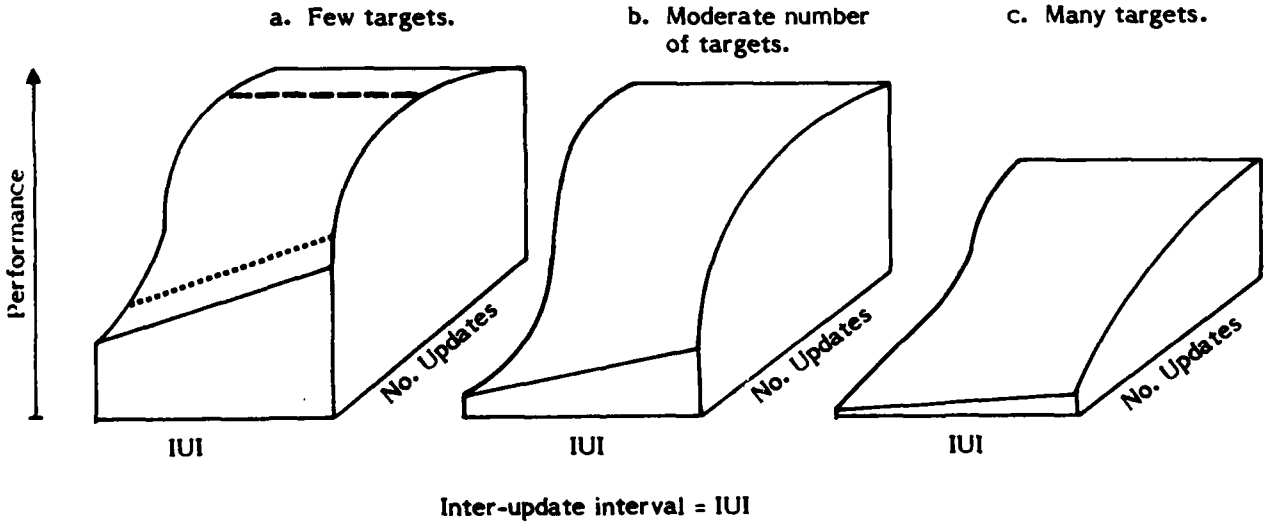


Figure 1. Hypothesized performance surface for multi-object tracking.

It is presumed that the two variables, IUI and the number of updates, interact. For example, Figure 1a shows that with few updates, increases in IUI enhance performance (e.g., the dotted line). However, with a large number of updates, the IUI effect vanishes (e.g., the dashed line). To understand this interaction, consider the situation with few updates. A certain amount of time is required to process an object's trajectory; that is, to perceive, encode, and establish it in memory. If the IUI is too short, several updates will be needed to process even a single object; with only a few updates available, overall tracking performance will be low. However, increasing the IUI allows more time for rehearsing the acquired trajectories and thus facilitates performance.

On the other hand, consider the situation with a large number of updates. Again, a short IUI limits encoding and rehearsal, but here more updates compensate by providing additional processing opportunities. Moreover, the short IUIs favor the perception of object motion. The result is a high, near-asymptotic level of tracking performance and, as the IUI increases, virtually no further effect is expected.

This hypothesized interaction is presumed to hold for any number of objects, but the performance asymptote decreases as the number of objects to be tracked increases (Figures 1b and 1c).

The present research seeks to verify and quantify a portion of the surface hypothesized in Figure 1 and to understand the basic information-processing and memory requirements for the tracking of moving objects by the unaided observer.

Objectives

The objectives of this research were:

1. To obtain basic performance data on the effect of critical task variables on multi-object tracking.

2. To develop and assess mathematical models that quantify human performance limits in a basic multi-object tracking task for possible application to more sophisticated operational tasks.

METHOD

Observers

Six civilian experimental psychologists (five males and one female) served as observers. Their ages ranged between 27 and 44 years, with a median of 31.5 years.

Apparatus

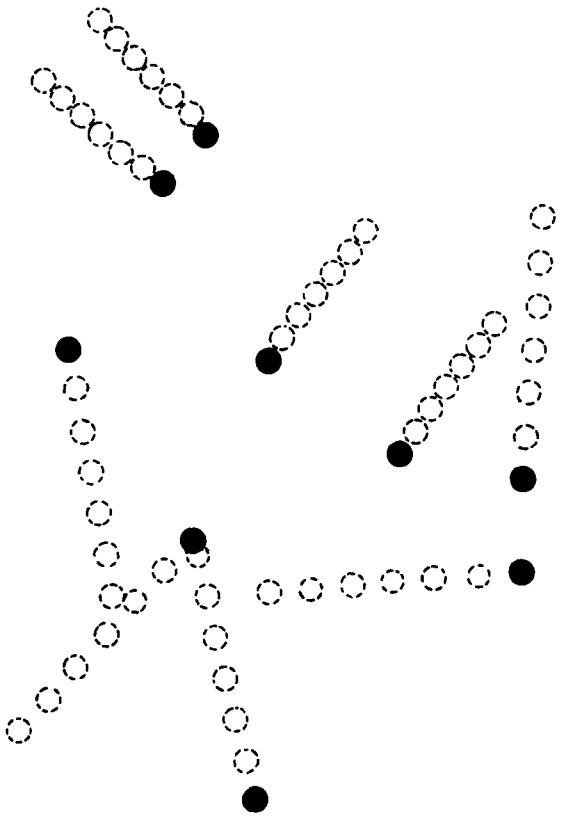
The tracking task was implemented on a Tektronix 4027 color graphics terminal driven by a Tektronix 4051 microcomputer. The observer viewed the 20x25 cm screen of the Tektronix 4027 from a distance of approximately 60 cm. The phosphor in the CRT caused erased stimuli to persist for approximately 0.5 second.

Tracking Task

The display presented a series of seven snapshots of idealized target objects moving on linear paths across the CRT. The objects were identical 6-mm pale yellow discs on a

black background and initially randomly located on the screen. At regular intervals, all object positions were updated simultaneously. Except for the brief persistence of the phosphor, no track history was displayed. Thus, to track the objects, the observer first had to derive their courses by integrating successive displays. Since no aids (e.g., grease pencil, pencil and paper, etc.) were allowed, this information had to be maintained in memory until tested.

Figure 2 illustrates a nine-object tracking problem to which the track history (the dashed circles) has been added for the reader's benefit. The observer viewed only the solid discs, which corresponded to the objects' present positions. Each object maintained a constant but randomly selected heading and moved at a constant display speed of either 6 or 12 mm per update. On the average, half of the objects moved at each speed. All objects were constrained to stay on the screen for at least eight updates. The paths of the objects were permitted to cross, but their positions could not overlap. The display, then, loosely resembled a noise-free radar screen with no external tracking aids.



Note. The dashed circles indicate previous object positions; only the current positions, shown as the solid discs, were visible to the observer.

Figure 2. Trajectories in a representative nine-object display.

One sequence of seven object-position snapshots comprised a single trial. After each trial, the screen was erased, which was the cue for the observer to respond on a hard copy facsimile of the most recent display. Observers were instructed to place an "X" at the position that each object would occupy if the display had been up !ated once more. They were encouraged to be as accurate as possible but to guess when uncertain. No feedback was provided. After 1 minute, a warning tone signaled the imminent start of the next trial.

Performance Measurement

A simple, three-valued scoring rule was used to measure tracking accuracy. If the observer's predicted position X was within ±15° of the true trajectory, two points were scored. If the absolute error in X was greater than 15°, but not greater than 45°, one point was assigned. A score of zero was given if the error in X exceeded 45°. The mean score for all objects was then taken as the observer's performance for the given trial.

Design

The observers were tested individually in six 30-minute sessions, one per day on consecutive weekdays. The first session served as practice, and these data were not analyzed. Each session consisted of 12 trials constructed from the factorial combination of two factors: number of objects (5, 7, or 9) and IUI (5, 8, 13, or 18 sec). The 12 types of trials were presented in random order in each session and for each observer. All observers viewed identical stimulus displays but in a different order. Speed of the objects was also varied to complete the 3 (number of objects) x 4 (IUI) x 5 (test session) x 2 (speed) design. All statistical analyses employed the .01 level of significance.

RESULTS AND DISCUSSION

Effects of Major Variables

A four-factor analysis of variance (ANOVA) was conducted on the mean tracking scores derived from the observers' responses. Tracking accuracy exceeded chance performance in all conditions. Speed was not a significant factor, $F(1,5) = .28$; $MSe = .09$; the mean tracking score was 1.50 for slow objects and 1.51 for fast ones. Moreover, speed did not interact with the number of objects, the IUI, or the test session. Since speed was not a significant variable, it was eliminated as a factor, and the collapsed data were reanalyzed. The ANOVA table is exhibited in Table 1.

Mean tracking accuracy, which is presented in Figure 3, improved significantly across the five test sessions ($F(4,20) = 5.92$). However, only the first and fifth test sessions differed significantly (Tukey test). There was no interaction with the number of objects or with the IUI. Thus, the effect of practice was uniform across task conditions.

The effects of IUI and the number of objects on tracking accuracy are presented in Figure 4. The number of objects had a significant effect, $F(2,10) = 18.42$. Accuracy was highest for five objects (mean score = 1.69 out of a possible 2.0) and then decreased for seven objects (mean = 1.52) and for nine objects (mean = 1.30). Although this effect is not surprising, it does reinforce the view that a memory limitation exists within this range.

Table 1

Analysis of Variance on the Effects of Test Sessions, Number of Objects, and Inter-update Interval

Source	df	SS	MS	F
Test sessions (T)	4	3.428	0.857	5.92*
Error	20	2.896	0.145	
Number of objects (O)	2	8.863	4.432	18.42*
Error	10	2.406	0.241	
Inter-update interval (IUI)	3	2.013	0.671	9.95*
Error	15	1.011	0.067	
T x O	8	0.376	0.047	0.44
Error	40	4.253	0.106	
T x IUI	12	1.229	0.102	0.98
Error	60	6.282	0.105	
O x IUI	6	3.180	0.530	4.96*
Error	30	3.206	0.107	
T x O x IUI	24	2.436	0.101	1.01
Error	120	12.000	0.100	
Subjects	5	0.350	0.070	
Total	359	53.929		

*$p < .01$.

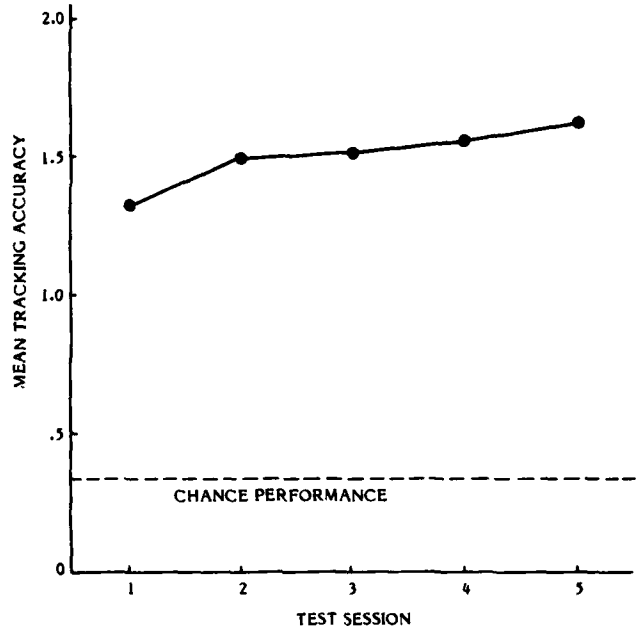


Figure 3. Tracking accuracy for the five test sessions.

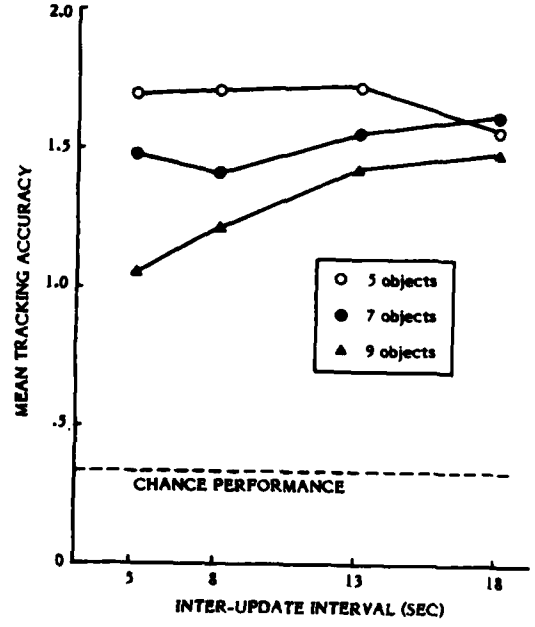


Figure 4. Tracking accuracy as a function of number of objects and inter-update interval.

Accuracy generally increased with the IUI. Mean tracking scores of 1.40, 1.46, 1.58, and 1.57 were obtained for IUIs of 5, 8, 13, and 18 seconds respectively. This effect was significant, $F(3,15) = 9.95$. Accuracy with IUI = 5 seconds was significantly poorer than with IUI = 13 or 18 seconds (the Tukey test). The longer IUIs may have facilitated recall by permitting greater opportunity for rehearsal of the object trajectories.

The effect of IUI interacted with the number of objects, $F(6,30) = 4.96$. Thus, IUI had virtually no effect with five objects, a moderate facilitative effect with seven objects, and a large facilitative effect with nine objects. When 18 seconds were available for rehearsal, the effect of the number of objects vanished. The additional time apparently helped compensate for the increased processing load.

Learning Strategies

The observers were free to employ any strategies or mnemonic techniques to learn the trajectories. Here, "chunking" (Miller, 1956), would be an appropriate strategy, because, if several objects can be grouped together into a single memory unit or chunk, fewer total items need to be remembered. If chunking were based, for example, on clusters of objects that share proximity and directional cues, the effective memory load would decrease and performance would improve.

Although the chunking process cannot be observed directly, a post hoc analysis revealed evidence of its use. For these purposes, the experimenter made direct measurements of clustering in the stimuli, the premise being that such clustering might provide a basis for the chunking phenomenon. The measurements were independent of the observers' responses and used the following definition of a cluster: A set of two or more objects is a cluster if (1) the trajectories of all its members fall within a 30° arc, (2) all of its members start within 5 cm of another object in the set, and (3) enlarging the set violates either (1) or (2). A single object is a cluster if it satisfies condition 3. For example, the stimulus in Figure 2 contains two clusters of two objects each (the objects moving in parallel at the upper left and the objects moving in parallel at the center right); and five single-object clusters. Thus, the display in Figure 2 contains seven clusters.

Only the nine-object displays were analyzed, since they would provide the greatest variability in chunking. Figure 5 shows the relation between tracking accuracy and the measured number of clusters in the display. Note that low cluster scores indicate a high potential for chunking. Results of the ANOVA, presented in Table 2, indicate that the relationship is significant, $F(4,99) = 5.29$. Accuracy was highest for those nine-target patterns that could be compressed to five or six clusters; accuracy decreased as the number of clusters increased. This suggests that observers may well have grouped nearby objects with similar trajectories into chunks to reduce their memory load.

As discussed previously, tracking accuracy generally improved with increases in IUI. In the present analysis, IUI tended to exhibit a similar effect ($F(3,99) = 3.29$, $p < .05$) and there was no interaction with the presumed number of chunks ($F(12,99) = .54$). Thus, the form of the relationship between IUI and performance is constant across the levels of the derived clustering factor. This suggests that chunking was performed by the observers early in the update, as the objects were acquired.

Debriefings gave evidence of marked individual differences in the way object trajectories were encoded in memory. Although there were no external cues to do so, most observers scanned the display in a clockwise pattern. They acquired the trajectories sequentially at the rate of one or two objects per update. The predominant encoding strategies used either a local verbal mediator or a global visual representation of the

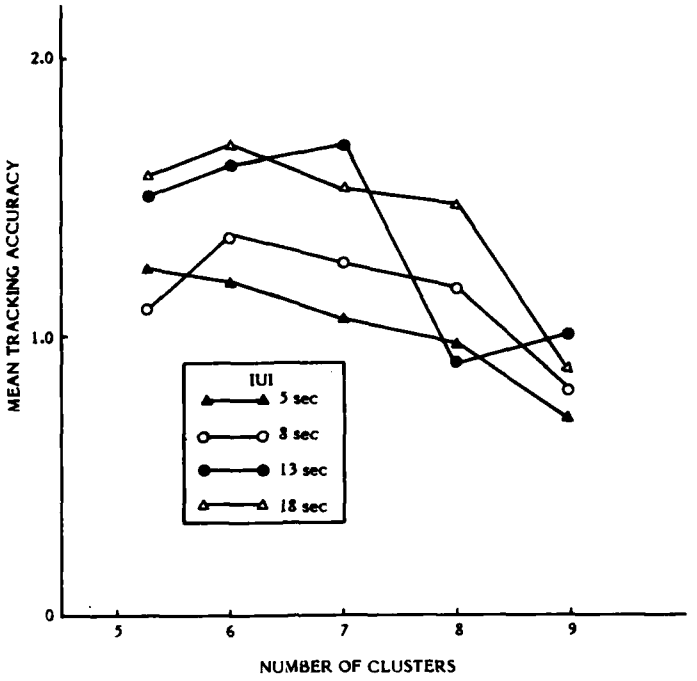


Figure 5. Tracking accuracy for nine objects as a function of the number of clusters in the display with inter-update interval (IUI) as a parameter.

Table 2

Analysis of Variance on the Effects of Clustering in the Nine-target Displays

Source	df	SS	MS	F
Inter-update interval (IUI)	3	1.962	0.654	3.29*
Number of clusters (C)	4	4.209	1.052	5.29**
IUI x C	12	1.297	0.108	0.54
Error	99	19.69	0.199	
Total	118	27.158		

Note. An unweighted means analysis was performed due to unequal cell frequencies.

*$p < .05$.
**$p < .01$.

entire display. Several studies have shown that associating verbal labels with visual material improves memory performance (e.g., Daniel, 1972). In the present experiment, some observers attached a "time of day" to each object based on its trajectory. An object moving horizontally from left to right was encoded as a "three" (for three o'clock), etc. The resulting numbers were then rehearsed sequentially according to the larger clockwise pattern. This verbal strategy could be disrupted by interference between clockface numerals and by changes in the serial order as objects would cross paths.

On the other hand, the visual strategy relied heavily on the spatial relations between objects. For example, Gestalt properties (e.g., symmetry) enhance memory for visual material (Howe, 1980). In the present task, some observers reported grouping nearby objects having similar, opposing, or crossing trajectories into chunks. In this way, the spatial relations shared by targets within a chunk would help the observer retain the individual trajectories. Other visual referents, such as the edges of the screen, were used as mnemonics when possible. Outlying objects that could not be easily grouped were attended to separately. Although these detailed strategies did not correlate with performance, their existence suggests further research and possible operator aids.

MODELS OF MULTI-OBJECT TRACKING

Information-processing Framework

The approach to the multi-object tracking problem was borrowed freely from Sperling's (1960, 1963) research on visual processes in memory, from the human information-processing concepts of Norman (1968) and Atkinson and Shiffrin (1968), and from the "levels of processing" formulation of Craik and Lockhart (1972). First, it is convenient to distinguish among three functional types of memory without regard to their possible structural properties:

1. Sensory memory refers to information carried "in a format faithful to its modality of arrival" (Crowder, 1976, p. 45). Such memory is "precategorical"; that is, it precedes the attachment of any linguistic category to the stimulus information. The sensory memory for the visual modality is called iconic memory, while the corresponding memory for audition is called echoic memory (Neisser, 1967).

2. Primary memory refers to "a mental process whereby a small amount of information can be held in a highly accessible state for transformation, rehearsal, recombination, or other operations" (Crowder, 1976, p. 156). Information so held is very susceptible to disruption so that rehearsal is a key factor in its maintenance. Although this information may take various forms or "encodings" (visual, phonological, etc.), the preeminence of echoic memory in the rehearsal process implies a major role for phonological encodings in the primary memory.

3. Secondary memory refers to information held in a more permanent, enriched form that can include semantic content and associations resulting from more elaborate processing. Access to such information is not immediate as in the case of primary memory; rather, its access requires active and deliberate retrieval processes.

The functional components just described are illustrated schematically in Figure 6. Although represented as separate "structures," they need not be interpreted as physically distinct memory stores. Rather, they may be viewed as distinct memory representations that result from different levels of processing. Craik and Lockhart (1972) have argued convincingly for a process-oriented concept of memory in which successively deeper

levels of encoding (e.g., sensory vs. phonological vs. semantic) produce increasingly more elaborate and persistent memories.

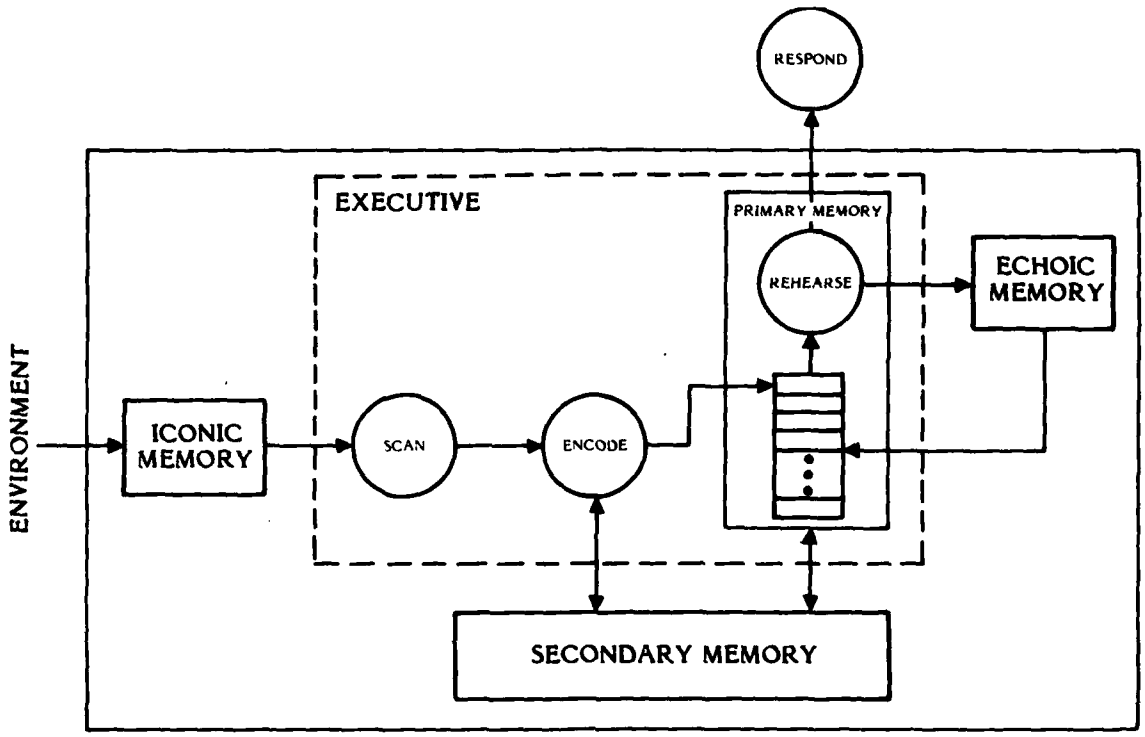


Figure 6. Schematic model of information-processing framework.

The level of processing to be employed is largely determined by conscious strategies, although the demands of the task are highly influential in this selection. These strategies (referred to as "control processes" by Atkinson & Shiffrin, 1968) include the processes that direct attention among sensory modalities, the processes that transform sensory information into selected encodings, the process of rehearsal that maintains information in primary memory, and a variety of elaborate imagery or other recoding operations that build up long-term representations in secondary memory. In Figure 6, the term "executive" is used to represent the general management function of supervising the various control processes, and its span of control is indicated by the dashed box.

In the specific context of the tracking task, it was assumed that, at the beginning of each update, the observer first attends to the displayed information about one or more of the object trajectories. The displayed motion of the object(s) is held in iconic memory, which has a large capacity but a very rapid decay, on the order of about 1 second (Sperling, 1960). To yield any more than this momentary storage, the information must be transformed so that it is compatible with the encoding characteristics of primary memory (which is largely phonological). To effect this encoding, the executive invokes a scanning and a search recognition process in which a familiar representation is sought for the object's apparent trajectory. Such a representation is presumed to exist independently in secondary memory, having been developed from the observer's prior experience. For instance, an object moving horizontally from left to right might be encoded as "three

o'clock," with its trajectory represented as a pointing clock hand. Other representations or encodings are, of course, possible.

The initial encoding of an object gives rise to learning (i.e., the observer's knowledge of the object's true trajectory is increased). Additional, more refined learning may occur on subsequent updates.

The foregoing processes of attending, encoding, and learning are extremely rapid. The bulk of each update is presumably devoted to rehearsal; that is, maintaining the learned trajectories in primary memory. Only about 7±2-object trajectories can be retained (Miller, 1956) and, unless maintained by a rehearsal process, the information is lost within about 20 seconds (Peterson & Peterson, 1959). Rehearsal is presumably effected by subvocalizations and is mediated by echoic memory. The set of object trajectories is presumed to be stored in a sequence, then rehearsed in that sequence, and later retrieved in that same sequence (Sperling, 1963). The locus of the rehearsal process is referred to as the rehearsal buffer (Atkinson & Shriffin, 1968). If the demand for processing exceeds the capacity of the buffer, the executive must decide whether to forestall further processing of new items or to replace old items in the buffer with newly encoded ones. In either case, performance will suffer if the rehearsal capacity is exceeded.

Since the tracking task tests for immediate recall, information can be output directly from primary memory at the time of response. Also, the nature of the task would seem to confine the observer to relatively shallow levels of processing (i.e., visual or phonological encodings), providing little opportunity for more permanent storage in secondary memory. Therefore, issues that relate long-term retention, retrieval, and forgetting are not critical here, and the present models make no provisions for such processing.[1]

Mathematical Models

Development

A family of four mathematical models of the observer's processing was developed and evaluated. Each model includes:

1. An encoding process in which the observer transforms "raw" visual information into a form more suitable for storage and further processing.

2. A learning process that provides for stepwise improvement in the accuracy of an object's perceived trajectory at each update.

3. A limited rehearsal process that maintains the information learned about target trajectories in primary memory.

4. A response process that sequentially outputs information about the objects at test time.

---

[1] A model with a long-term memory component was developed. Here, secondary memory served as an auxiliary store for information that was "bumped out" of a fixed-capacity rehearsal buffer. While its predictions were satisfactory over the range of the conditions studied, the model made unreasonable predictions as the number of objects and updates were increased.

The four models make identical assumptions for the learning and response processes. Two alternative assumptions are introduced for the number of new objects encoded at each update. Also, two different assumptions are considered for an object's required rehearsal time. These variations combined to yield the family of four models that were evaluated.

Encoding. As the observer attends to the display, information first enters iconic memory. Further processing requires encoding, which is the transformation of this "raw" data into a form compatible with primary memory and entails invoking a familiar representation for the visual datum. The details of this process are not stipulated here, but some encoding scheme is deemed essential to yield efficient processing for multiple objects.

In general, the objects to be tracked need not all be encoded at once because (1) this might be a large processing burden at a single update, and (2) the limited rehearsal capacity might not be adequate to maintain them in storage. This leads to the introduction of two alternative assumptions for the encoding of new (i.e., previously unencoded) objects. The first assumption, E1, says that the observer keeps the rehearsal buffer filled to its capacity. This notion presumes that the executive knows the available buffer capacity at every update and directs the encoding process to provide sufficient new objects (if any remain) to keep the buffer working at its maximum. The second encoding assumption, E2, which is less restrictive, denies such complete control by the executive and recognizes that variability and lapses in the observer's processing may yield fewer objects for rehearsal than the buffer might hold.

Definitions: Let there be Z objects in the tracking task. At the kth update, define E_k as the number of objects previously encoded and A_k as the number of additional objects that can be accommodated in the rehearsal buffer. Let $N_k = \min (A_k , Z - E_k)$.

Assumption E1: At update k, the number of new objects encoded is precisely N_k.

Assumption E2: At update k, the number of new objects encoded is a random variable X with the following distribution:

$$P(X = x) = TP(x) = (e^{-\lambda} \lambda^x) / x! (1 - e^{-\lambda}) \text{ for } x = 1, 2, \ldots, (N_k - 1) \quad (1)$$

$$P(X = N_k) = \sum_{x = N_k}^{\infty} TP(x), \quad (2)$$

where TP(x) is the Poisson distribution truncated at zero (cf. Haight, 1967).

Under assumption E1, then, the extent of new encoding is precisely tuned to equal N_k, which is the smaller of (1) the additional room in the buffer (A_k) and (2) the number of objects as yet uncoded ($Z - E_k$).

Under E2, encoding is probabilistic. TP is the truncated Poisson distribution that excludes zero from its range and has λ as its parameter ($\lambda > 0$). With this distribution, it is theoretically possible to sample more objects than the N_k that can be accommodated in the buffer; in such cases, the number encoded is limited to this maximum of N_k objects

(equation 2 above). Note that, under E2, no new objects will be encoded, if and only if $N_k = 0$ (which follows from equation 2). Likewise, it follows from equation 2 that, if $N_k = 1$, then precisely one new object is encoded. In all other cases (i.e., for $N_k \geq 2$), the number of newly encoded objects will be an integer between 1 and N_k.

It should be seen that E1 and E2 make identical assumptions for $N_k = 0$ or 1. If $N_k > 1$, the assumptions diverge. E1 always ensures that a full sample of N_k new objects will be encoded, while E2 allows for random processing effects to yield less than this maximum.

Learning. While knowledge about an object's trajectory can vary continuously from no information to virtually perfect information, it is convenient to reduce this continuum to discrete states. Consistent with the three-valued scoring procedure used in the experiment, the proposed models assume three learning states:

State U: Unencoded (nothing is known about the object's trajectory).

State 1: Partial learning ($15° < |$ tracking error $| \leq 45°$).

State 2: Refined learning ($|$ tracking error $| \leq 15°$).

All objects start and remain in state U unless encoded. Any changes in state occur only in the initial moments of each update as the new positions of the objects are detected. At such times, the observer may encode an object for the first time or refine the information for a previously-encoded object.

When newly encoded, it is assumed that an object's state changes from U to either state 2 (with probability α) or to state 1 (with probability $1 - \alpha$). An object in state 1 may, on any update, be upgraded to state 2 with probability β. This improvement in knowledge is taken to be stochastic for two reasons: (1) The observer may not attend to the object when its position is updated, and (2) any improvement in the encoded trajectory may be insufficient to yield a higher level of accuracy according to the scoring criteria. Formally, the changes in learning states are defined by the following transition matrix:

		New Learning State		
		2	1	U
Prior Learning State	2	1	0	0
	1	β	$1 - \beta$	0
	U	α	$1 - \alpha$	0

The matrix allows stepwise improvement, but no regression in learning. Figure 7 illustrates the results of this learning process as a sequence of K updates unfolds for nine objects to be tracked.

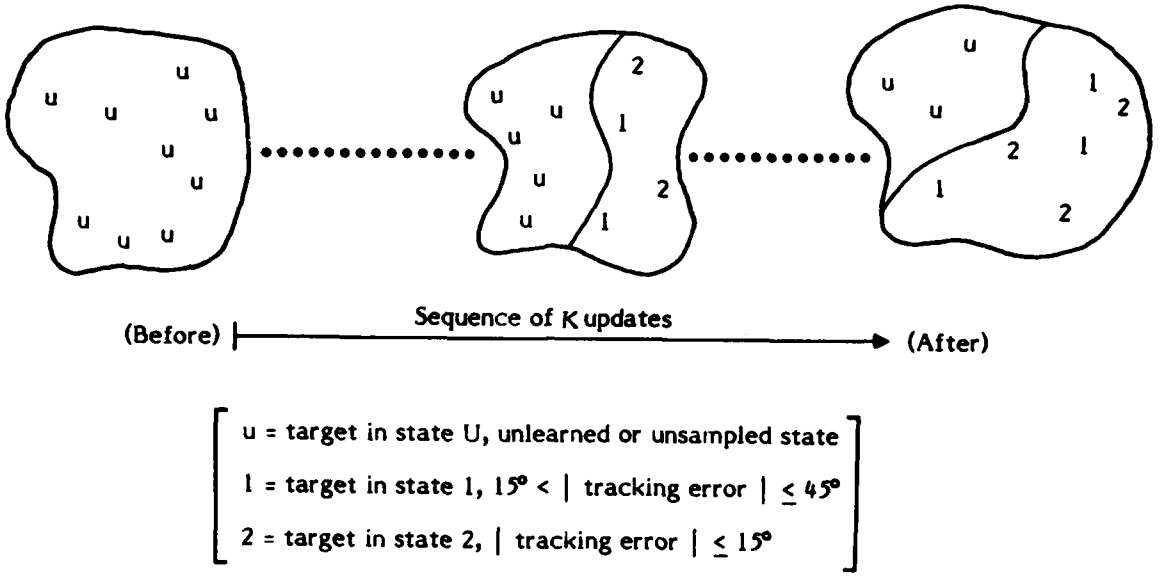


Figure 7. Illustration of learning states for nine objects before, during, and after a sequence of K updates.

If an object is encoded early in a sequence of updates (as compared with a later encoding), there will be additional opportunities to increase the level of learning. There is a greater probability that such an object will be "absorbed" in state 2. Similarly, increasing the total number of updates enhances learning and improves tracking performance.

Rehearsal.

1. Memory maintenance. Rehearsal serves as a memory maintenance function that refreshes the encoded representations that would otherwise be lost from primary memory. Restrictions on the rehearsal process ensure that only a limited number of objects can be retained. The following assumptions were made:

a. All encoded objects are rehearsed sequentially in the order of their encoding and then cyclically during the IUI.

b. In each cycle, each encoded object is rehearsed for its required rehearsal time, t.

c. There is a maximum time allowed (μ seconds) between an object's successive rehearsals.

Thus, objects are rehearsed in sequence and form cycles in which each object is rehearsed for its required t seconds. The rationale for the time requirement is that any rehearsal time that is less than some critical duration is ineffective in refreshing memory for the object's trajectory. A further restriction derives from the limited span of primary memory. Memory is presumed to decay unless an object is rehearsed at least once every μ seconds.

2. Required rehearsal time: fixed vs. variable. In addition to the μ restriction, the working capacity of the rehearsal buffer clearly depends on t, the time required for the rehearsal of one object. Two competing assumptions must be considered. The first holds that t is fixed for every update, while the second allows the required time to decrease as a function of previous rehearsals. These assumptions imply two different concepts for the capacity of the rehearsal buffer; namely, a fixed-size buffer vs. a variable-size buffer respectively.

Assumption R1: The required rehearsal time t for an object is fixed; that is, $t = \theta$ always.

Assumption R2: The required rehearsal time t for an object depends on the amount of its rehearsal during prior updates. In particular, t is a constant (θ) until at least τ seconds of prior rehearsal time have been accumulated, after which t decreases exponentially to ϵ. Formally, for an object with cumulative rehearsal time T on previous updates,

$$t = \begin{cases} \theta, & \text{if } T < \tau \\ \epsilon + (\theta - \epsilon)\exp\left[-\nu\ (T - \tau)\right], & \text{if } T \geq \tau \end{cases}$$

where $\tau, \nu > 0$; $\theta > \epsilon > 0$.

Assumption R1 clearly implies the familiar notion of a fixed-size buffer (Atkinson & Schiffrin, 1968); that is, the maximum number of objects that can be rehearsed on each update is constant. With assumption R2, however, a previously rehearsed object may require less time than one that is new and unrehearsed. Thus, as objects are rehearsed over several updates, they require less and less rehearsal time as illustrated in Figure 8. The effect is to make additional room available in the rehearsal buffer and effectively increase its capacity. Assumption R2, therefore, departs from other interpretations of buffer capacity (cf. Norman, 1968, Atkinson & Shiffrin, 1968).

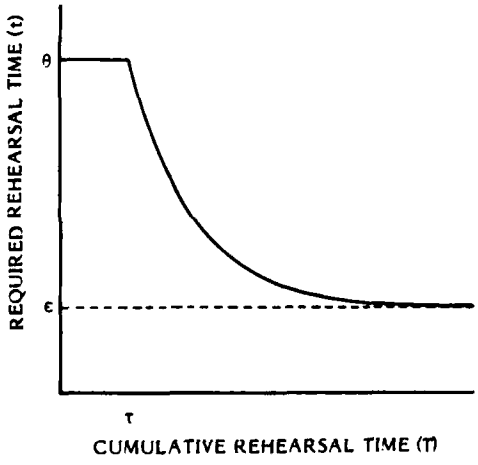


Figure 8. Required rehearsal time function proposed in assumption R2.

The assumed rehearsal process is illustrated in Figure 9, in which three encoded objects have required rehearsal times of t_1, t_2, and t_3. There is sufficient time in the IUI for three full rehearsal cycles, which consumes $3(t_1 + t_2 + t_3)$ seconds. Since the remaining time (say, r seconds) at the end of the IUI does not permit an additional cycle, it was arbitrarily assumed that the excess time is allocated equally to the three objects. The actual rehearsal time that object i would receive in this update is, therefore, $3t_i + r/3$. (Note that Figure 9 depicts variable rehearsal times that follow from assumption R2; assumption R1 would have fixed $t_i = \theta$ for all objects on all updates.)

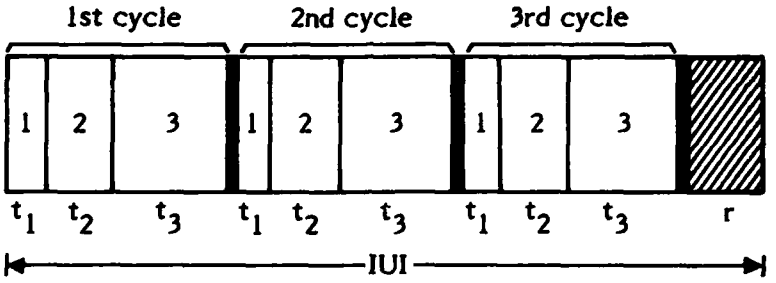


Figure 9. Example of rehearsal process during an update.

Response. After the last update, the observer's task is to indicate the next position for each object. It is assumed that the output of the rehearsal buffer is sequential and that the projected position for each object conforms to its last encoded trajectory. A chance response is made to all unencoded objects. When the three-valued scoring system is applied, each encoded object receives a score (1 or 2) that corresponds to its current learning state.

Summary of the Models. The four alternative models of performance in the multi-object tracking task are represented in Figure 10. The top of the figure shows that the models share identical assumptions for the probabilistic learning process (parameters α and β), the constraint on rehearsal maintenance (parameter μ), and the response process.

The first distinction among the models concerns the encoding of new objects. Models I and II (assumption E1) hold that enough new objects are encoded at each update to keep the rehearsal buffer filled. In constrast, Models III and IV (assumption E2) invoke a modified Poisson sampling process (with parameter λ) to govern the encoding of new objects.

The second distinction is the required rehearsal time for objects in the rehearsal buffer. Assumption R1 in Models I and III holds that the required rehearsal time is constant (parameter θ). Models of this type, when combined with the memory maintenance assumption, imply a conventional fixed-size rehearsal buffer. On the other hand, assumption R2 in Models II and IV postulates that required rehearsal time for an object varies as a function of its prior rehearsal (parameters θ, τ, ε, and ν). The effect is to create a variable-sized buffer as the required rehearsal times decrease over the sequence of updates.

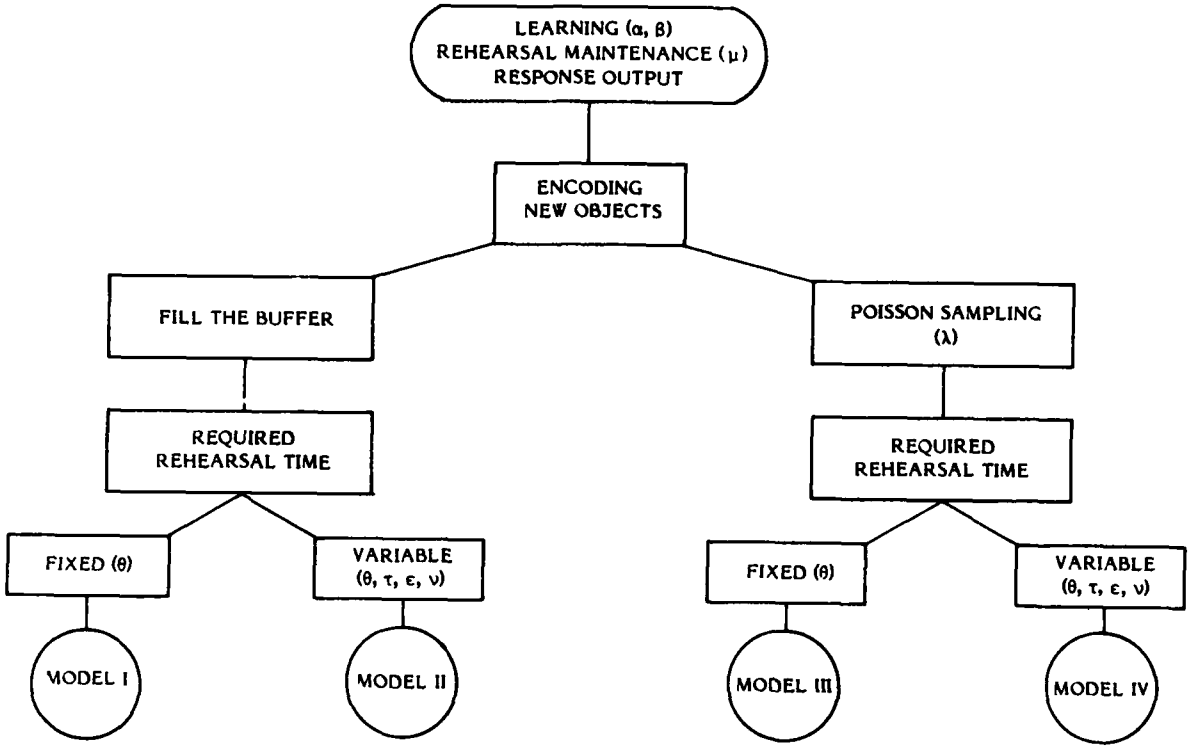


Figure 10. Alternative assumptions for Models I-IV.

Consider the effect of the different assumptions on the number of objects being tracked and rehearsed in the buffer as a sequence of six updates unfolds. For illustration, let the interval between updates be 8.0 seconds, let $\mu = 5.5$ (i.e., an object's trajectory must be rehearsed at least every 5.5 seconds), and let $\theta = 1.25$ seconds, so that Models I and III demand 1.25 seconds per rehearsal for each object. Then, let $\tau = 2.5$ seconds, $\varepsilon = 0.4$ seconds, and $\nu = 0.5$ govern the decreases in required rehearsal time that are assumed in Models II and IV.

For these parameter values, Table 3 presents the maximum number of objects that could be tracked and the actual number of objects tracked on each update for Models I-IV. ($\lambda \simeq 2$ and representative sampling in Models III and IV were assumed). It follows that, for the parameters selected, Model I predicts higher tracking performance than does Model III.[2] Although both have fixed buffers that can admit five objects, Model I immediately fills the buffer by encoding five objects at the first update. Model III, however, must obey the Poisson sampling assumption and here does not encode the fifth trajectory until the third update. Therefore, performance in Model I has the benefit of additional learning trials in which to refine the accuracy of the encoded trajectories.

[2]The two models will be equivalent if $\lambda >> 5$.

Table 3

Number of Objects Tracked for the Four Models on each Update

Number of Objects Tracked	Model I Update No.						Model II Update No.						Model III Update No.						Model IV Update No.					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Maximum	5	5	5	5	5	5	5	5	6	7	7	8	5	5	5	5	5	5	5	5	6	7	7	8
Actual[a]	5	5	5	5	5	5	5	5	6	7	7	8	2	3	5	5	5	5	2	3	5	6	7	8

Note. Parameter values used: $\mu = 5.5$, $\theta = 1.25$, $\epsilon = .4$, $\nu = .5$, $\lambda \simeq .2$. Interval between updates = 8.0 seconds.

[a] Models III and IV are not deterministic. Representative random sampling was assumed.

A similar argument leads to the conclusion that Model II predicts higher tracking performance than does Model IV.[3] Each has a variable-size buffer that can admit five objects on the first update and one object on updates 3, 4, and 6. But in Model II the buffer is always kept filled, while Model IV is constrained by its Poisson process.

Finally, with the parameters selected for this illustration, note that (1) Model II predicts better tracking than does Model I, and (2) Model IV predicts better tracking than does Model III. More objects are encoded in Models II and IV than in their counterparts.

The matter at issue is which of the four models more closely fits the performance data of the observers. Do the data support a choice between the two assumptions for encoding new objects and/or between the two assumptions for the required rehearsal time? Do the models yield plausible estimates for the several parameters?

Assessment

Three criteria were used to evaluate the models:

1. They must produce sensible predictions across broad ranges in the task variables (number of objects, number of updates, and length of IUI); that is, the predictions should agree qualitatively with the hypothetical performance surface in Figure 1.

2. There must be reasonable quantitative agreement between predictions and the data observed in the present experiment.

3. The best-fitting parameter values must be psychologically plausible.

[3] Ibid.

A Tektronix 4054 microcomputer was programmed in BASIC to yield, for each model, predictions of the mean tracking scores in the 12 experimental conditions. (Mathematical details are provided in the appendix.) Parameter estimation was based on a criterion that minimized the average absolute error in predicting the 12 means. A general "hill-climbing" FORTRAN algorithm (Wickens, 1967) was adapted to BASIC and used to search the parameter space.

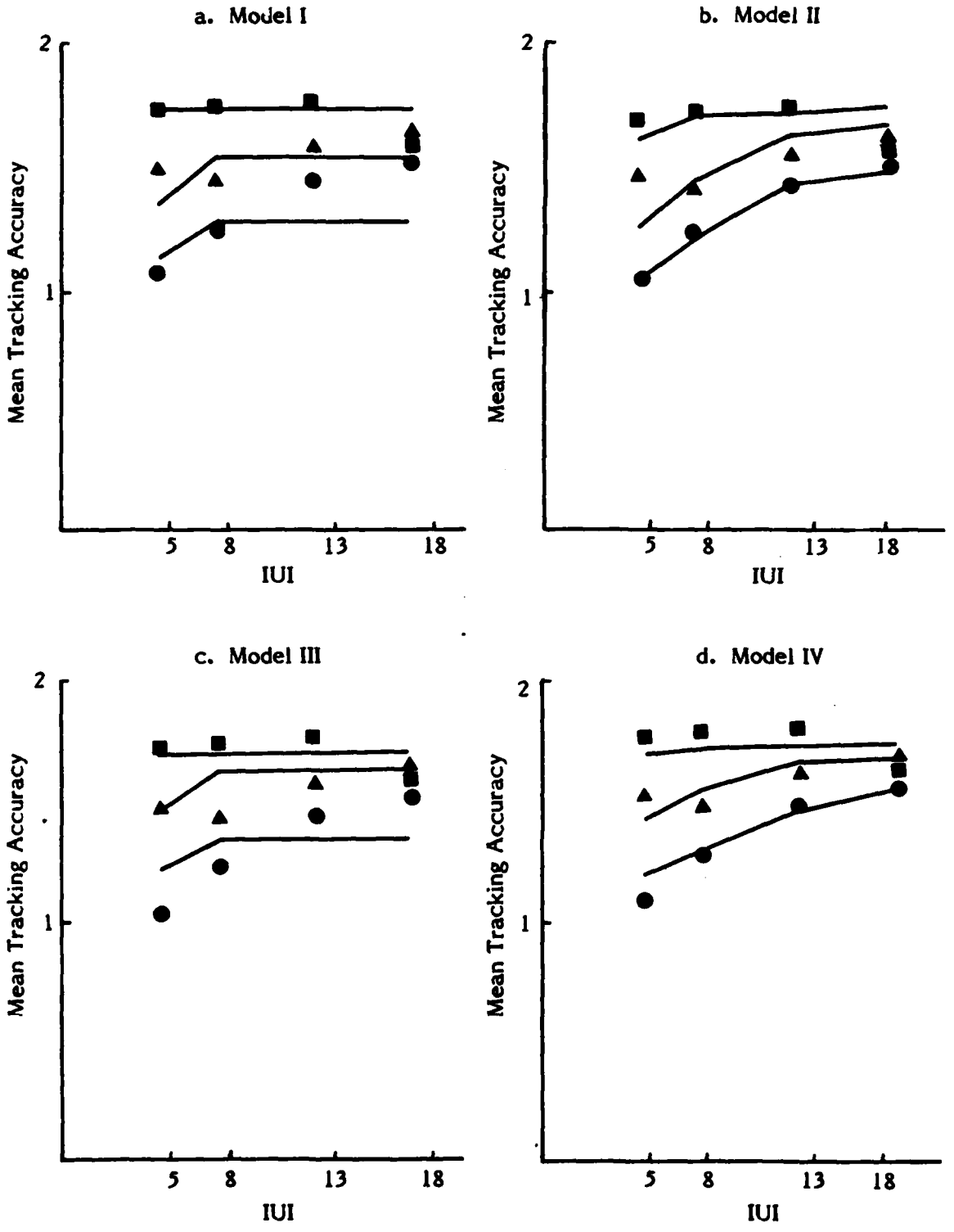
All four models make reasonable qualtitative predictions as the variables are allowed to take on extreme values. Thus, as the number of objects or updates or the IUI grows very large, predicted performance is nevertheless bounded within reasonable limits. Similarly, appropriate predictions are obtained with small values of the variables. Deciding among the models therefore requires more precise tests involving the fits to observed data.

The parameter estimates and the associated prediction errors for these quantitative tests of the four models are presented in Table 4. Models II and IV clearly provide better agreement with the observed data. The quality of the fits is also displayed in Figure 11, which plots the observed data and the predictions for Models I-IV.

Table 4

Parameter Estimates and Prediction Errors
for the Four Models

	Encoding	Learning		Rehearsal					
Model	λ	α	β	θ	μ	τ	ϵ	ν	Error
I	-	.58	.078	1.0	5.5	-	-	-	.089
II	-	.072	.233	1.4	4.5	2.80	.40	.50	.058
III	1.25	.083	.25	0.8	5.5	-	-	-	.084
IV	1.83	.20	.20	1.3	4.44	2.28	.40	.65	.054



Note. Solid lines are predictions. Plotted points are for five-object displays (■), seven-object displays (▲), and nine-object displays (●).

Figure 11. Observed data and predictions of the four models.

Thus, the evidence strongly favors the rehearsal assumption R2 in Models II and IV; viz., the time required to rehearse a given trajectory decreases as the prior rehearsal time increases. Models I and III are rejected. These assume fixed rehearsal times and imply rehearsal buffers with fixed capacity. However, they do not properly predict the improvement in performance observed as IUI increases.

There is no clear choice between the alternative encoding assumptions (E1 vs. E2) in Models II and IV. Model IV requires the additional Poisson parameter (λ) to govern the encoding of new objects, but its predictions were only trivially better than the deterministic "fill-the-buffer" assumption of Model II. The fits for these two preferred models are generally satisfactory, as shown in Figures 11b and 11d, and the values of the estimated parameters (Table 4) are reasonable.[4]

Human information processing in the tracking task may be interpreted as follows:

1. If there is room in the rehearsal buffer, the observer encodes about two objects on each update. In general, this is sufficient to fill or nearly fill the rehearsal buffer.

2. An object's first encoding is most likely (probability = $1 - \alpha \simeq .8$-$.9$) to produce marginal tracking accuracy (i.e., absolute error $\simeq$ 15-45°).

3. There is a moderate chance (probability = $\beta \simeq .2$) on each subsequent update of improving its accuracy.

4. An object's first rehearsal requires $\simeq 1.4$ seconds (parameter θ).

5. The rehearsal maintenance constraint demands that each trajectory be rehearsed at least every μ seconds, with $\mu \simeq 4.5$.

6. After two rehearsals of an object, sufficient rehearsal time is accumulated ($\tau \simeq 2.3 - 2.8$ seconds) to produce an exponential decrease in its required rehearsal times on succeeding updates. Rehearsal time then decreases to an asymptote = $\varepsilon \simeq 0.4$ seconds with rate = $\nu \simeq .5$ -.6).

The foregoing analysis implies that, with a sufficiently large number of updates, an unaided observer could maintain as many as 9 to 12 trajectories with an interval between updates of at least 5 seconds. If hardware/software aids were provided, the operator's capacity could, of course, be increased. Likewise, any demands of auxiliary tasks would result in lower performance limits due to disruption of processing.

---

[4]Also considered were two additional models that follow from a third encoding assumption; namely, that a constant (integer) number of objects is sampled on each update, subject to space in the buffer. Applying either assumption R1 or R2 yielded two alternative models, with prediction errors of .085 and .055 respectively. The best fit in each case was for a fixed sample size of two objects. However, the invariance of the sampling process in these models has less appeal. In any event, the general conclusions remain the same.

CONCLUSIONS

1. The unaided observer can keep track of, and later report the movements of, about seven objects. This constraint results from perceptual and memory limitations.

2. Tracking performance improves as the time between display updates is increased, up to about 13 seconds. This processing time is apparently required to establish and maintain suitable representations in memory. Performance can be improved by adding more updates; however, the several variables interact and may be traded off in a complex manner to affect tracking performance.

3. Short-term memory capacity can be enhanced by clustering strategies in which neighboring objects with similar trajectories are grouped together in memory.

4. Multi-object tracking performance is accounted for by two mathematical models that express human memory limitations in terms of encoding, learning, and rehearsal processes. Evidence supports the unconventional notion of a variable rehearsal capacity in short-term memory.

5. It is useful to analyze C^2 tasks that impose critical perceptual and memory requirements in terms of human information-processing characteristics.

RECOMMENDATION

Analysis of human memory and information-processing limitations should be applied to complex operational tasks to support system designers with quantitative estimates of operator performance.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (eds.), The psychology of learning and motivation (Vol. 2). New York: Academic Press, 1968, pp. 89-105.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.

Crowder, R. G. Principles of learning and memory. Hillsdale, NJ: Lawrence Erlbaum Associates, 1976.

Daniel, T. C. Nature of the effect of verbal labels on recognition memory for form. Journal of Experimental Psychology, 1972, 96, 152-157.

Haight, F. A. Handbook of the Poisson distribution. New York: John Wiley and Sons, Inc., 1967.

Howe, E. S. Effects of partial symmetry, exposure time, and backward masking on judged goodness and reproduction of visual patterns. Quarterly Journal of Experimental Psychology, 1980, 32, 27-55.

Mackworth, J. F. The visual image and the memory trace. Canadian Journal of Psychology, 1962, 16, 55-59.

Miller, G. A. The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Murray, D. J., & Newman, F. M. Verbal and visual coding in short-term memory. Journal of Experimental Psychology, 1973, 100, 58-62.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Norman, D. A. Toward a theory of memory and attention. Psychological Review, 1968, 75, 522-536.

Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. Journal of Experimental Psychology, 1959, 58, 193-198.

Pollack, I. Detection of changes in spatial position: Short-term visual memory or motion perception? Perception and Psychophysics, 1972, 11, 17-27.

Pollack, I. Detection of changes in spatial position: IV. Multiple display fields, display aiding, and interference. Human Factors, 1974, 16, 93-116.

Posner, M. I., & Boies, S. J. Components of attention. Psychological Review, 1971, 78, 391-408.

Sperling, G. The information available in brief visual presentations. Psychological Monographs, 1960, 74 (Whole No. 11).

Sperling, G. A model for visual memory tasks. Human Factors, 1963, 5, 19-31.

PRECEDING PAGE BLANK-NOT FILMED

White, C. T. Time compression of radar and sonar displays (Naval Electronics Laboratory Rep. 671). San Diego: Naval Electronics Laboratory, 18 February 1956.

Wickens, T. D. Parameters estimation in Markov chain learning models. Unpublished master's thesis, Brown University, 1967.

APPENDIX

DERIVATIONS OF THE PREDICTIONS FOR THE FOUR MODELS

DERIVATIONS OF THE PREDICTIONS FOR THE FOUR MODELS

The assumptions for encoding, memory maintenance, and rehearsal time determine strictly or govern probabilistically the updates at which objects get encoded; that is, they control the encoding schedules. These schedules will be described later, after a consideration of the tracking scores that the observer obtains at test time for a given schedule.

Tracking Score for an Encoded Object

Consider an arbitrary encoded object at test time. For each of the four models, the learning assumptions embodied in the transition matrix ensure that the object will be in either state 1 or 2. Its actual state depends only on (1) the update at which it was encoded and (2) the transition probabilities α and β. Suppose the object was encoded on update k ($1 \leq k \leq 6$). Then, at test time after the sixth update, the probability, $P_k(j)$, that the object is in state j, j = 1, 2 is given by

$$P_k(1) = (1-\alpha)\ (1-\beta)^{6-k},$$

$$P_k(2) = 1 - P_k(1).$$

This follows because, in order to terminate in state 1, the object must fail to enter state 2 on update k (probability = $1-\alpha$) and then must remain in state 1 (probability = $1-\beta$) on each of the (6 - k) updates that follow.

Next, recall that encoded objects receive a tracking score equal to their terminal encoding state, so that the expected score, say $\bar{x}_k$, for an object encoded on update k is

$$\begin{aligned} \bar{x}_k &= 2\,P_k(2) + P_k(1) \\ &= 2\,[\,1 - P_k(1)\,] + P_k(1) \\ &= 2 - P_k(1) \\ &= 2 - (1-\alpha)\ (1-\beta)^{6-k}. \end{aligned} \tag{A1}$$

Tracking Score for an Unencoded Object

If an object remains unencoded through the six updates, it terminates in state U and it is assumed that the observer makes a chance response at test time. Given the three-valued scoring rule tied to the error in its reported trajectory, its expected score, say $\bar{x}_U$, is given by

$$\begin{aligned} \bar{x}_U &= 2[\text{Prob}(|\text{error}| \leq 15^\circ)] + \text{Prob}(15^\circ < |\text{error}| \leq 45^\circ) \\ &= 2\ (30/360) + (60/360) \\ &= 1/3\,. \end{aligned} \tag{A-2}$$

Tracking Score for a Given Encoding Schedule

Let the total number of objects to be tracked be Z and consider an encoding schedule, E, in which e_k objects are newly encoded on the kth update. In general, E = (e_1,

$e_2, \ldots, e_k, \ldots, e_6)$, where $e_k = 0, 1, 2, \ldots$ and $\sum_{k=1}^{6} e_k \leq Z$. Again, for all four models, the expected tracking score, say $\bar{X}$, for the schedule $\mathcal{E}$ is the mean expected score for the Z objects that constitute the task; that is,

$$\bar{X}(\mathcal{E}) = \left[\sum_{k=1}^{6} e_k \bar{x}_k + (Z - \sum_{k=1}^{6} e_k) \bar{x}_u \right] / Z . \qquad (A3)$$

Here, each of the e_k objects encoded on the kth update has expected score $\bar{x}_k$ (given by equation A1), and each of the $(Z - \Sigma e_k)$ unencoded objects has an expected score = $\bar{x}_u$ = 1/3 (from equation A2). Note that equation A3 yields the (expected) tracking score, if the schedule $\mathcal{E}$ is known. (Of course, Z is under the control of the experimenter.) The schedules for Models I-IV must still be constructed.

Finding the Encoding Schedule

In general, there are three restrictions on e_k, the number of objects that are newly encoded on the kth update:

1. The number of unencoded objects. The number of objects newly encoded at a given update obviously cannot exceed the number of unencoded objects that remain. If E_k is defined as the number of objects encoded prior to update k, then there are $(Z - E_k)$ unencoded objects at the start of update k. It follows that

$$e_1 \leq Z - E_1 = Z$$

and

$$e_k \leq Z - E_k = Z - \sum_{j=1}^{k-1} e_i \text{ for } k > 1.$$

2. The duration of the IUI. Consider the E_k objects that had been encoded on prior updates; they must be rehearsed again on update k. Suppose that their required rehearsal times are $t_{1,k}, t_{2,k}, \ldots, t_{E_k,k}$ seconds, with sum $S_k = \sum_{i=1}^{E_k} t_{i,k}$. Then $(IUI - S_k)$ seconds remain to accommodate new objects, each of which requires θ seconds for its initial rehearsal. Therefore, $e_k \leq [INT(IUI - S_k) / \theta]$, where INT(a) is the largest integer $\leq a$.

3. The rehearsal maintenance constraint. This constraint demands that no object go unrehearsed for more than μ seconds. First, suppose that there are previously encoded objects (i.e., $E_k > 0$), and consider the oldest, first-encoded one. Under either assumption R1 or R2, its required rehearsal time, $t_{1,k}$, cannot exceed that of any other object. Therefore, any number of new encodings that satisfies the μ-constraint for the oldest object necessarily does so for all objects to be rehearsed. It then follows that e_k cannot exceed the largest integer having this property.

Formally,

$$e_k \leq \text{INT}\{[\mu - (S_k - t_{1,k})]/\theta\}, \text{ if } E_k > 0. \quad \text{(A4)}$$

If $E_k = 0$, there were no previous encodings and, in this case, the rehearsal constraint always permits an $e_k \geq 1$. In particular,

$$e_k = 1 + \text{INT}(\mu/\theta) = \text{INT}[(\mu + \theta)/\theta], \text{ if } E_k = 0. \quad \text{(A5)}$$

Since $E_k = 0$ implies $S_k = 0$, equations A4 and A5 can be combined as follows:

$$e_k \leq \text{INT}\{[\mu - (S_k - t_k^*)]/\theta\},$$

where

$$t_k^* = \begin{cases} \theta, \text{ if } E_k = 0 \\ t_{1,k}, \text{ if } E_k > 0. \end{cases}$$

Since each e_k must satisfy <u>jointly</u> the three restrictions (1-3 above), there is a fundamental upper bound for e_k:

$$e_k \leq \min\left(Z - E_k,\ \text{INT}[(IUI - S_k)/\theta],\ \text{INT}\{[\mu - (S_k - t_k^*)]/\theta\}\right). \quad \text{(A6)}$$

Let k=1. Then $E_k = S_k = 0$ and $t_k^* = \theta$, so that

$$e_1 \leq \min\left(Z,\ \text{INT}(IUI/\theta),\ \text{INT}[(\mu + \theta)/\theta]\right). \quad \text{(A7)}$$

Thus, equation A7 yields an upper bound for e_1, and upper bounds for e_2, e_3, etc. are given recursively by equation A6. The recursive burden is carried by the terms E_k (the number of objects encoded prior to the start of update k), S_k (the sum of their required rehearsal times, $t_{i,k}$), and $t_k^* = t_{1,k}$. Formulas for the encoding schedules produced by the four models are obtained by applying their specific sampling and rehearsal assumptions to equation A6.

<u>Model I</u>

Model I always encodes enough objects to fill the available room in the rehearsal buffer. This means that strict equality holds in equations A6 and A7. In particular, for Model I, equation A7 becomes:

$$e_1 = \min(Z, \text{INT}(IUI/\theta), \text{INT}[(\mu + \theta)/\theta]). \quad \text{(A8)}$$

Also, recall that Model I has fixed required rehearsal time (θ) for each encoded object on all updates. It follows that no new encodings are possible on updates $k > 1$. The buffer is already filled at update 1, and the required rehearsal times remain constant at θ thereafter. Thus, no additional room can be created in the buffer. Therefore, for Model I:

$$E = (e^*, 0, 0, 0, 0, 0)$$

where $e^* = \min(Z, \text{INT}(IUI/\theta), \text{INT}[(\mu + \theta)/\theta])$.

Model II

Model II also always encodes enough objects to fill the rehearsal buffer. However, unlike Model I, additional space may become available on later updates due to the reduced rehearsal times for previously-encoded objects. Since the buffer is always filled, strict equality again applies in equation A6. Thus, for Model II, e_1 is given by equation A8, and for $k > 1$

$$e_k = \min\left(Z - E_k,\ \mathrm{INT}[(|U| - \sum_{i=1}^{E_k} t_{i,k})/\theta],\mathrm{INT}\left\{[\mu - (\sum_{i=1}^{E_k} t_{i,k} - t_{1,k})/\theta]\right\}\right), \qquad \text{(A9)}$$

where $E_k = \sum_{i=1}^{k-1} e_i$.

By assumption R2, each encoded object i has a required rehearsal time $t_{i,k}$ on update k that depends on its prior cumulative rehearsal time and the four parameters θ, τ, ϵ, and ν. The encoding schedule, $\mathcal{E}$, is obtained by (1) fixing values for the parameters μ, θ, τ, ϵ, and ν; (2) applying equation A8 to yield e_1; and (3) recursively applying equation A9 to yield $e_2, e_3, \ldots, e_6$.

Model III

For Model III, the constant rehearsal times imply a fixed-size buffer, but limitations in the encoding process cannot guarantee to fill the buffer at each update. An upper bound for e_1 is given by equation A7. The inequality is preserved because Model III, unlike Models I and II, depends on statistical sampling to encode new objects. According to the fixed rehearsal time assumption, all $t_{i,k} = \theta$. Thus, $t_k^* = \theta$ for all k and $S_k = \theta E_k$. Thus, equation A6 becomes

$$e_k \leq \min\ Z - E_k,\ \mathrm{INT}[(|U| - \theta E_k)/\theta],\ \mathrm{INT}\{[\mu - \theta(E_k - 1)/\theta\} , \qquad \text{(A10)}$$

where again $E_k = \sum_{i=1}^{k-1} e_i$.

When μ and θ are given, Model III gives rise to a set of possible encoding schedules rather than a single $\mathcal{E}$. First, μ and θ determine an upper bound for e_1 via equation A7. The Poisson sampling scheme then determines the probability of the possible values of e_1, with a maximum value given by the right-hand side of equation A7. Each of these values, in turn, determines an upper bound for e_2 via equation A10; the truncated Poisson distribution is again applied to obtain the probability of the possible values for e_2. This recursion is continued for all six updates. The result is a set of possible encoding schedules $\{\mathcal{E}_i\}$ together with the probability of each schedule being realized, say, $P(\mathcal{E}_i)$.

For any one schedule, $\mathcal{E}_i = (e_1, e_2, \ldots, e_k, \ldots, e_6)$, and

$$P(\mathcal{E}_i) = \prod_{k=1}^{6} \mathrm{Prob}\,(X = e_k),$$

which is the probability given by assumption E2, with $x = e_k$ and N_k equal to the right-hand side of equation A6. Finally, the expected tracking score for Model III is given by:

$$\text{Expected tracking score} = \sum_i \bar{X}(\mathcal{E}_i)\ P(\mathcal{E}_i), \qquad \text{(A11)}$$

where $\bar{X}(\mathcal{E}_i)$ is that of equation A3.

Model IV

Model IV postulates Poisson sampling and variable rehearsal time. Here, the reduced rehearsal times generate a variable-sized buffer, and the encoding of new objects is governed by the probability distribution in Assumption E2. As in Model III, equation A7 yields an upper bound for e_1. For $k > 1$, we use equation A9 recursively except that strict equality does not hold in Model IV. Rather, the right-hand side gives an upper bound for each e_k.

The parameters μ, θ, τ, ε, and ν determine the overall limits for values of e_k. The actual e_k values again depend on statistical sampling. Also, as in Model III, the set of possible encoding schedules must be elaborated, together with the probability of each schedule's occurrence. The expected tracking score for Model IV is obtained as in equation A11.

Parameter Estimation

The criterion for "goodness-of-fit" was minimization of the absolute error in predicting the observed data. A standard procedure would be to submit each model to the parameter estimation program with all its parameters free to vary. This was done for Model I. For Models II-IV, however, the mapping from the set of rehearsal parameters $\{\mu, \theta, \tau, \varepsilon, \nu\}$ to the encoding schedule $\mathcal{E}$ (or to the set of schedules $\{\mathcal{E}_i\}$) is many-to-one. Accordingly, $\mathcal{E}$ or $\{\mathcal{E}_i\}$ was first fixed by choosing values for the appropriate rehearsal parameters. Then the model was submitted to the estimation program with free parameters α and β (and λ if applicable). This technique--selecting rehearsal parameters to yield distinct schedules and then optimizing the remaining parameters--was iterated to yield the overall best parameter set.

DISTRIBUTION LIST

Director of Manpower Analysis (ODASN(M))
Chief of Naval Operations (OP-01), (OP-11), (OP-12) (2), (OP-13), (OP-14), (OP-15), (OP-115) (2), (OP-140F2), (OP-987H)
Chief of Naval Material (NMAT 00), (NMAT 05)
Chief of Naval Research (Code 200), (Code 440) (3), (Code 442), (Code 448)
Chief of Information (OI-213)
Chief of Naval Education and Training (N-5)
Commander in Chief U.S. Atlantic Fleet
Commander in Chief U.S. Pacific Fleet
Commander Fleet Training Group, Pearl Harbor
Commander Naval Air Force, U.S. Atlantic Fleet
Commander Naval Air Force, U.S. Pacific Fleet
Commander Naval Surface Force, U.S. Atlantic Fleet
Commander Naval Surface Force, U.S. Pacific Fleet
Commander Naval Military Personnel Command (NMPC-013C)
Commander Submarine Force, U.S. Atlantic Fleet
Commander Submarine Force, U.S. Pacific Fleet
Commander Training Command, U.S. Atlantic Fleet
Commander Training Command, U.S. Pacific Fleet
Commanding Officer, Naval Regional Medical Center, Portsmouth, VA (ATTN: Medical Library)
Director, Naval Civilian Personnel Command
Director, Training Analysis and Evaluation Group (TAEG)
President, Naval War College (Code E114)
Superintendent, Naval Postgraduate School
Secretary Treasurer, U.S. Naval Institute
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria (PERI-ASL)
Chief, Army Research Institute Field Unit, Fort Harrison
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Scientific and Technical Information Office)
Commander, Air Force Human Resources Laboratory, Williams Air Force Base (AFHRL/OT)
Commander, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base (AFHRL/LR)
Superintendent, U.S. Coast Guard Academy
Defense Technical Information Center (DDA) (12)