1

PM-1 3 1/2"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services Branch

Direction des acquisitions et
des services bibliographiques

395 Wellington Street
Ottawa, Ontario
K1A 0N4

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Memory Load and the Rate of Forgetting

MEMORY LOAD AND THE RATE OF FORGETTING
OF UNIDIMENSIONAL PERCEPTUAL STIMULI

by

CRAIG LETH-STEENSEN

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of

MASTER OF ARTS

Department of Psychology

Carleton University

Ottawa, Ontario

August, 1994

copyright 1994, Craig Leth-Steensen

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

385 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

385, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

**THE AUTHOR HAS GRANTED AN
IRREVOCABLE NON-EXCLUSIVE
LICENCE ALLOWING THE NATIONAL
LIBRARY OF CANADA TO
REPRODUCE, LOAN, DISTRIBUTE OR
SELL COPIES OF HIS/HER THESIS BY
ANY MEANS AND IN ANY FORM OR
FORMAT, MAKING THIS THESIS
AVAILABLE TO INTERESTED
PERSONS.**

**L'AUTEUR A ACCORDE UNE LICENCE
IRREVOCABLE ET NON EXCLUSIVE
PERMETTANT A LA BIBLIOTHEQUE
NATIONALE DU CANADA DE
REPRODUIRE, PRETER, DISTRIBUER
OU VENDRE DES COPIES DE SA
THESE DE QUELQUE MANIERE ET
SOUS QUELQUE FORME QUE CE SOIT
POUR METTRE DES EXEMPLAIRES DE
CETTE THESE A LA DISPOSITION DES
PERSONNE INTERESSEES.**

**THE AUTHOR RETAINS OWNERSHIP
OF THE COPYRIGHT IN HIS/HER
THESIS. NEITHER THE THESIS NOR
SUBSTANTIAL EXTRACTS FROM IT
MAY BE PRINTED OR OTHERWISE
REPRODUCED WITHOUT HIS/HER
PERMISSION.**

**L'AUTEUR CONSERVE LA PROPRIETE
DU DROIT D'AUTEUR QUI PROTEGE
SA THESE. NI LA THESE NI DES
EXTRAITS SUBSTANTIELS DE CELLE-
CI NE DOIVENT ETRE IMPRIMES OU
AUTREMENT REPRODUITS SANS SON
AUTORISATION.**

ISBN 0-315-98506-2

Canada

Name

IRAIL LETH-STEESEN

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

PSYCHOLOGY

SUBJECT TERM

0623 UMI
SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS

Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION

General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0277
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523

Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0711
Tests and Measurements 0268
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language 0679
General 0289
Ancient 0290
Linguistics 0291
Modern 0401
Literature 0294
General 0295
Classical 0297
Comparative 0298
Medieval 0316
Modern 0591
African 0305
American 0352
Asian 0355
Canadian (English) 0593
Canadian (French) 0311
English 0312
Germanic 0315
Latin American 0313
Middle Eastern 0314
Romance 0314
Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy 0422
Religion 0318
General 0321
Biblical Studies 0319
Clergy 0320
History of 0322
Philosophy of 0469
Theology 0323

SOCIAL SCIENCES

American Studies 0324
Anthropology 0326
Archaeology 0327
Cultural 0310
Physical 0272
Business Administration 0770
General 0454
Accounting 0338
Banking 0385
Management 0501
Marketing 0503
Canadian Studies 0505
Economics 0508
General 0509
Agricultural 0510
Commerce-Business 0511
Finance 0358
History 0366
Labor 0351
Theory 0358
Folklore 0366
Geography 0351
Gerontology 0578
History 0578
General 0578

Ancient 0579
Medieval 0581
Modern 0582
Black 0328
African 0331
Asia, Australia and Oceania 0332
Canadian 0334
European 0335
Latin American 0336
Middle Eastern 0333
United States 0337
History of Science 0585
Low 0398
Political Science 0615
General 0616
International Law and Relations 0617
Public Administration 0814
Recreation 0452
Social Work 0626
Sociology 0627
General 0738
Criminology and Penology 0631
Demography 0628
Ethnic and Racial Studies 0629
Individual and Family Studies 0630
Industrial and Labor Relations 0700
Public and Social Welfare 0344
Social Structure and Development 0709
Theory and Methods 0999
Transportation 0453
Urban and Regional Planning 0453
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture 0473
General 0285
Agronomy 0475
Animal Culture and Nutrition 0476
Animal Pathology 0359
Food Science and Technology 0478
Forestry and Wildlife 0479
Plant Culture 0480
Plant Pathology 0817
Plant Physiology 0777
Range Management 0746
Wood Technology 0306

Biology

General 0306
Anatomy 0287
Biostatistics 0308
Botany 0309
Cell 0379
Ecology 0329
Entomology 0353
Genetics 0369
Limnology 0793
Microbiology 0410
Molecular 0307
Neuroscience 0317
Oceanography 0416
Physiology 0433
Radiation 0821
Veterinary Science 0778
Zoology 0472

Biophysics

General 0786
Medical 0760

EARTH SCIENCES

Biogeochemistry 0425
Geochemistry 0996

Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleocology 0426
Paleontology 0418
Paleozoology 0985
Polynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences 0768
Health Sciences 0566
General 0300
Audiology 0992
Chemotherapy 0567
Dentistry 0350
Education 0769
Hospital Management 0758
Human Development 0982
Immunology 0564
Medicine and Surgery 0347
Mental Health 0569
Nursing 0570
Nutrition 0380
Obstetrics and Gynecology 0354
Occupational Health and Therapy 0381
Ophthalmology 0571
Pathology 0419
Pharmacology 0572
Pharmacy 0382
Physical Therapy 0573
Public Health 0574
Radiology 0575
Recreation 0575

Speech Pathology 0460
Toxicology 0383
Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences 0485
Chemistry 0749
General 0486
Agricultural 0487
Analytical 0488
Biochemistry 0738
Inorganic 0490
Nuclear 0491
Organic 0494
Pharmaceutical 0495
Physical 0754
Polymer 0405
Radiation 0605
Mathematics 0986
Physics 0606
General 0608
Acoustics 0748
Astronomy and Astrophysics 0607
Electronics and Electricity 0798
Elementary Particles and High Energy 0759
Fluid and Plasma 0609
Molecular 0611
Nuclear 0752
Optics 0756
Radiation 0611
Solid State 0463
Statistics 0346
Applied Sciences 0984
Applied Mechanics 0984
Computer Science 0984

Engineering 0537
General 0538
Aerospace 0539
Agricultural 0540
Automotive 0541
Biomedical 0542
Chemical 0543
Civil 0544
Electronics and Electrical 0348
Heat and Thermodynamics 0545
Hydraulic 0546
Industrial 0547
Marine 0794
Materials Science 0548
Mechanical 0743
Metallurgy 0551
Mining 0552
Nuclear 0549
Packaging 0765
Petroleum 0554
Sanitary and Municipal 0790
System Science 0428
Geotechnology 0796
Operations Research 0795
Plastics Technology 0994
Textile Technology 0994

PSYCHOLOGY

General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

Memory Load and the Rate of Forgetting

Abstract

The effects of memory load (i.e., the number of stimuli stored in memory) on the individual rates of forgetting for stimuli in short-term perceptual memory was examined. Previous experiments (Palmer, 1990) suggested that set-size effects of the number of to-be-remembered stimuli on performance in a successive-comparison task were limited to attentional processes alone and did not affect memory processing (i.e., the rate of forgetting). In that research, subjects were asked to study and remember the lengths of either one, two, or four lines presented in a multiple-object display. The subjects then compared the length of a successively-presented comparison line with the length of the study line which had been previously presented in that same location. A similar paradigm was employed in the present research with the exception that subjects (two in Experiment 1 and two in Experiment 2) performed a two-alternative forced-choice discrimination task, longer acquisition times were provided, three ISI's (0.5, 2, and 5 seconds) were used in order to obtain true forgetting curves, and there was an attempt to generalize the results to absolute position discrimination (Experiment 2).

An extant quantitative model of perceptual memory,

Memory Load and the Rate of Forgetting

the diffusion model (Kinchla & Symzer, 1967), was applied to the observed d' measures of perceptual-memory performance. Within this model, the forgetting of a perceptual stimulus was regarded as the representational diffusion of the remembered value of that stimulus. The assumption was made that the rate of diffusion can depend on the amount of available attention and hence on the number of stimuli which were required to be maintained in memory. This assumption was supported (but not conclusively) by the data.

Furthermore, large time-order errors were found in the response data of all four subjects. The potential origin of these errors as well as their potential effects on the observed d' measures were discussed.

Memory Load and the Rate of Forgetting

Table of Contents

	Page:
Abstract	ii
Table of Contents	iv
List of Tables	v
List of Figures	vi
List of Appendices	vii
Introduction	1
Attentional limitations on perception	5
The diffusion model	11
Attentional limitations in memory	14
The present research	16
Experiment 1	19
Method	19
Results and Discussion	24
The diffusion model estimates	26
Time-order error	32
Experiment 2	46
Method	49
Results and Discussion	51
The diffusion model estimates	52
Time-order error	56
General Discussion	58
References	63

Memory Load and the Rate of Forgetting

List of Tables

Table:	Title:	Page:
1	Percent correct for Experiment 1	69
2	Diffusion Model Estimates for Experiment 1	70
3	TOE Indices for Experiment 1	71
4	Percent correct for Experiment 2	72
5	Diffusion Model Estimates for Experiment 2	73
6	TOE Indices for Experiment 2	74

Memory Load and the Rate of Forgetting

List of Figures

Figure:	Description:	Page:
1	Observed and predicted d' scores for Experiment 1.	75
2	Observed and predicted d' scores for Experiment 2.	76

Memory Load and the Rate of Forgetting

List of Appendices

Title:	Page:
Appendix A	77
Appendix B	79
Appendix C	81
Appendix D	84
Appendix E	85

Memory Load and the Rate of Forgetting
of Unidimensional Perceptual Stimuli

After years of memory research, few researchers would be willing to dispute that short-term memory, for most types of stimuli, deteriorates monotonically over the time between the presentation of a to-be-remembered study stimulus and its subsequent recall or recognition. However, memorability for (non-degraded) stimulus items, over blank retention intervals, can be influenced by certain factors such as the number of stimuli to be remembered (Miller, 1956), the level of processing of the stimulus items in memory (Craik & Lockhart, 1972), the ease with which the stimulus items may be verbally encoded and rehearsed (Peterson & Peterson, 1959), and the original presentation, or level of acquisition, of each stimulus in memory (Norman, 1966). Measures of memory performance can also be influenced at the response stage by any response biases which might be present as well as by the number of available competing responses (Norman, 1966; Norman & Wickelgren, 1969).

An examination of these factors in turn, from the latter to the former, reveals that it is possible to control for their influence on memorability by the appropriate choices of experimental task, performance measure, procedure, and type of stimulus. In order to limit the number of responses, each of which contributes a potential source of noise to the decision process, recognition tasks have been employed (in which the

response would be limited to either "yes", a test stimulus was presented previously , or "no", a test stimulus was not previously presented), instead of recall tasks (in which the appropriate response could be any one of a number of relevant stimuli). In order to bypass the effects of response-criterion biases on the observed data, previous researchers have relied on d' , the distance of the receiver operating characteristic (ROC) from the chance diagonal (signal-detection theory, Green & Swets, 1966), rather than simple percent correct, as a measure of memory performance. To control for differences in acquisition, or input processes, the duration of the presentation of a single study stimulus can be held constant over all conditions of the experiment (Kinchla & Symzer, 1967).

Next, perceptual stimuli can be used as study stimuli. As opposed to more discrete stimuli, such as words or numbers, perceptual, or sensory-based, stimuli can take any one of a continuum of possible values. For this reason, memory studies using perceptual stimuli have several advantages over the study of memory for discrete, verbal stimuli. First, because memory for perceptual stimuli does not involve explicit verbal coding, it has been assumed that effective rehearsal strategies are minimized in such memory tasks (Massaro, 1970b). Secondly, it can be assumed that perceptual stimuli have minimal related semantic associations (Laming & Scheiwiller, 1985). Hence, if perceptual stimuli are

employed, control for these two factors is unnecessary.

Previous researchers have shown in successive-comparison tasks that short-term memory for unidimensional perceptual stimuli degrades over the time between the presentation of a study stimulus and a corresponding test stimulus (the inter-stimulus interval, or ISI). This deterioration of memory has been observed for the absolute positions of points of light, loudness of tones (Kinchla & Smyzer, 1967; Kinchla & Allan, 1969, 1970), pitch discrimination (Massaro, 1970a, 1970b, 1970c; Wickelgren, 1969), and line lengths (Palmer, 1991). Under certain conditions, however, some visual stimuli have shown very little forgetting over time. These types of stimuli include the recall of the relative locations of points positioned on a line, when subjects were not required to perform an interpolated task over the ISI (Posner & Konick, 1966; see also Posner, 1967), and the recognition of matrices filled with black and white squares over small ISIs of less than a second, when the matrices were moved slightly from the position of the initial study matrices (Phillips, 1974; see also Purdy, Eimann, & Cross, 1980).

Lastly, due to the limited capacity available in short-term memory, it should be desirable to limit the number of to-be-remembered stimuli. Typically, the upper limit of available short-term memory span for discrete verbal stimulus items (i.e., the number of stimuli which can be kept active in immediate memory) is about seven distinct items (Miller,

1956). For unidimensional perceptual stimuli, such as brightness or line length, Miller (1956) reported data that indicated that the span of absolute judgment (i.e., the number of stimuli that could be accurately discriminated in an absolute judgment task) was about five or six items. While it cannot be assumed that absolute judgment span and short-term memory span represent different aspects of the same underlying process (because these two types of spans were shown to behave differently as the stimuli became more dimensionally complex, a phenomenon which suggested to Miller, 1956, that people can "chunk" information in memory), their operational similarity suggests that the span for any unidimensional perceptual stimulus can be approximated by its span of absolute judgment.

One issue that remains to be adequately resolved, however, concerns the dependence over time of short-term memory forgetting on the number of to-be-remembered stimuli, i.e., memory load. In other words, assuming that the number of stimulus items is within the appropriate memory span for that stimulus, is the rate of forgetting for each individual stimulus, over relatively short retention intervals, dependent on the number of stimuli which are required to be kept active in memory? Laming and Scheiwiller (1985) pointed out that a small set of discrete chunks of verbal information can be remembered virtually error-free over relatively short retention intervals. This level of performance can be accomplished due to both extremely effective rehearsal

strategies and subjects' knowledge of the discrete neighbourhood of possible responses (e.g., letters of the alphabet). For example, Williams, Beaver, Spence, and Rundell (1969) performed two experiments which required subjects to remember six single digits presented at a rate of one digit per second. After a 30 second rest period, the average recall (over both experiments) was 5.73 digits, with an average standard deviation of .36 digits. Hence, in order to control for this ceiling effect of memory performance within the memory span for discrete verbal items, it seems more appropriate to examine the dependence of short-term memory forgetting on memory load within a paradigm which employs purely perceptual stimuli.

Attentional limitations on perception

The only experiment that explored short-term forgetting in this manner was in a study by Palmer (1990). Palmer (1990; see also Palmer, 1988, 1991, and Palmer, Ames, & Lindsey, 1993) manipulated both ISI and study-stimulus set size, over a constant 100 ms acquisition interval, by presenting 2 subjects with either two or four line lengths. These study stimuli were arranged on an "imaginary" circle 180 arc min around a central fixation point (Experiment 4; a study set size of one was also included in other experiments). After either a 0.5 or an 8 second retention interval (ISI), a test line was presented for 100 ms. The test line was located in the same position as one of the previous study stimuli.

Subjects were required to decide whether the test stimulus was longer or shorter than the corresponding study stimulus at that same location. By using the test stimulus as the standard and the study stimuli as the comparisons in the method of constant stimuli, Palmer calculated psychometric functions (using Probit analysis to fit the cumulative normal curves) and difference limens (DLs) for subjects at each set-size condition at each retention ISI. The DLs provided a measure of the discriminability of the remembered stimulus from the test stimulus at the time of response. Hence, in Palmer's experiments (1990) enhanced memory degradation, or forgetting, was operationalized as the presence of larger DLs. (DLs are inversely related to d' , Laming & Scheiwiller, 1985).

This paradigm, which Palmer (1990, see also Palmer, 1988, 1991) referred to as partial discrimination due to its general similarity to Sperling's (1960) partial report paradigm, had several important experimental advantages. First, because the stimuli were equally spaced around the fixation point, possible sensory effects due to lateral masking were controlled (see also Grindley & Townsend, 1968). Second, the 100 ms presentation time ensured that the study stimuli were only available to be viewed peripherally and controlled for set-size effects due to serial eye movements (which were not possible at such short presentation times; see also Palmer, Ames, & Lindsey, 1993). Third, because the comparison task itself was based on only one of the initially-presented study

stimuli the decision process was assumed to be independent of the number of study stimuli (see also Lindsay, Taylor, & Forbes, 1968). In other words, any set-size effects shown by differences in the DLs must have been due to processing in either memory or perception alone.

Palmer's (1990) results indicated that memory was degraded when subjects had four line lengths to remember as compared to when they had only two line lengths as study stimuli. He also found a degradation of memory as the retention ISI increased from 0.5 to 8 sec. However, the simple main effects of ISI, for DLs measured in log arc min, were the same for each set size, i.e., no interaction of set size and ISI was observed.

The fact that these additive results were found for DLs which had been log transformed was not trivial. Previous researchers (Massaro, 1970a; see also Norman, 1966, for memory for verbal stimuli) have assumed that perceptual information stored in memory is lost, from some initial level of acquisition (which should depend, in this current paradigm, on the study set size), according to an exponential decay process:

$$\underline{d'} = \underline{A}e^{-\underline{h}\underline{t}} , \quad (1)$$

where d' represents the strength of the stimulus trace in memory, A represents the initial acquisition level, and h is the appropriate decay parameter over t (measured in length of ISI). Taking the natural logarithms of both sides of (1)

gives:

$$\ln d' = \ln A - ht, \quad (2)$$

which is a linear function of $\ln d'$ against t , the slope (h) of which represents the rate of forgetting. Hence, when DLs (which are inversely related to d') for each study set size are plotted in log arc min as a function of ISI, the contributions of both acquisition ($\ln A$) and the rate of decay (h) of the stimulus items in memory can be separated and examined in isolation. Because the slopes of these decay functions, in Palmer's (1990) Experiment 4, were parallel between study set sizes, these results strongly suggested that the actual rates of forgetting of the individual stimuli in each study set size were the same. (See also Loftus, 1985, for an excellent treatment of issues involved in the evaluation of forgetting data.)

Consequently, Palmer (1990) proposed that the set-size effects on memory for the line lengths in his experiments were due simply to limitations in perceptual attentional processing during acquisition. As more study stimuli were presented over the same acquisition presentation time, attention must have been allocated over a greater number of percepts. Less attention could be devoted to each study stimulus which, in turn, affected their subsequent storage in short-term memory. Attentional control, defined in terms of selectivity, was regarded as the major determinant of performance in Palmer's experiments because he found similar effects for conditions in

which only two line lengths were presented in the study display and conditions in which four lines were presented but subjects were instructed to attend only to two "relevant" line lengths (Palmer, 1990, Experiment 5). He also found that subjects performed poorly when they were asked to make comparisons between the test stimulus and to-be-ignored study stimuli when subjects were given non-valid attentional cues in the study display (Palmer, 1990, Experiment 7).

Several artifacts, which could have possibly accounted for the effect of set size were examined and ruled out as mediating factors in Palmer's (1990) experiments. Potential sensory artifacts (Experiment 1) included foveal viewing of the test stimuli through anticipatory eye movements, illusory effects of the distraction stimuli, and the absence of configuration cues for the test stimulus. As well, an examination of the psychometric functions in his Experiment 2 showed that performance for study set sizes of four was much better than would be expected if subjects had found it necessary to guess on a high proportion of trials due all-or-none attention-switching process during acquisition (see also Shaw, 1980).

One model of attentional processing during perception which was deemed compatible with Palmer's findings was the sample size model of attention such as that proposed by Lindsay, Taylor, and Forbes (1968). If the initial perception of the value of a unidimensional stimulus can be viewed as an

aggregate, or mean, of a number of independent samples, whose values fluctuate statistically due to both the external and internal noise inherent in processing that stimulus, then the variability, over comparison trials, associated with that aggregate value should increase as the number of available samples decreases (in accordance with basic statistical principles). In a typical signal-detection model of the discrimination process, the internalized psychological values of any two stimuli (x_1 and x_2), which have been repeatedly presented over many trials, are represented as random variables whose distributions have a mean and variance which have been determined by the mean and variance of the perceived (i.e., aggregate) input values of those stimuli. Assuming that these internalized distributions are normally distributed and have equal variance, the measure d' , which has been referred to as the discriminability index, is defined as the distance between the means of these distributions ($E[x_1]$ and $E[x_2]$) scaled in units of their common standard deviation ($[\text{Var}(x)]^{1/2}$):

$$d' = E(x_2) - E(x_1) / [\text{Var}(x)]^{1/2}. \quad (3)$$

Hence, d' should vary inversely with increases in input, or acquisitional, variance. In Palmer's experiments (1990), because observers were required to share attention over more stimuli as set sizes were increased the number of samples that could be accumulated was reduced for each stimulus. Accordingly, for larger set sizes, the variance associated

with the means of these sample values on each comparison trial would have increased, a measure of d' would have been assumed to decrease, and hence the DLs (which vary inversely with d') were observed to increase. Furthermore, if the sample size model alone was enough to account for the set-size results in these experiments, then it would explicitly predict that the size of the observed DLs should be proportional to the square root of the study set size (this quantitative relation is derived in Appendix A). In Palmer's (1990) Experiment 1, with single ISIs of two seconds, exactly this relation was found. DLs for set sizes of two and four were inflated from the set-size one DL by a factor of 1.41 and 2, respectively. (It should be noted [see Appendix B] quantitative predictions involving relations between d' s change slightly if the d' measure has been calculated under two-alternative forced-choice task conditions.)

The diffusion model

While the results in Palmer's (1990) Experiment 4 represented evidence for a memory process in which the rate of forgetting was independent of memory load, there were several reasons why it was felt that a replication of these findings was warranted. First, however, a theoretical viewpoint which more fully involves the role of memory, as well as possible attentional dependencies within the memory process, is considered as a possible alternative hypothesis.

In order to examine in greater detail the role of memory

it is necessary to outline a theory (the diffusion model, Kinchla & Symzer, 1967; Kinchla & Allan, 1969, 1970) which has been shown to model, quite adequately, short-term retention and forgetting of perceptual stimuli. As in signal-detection theory, this model assumes that the perceived, and now also the remembered, value of a presented stimulus is represented psychologically, in analog fashion, as a normally-distributed random variable with a variance and an expected value, or mean, which should be the actual physical value of that stimulus. It is further assumed, that the decision process involved in a comparative judgment uses the difference, as in typical two-alternative forced-choice signal-detection theory, between the sampled values, at time t (ISI), of the perceived test stimulus (x_t) and the remembered study stimulus (m_t), i.e., $(x_t - m_t)$.

Under these conditions, the d' measure is defined as the distance (scaled in standard deviation units) between the means of the difference distributions for pairs of stimuli when the test stimulus was the larger stimulus in the comparison pair ($E[x_t - m_t] > 0$), and when the corresponding study stimulus was the larger of the two stimulus ($E[x_t - m_t] < 0$). In order to make discriminations between test and study stimuli the observer sets an internal criterion (C) somewhere between the means of these difference distributions (e.g., at zero if the criterion is unbiased). For a test stimulus which is actually larger in value than its corresponding remembered

study stimulus, the observer should respond correctly if the sampled difference value on that trial is larger than C .

A key assumption in the diffusion model is that over the ISI the remembered value of the study stimulus undergoes a type of random-walk diffusion process (Kinchla & Symzer, 1967). While the overall mean value of the remembered stimulus is expected to remain the same over time (i.e., a symmetric walk) the diffusion process serves to increase its representational variance. Hence, the variance of the difference distributions (which is simply the sum of the variance of both the perceived test stimulus and the remembered study stimulus) increases proportionately as ISI increases. Because d' is scaled in units of difference-distribution variance it will decrease accordingly (indicating an attenuation of the discriminative sensitivity between the two difference distributions). Hence, in this model, degradation of perceptual memory involves not an exponential decay of trace strength (an operational concept in strength theory whose exact nature remains undefined), but a reduction in the clarity, or "grain", of the stimulus representation in memory. In other words, over the ISI the representation becomes noisier.

Kinchla and Allan (1969) reported that their model accounted for at least 90% of the variance in the obtained d' scores over all the ISI and stimulus difference manipulations in all of the experiments which they examined. In addition to

this excellent empirical fit, the diffusion model has several important advantages. First, it provides an explicit theoretical and mathematical description of memory retention over time which does not rely on any of the concepts proposed by strength theories (Massaro, 1970a, 1970c, 1970d; Wickelgren, 1969; see also Norman, 1966, for memory for verbal stimuli). Second, as such, the diffusion model can be theoretically applied to any type of discrimination task, whereas strength models are limited to a description of the decision and retention processes involved in similarity judgments (i.e., same-different recognition tasks; for a short review of strength theory see Appendix C). Third, as will be shown, the mathematics of the diffusion model also provides an analysis which separates the relative contributions of both the acquisition and diffusion processes to forgetting in short-term memory.

Attentional limitations in memory

There have been some studies whose results have indicated that the forgetting of stimulus items can depend on attentional capacity limitations. Lindsay and Norman (1969) showed that recognition-memory performance for six aurally-presented three-digit numbers was effected by a simultaneous signal-detection task, but only when both of the tasks were fairly difficult. In the Williams, Beaver, Spence, and Rundel? (1969) experiments mentioned earlier, the recall of the six single-digit stimuli over a 30 second retention

interval became increasingly impaired as the subjects were required to perform interpolated tasks which involved increasing amounts of information reduction (i.e., the recording, adding, or classifying of pairs of digits). More relevant to the present concerns, a study involving perceptual memory (Posner & Konick, 1966) showed impairment in the recall of the relative locations of points on a line which depended on the performance of an interpolated task. In that study, the mean absolute error of recall, measured in mm, was not effected by blank ISIs of 5, 10, and 20 seconds. However, this measure of memory performance became monotonically impaired when the subjects were required to either record, add, or classify pairs of digits during the retention interval (see also Dale, 1973, concerning memory for the relative locations of points within a square).

Based on the results of their experiments (which included an appropriate verbal control group), Posner and Konick (1966; see also Posner, 1969) proposed that non-verbal location information was being stored in the form of a visual code (which was not necessarily a visual image). More importantly, because the retention of this visual code was shown to require attentional capacity they also suggested that it was meaningful to assume that the visual code was being "rehearsed", or refreshed, in some manner. In the context of the diffusion model, it would be plausible to propose that the diffusion of any visual code in memory is dependent on

attention (as opposed to its explicit rehearsal). More specifically, if attention is drawn away from the retention of a visual code, or as attention is required to be split among several visual codes, diffusion proceeds at an accelerated rate. In the present paradigm, this viewpoint predicts that the rate of diffusion, or forgetting, should increase as the number of to-be-remembered study stimuli, or memory load, is increased. Furthermore, it should be more likely that an acceleration of the diffusion process would occur as the total amount of attention required to maintain a given set of study stimuli in memory begins to exceed the available capacity. In other words, attentional limitations in memory should only show up in diffusion-rate measures as the number of to-be-remembered stimuli becomes fairly large.

The present research

For reasons discussed earlier, the partial discrimination task developed by Palmer for line-length stimuli was regarded as an excellent paradigm in which to study the relation between diffusion rates and perceptual-memory load in the present experiment. However, several changes to Palmer's Experiment 4 (1990) were implemented (precluding an exact replication) in order to obtain conditions in which this relation could be examined more explicitly.

First, although Palmer (1990) found ISI effects in his Experiment 4 for study set sizes of two and four stimuli, it was felt that these effects should be demonstrated under the

same experimental conditions for a set size of one. Hence, in the present experiments memory performance was observed for study set sizes of one, two, and four stimuli.

Second, it was deemed necessary to use at least three ISIs in order to obtain actual forgetting functions. Hence, in the present experiments memory performance over all study set sizes was observed at ISI's of 0.5, 2, and 5 seconds.

Third, in Palmer's experiments (1990) he used horizontal line lengths as study stimuli which ranged from being shorter to being longer than a standard of 60 arc min (in steps of 2 arc min). For all conditions the test line was equal to the standard length. As well, in Experiment 4, the 2 subjects (Observers 11 & 13) had extensive previous experience with identical stimulus sets in two previous experiments (Experiment's 1 & 2) of this same study. Therefore it was quite possible, indeed very likely, that this standard test length would have become stored in long-term memory after several thousand trials. Hence, during the acquisition of the study stimulus (i.e., during their presentation and retention in iconic memory) subjects could immediately compare each study stimulus in the display with the standard stimulus retrieved from long-term memory. Massaro (1970a, 1970b) explicitly warned against using only one standard in a short-term perceptual memory task for this very reason. Harris (1952, cited in Massaro, 1970a) suggested that "long-term memory for the [standard stimulus] ... could possibly

eliminate any short-term deficit caused by increasing the [ISI]" (p. 33). As well, Baranski and Petrusic (1992) obtained reaction-time data which suggested that this strategy was being used by subjects in successive line-length comparisons of variable physical stimuli and (previously learned) standards which had to be retrieved from long-term memory. Subjects were faster to make a comparison between a remembered standard and an actual line length when the remembered standard was presented second in the pair.

If this strategy was being used, even imperfectly, by Palmer's (1990) subjects, then it was probable that on some trials they were actually making absolute judgments concerning relation of the study stimuli to the remembered test stimulus. Information about whether any particular study stimulus was either "longer" or "shorter" than the test stimulus (a one-bit classification) could have been encoded and stored verbally. In essence, the perceptual-memory task may have been reduced the remembering of a relatively small set of categorization labels that allowed the subjects to invoke extremely effective verbal rehearsal strategies. Palmer (1990) reported that some of his subjects had used this categorization strategy.

Under the conditions of Palmer's Experiment 4 (1990) perfect employment of this strategy would have predicted that very little ISI effect be observed. As well, because four verbal labels can be remembered as accurately as two, it would also have predicted that even if some forgetting occurred no

set-size effects due to memory processes could have occurred (i.e., that any effect of the ISI be the same for each set size). While the presence of robust ISI effects in Palmer's Experiment 4 precluded the perfect use of this strategy by his subjects, there is still the possibility that any set-size effects due to processes in memory could have been attenuated if this strategy was used on some of the trials. Hence, in the present study, it was deemed desirable to employ a method in which subjects would be unable to predict the length of the test stimulus in advance of its being presented. Instead, they would have to encode and store true perceptual memories of the actual lengths of the lines in the study display in order to perform the task.

Lastly, study-stimulus acquisition time was increased from 100 ms to 1000 ms. Although eye movements and foveation of the study stimuli could not be controlled during a presentation time of this length, it was felt that it was necessary to allow subjects sufficient processing time to acquire fairly accurate representations of the study line lengths in memory, at all set-size conditions, in order to properly perform the discrimination task. This manipulation would, hopefully, avoid floor effects of accuracy at set-size four in which memory performance would be so poor that it could not get much worse over longer ISIs.

Experiment 1

Method

Subjects

Two volunteer graduate students at Carleton University (both a psychology and a non-psychology student) were used as subjects. Each subject was required to participate in 12 experimental sessions each of which lasted one and one-half hours. Subjects were paid \$10 per hour for their participation. They had normal, or corrected-to-normal, vision and were naive with respect to the aims of this experiment.

Apparatus

Stimuli were presented on a Goldstar Model 1460 Plus VGA video monitor. High-resolution graphics were permitted with a VGA card (640 x 480) and MetaWindows (version 5) graphics under Turbo Pascal software control. Timing, accurate to within 1 msec, was possible with a Data Translation clock board and extensive software development. Graphics production, stimulus presentation, event sequencing, and the recording of responses and response times were controlled by an IBM PC 486/DX50 clone computer. Responses were made using the buttons on an IBM PC mouse, and confidence levels were registered using the PC keyboard.

Stimuli

The stimuli used were horizontal line lengths, one pixel in height, which appeared in white against a dark screen background. The line lengths appeared on an imaginary circle, 71 pixels (approximately 3 degrees or 180 arc min) in radius,

centred on the computer screen. The centres of each line were located on this circle at either 1, 4, 7, or 10 o'clock.

On each trial a "target" study line, which was either a standard or a corresponding comparison line, was presented at one of the four possible stimulus locations. The standard values were chosen from one of two possible stimulus line-length distributions: (a) 25, 26, 27, 28, or 29 pixels in length, i.e., the smaller distribution, or (b) 39, 40, 41, 42, or 43 pixels in length, i.e., the larger distribution. For each stimulus distribution the probability that a particular standard length was selected for display was .1, .2, .4, .2, and .1, respectively for each increasing line length in the distributions. The average values, in visual angle, of the standard line lengths were 1.14 degrees (68 arc min) for the smaller distribution, and 1.68 degrees (101 arc min) for the larger distribution (one pixel equals .03672 cm). On half of the trials the comparison lines were 8% longer than the standard, and on the remaining trials they were 8% shorter than the standard (a difference of two, or three, pixels in the smaller, or larger, stimulus distributions respectively).

For the study set size of one stimulus the target line was presented alone. For the study set size of two stimuli the target line was accompanied by another "non-target", or distracter, line located directly opposite it in the stimulus display (e.g., if the target line was presented at one o'clock then the distracter line was shown at seven o'clock). For a

study set size of four stimuli the target line was accompanied in the display by three other distracter lines located at each of the remaining three stimulus locations. All distracter lines were chosen at random from either of the previously defined stimulus distributions. For each set size the target lines were presented equally often at each stimulus display location.

After the ISI, a line (to be compared with the target line) appeared in the same location as the previous target line. Half of the time a standard was the target study line and its corresponding, longer or shorter, comparison was the test stimulus (presented post-ISI). The other half of the time the comparison was the target in the study display and the standard was the post-ISI test stimulus (i.e., symmetric order of presentation).

Three ISIs were used: 500, 2000, and 5000 ms. These ISIs were counterbalanced over the two standard distributions, three study set sizes, four target locations, two longer-shorter comparison types, and two standard-comparison orders of presentation. This counterbalancing generated $3 \times 2 \times 3 \times 4 \times 2 \times 2 = 288$ possible stimulus configurations.

For set sizes of two and four the observers did not know which of the initial study lines was the target line until the test line had been presented. Hence, he/she was required to acquire and retain the lengths of all of the lines presented in the study display during each comparison trial.

Procedure

At the start of each comparison trial either the instruction "LONGER" or "SHORTER" appeared at the top of the computer screen along with a fixation point positioned in the center of the screen. After 1000 ms this fixation point was replaced with a study display (of either one, two, or four line lengths) which remained in view for 1000 ms (the instruction remained on the screen until the completion of the response). After the ISI, during which the screen was blank (except for the instruction), a single line appeared in the same location as one of the previous study lines, i.e., the target line. Subjects were required to decide which line was longer, or shorter, depending on the relevant instruction, either the target line in the study display (the stimulus presented first) or the post-ISI test line (the stimulus presented second). The left and right buttons on the computer mouse were used to respond "first stimulus" and "second stimulus", respectively. Subjects were required to keep their index and middle fingers on these buttons at all times during the course of the experiment. The test stimulus remained in view until a response has been made (allowing for complete foveation, through eye movements, of the test stimuli in all conditions). No "equal" or "don't know" responses were allowed. The subjects were asked to be as accurate as they could be without taking too much time to respond.

Following each response the screen was cleared and the

subjects were given a visual prompt to provide a confidence rating indicating how certain they were that they had just made the correct response. Three confidence categories were available: ("3") high certainty, ("2") moderate certainty, and ("1") low certainty. Subjects were to type confidence values on the computer keyboard and then depress "Enter" to record the confidence report. After 2000 ms a new instruction appeared, indicating the start of the next comparison trial. No accuracy feedback was provided at any time.

The experimental room was darkened except for a small lamp positioned behind the computer terminal. Subjects were seated with their eyes approximately 50 cm from the computer screen. To respond they used their dominant hand. Twelve randomly-chosen practice trials were given at the start of each session.

Hence, the task was a two-alternative forced-choice discrimination of line length (i.e., was the longer, or shorter, line presented "first" or "second"). Taking each instruction into account, a completely counterbalanced block of comparisons was $2 \times 288 = 576$, randomized, trials in length. Each subject performed approximately six complete blocks of trials during this experiment. The first two complete blocks were regarded as practice. Due to the length of each block, the subjects were required only to perform one-half of a block in each daily session.

Results and Discussion

Because the standard and comparison orders of presentation were counterbalanced and confidence measures were collected it was possible to obtain a two-alternative forced-choice Type-2 ROC and corresponding d'_z measure of discriminative memory "strength" (see Macmillan & Creelman, 1991, chapters three and five; d'_z measures do not require that the ROCs have unit slope on normal-normal coordinates) for the line lengths in each of the two stimulus distributions, at each of the three ISI's, for each of the three study set sizes. Hence, 18 d'_z measures were obtained per subject.

The ROCs were calculated from the observed response probabilities, conditionalized on each confidence-rating category, for each of the two response possibilities (a "first stimulus" response to SHORTER [or equivalently a "second stimulus" response to LONGER], and a "second stimulus" response to SHORTER [or equivalently a "first stimulus" response to LONGER]), for each of the two stimulus presentation possibilities (the shorter line presented first and the shorter line presented second). By collapsing these observed response probabilities for each stimulus distribution over instruction type (LONGER, SHORTER), comparison type (longer or shorter than the standard), study-stimulus location (1, 4, 7, and 10 o'clock), and number of non-practice blocks (four), each ROC was based on 64×2 (shorter line first, shorter line second) = 128 observations. Due to computer

error Subject 2 was unable to complete the trials in the second half of the third block and the first half of that block was also regarded as practice for that subject. Furthermore, due to the presence of non-monotonic d' scores over ISIs (at study set sizes of two and four of the smaller stimulus distribution) this same subject was willing to perform an extra block of trials after four non-practice blocks. This extra block yielded ROCs, for her, which were based on $80 \times 2 = 160$ observations.

The diffusion model estimates

Because actual forgetting functions (for d' scores over ISIs) were obtained in this experiment it was possible to calculate quantitative measures (developed by Kinchla & Symzer, 1967, and Kinchla & Allan, 1969, in conjunction with their diffusion model of perceptual memory) of the rate of diffusion of the stimuli in memory for each study-stimulus set size. In a more formal mathematical description of their model under these task conditions:

$$d' = [(X_{t\text{long}} - m_{t\text{short}}) - (X_{t\text{short}} - m_{t\text{long}})] / [\text{Var}(X_t - m_t)]^{1/2}, \quad (4)$$

where m_t refers to the stimulus presented first in the comparison pair (i.e., the remembered stimulus). By defining $m_t = x_0 + d_t$ (where x_0 is the value of the study stimulus immediately after acquisition and d_t is the diffusion effect on x_0 of the random walk after an ISI of time t), then the expected value of each $(X_t - m_t)$ in equation (4) is:

$$E(X_t - m_t) = E(X_t - x_0 - d_t),$$

$$\begin{aligned}
 &= [E(X_t) - E(X_0) - E(d_t)], \\
 &= E(X_t - X_0), \quad \text{since } E(d_t) = 0,
 \end{aligned} \tag{5}$$

which is simply the difference between the study and the test stimuli (d_g). Hence the numerator in equation (4) is:

$$\begin{aligned}
 (X_{t\text{long}} - X_{0\text{short}}) - (X_{t\text{short}} - X_{0\text{long}}) &= 2(X_{\text{long}} - X_{\text{short}}), \\
 &= 2d_g,
 \end{aligned} \tag{6}$$

after dropping the identifying subscripts, t and 0 .

Similarly, the expected value of the square of the denominator in equation (4), i.e., the variance of each difference distribution, is:

$$\begin{aligned}
 \text{Var}(X_t - m_t) &= \text{Var}(X_t - X_0 - d_t), \\
 &= \text{Var}(X_t) + \text{Var}(X_0) + \text{Var}(d_t), \\
 &= \text{Var}(X_t) + \text{Var}(X_0) + \Phi t,
 \end{aligned} \tag{7}$$

where Φ is defined as the rate of diffusion of the study stimuli in memory over the ISI's (for a full mathematical description of the Φ parameter see Kinchla & Symzer, 1967). Note that this formulation of the diffusion model assumes, as in Thurstone's Case V, the independence of X_t and X_0 (i.e., $\text{Cov}[X_t, X_0] = 0$). Hence:

$$(d')^2 = (2d_g)^2 / \text{Var}(X_t) + \text{Var}(X_0) + \Phi t, \tag{8}$$

and estimates of the difference distribution variance can be obtained empirically by manipulating this equation:

$$\text{est}[\text{Var}(X_t) + \text{Var}(X_0) + \Phi t] = (2d_g)^2 / (\text{est}d')^2, \tag{9}$$

where $\text{est}d'$ are the observed d'_t scores. When plotted against t these variance estimates should yield a linear function in which the intercept on the ordinate estimates the amount of

acquisitional variance attributed to the input processes (i.e., $\text{Var}[x_t] + \text{Var}[x_0]$), and the slope of which represents the rate of diffusion (i.e., Φ). These parameters were obtained through a least-squares linear regression of the observed variance estimates (from equation [9]) against t .

The observed values of each subject's d'_t scores, for each stimulus distribution, are shown in the forgetting functions in Figure 1 (the corresponding percentage correct

 Insert Figure 1 about here

are shown in Table 1). When the least-squares predictions of

 Insert Table 1 about here

the difference-distribution variance at each ISI were substituted back into equation (8) predicted values of d'_t were obtained. These predicted values, for each experimental condition, are presented graphically (as the solid lines) in Figure 1. The average variance in the d'_t scores observed in this experiment which is accounted for by the diffusion model (average R^2) was 88%.

For each subject the estimated values obtained from the model of both the acquisitional variance (in pixels squared) and the corresponding rate of diffusion for each study-

stimulus set size are shown in Table 2. Upon examination of

Insert Table 2 about here

the estimated diffusion rates in memory for stimuli from the larger stimulus distribution it was obvious that the earlier predictions concerning the dependence of the rate of diffusion on study-stimulus set size were fulfilled exactly. While there was no real change in the rate of diffusion between study set sizes of one and two (6.75 vs 6.17 and 1.32 vs 1.39 for Subjects 1 and 2, respectively), the rate of diffusion began to accelerate as the number of to-be-remembered stimuli was increased to four (to 27.91 and 2.64, respectively).

In the smaller stimulus distribution the diffusion measure for Subject 1 was also observed to increase (quite drastically) from a study set size of one to a set size of four remembered line lengths (5.01 as compared to 91.29). However, it was not possible to obtain a reliable estimate of the diffusion rate for the set-size two stimuli because d' scores surprising increased over the ISI. It is not at all clear why a reminiscence-like effect occurred in this condition. An observation of this sort cannot be explained within any model of perceptual memory that these authors are familiar with. (Laming and Scheiwiller, 1985 [see also Laming & Wightman, 1992], reported the possible existence of reminiscence effects, but these occurred at very short ISIs of

250 ms.) For Subject 2 the diffusion measure was observed to decrease monotonically over increases in study set sizes in the smaller distribution (0.58, 0.26, and 0.20, respectively). However, because these diffusion rates in memory were all so small it could be assumed that these results indicated that very little diffusion occurred for any of those stimuli. As well, while the d_s' scores for this subject showed a large effect at 500 ms when study set sizes were increased from one to two stimuli, curiously, no such set-size effect was found at ISIs of 2000 or 5000 ms.

Because memory performance was so disparate between the two subjects an averaging of their data was not feasible (the actual d_s' scores are shown in Table 1, Appendix E). However, as an internal test of the diffusion model, each subjects' rates of diffusion, at each study set size, should be similar for the stimuli in both the smaller and larger stimulus distributions. For the present data, however, this comparison could only meaningfully be made for study set sizes of one in which the diffusion rates were observed to have been fairly similar for each subject over the stimulus distributions (6.74 vs 5.01 and 1.31 vs 0.58 for Subjects 1 and 2, respectively).

The estimates of the acquisitional variance provided by the diffusion model (in Table 2) allowed for a direct comparison with the predictions of the sample size model concerning the quantitative relations between the levels of input variance observed for the stimuli perceived under each

study set-size condition. These predictions were somewhat complicated, however, by the stimulus presentations conditions in the present experiment because the test stimulus was always held in view until a response was made (unlimited hold). But as shown in Appendix D, it was reasonable to assume that the sample size model predictions could be bounded above by the prediction that acquisitional variance for set sizes of two and four should have increased from the level of acquisitional variance observed for a set size of one by a factor of two and four, respectively.

Upon examination of these aquisitional-variance estimates, it was first noted that they did in fact increase monotonically, in all four cases (except at a set size of two for Subject 1) in which an examination of this relation is possible (i.e., both subjects and both stimulus distributions), as the study-stimulus set sizes increased. However, in almost all cases this increase was larger, especially for Subject 1, than the maximum possible increase predicted by the sample size model alone (the only exception was the increase in this estimate from 1.05 for set-size one to 2.08 for set-size two [a factor of two], for Subject 2 in the smaller distribution). There are several possible explanations for this finding. It might have occurred due to differences in the foveation of the study stimuli between set sizes, which was made possible in the present experiment by the length of the study-stimulus presentation time (1000ms).

Hence, the actual quality, in terms of the variance, of the sampled values during the perception of the stimuli might have differed for the stimuli in study set sizes of one (which it could be assumed were easily foveated) as opposed to the perception of the stimuli in larger study set sizes. As well, foveation of stimuli in study set sizes of two and four would have necessitated eye movements. Hence, it is possible that perceptual sampling "time" was lost during the acquisition of study stimuli for these larger set sizes due to both the execution of these eye movements as well as the attentional requirements involved in their initial generation.

Time-order error

The previous results showed that the rate of forgetting in short-term perceptual memory (as measured by the diffusion model parameter) increased in three out of four cases as the to-be-remembered stimulus set size was increased from one to four. However, the use of line lengths as stimuli raised the possibility that time-order errors (TOEs) might have occurred. The TOE has been recognized as being a very pervasive psychophysical phenomenon which arises in comparisons of extent between perceptual stimuli which are presented successively (Fechner, 1860; Needham, 1934; Woodrow, 1933, 1935; Jamieson & Petrusic, 1975a, 1975b, 1976, 1978; Hellstrom, 1977, 1979, 1985; Baranski & Petrusic, 1992). It has been observed that subjects tend either to underestimate (a negative TOE) or to overestimate (a positive TOE) the

actual extent of the stimulus which has been presented first in a successive comparison task.

One measure of the TOE, which is particularly appropriate to use when the stimulus pairs have been presented in symmetric orders over the course of an experiment, has been the TOE index developed by Jamieson and Petrusic (1975a, 1975b). This TOE index is:

$$1/2(p_{\text{correct}}[\text{long-short}] - p_{\text{correct}}[\text{short-long}]) \times 100\%, \quad (10)$$

where long-short and short-long refer to the order of presentation of the stimuli in a stimulus pair (note that the actual physical difference between the stimuli remains the same for both orders). TOE indices (shown in Table 3) were calculated for each subject under each experimental condition

Insert Table 3 about here

in which d'_s scores were obtained. Three things should be noted about these observed TOEs. First, their occurrence was very robust in almost every condition for Subject 1 and in the larger stimulus distribution for Subject 2. (Note, as well, that a negative, or positive, TOE of about 25% indicted that there was a full 50% discrepancy between correct responses for the exact same pair of stimuli which arose simply by reversing it's order of presentation.) Second, for these subjects there was a general tendency for TOEs to become more extreme in the larger stimulus distribution as ISIs increased from 500 ms to

2000ms. Third, the nature of the TOEs was markedly different for each subject. While a preponderance of negative TOEs was observed for Subject 1 under all experimental conditions, highly positive TOEs were found for Subject 2 at set sizes of two and four in the larger stimulus distribution. These conflicting findings over both subjects really highlighted the problems which have beset previous research on this "particularly elusive phenomenon" (Franzen, Nordmark, and Sjöberg, 1972, cited in Hellstrom, 1979).

Previous TOE research. Hellstrom (1977) and Jamieson and Petrusic (1975a) have presented evidence that TOE's are a perceptual-memory phenomena (that is, an actual illusion) and are not due simply to response biases. These researchers found that TOE indices were independent of four different modes of judgment which included: (a) judging the first stimulus as "LONGER" or "SHORTER", (b) judging the second stimulus as "LONGER" or "SHORTER" (Hellstrom, 1977), (c) judging as "LONGER" either the first or the second stimulus, and (d) judging as "SHORTER" either the first or the second stimulus (Hellstrom, 1977; Jamieson & Petrusic, 1975a). Furthermore, both the direction and magnitude of the TOE has been observed to depend on various factors such as the within-session range of stimuli, amount of practice, feedback, and ISI. More specifically, the direction of the TOE can switch for within-session standards which are high and low in value (Woodrow, 1933; Jamieson & Petrusic, 1975b; Hellstrom, 1977,

1979), and its magnitude has been shown to decrease (or reverse) with practice (Woodrow, 1935), feedback (Jamieson & Petrusic, 1976, 1978, who viewed the role of feedback as the enhancement of a bias against faulty, illusion-based responding), or increases in ISI (Jamieson & Petrusic, 1975a, 1975b, 1976; Hellstrom, 1979).

TOEs and d' . Several psychophysical theories have arisen to explain the existence of TOEs in successive-comparison data. For the purposes of the present study it was deemed necessary to determine theoretically the extent to which the observed d' s might depend on TOEs. In other words, given a specific cognitive process(es) that could have been responsible for the pattern of TOEs found over subjects, stimulus distributions, ISIs, and study set sizes in the present experiment, how should this process(es) have effected the measures of d' . While the effect of TOEs on d' (or equivalently, measures of threshold) should be an important theoretical consideration in any perceptual-memory experiment, it has been habitually ignored (e.g., Palmer, 1990).

First, it was not possible that a simple response-bias preference for either the response "first" or "second" was responsible for the TOEs which were observed. Any influence of a response bias of this sort should have had an opposite effect over both types of instructions that were employed (i.e., choose the LONGER or choose the SHORTER; see Jamieson & Petrusic, 1975b). (TOEs for each subject conditionalized on

the mode of instruction are shown in Table 3, Appendix E.)

Second, it was possible that the negative TOEs in Subject 1's data could have arisen due to a differential preference for the response "first" (the remembered stimulus) when the instruction was SHORTER, and for the response "second" (the perceptual stimulus) when the instruction was LONGER. A bias of this type can be conceptualized within the decision process of the diffusion model as a negatively-biased difference-criterion value (see also Allan, 1979, cited in Hellstrom, 1985). Hence, the perceived difference value ($x_t - m_t$) need only to have been larger than some $-c$ (rather than zero) in order for the subject to have decided either that m_t was the "SHORTER" or that x_t was the "LONGER" of the two presented stimuli.

Conceptualizing the TOE in this manner, as an idiosyncratic criterion bias, has been an attractive way to explain differences in observed TOEs between different subjects under identical experimental conditions (e.g., Allan & Kristofferson, 1974, cited in Jamieson & Petrusic, 1975b). For example, in the present experiment the existence of a positively-biased criterion could have explained the predominantly positive TOEs found in Subject 2's data. However, as pointed out by Hellstrom (1985; and Jamieson & Petrusic, 1975a, 1975b), TOEs have been observed to depend on both within-subject ISI and stimulus-range manipulations. These findings have had the implication that the decision

criterion, which has usually been assumed to be preset and to remain constant over an experimental session, would have had to shift over comparison trials depending on the relevant stimulus and ISI conditions.

Furthermore, it was not clear why such criterion biases might have arisen in the present experiment. What was clear, however, was that if TOEs were due to response biases (of any kind) then d' measures, by definition, were independent of their effects.

Third, it was possible that TOE's could have occurred due, not to response biases, but to a bias in the memory process. Kinchla and Symzer (1967; Kinchla & Allan, 1969) considered the fact that the diffusion process might be asymmetrical. Hence, the expected value of the diffusion process ($E[dt]$ in equation [5]) would not equal zero if asymmetry was present but would be either a negative or positive value depending on the systematic nature of the drift that occurred during the diffusion process. (A negative or positive drift of the remembered stimulus value directly implies the presence of a negative or positive TOE, respectively.)

These researchers likened the notion of a negative bias in memory to a "fading trace" (a popular interpretation of TOE's throughout psychophysical history, which was first introduced by Kohler, 1923, cited in Hellstrom, 1985; see also Needham, 1934). According to Kinchla and Symzer (1967), this

"fading trace" could be viewed as the assimilation of the remembered stimulus to an adaption level (e.g., Helson, 1947) that was smaller in value than the values of the presented stimuli. Hence, in the present experiment, this explanation could plausibly account for Subject 1's negative TOEs. However, the positive TOEs observed for Subject 2 (an "expanding trace"?), under identical experimental conditions, would now have to be explained, in terms of a drift viewpoint, as either assimilation to a higher-valued adaption level or perhaps as contrast from some lower-valued adaption level.

In terms of the diffusion model, Kinchla and Allan (1969) suggested that the presence of a systematic drift in memory serves to add a constant (which depends on the length of the ISI, t) to the value of each x_0 stimulus stored in memory. While the variance of the difference distributions would not be altered by the addition of a constant memory bias, the expected values of the difference distributions themselves (i.e., $E[x_t - m_t]$ for both orders of presentation of a stimulus pair) would be shifted around the response criterion. Hence, the result of this process on responding, and subsequently the TOEs, should be similar to that of an actual response-criterion shift. In fact, it would be very difficult to distinguish between these two types of biases. Furthermore, because each of the subjective difference distributions are shifted by this process in a similar manner, d' scores remain unaffected (it can be shown that a bias added to each x_0

cancels out in the numerator of equation [4]) as long as each stimulus value in memory, within each experimental condition, has been biased to the same degree.

Finally, Hellstrom (1977, 1979, 1985) developed a model which has provided a detailed description of the processes that are presumably involved in the comparison of successively-presented stimuli. Hellstrom (1977, 1979) showed that his model (the differential sensation weighting model) explained, in terms of processes in perception and memory, all of the TOE phenomena observed in his experiments without a reliance on any kind of bias.

Hellstrom determined that the subjective values of the stimuli in a comparison pair are not compared with each other directly, but that each stimulus interacts (through either assimilation or contrast) with a series adaption level (SAL) to form a subjective magnitude corresponding to a comparative adaption level (CAL; see Michels & Helson, 1954). (This SAL has been assumed to represent some form of weighted average of all the stimuli to which the observer has been exposed prior to the present stimulus.) Hence, the subjective comparative difference between any two successively-presented stimuli (x_0 and x_t , as before) can actually be represented (Hellstrom, 1979, Model 3) as:

$$\begin{aligned} (x_t - x_0) &= \text{CAL}_t - \text{CAL}_0 \\ &= (g_t[x_t - \text{SAL}_t] + \text{SAL}_t) - (g_0[x_0 - \text{SAL}_0] + \text{SAL}_0), \end{aligned} \quad (11)$$

where g_t and g_0 are the sensation weighting parameters which

indicate either assimilation (if $s < 1$) or contrast (if $s > 1$) of the subjective value to the appropriate SAL in forming the CAL. It should be noted that if $s = 0$ then $CAL_i = SAL$, and if $s = 1$ then the CAL value depends only on the subjective value of the stimulus.

When the simplifying assumption that $SAL_t = SAL_0 = SAL$ is made, then Hellstrom's Model 2 (1977, 1979) is derived:

$$(X_t - X_0) = W_t(X_t - SAL) - W_0(X_0 - SAL), \quad (12)$$

where W_t and W_0 are very closely related to g_t and g_0 . (Note that in Hellstrom's calculations [1977, 1979] he assumed that the observer takes the difference value $[X_0 - X_t]$. This relation was reversed in the present paper in order to remain consistent with the assumptions of the diffusion model.) The basis for the TOE can be directly inferred from equation (12) when two stimuli of initially equal subjective magnitude (i.e., before the CAL) are compared. By letting $X_t = X_0 = X$, then:

$$(X_t - X_0) = (W_t - W_0)(X - SAL). \quad (13)$$

Hence, both the direction and the magnitude of the TOEs in any experiment depend on both the relation between the differential weights (W_t and W_0) as well as the levels of the stimuli used. For example, if ($W_t > W_0$) and ($X > SAL$) then the actual comparative difference value ($X_t - X_0$, where $X_t = X_0$ initially) is positive, implying that X_0 had actually been underestimated relative to X_t (a negative TOE).

The differential sensation weighting model was very

effective in accounting for the results found in Hellstrom's experiments. TOEs which change from positive, for comparisons involving a low standard stimulus (below the SAL), to negative, for comparisons involving a higher standard stimulus (above the SAL), are apparent from equation (13) if $W_t > W_0$ (Hellstrom, 1977). Changes in TOEs over ISIs were the direct result of changes in the relations between the sensation weights as ISI increased. More specifically, Hellstrom (1979) found that $W_t < W_0$ for smaller ISIs and that $W_t > W_0$ for longer ISIs. In general, the estimated values of the sensation weights increased as either the initial stimulus presentation or the ISI was increased. These parameters were estimated through the use of an iterative technique which was applied to calculations of the scaled mean subjective difference (Hellstrom's TOE index, 1977, 1979) between the stimuli in each comparison pair at each experimental condition.

Hellstrom (1979) suggested that the nature of the weighting of any stimulus sensation (x) depends on both the distinctiveness and the effective information content present in that stimulus at the time of the formation of the CAL. A lack of either of these qualities would result in a lower sensation weighting (making it more likely that the subjective value of that stimulus is assimilated to the SAL in equation [11] or, equivalently, less likely that it is contrasted with the SAL). For example, it is likely that longer ISIs increase the distinctiveness of each stimulus in a comparison pair but

less so for the stimulus which must be remembered (for a more complete explanation of the cognitive basis for sensation weighting, see Hellstrom, 1979, 1985).

In the present experiment it can be assumed that due to the extremely large number of both practice and measured experimental trials performed by each subject the SAL was heavily influenced by the actual stimuli in the experiment (as opposed to the influence of other background factors, see Hellstrom, 1979). Hence, it was likely that the SAL was located somewhere near the mean (either the arithmetic or geometric mean) of those stimuli (i.e., between the values of the smaller and larger distribution stimuli). Therefore, given this SAL, the TOEs should have been dependent on the sensation weightings that were applied to x_t and x_0 , under the various experimental conditions.

Various predictions concerning the relation of the weights (W_t and W_0) were inferred in order to account for the pattern of TOEs observed in the present data. Two of these possible predictions stood out. First, in order to account for the presence of uniformly negative TOEs in Subject 1's data it would have to be inferred that $W_t > W_0$ for the comparisons involving larger-distribution stimuli, but that this relation switched (i.e., $W_t < W_0$) for comparisons involving smaller-distribution stimuli. Second, the TOE's for Subject 2 (mostly positive for the larger-distribution stimuli and slightly negative or close to zero for the smaller-

distribution stimuli) implied that, in general, $W_t < W_0$. However, under the present conditions it seemed unlikely that the first stimulus should have been weighted more than the second given that: (a) it could be assumed that the first stimulus loses information as it resides in memory over the ISI, (b) the first stimulus was often presented within a set of two or four line lengths (reducing its distinctiveness), and (c) the second stimulus was available to the observer on unlimited hold.

It might have been possible in the present experiment to derive estimates of these weighting parameters by having regarded the responses in the "1", or "guess", confidence category as similar to a decision that the stimuli were equal (Hellstrom's model calculations, similar to that of Greenberg [1965], requires that an "equal" response category be included in the task). Presumably, these estimates would have shed some light on the nature of the sensation weighting which occurred for each subject under the present experimental conditions. What was of major concern, however, was the effect that processes involved in differential sensation weighting would have had on the observed measures of d' .

First, assuming the more general case of Hellstrom's Model 2 (equation [12]), the numerator of d' would reduce to:

$$\begin{aligned} & (X_{t\text{long}} - X_{0\text{short}}) - (X_{t\text{short}} - X_{0\text{long}}) \\ &= (W_t [X_{\text{long}} - \text{SAL}] - W_0 [X_{\text{short}} - \text{SAL}]) \\ &\quad - (W_t [X_{\text{short}} - \text{SAL}] - W_0 [X_{\text{long}} - \text{SAL}]), \end{aligned}$$

$$\begin{aligned}
&= (\underline{W}_t \underline{X}_{long} - \underline{W}_t \text{SAL} - \underline{W}_0 \underline{X}_{short} + \underline{W}_0 \text{SAL}) \\
&\quad - (\underline{W}_t \underline{X}_{short} - \underline{W}_t \text{SAL} - \underline{W}_0 \underline{X}_{long} + \underline{W}_0 \text{SAL}), \\
&= \underline{W}_t \underline{X}_{long} - \underline{W}_0 \underline{X}_{short} - \underline{W}_t \underline{X}_{short} + \underline{W}_0 \underline{X}_{long}, \\
&= \underline{W}_t (\underline{X}_{long} - \underline{X}_{short}) + \underline{W}_0 (\underline{X}_{long} - \underline{X}_{short}), \\
&= (\underline{W}_t + \underline{W}_0) (\underline{X}_{long} - \underline{X}_{short}). \tag{14}
\end{aligned}$$

Second, under the simplifying assumptions that $\underline{x}_t$ and $\underline{x}_0$ are independent (i.e., $\text{Cov}[\underline{x}_t, \underline{x}_0] = 0$; it should be noted that Hellstrom [1979] might question this assumption), and that the input variances are equal (i.e., $\text{Var}(\underline{x}_t) = \text{Var}(\underline{x}_0) = \text{Var}(\underline{x})$, a reasonable assumption at a study set size of one only), then the variances of each difference distribution can be represented as:

$$\begin{aligned}
&\text{Var}(\underline{W}_t [\underline{x}_t - \text{SAL}] - \underline{W}_0 [\underline{x}_0 - \text{SAL}]) \\
&= \text{Var}(\underline{W}_t \underline{x}_t) + \text{Var}(\underline{W}_t \text{SAL}) + \text{Var}(\underline{W}_0 \underline{x}_0) + \text{Var}(\underline{W}_0 \text{SAL}), \\
&= \underline{W}_t^2 \text{Var}(\underline{x}_t) + \underline{W}_t^2 \text{Var}(\text{SAL}) + \underline{W}_0^2 \text{Var}(\underline{x}_0) + \underline{W}_0^2 \text{Var}(\text{SAL}), \\
&= (\underline{W}_t^2 + \underline{W}_0^2) (\text{Var}[\underline{x}] + \text{Var}[\text{SAL}]). \tag{15}
\end{aligned}$$

Hence, based on these simplifying assumptions $\underline{d}'$ (from equations [14] and [15]) is:

$$d' = (\underline{W}_t + \underline{W}_0) (\underline{x}_{long} - \underline{x}_{short}) / ([\underline{W}_t^2 + \underline{W}_0^2] [\text{Var}\{\underline{x}\} + \text{Var}\{\text{SAL}\}])^{1/2}. \tag{16}$$

In Hellstrom's model, as it stands so far in equation (16), the only consequence of the actual memory process itself should be the effect of ISI on the sensation weights, $\underline{W}_0$ and $\underline{W}_t$. Therefore, according to this interpretation of Hellstrom's model, any changes in $\underline{d}'$ over ISI's, for a study set size of one, should have depended only on changes in the values of the sensation weights according to the function :

$(W_t + W_0) / (W_t^2 + W_0^2)^{1/2}$. However this weighting function is bounded above by:

$$\begin{aligned} \text{limit}(W_t \rightarrow W_0 = W) &= 2W / (2W^2)^{1/2}, \\ &= 2 / (2)^{1/2}, \\ &= 1.4142, \end{aligned} \quad (17)$$

and bounded below by:

$$\begin{aligned} \text{limit}(W_t \rightarrow 0, W_0 \rightarrow \infty) &= ([W_t/W_0] + 1) / ([W_t^2/W_0^2] + 1^2)^{1/2}, \\ &= 1 = \text{limit}(W_0 \rightarrow 0, W_t \rightarrow \infty). \end{aligned} \quad (18)$$

Hence, d' 's for a study set size of one, which decrease over ISI by more than a factor of $1/1.4142 = 0.7071$ can not be accounted for by this model alone (as it stands in equation [16]). Furthermore, even observed decreases in d' within this factor would have to be accounted for by drastic changes in the relation between W_t and W_0 . For example, a decrease in d' over ISI of a factor of $1.1767/1.4142 = 0.832$, would have required that the relation between the weights change from, say, $W_0/W_t = 1$ to $W_0/W_t = 5$ (or $1/5$). This represents a change in weighting that does not seem very likely given the sensation weighting estimates (g_1 and g_2 for loudness comparisons with ISI's of 500 ms or greater, found in Hellstrom, 1979) never exceeded $g_2/g_1 = 1.5$ (or similarly fell below $1/1.5$).

Hence, it was the view of this researcher that the differential sensation weighting model must also include some process, such as diffusion of the remembered stimulus, in order to account for the observed changes in d' . The extent

to which increases in diffusion observed at set-size four in the present data were actually due to changes in W_t and W_0 over t is a difficult question to answer at the present time. However, for reasons discussed earlier, it was not at all certain that the differential sensation weighting model, or in fact any of the models, was actually the appropriate model for the TOEs observed in this experiment. The only things that could be concluded with certainty about these TOEs was that they were very robust in both their magnitude and their inconsistency.

Experiment 2

Evidence in support of the diffusion model was obtained in Experiment 1. These results, however, were accompanied by the presence of huge TOEs in several of the experimental conditions. The previous discussion suggests that a convincing demonstration of the effects that TOEs have on measures of d' , and hence the acquisition and retention estimates derived using the diffusion model, cannot be determined until a convincing explanation of TOEs has been obtained.

At the very least, however, the occurrence of a large TOE suggests that the mean subjective difference between the stimuli, in the comparison pairs under that experimental condition, was shorter for one order of presentation (for example, the long-short order if the TOE was negative) than for the other order of presentation. Evidence for this

supposition can be found in an experiment performed by Baranski and Petrusic (1992) in which subjects were required to perform successive line-length comparisons under task conditions that were based on the method of constant stimuli. Their results showed that the slopes of the psychometric functions (calculated using effect coding of the comparison stimuli), for each of three standard lengths, depended on both the order of presentation of the stimuli in any comparison pair and the magnitude of the TOEs (almost all of which were negative) observed for that standard.

More specifically, while the slopes of the psychometric functions calculated for the line lengths which had been presented in the short-long order did not depend on the corresponding TOE magnitudes, the slopes for the line lengths presented in the long-short order were observed to decrease with increases (negatively) in TOE magnitudes. A decrease in the slope of a psychometric function indicated a corresponding increase in the difference threshold which, in turn, represented a decrease in the discriminability between the standard and the stimuli with which it was being compared. It is important to note that this dependency of slope on the TOE was asymmetric over presentation orders (i.e., the loss in discriminability between the stimuli in the long-short order of presentation was not accompanied by a corresponding increase in discriminability for the short-long order). Hence, differences in relative threshold values (in terms of

Weber fractions), after collapsing over both orders of presentation, depended largely on differences in TOEs.

In the present Experiment 1 it was possible to examine the empirical relation between TOEs and d' by calculating d' for the responses to each instruction within each set-size, ISI, and stimulus-distribution treatment combination condition. There seemed to be a consistent inverse relation (see Table 3, Appendix E) between the absolute size of the TOE and these d' estimates. As well, any effect of the TOEs on the variabilities of the difference distributions for each order of presentation in Experiment 1 should have been reflected in the slopes of the ROCs. Equal variances of these difference distributions over presentation orders would be indicated by slopes that were close to one. While this ROC slope criteria was violated about half of the time in Experiment 1 (see Table 2, Appendix E), no systematic pattern of deviation or relation to the observed TOEs could be determined.

It seemed desirable to try and re-examine the diffusion hypothesis of forgetting in short-term perceptual memory under stimulus conditions which minimized any potentially confounding effects of TOEs. Hence, it was necessary to repeat Experiment 1 using perceptual stimuli which could be simultaneously presented in the various study set sizes and were also known to be relatively immune to the effects of TOEs. It was decided that memory for the absolute position of

"points" on the computer screen should meet these requirements.

Method

Subjects

Two new subjects were used (Subjects 3 and 4). The conditions under which these subjects participated in this experiment were identical to Experiment 1 except for the use of only eight experimental sessions per subject.

Apparatus

The apparatus was the same as in Experiment 1.

Stimuli

The stimuli were points on the computer screen 1 pixel in diameter (0.037cm or 0.042 degrees in visual angle) appearing in white against a dark screen background. These stimuli were almost identical to those used in studies of perceptual memory for absolute position performed by Kinchla and Symzer (1967; Kinchla & Allan, 1969, 1970).

The standard points (and distracter points in the study display) appeared on the computer screen in the same four imaginary circle locations as in Experiment 1. However, these points were horizontally displaced either -10, -5, 0, +5, or +10 pixels (with probability .1, .2, .4, .2, and .1, respectively) from these locations. Because subjects could not predict perfectly the exact position of the study stimuli, at each possible circle location, this manipulation helped to ensure that true absolute position discrimination occurred on

each trial.

On each trial a "target" study point appeared at one of the possible stimulus locations. Exactly as in Experiment 1, this target point appeared equally often at each stimulus location, it was accompanied by either zero, one, or three distracter points, and it was either in a standard position or in a comparison position. Half of the time the position of the comparison points were to the right of the standard, and the remaining half of the time they were to the left of the standard. The actual horizontal displacement distance, to the right or left, of the comparison point from the standard point was 10 (0.42 deg.) pixels.

As in Experiment 1 the location of the target point in the study display was indicated by the location of the point (whose position was to be compared with the remembered position of the target point) which was presented after the appropriate ISI (which were the same as in Experiment 1). Counterbalancing generated $3 \text{ (ISI)} \times 3 \text{ (study set size)} \times 4 \text{ (target locations)} \times 2 \text{ (left-right comparison types)} \times 2 \text{ (standard-comparison orders of presentation)} = 144$ possible stimulus configurations.

Procedure

The details of the procedure were identical to Experiment 1, except for the nature of the comparison. The displayed instructions (which only appeared on the screen during the presentation of the fixation point and the test stimulus) were

"ON THE LEFT" or "ON THE RIGHT". Subjects were required to respond which stimulus, either the "first" or the "second", they thought was either the leftmost or the rightmost (depending on the relevant instruction).

As well, for this experiment a "0" confidence category was available to the subjects. This category was to be used when a subject was certain that he/she had pressed the wrong response button by mistake. Such a mistake could occur either due to motor error or to confusion between the appropriate buttons used to respond (for example, it might have been necessary to press the left button on the mouse to respond that the "first" stimulus was "ON THE RIGHT"). Subjects often know when they have made a mistake of this kind and it seemed desirable to have these mistakes tagged as such in order to remove them from the analysis.

Taking instructions into account, a completely counterbalanced block of comparisons was $2 \times 144 = 288$ randomized trials in length. Each subject performed eight complete blocks of trials. The first three blocks were regarded as practice.

Results and Discussion

The introduction of the "0" confidence category resulted in the removal of 33 (2.3%) and 2 (0.1%) responses from the analysis of the non-practice blocks of Subjects 3 and 4, respectively. Because the nature of the task was similar to that of Experiment 1, two-alternative forced-choice

discriminations (of the horizontal positions of points in this experiment), it was also possible to obtain d'_s measures of discriminative memory strength for position. Nine d'_s scores, one for each set size and ISI treatment combination, were obtained for each of the subjects.

As in Experiment 1, the ROC's were calculated from the observed response probabilities, conditionalized on each confidence-rating category, for each of the two response possibilities (a "first stimulus" response to ON THE LEFT [or equivalently, a "second stimulus" response to ON THE RIGHT], and a "second stimulus" response to ON THE LEFT [or equivalently, a "first stimulus" response to ON THE RIGHT]), for each of the two stimulus presentation possibilities (the leftmost point presented first and the leftmost point presented second). By collapsing these observed response probabilities over instruction type (ON THE LEFT, ON THE RIGHT), comparison type (to the left or to the right of the standard), study-stimulus location (1, 4, 7, or 10 o'clock), and number of non-practice blocks (five), each ROC was based on approximately 80×2 (leftmost point first, leftmost point second) = 160 observations.

The diffusion model estimates

The observed, and predicted (solid lines), values of each subject's d'_s scores are shown in the forgetting functions in Figure 2 (the corresponding percentage correct are shown in

Insert Figure 2 about here

Table 4). The average variance accounted for by the diffusion

Insert Table 4 about here

model (average R^2) for these data was 85%. For each subject the estimated values obtained from the model of both acquisitional variance (in pixels squared) and the corresponding rate of diffusion for each study-stimulus set size are shown in Table 5.

Insert Table 5 about here

The parameter representing the diffusion of the position of the study points in memory increased, for both subjects, as the study set size was increased from one to two remembered stimuli. According to the earlier interpretation of the diffusion process concerning dependencies between the rate of the diffusion and the available attentional capacity, this result suggested that more attentional resources were required to store absolute-position information than was required to store line-length information (hence, an acceleration in the diffusion rate occurred for a smaller number of remembered stimuli).

Surprisingly, diffusion rates decreased, for both

subjects, as the study set size was increased from two to four remembered stimuli. It may have been the case for study set sizes of four, however, that subjects were not remembering the absolute positions of each study point independently of the others, but that they were actually remembering the relations of the positions of each point to the other points in the study display. In fact, it was possible to remember the positions of the four study points by viewing them as corners of a large four-sided geometric configuration around the fixation point, which in essence reduced the memory task to the remembering of the shape and relative location, on the screen, of a geometric figure. Presumably, it would have been easier to remember, with less of a loss of information over time, the relative location and shape of a single dimensionally-complex figure than the absolute positions of four independent points. This interpretation was bolstered by the presence of d_s' scores in Subject 3's data at ISIs of 2000 and 5000 ms which were surprisingly higher for a study set size of four than for even the study set size of one (although this subject did report that she was paying more attention to this condition because it seemed more difficult).

It was also possible that this kind of strategy could have also been used for study set sizes of two. For example, the two study points could have been viewed as the ends of an imaginary line with a certain length, orientation and relative location on the screen. If this was indeed the case, then the

remembering of the positions of the points would have actually involved a different form of memory representation for each study set size.

Furthermore, pure absolute position discrimination would involve, for example, "the movement of an object seen in an otherwise homogeneous visual field (Ganzfeld)" (Kinchla & Allan, 1969, p. 537). Kinchla and Allan (1969) suggested that, based on previous findings, it is unlikely that subjects will use relative position discrimination to perform this type of task if useful points of reference (e.g., the edge of the computer screen) are located more than 8-10 degrees in visual angle from the stimulus. Because this was not possible on the available computer screen, it could not actually be assumed that any of these results represented "pure" absolute position discrimination.

The estimates of the acquisitional variance (in Table 5) showed that for both subjects the proportionate increase in acquisitional variance for the comparisons involving a study set size of one to the comparisons involving a study set size of two was much less than would have been expected based on the results observed in Experiment 1 (these parameters only increased by factors of 1.61 and 1.28 for Subjects 3 and 4, respectively). This supported the previously advanced notion that the subjects performed the task differently for study set sizes of one and two. If this was the case, then the increases in the diffusion parameter mentioned earlier, for

study set sizes of two over set sizes of one, would have indicated only that the specific memory representation used for set sizes of two diffused in memory at a greater rate than the memory representation which was used to store the position of a single point for study set sizes of one.

At study set sizes of four there was no increase in acquisitional variance for Subject 3 from that observed for the study set size of two. This finding would not be surprising considering the possible use of the set-size four memory strategy mentioned earlier. However, Subject 4 showed a massive increase in acquisitional variance over these same set size conditions. Input representations which seemed to be very clear for this subject at set sizes of one and two (estimated positional standard deviations of 2.10 and 2.37 pixels, respectively) suddenly became very noisy for the positions of the points within study set sizes of four (a standard deviation of 7.05 pixels).

Time-order error

The TOEs indices (see Table 6) that were calculated

Insert Table 6 about here

for each subject were based on:

$$1/2(p_{\text{correct}}[\text{right-left}] - p_{\text{correct}}[\text{left-right}]) \times 100\%, \quad (19)$$

where right-left and left-right refer to the order of presentation of the stimuli in a stimulus pair. As shown in

Table 6, TOEs were not eliminated by the use of these stimuli. Subject 3's data showed moderately negative TOEs at study set sizes of one and two stimuli (which decreased slightly at the longest ISI) and, curiously, a fairly large positive TOE for the 2000 ms ISI at a study set size of four. Subject 4's data showed a fairly large negative TOE at the longest (5000 ms) ISI for set-size two and large negative TOEs which increased over ISI for the study set size of four. Interestingly, within each treatment combination condition in Experiment 2 there seemed to be some relation between the direction of the TOE and the actual location of the target stimulus in the study display (see Table 4, Appendix E). For example, in Subject 3's data large negative TOEs for study stimuli that were located at "7:00" in the study display were often accompanied by large positive TOEs for study stimuli located at "1:00".

A negative TOE, as it was calculated in this experiment, implied a "drift" to the left of the position of the remembered stimulus. If this TOE was due to a leftward, or negative, drift in the diffusion process it was not clear either why such a drift would have taken place or why it might have depended on the location of the stimulus in the study display. It should be noted, however, that as indicated by Kinchla and Symzer (1967; Kinchla and Allan, 1969) the variance, and hence drift, in memory for absolute position can depend on factors such as eye or head movements which are not

relevant to the perceptual memory processes involved in the comparison of the extents of stimuli such as the line lengths in Experiment 1.

It was also possible that negative TOEs could have occurred in these data due to a preferential response bias for the response "first" when the instruction was ON THE LEFT and for the response "second" when the instruction was ON THE RIGHT. A response bias of this kind seemed more plausible in Experiment 2 because (even after explicit instructions to respond carefully, the performance of three practice blocks, and the advantage of having a "0" confidence category in which to indicate erroneous responses) subjects might sometimes have mistakenly pressed the left button on the mouse (an actual choice for the response "first") when the instruction was ON THE LEFT or the right button on the mouse (an actual "second" response) when the instruction was ON THE RIGHT. It was also plausible that such a bias might have been more likely to effect responding as the task became more difficult (i.e., as ISI and study set sizes increased; see Subject 4).

Furthermore, in Experiment 2 no empirical relation (see Table 3, Appendix E) was observed between d' scores and TOEs which were calculated over instruction within each treatment combination condition. This finding also suggested that it was more likely that the TOEs which occurred in Experiment 2 were due to some form of bias.

General Discussion

The present results provided some important evidence that the rate of forgetting over time of individual stimuli in short-term memory, defined as the rate of diffusion within the present theoretical framework, was influenced by the number of stimuli which were to be remembered. This dependence of forgetting and short-term memory load, defined as the number of stimuli in a to-be-remembered set of stimuli, was observed in the present Experiment 1 for three out of four possible cases. It was also proposed that this dependence occurred because the attention required to maintain a number of distinct perceptual stimuli (e.g , four) at rates of diffusion similar to the rate of diffusion at which a single stimulus in memory could have been maintained, exceeded the limited attentional capacity available in memory.

Similar results were found when subjects were required to remember the positions of points in Experiment 2. In that experiment diffusion rates increased as the study set size was increased from one to two stimuli. However, as discussed earlier, there was reason to believe that subjects were not performing the memory task in the same manner for each study set size. Hence, it might not have been valid to compare diffusion rates across set-size manipulations in that experiment.

Nevertheless, for the first time an experimental and theoretical paradigm was employed which allowed for the examination of stimulus set-sizes effects on memory processes

alone. A procedure (Palmer's partial discrimination paradigm) was used which eliminated any set-size effects that could have been attributed to decision processing. A task (two-alternative forced-choice) was employed which provided a performance measure (d_p') that was free from any effects due to differential response-criterion biases over the set-size and ISI conditions of the present experiment. A class of stimuli (line lengths) was used which minimized any effects in memory due to rehearsal or semantic association. Experimental stimuli (the stimulus distributions) were chosen and displayed over comparison trials (symmetric presentation of the standard and comparison lengths) in such a manner as to ensure that only pure perceptual memories were being retained. Furthermore, a method of analysis (the diffusion model) was employed that allowed for the separation of set-size effects which were due to either sensory effects or perceptual processing limitations during acquisition, and set-size effects which acted on the memory retention processes themselves. As well, the diffusion model provided a compelling theoretical description of memory retention and the processes responsible for forgetting.

It was not clear in the present research why the increases in diffusion observed over set sizes for the larger distribution stimuli did not extend to the smaller distribution stimuli. If this finding could be shown to be reliable in future research, it would be important to

determine how it arises that set-size effects in memory can be mediated by the within-display range of study stimuli. It was very possible, however, that increases in the rates of diffusion over set size for the smaller distribution stimuli would have been found in the present Experiment 1 if more reliable estimates of d' could have been obtained. Because the collection of more reliable estimates would have required many more observations at each set-size, ISI, and stimulus-distribution condition, the sheer size of the design restricted such an endeavour.

Furthermore, due to the changes implemented to Palmer's paradigm in the present experiment it was not possible from the present results to say that increased rates of forgetting for larger stimulus set sizes in memory should necessarily be observed under experimental conditions identical to Palmer (1990). A more exact replication of Palmer's (1990) Experiment 4 might be in order. While it was stressed earlier that the maintenance of a constant test stimulus (as in Palmer's study, 1990) might not be desirable if ISI manipulations are used, it would seem to be necessary, however, to extend the results of the present experiments to shorter acquisition times such as that used by Palmer (100 ms; 1990).

The fact that a dependence of short-term forgetting on memory load can be demonstrated, given the appropriate experimental conditions and theoretical analysis, should not

be surprising. In fact, this dependence is explicitly implied by Miller's (1956) span of memory. If forgetting occurs at a rate which is independent of the number of remembered stimuli, then it would logically follow that the number of distinct items which can be accurately remembered in the short-term would be limited only by the number of items which can be adequately perceived and not by some magic number. The presence of a measurable span of memory should indicate that a set of to-be-remembered stimuli which is larger than that span is forgotten extremely rapidly.

References

- Baranski, J. V., & Petrusic, W. M. (1992). The discriminability of remembered magnitudes. Memory and Cognition, 20, 254-270.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behaviour, 11, 671-684.
- Dale, H. C. A. (1973). Short-term memory for visual information. British Journal of Psychology, 64, 1-8.
- Fechner, G. T. (1966). Elements of Psychophysics (H. E. Adler, Trans.). New York: Harper & Row. (Original work published in 1860).
- Green, D. M., & Swets, J. A. (1966). Signal Detection Theory and Psychophysics. New York : Wiley.
- Greenberg, M. G. (1965). A modification of Thurstone's law of comparative judgment to accommodate a judgment category of "equal" or "no difference". Psychological Bulletin, 64, 108-112.
- Grindley, G. C., & Townsend, V. (1968). Voluntary attention in peripheral vision and its effects on acuity and differential thresholds. Quarterly Journal of Experimental Psychology, 20, 11-19.
- Hellstrom, A. (1977). Time errors are perceptual: An experimental investigation of duration and a quantitative successive comparison model. Psychological Research, 39, 345-388.

- Hellstrom, A. (1979). Time errors and differential sensation weighting. Journal of Experimental Psychology: Human Perception & Performance, 5, 460-477.
- Hellstrom, A. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. Psychological Bulletin, 97, 35-61.
- Helson, H. (1947). Adaption-level as a frame of reference for prediction of psychological data. American Journal of Psychology, 60, 1-29.
- Jamieson, D. G., & Petrusic, W. M. (1975a). Presentation order effects in duration discrimination. Perception & Psychophysics, 17, 197-202.
- Jamieson, D. G., & Petrusic, W. M. (1975b). The dependence of time-order direction on stimulus range. The Canadian Journal of Psychology, 29, 175-182.
- Jamieson, D. G., & Petrusic, W. M. (1975c). Pairing effects and time-order errors in duration discrimination. Perception & Psychophysics, 18, 107-113.
- Jamieson, D. G., & Petrusic, W. M. (1976). On a bias induced by the provision of feedback in psychophysical experiments. Acta Psychologica, 40, 199-206.
- Jamieson, D. G., & Petrusic, W. M. (1978). Feedback versus an illusion in time. Perception, 7, 91-96.
- Kinchla, R. A., & Allan, L. G. (1969). A theory of visual movement perception. Psychological Review, 76, 537-558.
- Kinchla, R. A., & Allan, L. G. (1970). Visual movement

- perception: A comparison of sensitivity to vertical and horizontal movement. Perception and Psychophysics, 8, 399-405.
- Kinchla, R. A., & Symzer, F. (1967). A diffusion model of perceptual memory. Perception and Psychophysics, 2, 219-229.
- Laming, D., & Scheiwiller, P. (1985). Retention in perceptual memory: A review of models and data. Perception and Psychophysics, 37, 189-197.
- Laming, D., & Wightman, D. (1992). Perceptual memory over very short interstimulus intervals. Bulletin of the Psychonomic Society, 30, 170-172.
- Lindsay, P. H., & Norman, D. A. (1969). Short-term retention during a simultaneous detection task. Perception and Psychophysics, 5, 1969.
- Lindsay, P. H., Taylor, M.M., & Forbes, S. M. (1968). Attention and multidimensional discrimination. Perception and Psychophysics, 4, 113-117.
- Loftus, G. R. (1985). Evaluating forgetting curves. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 397-406.
- Macmillan, N. A., & Creelman, C. D. (1991). Detection Theory: A User's Guide. New York: Cambridge University Press.
- Massaro, D. W. (1970a). Retroactive interference in short-term recognition memory for pitch. Journal of

Experimental Psychology, 83, 32-39.

Massaro, D. W. (1970b). Forgetting: Interference or Decay?

Journal of Experimental Psychology, 83, 238-243.

Massaro, D. W. (1970c). Consolidation and interference in the perceptual memory system. Perception and Psychophysics, 7, 153-156.

Massaro, D. W. (1970d). Perceptual processes and forgetting in memory tasks. Psychological Review, 77, 557-567.

Michels, W. C., & Helson, H. (1954). A quantitative theory of time-order effects. American Journal of Psychology, 67, 327-334.

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

Needham, J. G. (1934). The time error in comparison judgments. Psychological Bulletin, 31, 229-243.

Norman, D. A. (1966). Acquisition and retention in short-term memory. Journal of Experimental Psychology, 72, 369-381.

Norman, D. A., & Wickelgren, W. A. (1969). Strength theory of decision rules and latency in retrieval from short-term memory. Journal of Mathematical Psychology, 6, 192-208.

Palmer, J. (1988). Very short-term visual memory for size and shape. Perception and Psychophysics, 43, 278-286.

Palmer, J. (1990). Attentional limits on the perception and

- memory of visual information. Journal of Experimental Psychology: Human Perception and Performance, 16, 332-350.
- Palmer, J. (1991). Isolating the components of very- short-term visual memory. Bulletin of the Psychonomic Society, 22, 399-402.
- Palmer, J., Ames, C. T., & Lindsey D. T. (1993). Measuring the effect of attention on simple visual search. Journal of Experimental Psychology: Human Perception and Performance, 19, 108-130.
- Peterson, L.R, & Peterson, M.J. (1959). Short-term retention of individual verbal items. Journal of Experimental Psychology, 58, 193-198.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. Perception and Psychophysics, 16, 283-290.
- Posner, M. I. (1967). Characteristics of visual and kinaesthetic memory codes. Journal of Experimental Psychology, 75, 103-107.
- Posner, M. I., & Konick, A. F. (1966). Short-term retention of visual and kinaesthetic information. Organizational Behaviour and Human Performance, 1, 71-86.
- Purdy, J. E., Eimann, D. G., & Cross, H. A. (1980). Persistence of a briefly presented visual stimulus in sensory memory. Bulletin of the Psychonomic Society, 16, 374-376.

- Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. In R. S. Nickerson (Ed.), Attention and performance VIII (pp. 277-296). Hillsdale, NJ: Erlbaum.
- Sperling, G. (1960). The information available in brief visual presentations. Psychological Monographs, 74 (11, Whole No. 498).
- Wickelgren, W. A. (1968). Unidimensional strength theory and component analysis of noise in absolute and comparative judgments. Journal of Mathematical Psychology, 5, 102-122.
- Wickelgren, W. A. (1969). Associative strength theory of recognition memory for pitch. Journal of Mathematical Psychology, 6, 13-61.
- Wickelgren, W. A., & Norman, D. A. (1966). Strength models and serial position in short-term recognition memory. Journal of Mathematical Psychology, 3, 316-347.
- Williams, H. L., Beaver, W. S., Spence, M. T., & Rundell, O. H. (1969). Digital and kinaesthetic memory with interpolated information processing. Journal of Experimental Psychology, 80, 530-536.
- Woodrow, H. (1933). Weight discrimination with a varying standard. American Journal of Psychology, 45, 391-416.
- Woodrow, H. (1935). The effect of practice upon time-order errors in the comparison of temporal intervals. Psychological Review, 42, 127-152.

Table 1

Percent Correct for Experiment 1

Subject 1

Smaller stimulus distribution

	500 ms	2000 ms	5000 ms
Set 1	66	62	62
Set 2	43	53	64
Set 4	52	50	48

Larger stimulus distribution

	500 ms	2000 ms	5000 ms
Set 1	78	66	66
Set 2	67	58	59
Set 4	58	61	54

Subject two

Smaller stimulus distribution

	500 ms	2000 ms	5000 ms
Set 1	96	90	84
Set 2	90	90	83
Set 4	78	74	79

Larger stimulus distribution

	500 ms	2000 ms	5000 ms
Set 1	94	93	83
Set 2	86	80	79
Set 4	71	70	66

Table 2

Diffusion Model Estimates for Experiment 1

Subject 1 Smaller distribution: Larger distribution:

	Acquisition variance	Diffusion parameter	Acquisition variance	Diffusion parameter
Set-size 1	4.68	5.01	3.37	6.75
Set-size 2	412.69	- 79.98	35.32	6.17
Set-size 4	106.36	91.30	84.01	27.91

Subject 2 Smaller distribution: Larger distribution:

	Acquisition variance	Diffusion parameter	Acquisition variance	Diffusion parameter
Set-size 1	1.06	.58	2.51	1.32
Set-size 2	2.09	.26	6.31	1.39
Set-size 4	5.60	.20	15.69	2.64

Table 3

TOE Indices for Experiment 1

Subject 1

Smaller stimulus distribution

	500 ms	2000 ms	5000 ms
Set 1	- 23	- 29.5	- 8.5
Set 2	- 29	- 22.5	- 21.5
Set 4	- 22	- 17	- 17.5

Larger stimulus distribution

	500 ms	2000 ms	5000 ms
Set 1	- 10.5	- 27	- 26
Set 2	- 14	- 26.5	- 28
Set 4	- 16.5	- 20.5	- 27.5

Subject 2

Smaller stimulus distribution

	500 ms	2000 ms	5000 ms
Set 1	+ 1	+ 2.5	- 5
Set 2	+ 2	+ 2.5	+ 0.5
Set 4	- 0.5	- 1	- 7

Larger stimulus distribution

	500 ms	2000 ms	5000 ms
Set 1	+ 0	+ 0.5	+ 5.5
Set 2	+ 10.5	+ 16.5	+ 10.5
Set 4	+ 19	+ 25	+ 25.5

Table 4

Percent Correct for Experiment 2

Subject 3

	500 ms	2000 ms	5000 ms
Set 1	87	79	74
Set 2	78	79	68
Set 4	81	89	80

Subject 4

	500 ms	2000 ms	5000 ms
Set 1	98	96	82
Set 2	89	81	66
Set 4	76	66	65

Table 5

Diffusion Model Parameters for Experiment 2

Subject 3 :

	Acquisition variance	Diffusion parameter
Set-size 1	18.58	13.03
Set-size 2	29.96	18.01
Set-size 4	29.84	2.36

Subject 4 :

	Acquisition variance	Diffusion parameter
Set-size 1	4.40	4.24
Set-size 2	5.63	14.85
Set-size 4	49.72	10.02

Table 6

TOE Indices for Experiment 2

Subject 3

	500 ms	2000 ms	5000 ms
Set 1	- 6	- 6.5	- 3
Set 2	- 5.5	- 5	- 4
Set 4	- 1.5	+ 7	- 1

Subject 4

	500 ms	2000 ms	5000 ms
Set 1	+ 1	- 3.5	- 2.5
Set 2	- 5	- 2.5	- 9.5
Set 4	- 8	- 12	- 14.5

Figure 1. Observed (symbols) and predicted (solid lines) d' scores for the subjects in Experiment 1.

Figure 2. Observed (symbols) and predicted (solid lines) d' scores for the subjects in Experiment 2.

Appendix A

In order to show how the sample size model can be applied to Palmer's data, first consider a simple detection task in which either of two line lengths (x_1 and x_2 , $x_2 > x_1$) is presented and the observer is required to respond "yes" or "no" as to whether he/she thinks that the presented line is the longer line. After many trials, in which both line lengths have been presented equally often, the observer internalizes the lengths of the lines representationally as a distributions of random variables, whose means are expected to equal the actual physical values of each stimulus length. In order to make a "yes-no" response, classical signal detection theory (Green and Swets, 1966) would assume that the observer would set a psychological criterion value (c) somewhere between the means of these two internalized stimulus distributions. On any particular trial they would then respond "yes" if their perception of the stimulus (i.e., the aggregate value in the sample size model) is higher than that criterion.

The d' signal detection measure can be calculated from the observed "hit" and "false alarm" probabilities by defining a "hit" as a "yes" response when the presented stimulus was actually the longer of the two stimuli and a "false alarm" as a "yes" response when the presented stimulus was actually the shorter stimulus. A d' obtained in this manner is called the discriminability index. According to the sample size model:

$$\underline{d'} = E(\underline{x}_2) - E(\underline{x}_1) / (\text{Var}[\underline{x}] / \underline{n})^{1/2}, \quad (1)$$

where $\underline{n}$ is the number of perceptual samples which can be accumulated during the available acquisition time. In Palmer's (1990) experiments, the variability associated with the perception of each line length on any trial, the square of the denominator in (1), should depend on the study set size (i.e., $\text{Var}[\underline{x}] / [\underline{n}/\underline{N}]$, where $\underline{N}$ is the set size), because subjects must share the number of possible perceptual samples over $\underline{N}$ stimuli. Hence:

$$\underline{d'_N} = \underline{d'_1} (\underline{1}/\underline{N})^{1/2}, \quad (2)$$

or similarly (because DL's are inversely related to the discriminability index, $\underline{d'}$, Laming & Scheiwiller. 1985):

$$DL_N = DL_1 (\underline{N})^{1/2}. \quad (3)$$

Appendix B

In a two-alternative forced-choice comparison task, two stimuli are presented, usually successively, and the task is to decide which stimulus is larger (or smaller). It is assumed, in these circumstances, that in order to perform the task the observer takes the difference between the two perceived values ($x_2 - x_1$). Accordingly, he/she will respond that x_2 is larger if this perceived difference value is larger than some preset internal criterion (e.g., if it is positive).

Hence the d' measure is now defined as the distance between the means of the two difference distributions, scaled in units of the standard deviation of these difference distributions, for $(x_2 - x_1)$ and $(x_1 - x_2)$. Mathematically:

$$d' = E(x_2 - x_1) - E(x_1 - x_2) / (\text{Var}[x_2 - x_1])^{1/2}. \quad (1)$$

(It can be shown that d' obtained in a two-alternative forced-choice task is equal to $(2)^{1/2}d'$ obtained in a yes-no task.) More importantly, for the present purposes, according to the sample size model the variability, the square of the denominator in (1), for a study set size of N is equal to $\text{Var}(x)(1/\underline{n} + N/\underline{n})$, where $\underline{n}$ is the number of perceptual samples acquired during study and test stimulus presentation conditions similar to that of Palmer (1990), and the variabilities of x_1 and x_2 are assumed to be equal. Hence, it can be shown that:

$$d'_{\text{sc}} = d'_{\text{sc}} (2/[1 + N])^{1/2}. \quad (2)$$

Under these task conditions the sample size model would predict that observed d' 's for study set sizes of two, and four, should be reduced from the observed d' at a set size of one by a factor of 0.82, and 0.63, respectively (in comparison to a reduction by a factor of 0.71, and 0.5, respectively, in a "yes-no" task such as that described in Appendix A).

Appendix C

Typically, perceptual memory tasks have involved judgments of similarity, in which subjects have been asked to decide whether a test or comparison stimulus was either the "same" or "different" from a previously-presented study stimulus. Half of the time the test stimulus and the study stimulus were identical, and the remaining half of the time the test stimulus differed from the study stimulus by some constant amount. The signal-detection measure, d' , was calculated from the observed hit and false alarm probabilities by defining, for example, a "hit" as the proportion of trials in which the observer reported a stimulus difference when in fact there was a stimulus difference and a "false alarm" as the proportion of trials in which a stimulus difference was reported when in fact none had occurred (Kinchla & Symzer, 1967).

Two major classes of quantitative models have arisen in the modern perceptual-memory literature. The first class can be described as strength models (Wickelgren, 1968, 1969; Massaro, 1970a, 1970c, 1970d; see also Norman & Wickelgren, 1969; Wickelgren & Norman, 1966); the second class being the diffusion model. These models assume that perceptual-memory tasks based on similarity judgments can be regarded as recognition tasks (i.e., a new test stimulus is to be recognized as being either the same or different than a previously-presented study stimulus). Through analyses of

observed receiver operating characteristics (ROCs), Wickelgren (1969) attempted to show that observers perform such a task by way of a familiarity judgment between the perception of the test stimulus and the memory trace of the study stimulus. Familiarity is assumed to be a dimension of the perception of the test stimulus which should depend on both its actual closeness in physical value to the study stimulus, as well as the remaining strength, or consolidation, of the study-stimulus trace in memory.

In these models, the strength of the remaining memory trace is operationally defined as the psychological distance (scaled in variance units, i.e., d') between the means of the representative familiarity distributions which arise when the test and the study stimuli are equal, and when the test and the study stimuli are different. If the perception of familiarity of the test stimulus is lower than some criterion (β) then the observer responds that both the test and study stimuli were "different". As memory strength decreases over the ISI, this distance between the representative familiarity distributions (d') decreases and, as a result, the discriminative sensitivity between them becomes degraded.

The nature of this decrease in memory strength over ISI has been attributed to both temporal decay (Wickelgren, 1969) and interference from other traces in memory (Massaro, 1970b). In order to model the acquisition and retention of memory traces, observed forgetting curves were fit, by these

researchers, with functions which assumed that: (a) acquisition involves an exponential growth process of trace strength over the presentation of the study stimulus, and (b) retention involves an exponential decay process of trace strength over the ISI. While Massaro (the storage forgetting model, 1970a, 1970c, 1970d) proposed that the reduction in strength of a single memory trace is responsible for observed perceptual memory performance, Wickelgren (dual-trace model, 1969) proposed the existence of two separate memory traces (both short- and intermediate-term traces).

Appendix D

According to the sample size model, the sum of the input variance of the perceived values of the test (x_t) and study (x_0) stimuli, under the stimulus presentation conditions of the present experiment would be (see also Appendix B):

$$\text{Var}(x_t) + \text{Var}(x_0) = \text{Var}(x) \left(\frac{1}{n_c} + \frac{N}{n_s} \right), \quad (1)$$

where N is the study-stimulus set size, n_s is the number of perceptual samples acquired during the study stimulus presentation time (1000 ms), and n_c is the corresponding number of perceptual samples which can be acquired during the presentation of the test stimulus (unlimited hold). Hence, the sample size model would predict, under two-alternative forced-choice task conditions, that the acquisitional variance observed (independent of variance due to diffusion in memory) for a study set size of N should be related to the acquisitional variance observed for a study set size of one by the relation:

$$\begin{aligned} (\text{Var}[x_t] + \text{Var}[x_0])_N &= \\ (\text{Var}[x_t] + \text{Var}[x_0])_1 \left(\frac{1}{n_c} + \frac{N}{n_s} \right) / \left(\frac{1}{n_c} + \frac{1}{n_s} \right). \end{aligned} \quad (2)$$

If it is assumed, as an upper bound, that n_c goes to infinity (i.e., unlimited sampling of x_t) then $(1/n_c)$ goes to 0 and:

$$\begin{aligned} \left(\frac{1}{n_c} + \frac{N}{n_s} \right) / \left(\frac{1}{n_c} + \frac{1}{n_s} \right) &\Rightarrow \left(\frac{N}{n_s} \right) / \left(\frac{1}{n_s} \right) \\ &= N. \end{aligned} \quad (3)$$

Appendix E

Table 1

D.' Scores, Variance Estimates (in brackets), and Model R²'s

Experiment 1 :

Subject 1 - Smaller distribution

	500 ms	2000 ms	5000 ms
Set-size 1 R ² = .979	1.64 (5.92)	.98 (16.60)	.74 (29.12)
Set-size 2 R ² = .833	.19 (434.81)	.32 (159.57)	.60 (43.85)
Set-size 4 R ² = .824	.45 (78.83)	.20 (398.73)	.17 (526.26)

Subject 1 - Larger distribution

	500 ms	2000 ms	5000 ms
Set-size 1 R ² = 1.000	2.05 (8.53)	1.64 (13.33)	1.23 (23.63)
Set-size 2 R ² = .989	.95 (39.53)	.88 (45.96)	.73 (66.72)
Set-size 4 R ² = .968	.58 (106.82)	.53 (126.54)	.40 (227.97)

Subject 2 - Smaller distribution

	500 ms	2000 ms	5000 ms
Set-size 1 R ² = .979	3.65 (1.20)	2.56 (2.45)	2.03 (3.90)
Set-size 2 R ² = .884	2.60 (2.38)	2.60 (2.36)	2.15 (3.46)
Set-size 4 R ² = .618	1.72 (5.43)	1.58 (6.43)	1.57 (6.49)

Subject 2 - Larger distribution

	500 ms	2000 ms	5000 ms
Set-size 1 $R^2 = .967$	3.17 (3.58)	2.82 (4.53)	1.97 (9.29)
Set-size 2 $R^2 = .540$	2.74 (4.78)	1.70 (12.46)	1.72 (12.19)
Set-size 4 $R^2 = .940$	1.41 (18.17)	1.37 (19.25)	1.11 (29.48)

Experiment 2 :

Subject 3

	500 ms	2000 ms	5000 ms
Set-size 1 $R^2 = .998$	1.87 (26.66)	1.54 (42.29)	1.09 (84.52)
Set-size 2 $R^2 = .906$	1.37 (53.57)	1.51 (44.07)	.89 (127.29)
Set-size 3 $R^2 = .497$	1.62 (38.15)	2.05 (23.87)	1.49 (45.20)

Subject 4

	500 ms	2000 ms	5000 ms
Set-size 1 $R^2 = .998$	3.77 (7.03)	2.87 (12.11)	1.97 (25.85)
Set-size 2 $R^2 = .993$	2.49 (16.12)	1.80 (30.73)	1.11 (81.41)
Set-size 4 $R^2 = .700$	1.69 (35.04)	1.03 (94.30)	1.06 (88.98)

Table 2

Slopes of the ROCs

Experiment 1 :

Subject 1 - Smaller distribution

	500 ms	2000 ms	5000 ms
Set-size 1	.36 ^a	.85	.91
Set-size 2	.56	.77	1.08
Set-size 4	.79	1.02	.90

Subject 1 - Larger distribution

	500 ms	2000 ms	5000 ms
Set-size 1	.95	.51	.76
Set-size 2	1.27	.94	.95
Set-size 4	.99	1.23	.90

Subject 2 - Smaller distribution

	500 ms	2000 ms	5000 ms
Set-size 1	1.00	.66	1.22
Set-size 2	1.31	1.41	1.03
Set-size 4	.91	1.08	.85

Subject 2 - Larger distribution

	500 ms	2000 ms	5000 ms
Set-size 1	1.28	.96	.92
Set-size 2	1.50	.65	1.38
Set-size 4	1.07	.84	.94

2 of/de 2

PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISIONSM RESOLUTION TARGETS

Experiment 2 :

Subject 3

	500 ms	2000 ms	5000 ms
Set-size 1	1.87 ^b	1.25	1.39
Set-size 2	1.15	1.00	1.19
Set-size 4	.98	.39	.97

Subject 4

	500 ms	2000 ms	5000 ms
Set-size 1	1.28	1.83	.75
Set-size 2	.75	.76	1.06
Set-size 4	.90	.84	.71

* Slopes represent the ratio of the difference-distribution standard deviation when the stimuli were presented in the long-short order of presentation over the short-long order of presentation.

^b Slopes represent the ratio of the difference-distribution standard deviation when the stimuli were presented in the right-left order of presentation over the left-right order of presentation.

Table 3

D.' Scores, Percent Correct (in brackets), and TOEs* (in square [] brackets) for each Instruction

Experiment 1 :

Subject 1 - Smaller distribution

SHORT ER?	LONG ER?	500 ms		2000 ms		5000 ms	
Set-size 1		1.30 (58) (-29)	2.26 (77) (-15)	.86 (58) (-32)	1.10 (66) (-27)	.68 (58) (-21)	.83 (65) (+ 4)
Set-size 2		.13 (40) (-35)	.27 (46) (-23)	.11 (45) (-23)	.53 (60) (-22)	.43 (62) (-30)	.72 (66) (-13)
Set-size 4		.23 (45) (-29)	.70 (59) (-16)	-.06 (44) (-31)	.38 (57) (- 5)	.08 (46) (-24)	.26 (51) (-12)

Subject 1 - Larger distribution

SHORT ER?	LONG FR?	500 ms		2000 ms		5000 ms	
Set-size 1		2.17 (80) (-11)	1.94 (77) (-11)	1.48 (60) (-29)	1.82 (71) (-26)	.93 (66) (-26)	1.58 (67) (-27)
Set-size 2		.98 (66) (-19)	.94 (69) (-10)	.95 (58) (-27)	.80 (68) (-27)	.72 (55) (-26)	.77 (63) (-31)
Set-size 4		.38 (52) (-21)	.80 (65) (-14)	.52 (66) (-22)	.57 (56) (-19)	.52 (50) (-29)	.23 (58) (-27)

Subject 2 - Smaller distribution

SHORT ER?	LONG ER?	500 ms		2000 ms		5000 ms	
Set-size 1		3.74 (96) (+ 2)	3.50 (96) (+ 2)	3.59 (96) (- 2)	2.04 (84) (+ 6)	1.87 (81) (- 7)	2.18 (87) (- 4)
Set-size 2		2.80 (93) (0)	2.42 (87) (+ 4)	2.75 (93) (0)	2.51 (88) (+ 5)	1.95 (81) (+ 2)	2.38 (85) (0)
Set-size 4		2.12 (87) (- 1)	1.38 (70) (0)	1.42 (70) (- 5)	1.80 (78) (+ 3)	1.58 (79) (- 9)	1.61 (79) (- 6)

Subject 2 - Larger distribution

SHORT ER?	LONG ER?	500 ms		2000 ms		5000 ms	
Set-size 1		3.11 (93) (0)	3.27 (95) (0)	3.30 (95) (0)	2.47 (91) (+ 2)	2.20 (85) (+ 5)	1.79 (81) (+ 7)
Set-size 2		2.62 (87) (+ 9)	3.05 (85) (+13)	1.96 (81) (+14)	1.50 (78) (+20)	1.77 (76) (+12)	1.72 (83) (+10)
Set-size 4		1.64 (74) (+19)	1.31 (68) (+19)	1.50 (70) (+23)	1.28 (70) (+28)	1.30 (70) (+19)	.95 (61) (+33)

Experiment 2 :

Subject 3

ONTHE LEFT? ONTHE RIGHT?	500 ms		2000 ms		5000 ms	
Set-size 1	1.48 (84) (- 9)	2.28 (90) (- 4)	1.50 (81) (- 5)	1.54 (76) (- 9)	1.21 (76) (- 5)	.98 (72) (- 1)
Set-size 2	1.20 (75) (- 3)	1.46 (81) (- 9)	1.48 (82) (-12)	1.45 (77) (+3)	.98 (65) (- 2)	.78 (71) (- 7)
Set-size 4	1.56 (83) (- 5)	1.69 (80) (+ 3)	1.76 (86) (+10)	2.40 (93) (+ 5)	1.31 (77) (+ 1)	1.59 (82) (- 3)

Subject 4

ONTHE LEFT? ONTHE RIGHT?	500 ms		2000 ms		5000 ms	
Set-size 1	3.22 (97) (0)	- ^b (99) (+ 2)	- ^b (95) (- 5)	2.99 (96) (- 2)	1.73 (79) (- 4)	2.31 (86) (- 2)
Set-size 2	2.16 (84) (- 6)	3.12 (94) (- 4)	2.03 (82) (- 8)	1.74 (79) (+ 2)	1.25 (65) (-17)	1.01 (67) (- 2)
Set-size 4	2.10 (77) (-17)	1.56 (76) (0)	1.11 (67) (-16)	.95 (64) (- 9)	.93 (59) (-25)	1.28 (71) (- 4)

* TOE indices have been rounded up to the nearest full percentage point

^b ROC's could not be obtained for this data

Table 4

Percentage Correct and TOEs^a (in brackets) for each
Stimulus Location

Experiment 1 :

Subject 1 - Smaller distribution

10:00 7:00	1:00 4:00	500 ms	2000 ms	5000 ms
Set-size 1		50 (-30)	53 (-41)	61 (- 5)
		69 (-31)	73 (-27)	63 (-13)
		69 (-12)	61 (-32)	66 (-16)
		72 (-22)	61 (-20)	59 (- 3)
Set-size 2		29 (-21)	66 (-29)	59 (-29)
		47 (-34)	50 (-31)	66 (-35)
		41 (-28)	45 (- 8)	66 (-10)
		50 (-31)	59 (-22)	64 (-13)
Set-size 4		56 (-19)	46 (-25)	59 (-22)
		47 (-34)	63 (-32)	41 (-28)
		41 (-10)	41 (+ 3)	45 (- 5)
		64 (-25)	53 (-16)	49 (-17)

Subject 1 - Larger distribution

10:00 7:00	1:00 4:00	500 ms	2000 ms	5000 ms
Set-size 1		81 (-19)	61 (-30)	64 (-36)
		73 (-15)	64 (-30)	69 (-13)
		82 (- 1)	61 (-30)	53 (-35)
		78 (-10)	78 (-22)	79 (-22)
Set-size 2		66 (- 3)	59 (-29)	73 (-28)
		66 (-22)	59 (-29)	50 (-38)
		69 (-19)	56 (-25)	52 (-23)
		69 (-13)	56 (-25)	62 (-23)
Set-size 4		64 (-11)	61 (-22)	56 (-44)
		52 (-32)	63 (-13)	55 (-23)
		69 (-13)	66 (-22)	50 (-20)
		50 (-13)	56 (-25)	53 (-22)

Subject 2 - Smaller distribution

10:00 1:00 7:00 4:00	500 ms	2000 ms	5000 ms
Set-size 1	100 98 (0) (- 3)	98 90 (+ 3) (- 5)	88 80 (- 8) (-10)
	98 90 (+ 3) (+ 5)	86 88 (+10) (+ 3)	78 90 (+ 2) (-10)
Set-size 2	97 90 (- 4) (0)	95 83 (0) (+ 3)	83 85 (+ 3) (0)
	85 90 (+10) (0)	90 93 (+ 5) (+ 3)	90 75 (+ 5) (- 5)
Set-size 4	80 70 (+ 5) (+ 5)	70 80 (+ 5) (-10)	68 78 (- 3) (-13)
	78 85 (- 6) (+ 5)	65 80 (+ 5) (- 5)	85 85 (-10) (-10)

Subject 2 - Larger distribution

10:00 1:00 7:00 4:00	500 ms	2000 ms	5000 ms
Set-size 1	90 88 (+ 5) (-13)	93 95 (- 3) (+ 5)	85 83 (+ 5) (+ 3)
	98 100 (+ 3) (0)	90 95 (- 5) (+ 5)	78 88 (+13) (+ 3)
Set-size 2	95 88 (+ 5) (+13)	85 80 (+15) (+20)	83 73 (-13) (+28)
	80 80 (+10) (+20)	83 70 (+13) (+20)	88 75 (+ 8) (+15)
Set-size 4	71 70 (+10) (+25)	73 63 (+23) (+28)	71 63 (+24) (+28)
	69 75 (+15) (+25)	74 73 (+23) (+28)	55 72 (+18) (+28)

Experiment 2 :

Subject 3

10:00 7:00	1:00 4:00	500 ms	2000 ms	5000 ms
Set-size 1		93 (- 3)	87 (+12)	90 (+ 6)
		83 (+18)	72 (+28)	65 (+34)
Set-size 2		78 (-21)	89 (-12)	66 (-25)
		77 (-23)	72 (-28)	71 (-28)
Set-size 4		83 (- 3)	74 (+24)	89 (- 4)
		74 (+24)	70 (+ 4)	63 (+38)
Set-size 4		69 (-33)	85 (-11)	71 (-28)
		83 (-13)	63 (-36)	77 (-20)
Set-size 4		74 (- 5)	89 (+12)	98 (+ 3)
		76 (+22)	76 (-13)	74 (+28)
Set-size 4		75 (-27)	87 (+13)	93 (- 3)
		90 (+ 6)	79 (-21)	90 (+ 5)

Subject 4

10:00 7:00	1:00 4:00	500 ms	2000 ms	5000 ms
Set-size 1		100 (0)	100 (0)	98 (+ 3)
		95 (+ 6)	95 (- 6)	90 (0)
Set-size 2		98 (- 3)	100 (0)	83 (+11)
		85 (-16)	72 (- 7)	72 (- 7)
Set-size 2		93 (+ 3)	89 (+ 5)	76 (- 3)
		93 (- 3)	76 (- 3)	60 (-23)
Set-size 4		88 (-11)	85 (-15)	73 (-22)
		77 (- 9)	76 (+ 4)	55 (+16)
Set-size 4		71 (-28)	84 (- 8)	67 (-29)
		74 (- 7)	63 (-22)	78 (-16)
Set-size 4		73 (- 8)	77 (+ 9)	60 (-12)
		62 (- 2)	54 (-17)	66 (- 6)

* TOE's indices have been rounded up to the nearest full

Memory Load and Rate of Forgetting
percentage point

95

Table 5

End Stimulus Pairs: The Actual Percentage^a of Times That They Were Presented and the Percent of Correct Responses (in brackets) To Those Stimuli

Experiment 1 :

Subject 1 - Smaller distribution

LOW 23,25	HIGH 29,31 ^b	500 ms		2000 ms		5000 ms	
Set-size 1		9.3 (64)	2.5 (67)	4.6 (83)	3.1 (100)	15.3 (65)	2.3 (67)
Set-size 2		3.4 (50)	1.7 (100)	13.0 (47)	6.1 (75)	18.6 (75)	0
Set-size 4		7.0 (33)	4.7 (50)	11.5 (40)	8.4 (55)	3.8 (40)	3.1 (75)

Subject 1 - Larger distribution

LOW 37,39	HIGH 43,45 ^b	500 ms		2000 ms		5000 ms	
Set-size 1		3.8 (80)	3.8 (100)	8.4 (73)	4.6 (83)	7.6 (50)	3.1 (100)
Set-size 2		5.5 (57)	3.1 (75)	10.2 (62)	3.1 (75)	2.3 (33)	3.0 (75)
Set-size 4		3.8 (60)	5.4 (57)	8.5 (64)	3.9 (80)	8.5 (60)	6.0 (43)

Subject 2 - Smaller distribution

LOW 23,25	HIGH 29,31 ^b	500 ms		2000 ms		5000 ms	
Set-size 1		7.3 (100)	1.3 (100)	3.7 (100)	0	1.2 (100)	5.6 (100)
Set-size 2		9.9 (87)	2.0 (100)	6.2 (100)	3.1 (100)	5.6 (89)	3.1 (100)
Set-size 4		3.1 (60)	4.3 (57)	13.1 (81)	3.8 (33)	8.0 (92)	8.6 (79)

Subject 2 - Larger distribution

LOW 37,39	HIGH 43,45 ^b	500 ms		2000 ms		5000 ms	
Set-size 1		4.3 (100)	2.5 (100)	7.3 (100)	3.7 (83)	8.8 (79)	4.4 (100)
Set-size 2		5.6 (89)	6.2 (60)	8.1 (69)	1.9 (67)	6.9 (82)	5.0 (88)
Set-size 4		4.9 (75)	1.2 (50)	3.7 (83)	4.3 (86)	7.2 (82)	5.9 (67)

Experiment 2 :

Subject 3

LOW -20,-10	HIGH 10,20 ^b	500 ms		2000 ms		5000 ms	
Set-size 1		4.5 (86)	4.5 (86)	2.5 (75)	6.1 (80)	9.6 (53)	1.9 (67)
Set-size 2		2.7 (100)	6.0 (67)	5.9 (78)	4.6 (57)	4.5 (86)	7.7 (58)
Set-size 4		6.2 (90)	5.6 (56)	3.2 (100)	3.8 (100)	3.8 (100)	5.7 (89)

Subject 4

LOW -20,-10	HIGH 10,20 ^b	500 ms		2000 ms		5000 ms	
Set-size 1		3.3 (100)	5.2 (100)	5.1 (88)	2.5 (100)	3.6 (100)	3.6 (67)
Set-size 2		5.7 (78)	5.7 (89)	9.4 (80)	5.7 (78)	6.3 (80)	3.2 (60)
Set-size 4		4.8 (63)	7.3 (75)	3.6 (67)	3.6 (33)	10.9 (71)	6.4 (50)

^a The percentage of times that these stimulus pairs were
Expected to be presented was 5%

^b In pixels

END

2 9-05-95

FIN